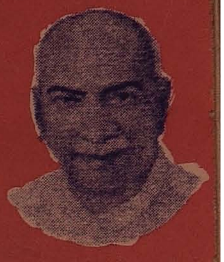
431



MADURAI KAMARAJ UNIVERSITY

(University with Potential for Excellence)

DISTANCE EDUCATION

www. mkudde. org

MBA

SECOND YEAR

COMPUTER NETWORKS

431

DIRECTORATE OF
DISTANCE EDUCATION

MASTER OF BUSINESS ADMINISTRATION
(Distance Learning Programme)

SECOND YEAR

COMPUTER NETWORKS

MADURAI KAMARAJ UNIVERSITY
MADURAI - 625 021. INDIA.

Study Materials Prepared by :

Mr. S. MANIMARAN
Feculty Member
Department of Management Studies,
PSNA College of Engineering & Technology,
Dindigul - 624622.

Anbagam Printers, Madurai - 7
Copies - 3000 / 2006 - 07

MBA/DLP Third Year

COMPUTER NETWORKS

CONTENTS

SYLLABUS

MBA - DLP III YEAR

COMPUTER NETWORKS

I. Introduction goals and application of Network-Network structure and architecture- OSI reference model-various layers-services - Network, standardization- ARPANET, MAP and TOP, USENET, CSNET, BINET, SNA and public networks.

II. The physical layer- Fourier analysts- bandwidth- Limited signals- maximum data rate of a channel - Transmission media - magnetic media - twisted pair - Baseband - and boardband cable- Fibre optics-Line of sight transmission - telephone systems - modern Rs-232-C and RS-419. Medium access sublayer - Local and metropolitan networks - ALDHA protocols - LAN protocols-IEEE standard 802 dor LAN-fibre optic networks.

III. Data link layer -design issues - Error detection and correction - dat a link protocol-network layer-layer design issues - routing algorithms- internet working - examples.

IV. Transport layer- design issues-connection management- simple transport protocol on top of X.25 - samples session layer - design issue - remote procedure call.

V. Presentation layer- design issues- application layer - Design issues.

Text

Computer Networks- A.S. Tanenbaum, Second edition, PHI Private Ltd., New Delhi 1990.

Reference

1. Computer Communication and Network, John freer, Fitman Computer system Series, 1980.

2. Computer Network & Simulation III, Scnemaker, Elacvier Science Publications, 1986.

LESSON – 1

INTRODUCTION

Early computers were being used as "stand-alone" systems in organizations fulfilling their own requirements. With widespread use of computers there was a realization that it would be advantageous in many situations to use computers from remote points. It was also felt that connecting computers together via telecommunication lines will lead to widespread availability of powerful computers. Advances in computer technology also made these interconnections possible. In this subject we will discuss various aspects of communication technology and examine how this technology can be used along with computer technology to provide powerful networks of computers.

Throughout the subject we will use the term "Computer Network" to mean an interconnected collection of autonomous computers. Two computers are said to be interconnected if they are able to exchange the information. Users would prefer to have access to a computer from their place of work or even their homes without having to go to the computer centre. Connecting the users' terminals can provide such access by communication lines to the computer.

Goals and Applications of Network

Before we see the technical issues of the computer network, we will see about why people are interested in computer networks and where it is actually used.

Networks for Companies

Some companies have a substantial number of computers in operation, often located far apart. For example, a company may have more than one branch offices in various places. At that time management may have decided to connect all the computers which are available in various offices to extract and correlate the information about the entire company. The main goals are

1. Resource Sharing
2. High Reliability
3. Saving Money
4. Scalability

Resource Sharing means any one can share all details about the company from anywhere if the computers are connected in a network. The goal is to make all programs, equipment, and especially data available to anyone on the network without regard to the physical location of the resource and the user.

A second goal is to provide high reliability by having alternative sources of supply. If any one of the systems is unavailable due to a hardware failure, then we can use the other copies in that network.

Another goal is saving money. Small computers have a much better price ratio than large ones.

Another networking goal is scalability, the ability to increase system performance gradually as the workload grows just by adding more processors.

A Computer network can provide a powerful communication medium among widely separated employees.

Network for People

We will sketch three of the more exciting ones that are starting to happen.

1. Access to remote information
2. Person-to-person communication
3. Interactive entertainment

Access to Remote Information

Home shopping is becoming popular, with the ability to inspect the on-line catalogs of thousands of companies. People manage their bank accounts electronically. All these type of applications involve interactions between a person and a remote database. Another example for access to remote information is World Wide Web. This contains information about arts, business, science and sports etc.

Person-to-Person Communication

Electronic mail (e-mail) is the best example for this type. Using e-mail, you can transfer audio and video as well as text. This technology makes it possible to have virtual meetings, called videoconference.

Interactive Entertainment

This is a huge and growing industry. The killer application here is video on demand. New films may become interactive, where the user is occasionally prompted for the story direction with alternative scenarios provided for all cases.

Game playing is also one of the entertainment in network. Already we have multiperson real-time simulation games.

Network Structure

We will see the social aspects of networking to the technical issues involved in network design. There are two types of transmission technology.

1. Broadcast networks
2. Point-to-point networks

Broadcast Networks

It has a single communication channel that is shared by all the machines on the network. All the others receive short messages, called packets in certain contexts, sent by any machine. An address field within the packet specifies for whom it is intended.

Broadcast systems generally also allow the possibility of addressing a packet to all destinations by using a special code in the address field. When a packet with this code is transmitted, it is received and processed by every machine on the network. This mode of operation is called broadcasting. Some broadcast systems also support transmission to a subset of the machines, something known as multicasting.

Point-to-Point Networks

It consists of many connections between individual pairs of machines. To go from the source to the destination, a packet on this type of network may have to first visit one or more intermediate machines. Routing algorithms play an important role in point-to-point networks.

Classification of Interconnected Processors by Scale

Interprocessor Distance	*Processors located in same*
0.1 m	Circuit Board
1 m	System
10 m	Hall
100 m	Block
1 km	Campus
10 km	Town
100 km	Country
1,000 km	Continent
10,000 km	Planet

Local Area Network

A Local Area Network, or LAN as it is more widely called, is a group of computers in a localized area. The term-localized area could mean a small room, twenty feet by ten feet, or it could mean a factory spanning several acres. Another definition, widely accepted, states that a LAN is a computer network that is confined to a building or a cluster of buildings. A LAN is a network that is typically personal to an organization and is installed for the exclusive use of a particular office or factory of a given organization. It is not often that you will come across two or more organizations in an office complex sharing a LAN.

LANs are privately owned networks within a single building or campus of up to a few kilometers in size. They are widely used to connect personal computers and workstations in company offices and factories to share resources (e.g. printers) and exchange information.

Traditional LANs run at speeds of 10 to 100 Mbps, have low delay (tens of microseconds), and make very few errors.

A LAN is a network allowing easy access to other computers or peripherals. The typical characteristics of a LAN are,

- Physically limited (<2km)
- High bandwidth (>1mbps)
- Inexpensive cable media (coax or twisted pair)
- Data and hardware sharing between users
- Owned by the user

Benefits of LAN

1. Resource Sharing

Using a LAN, expensive resources like laser printers, modems, graphic devices and data storage units can be shared. This enables several users to access these resources at the same time. Software and programs can be stored at a common location where every user who has the need can access them.

2. Communication

Another use of LAN is that it can help you make the computer do the job of an office intercom. You can use the computer to flash messages on the screen of other computers in this office. This would save employee the time they would spend in going to someone on another department, or some other floor to deliver a message or a memo.

3. Security

The PC is capable of storing a fair amount of information, but it is not a very secure place to store data. Using a LAN, users can store their files on a computer that is a part of the LAN. A LAN has built-in security features such that it would be virtually impossible for anyone to get hold of these files.

Some typical hardware components of a LAN are :

- Workstations
- Server
- Network Interface Unit
- Communication Channel

Advantages of LAN

- Local area networks are the best means to provide a cost effective multi-user computer environment
- A LAN can fit any site requirements.
- Any number of users can be accommodated.
- It offers sharing of peripherals.
- It is flexible and growth oriented.
- It offers existing single users a familiar Disk Operating System (DOS) environment.
- It can minimize adverse effect of loss of any one system.
- Increase system performance through distribution of tasks and equipment.

Metropolitan Area Network

A metropolitan area network, or MAN is basically a bigger version of a LAN and normally uses similar technology. It might cover a group of nearby corporate offices or a city and might be either private or public. A MAN can support both data and voice, and might even be related to the local cable television network. A MAN just has one or two cables and does not contain switching elements, which shunt packets over one of several potential output lines.

A key aspect of a MAN is that there is a broadcast medium to which all the computers are attached. This greatly simplifies the design compared to other kinds of networks.

Wide Area Network

Wide Area Networks are the set of connecting links between local area networks. These links are made over telephone lines leased from the various telephone companies. In rare instances, WANs can be created with satellite links, packet radio, or microwave transceivers. These options are generally far more expensive than leased telephone line, but they can operate in areas where leased lines are not available.

A wide area network, or WAN, spans a large geographical area, often a country or continent. It contains a collection of machines intended for running user programs. These machines are called as hosts. The hosts are connected by a communication subnet, or just subnet for short.

Most WANs are private and owned by the business that operates with them. Recently, however, the Internet has emerged as both the largest and the least expensive WAN through encrypted communications over the Internet.

In most wide area networks, the subnet consists of two distinct components: transmission lines and switching elements. Transmission lines (also called circuits, channels, or trunks) move bits between machines. The switching elements are specialized computers used to connect two or more transmission lines.

Internetworks

A collection of interconnected networks is called an internetwork or just internet. A common form of internet is a collection of LANs connected by a WAN. An internetwork is formed when distinct networks are connected together. In our view, connecting a LAN and a WAN or connecting two LANs forms an internetwork.

Computer Network Topologies

When computers at different locations are to be interconnected one may do it in a number of ways. For example, if five computers A, B, C, D, E are to be interconnected we may do it as shown in Fig. 1.1. In this case there are physical links between A-C, A-E, D-C, B-E and B-D. Assuming full duplex links, A can communicate with C and E, B with E and D, C with A and D, D with B and C, and E with A and B. Direct communication between A and B and A and D is not possible. If, however, C can route a message from A to D then there would be a logical connection between A and D. Similarly E can communicate with D via B and C with B via D. Each computer in the network will be called a node.

This interconnection pattern (Fig. 1.1) is known as a ring network and AEBDC form a ring.

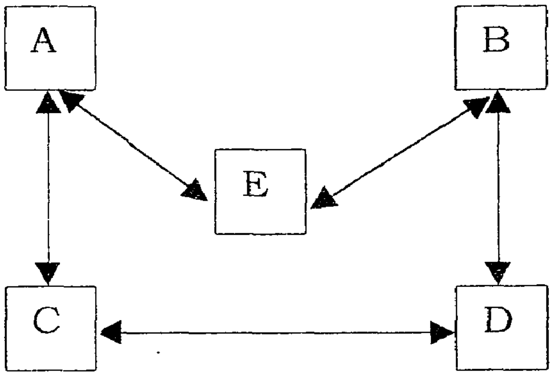


Fig. 1.1 A Ring Connection of Computers

Two other interconnection patterns are shown in Figs. 1.2 and 1.3. The pattern of Fig. 1.2 is called a star network and that of Fig. 1.3 a fully interconnected network. Different patterns of interconnections are known as network topologies.

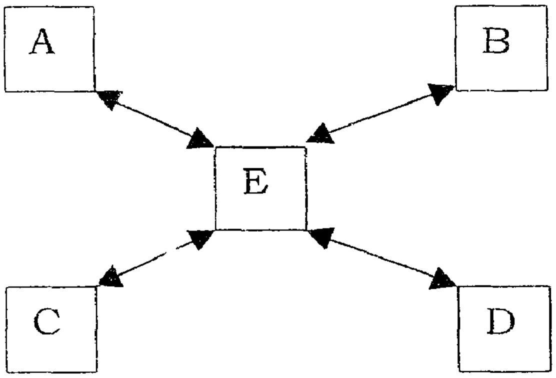


Fig. 1.2 A Star Connection of Computers

The main considerations in selecting a particular topology are:

(i) The availability and cost of physical communication lines between nodes and line bandwith.

(ii) The capability of a node to route information to other nodes.

(iii) Delays due to routing of information.

(iv) Reliability of communication between nodes when there is a breakdown of a line or a node.

(v) Strategy of controlling communication between nodes in the network – centralized or distributed.

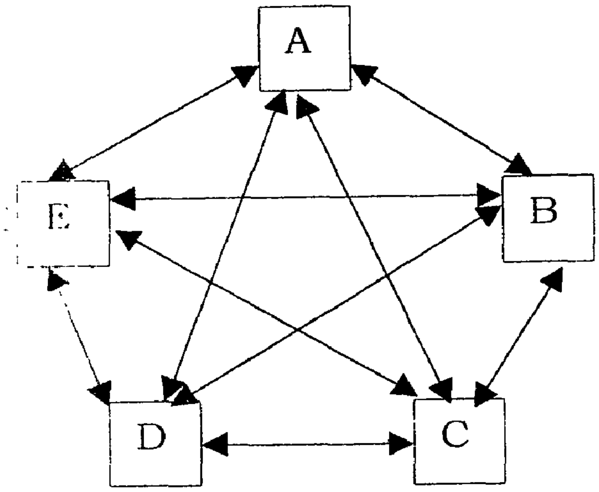


Fig. 1.3 A Fully Interconnected Network

The fully connected topology of Fig. 1.3 has a separate physical connection for connecting each node to any other node. It is the most expensive system from the point of view of line costs, as there are 10 separate point-to-point lines. It is, however, very reliable as any line breakdown will affect only communication between the connected machines. Each node need not have individual routing capabil'ty. Communication is very fast between any two nodes. The control is distributed, with each computer deciding its communication priorities.

The star topology of Fig. 1.2 has minimum line cost as only 4 lines are used. The routing function is performed by E which centrally controls communication between any two nodes by establishing a logical path between them. Thus if A wants to communicate with D,E would receive this request from A and set up the logical path A-E-D based on line availability. Delays would not increase when new nodes are added as any two nodes may be connected via two links only. The system, however, crucially depends on E. If E breaks down the whole network would break down.

The ring topology of Fig. 1.1 is not centrally controlled. Each node must have simple communication capability. A node will receive data from one of its two neighbours. The only decision the node has to take is whether the data is for its use or not. If the data is not addressed to it, it merely passes it on to its other neighbour. Thus if E receives data from B (see Fig. 1.1) it examines whether it is addressed to itself. If it is, then it uses the data, else it passes the data to A.

The main disadvantage of a ring is larger communication delays if the number of nodes increases. It is, however, more reliable than a star network because communication is not dependent on a single computer. If a line between any two computers breaks down, or if one of the computers breaks down, alternate routing is possible.

One may use hybrid approach to interconnection. In other words, the interconnections may not be a pure star, loop or full interconnection. The physical links may be set up based on the criteria specified at the beginning of the section to have an optimal communication capability for the specified network functions.

Another interconnection method is a multipoint or multidrop linkage of computers shown in Fig. 1.4. The main advantage of this method is the reduction in physical lines. One line is shared by all nodes. If computer A wants to communicate with E then it first checks whether the communication line is free. When the line becomes free it transmits the message addressed to E on it. As the message travels on the line, each computer checks whether it is addressed to it. In this case when E finds its "address" in the message it accepts it, sends an acknowledgement to A and frees the line. Thus each computer connected to the line ~~must~~ have good communication and decision making capability.

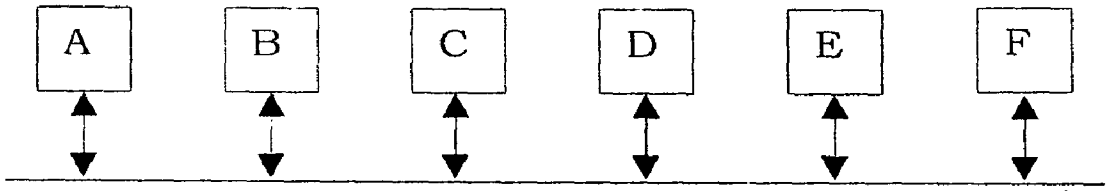


Fig. 1.4 A multidrop configuration

An alternate approach which can free each machine of this task is to have one master computer overseeing communications on the line. The master would receive all messages and route them to appropriate machine. This approach would however create a bottleneck when computers connected to the link increase and consequently the master computer becomes too busy.

The method whereby each computer in a multidrop configuration places a message with the source and destinations addresses, to be picked up by the addresses, is known as a broadcast scheme. This method is appropriate for use in a local area network where a high speed communication channel is used and computers are confined to a small area. This method is also appropriate when satellite communication is used as one satellite channel may be shared by many computers at a number of geographical locations. In this method it is easy to add new computers to the network. The reliability of the network will be high with distributed control because the failure of a computer in the network will not affect the network functioning for other computers.

Network Architecture

To reduce their design complexity, most networks are organized as a series of layers or levels, each one built upon the one below it. The number of layers, the name of each layer, the contents of each layer, and the function of each layer differ from network to network.

Layer n on one machine carries on a conversation with layer n on another machine. The rules and conventions used in this conversation are collectively known as the layer n protocol. A five-layer network is illustrated in Fig 1.5.

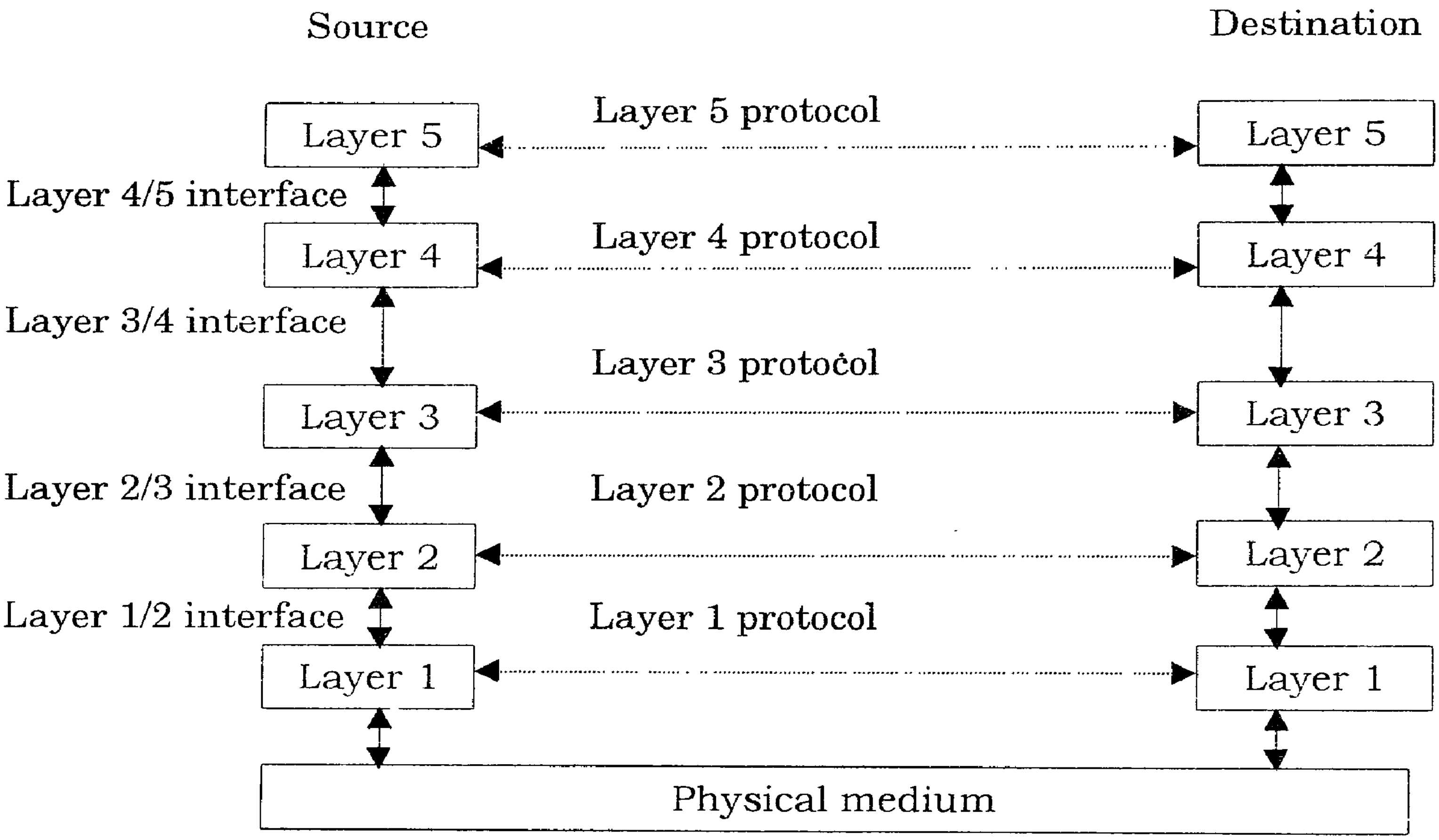


Fig 1.5 Five-layer Network

The active elements in each layer are often called entities.

The entities comprising the corresponding layers on different machines are called peers.

Between each pair of adjacent layers there is an interface.

A set of layers and protocols is called a network architecture.

A list of protocols used by a certain system, one protocol per layer, is called a protocol stack.

Some of the key design issues that occur in computer networking are present in several layers. Every layer needs a mechanism for identifying senders and receivers. Since a network normally has many computers, some of which have multiple processes, a means is needed for a process on one machine to specify with whom it wants to talk.

In some systems, data travel in one direction (simplex communication). In others they can travel in either direction, but not simultaneously (half – duplex communication). In still others they travel in both directions at once (full – duplex communication).

Error control is an important issue because physical communication circuits are not perfect. An issue that occurs at every level is how to keep a fast sender from swamping a slow receiver with data. One of the solution for this is feedback from the receiver to the sender.

Connection-Oriented and Connectionless Services

Connection-oriented service is modelled after the telephone system. To talk to someone, you pick up the phone, dial the number, talk, and then hang up. Similarly, to use a connection-oriented service, the service user first establishes the connection, uses the connection, and then releases the connection.

Connectionless service is modelled after the postal system. Each message (letter) carries the full destination address, and each one is routed through the system independent of all the others. Normally, when two messages are sent to the same destination, the first one sent will be the first one to arrive. However, it is possible that the first one sent can be delayed so that the second one arrives first. With a connection-oriented service this is impossible.

Lesson - 2

THE OSI REFERENCE MODEL

It is called as ISO OSI (Open System Interconnection) reference model - because it deals with connecting open systems - i.e., systems that are open for communication with other system.

* It has seven layers

# Principles that are applied to arrive at 7 layers:-

1. A layer should be created where a different level of abstraction is needed.

2. Each layer should perform a well defined function.

3. The function of each layer should be chosen in such a way that it satisfies the internationally standardized protocols.

4. The layer boundaries should be chosen to minimize the information flow across the interface.

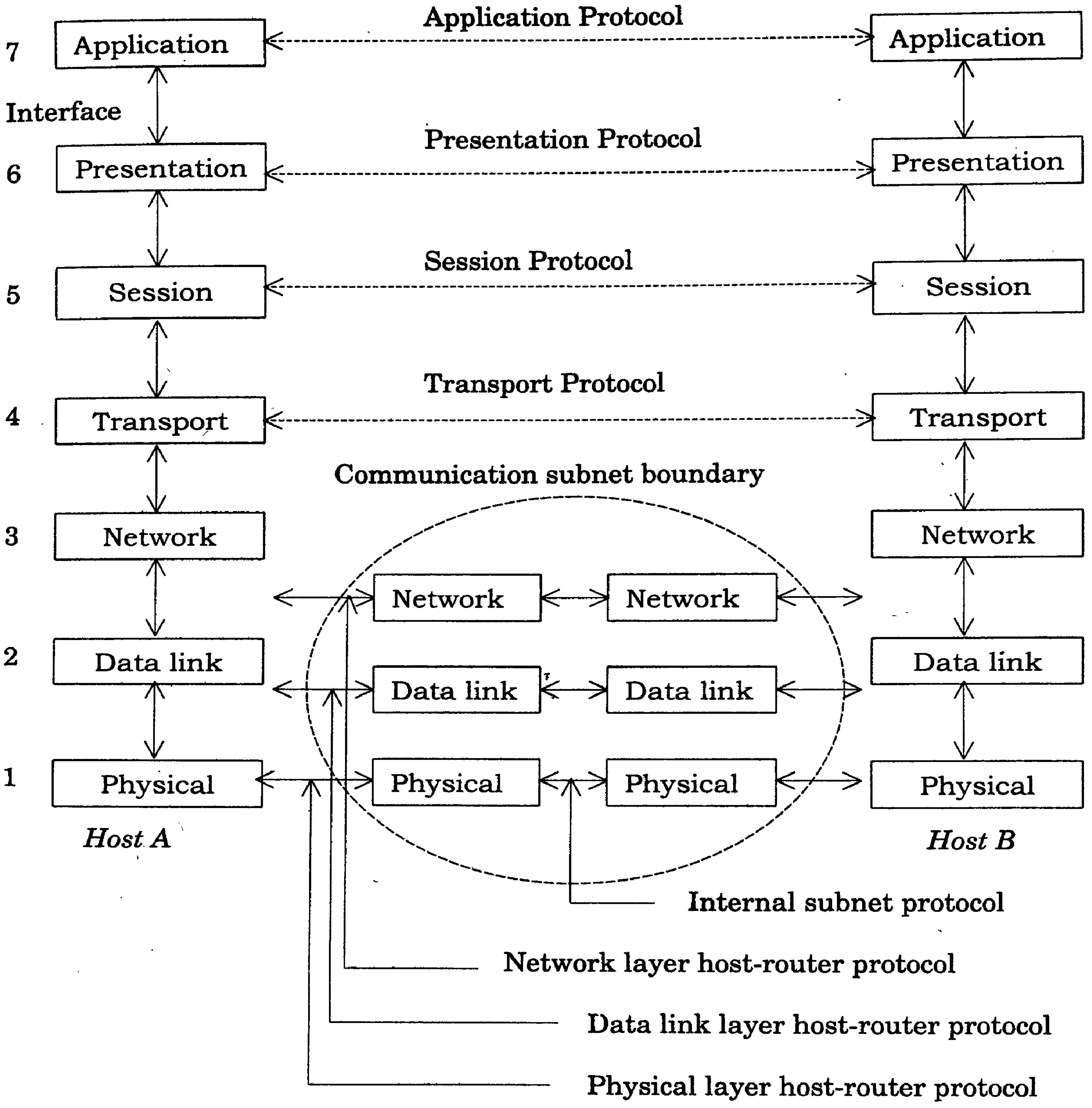
5. The number of layers should be large - so that distinct functions are not thrown together & small enough - so that the architecture - is not wide.

The Physical Layer

❖ It is concerned with Transmitting "raw bits" over a communication channel. The design issues has to do with making sure that when 1 is *transmitted* on the side, it is *received* as 1 and not 0 by the other side.

Typical questions here are

❖ How many volts should be used to represent 1 & 0.
❖ How many microseconds a bit lasts.
❖ Whether transmission takes place in both directions.
❖ How the initial connection is established.
❖ How many pins the n/w connector has & what each pin is used for.
❖ It deals with mechanical, electrical, & procedural interfaces and physical transmission medium.



The Data Link Layer

- The main task is to take a raw transmission facility & transform it into a line that appears free of transmission errors.

- It carries out the task by breaking the input data in to "data frames" (Few hundred bytes).

- It transmits the frame sequentially & process the "acknowledgement frames"

- Since physical layer just puts in the data, the data link layer creates & recognize frame boundaries - It is achieved by adding bit patterns at beginning & end of frame.

- Data - care has to be taken to avoid confusion.

- Noise burst - destroys the frame. The software of Data link layer on source machine retransmits it.

- Duplicate frames are used when acknowledgement signal is not received.

- It is the layer which has to take care about the damaged, lost & duplicate frames.

- To keep transmitter from drowning a slow receiver, some traffic regulation mechanism is applied which tells how much buffer space the receiver has at the moment.

- Flow regulation & error handling - integrated.

- This is handled by a solution called - piggy backing.

The Network Layer

- Concerned with controlling the operation of the subnet.

- Key design issue is to determine how the packets are routed from source to destination.

- It can be based on static tables or dynamic & it is determined at the start of each conversation.

- Also the problems in routing should be determined or solved by this network layer.

  For eg. If too many packets are present at the same time - they form bottlenecks.

- Accounting function are built to reward the operator.

- The software should count how many packets, or characters or bits sent by each customer, to produce billing information.

- Difficulty - when packets crosses a national border with different rates on both sides - complicated.

- At the time of sending the information, the following problems may arise.

  1) The addressing used by second network may differ from first.

  2) The second network may not accept the packet because it is too large.

- The network overcomes all the problems & this layer is responsible for heterogeneous networks to be interconnected.

- In broadcast network, the routing algorithm is simple so the layer is often thin or non existent.

The Transport Layer

- The basic function is to accept data from session layer, split up into smaller units and pass it to network layer.

- Transport layer creates a distinct network connection for each transport connection required by the session layer.

- It create multiple network connections, dividing the data among the N/W connections to improve throughout.

- It is required to make the multiplexing transparent to the session layer.

- The transport layer is a true source - to destination or end-to-end layer. A program on source machine on with similar program on destination machine using the message header, control messages.

The Session Layer

- It allows users on different machines to establish sessions between them.

- It uses ordinary data transport.

- A session allows a user to log in to remote-time sharing system or to transfer a file between 2 machines.

- A related session service is token management. The session layer provides tokens that can be exchanged.

- Another session service is synchronization.

The Presentation Layer

- It is concerned with the syntax & semantics of the information transmitted. i.e. user program do not exchange binary bit strings.

  It exchange - names, dates, amounts of money, & invoices as character strings, in layers Floating point, numbers & data structures.

- The presentation layer manages the abstract data structures and converts them the representation used inside the computer to the network standard representation and back.

The Application Layer

- It contains a variety of protocols.

- All the virtual terminal software is in the application layer.

- Another function is file transfer.

- Transferring a file between systems requires handling these and other incompatibilities.

- It is used for electronic mail, remote job entry and special purpose facilities.

Network Standardization

Many network vendors and suppliers exist, each with their own ideas of how things should be done. Without coordination, there would be complete chaos, and users would be able to get nothing done. The only way out is to agree upon some network standards.

Standards fall into two categories : de facto (Latin form "from the fact") standards are those that have just happened, without any formal plan. De jure (Latin form"by law")standards, in contrast, are formal, legal standards adopted by some authorized standardization body. In the area of computer network standards, there are several organizations of each type, which are discussed below.

Who's Who in the Telecommunications World

Companies in the United States that provide communication services to the public are called common carriers. Their offerings and prices are described by a document called a tariff, which must be approved by the Federal Communications Commission for the interstate and international traffic, and by the state public utilities commissions for intrastate traffic.

In some cases the telecommunication authority is a nationalized company, and in others it is simply a branch of the government, usually known as the PTT (Post, Telegraph & Telephone administration).

ITU's (International Telecommunication Union) job was standardizing international telecommunications, which in those days meant telegraphy. In 1947, ITU became an agency of the United Nations.

ITU has Three Main Sectors

1. Radio communications Sector (ITU-R)

2. Telecommunications standardization Sector (ITU-T)

3. Development Sector (ITU-D)

ITU has Five Classes of Members

1. Administrations (national PTTs)

2. Recognized private operators (e.g. AT&T, MCI, British Telecom)

3. Regional telecommunications organizations (e.g. the European ETSI)

4. Telecommunications vendors and scientific organizations.

5. Other interested organizations (e.g. banking and airline networks)

Who's Who in the International Standards World

International standards are produced by ISO (International Standards Organization) a voluntary, nontreaty organization founded in 1946. ISO issues standards on a vast number of subjects, ranging from nuts and bolts (literally) to telephone pole coatings. Over 5000 standards have been issued, including the OSI standards.

The U.S. representative in ISO is ANSI (American National Standards Institute) which despite its name, is a private, nongovernmental, non-profit organization. ANSI standards are frequently adopted by ISO as international standards.

NIST (National Institute of Standards and Technology) is an agency of the U.S. Dept. of Commerce. It was formerly known as the National Bureau of Standards.

Another major player in the standards world is IEEE (Institute of Electrical and Electronics Engineers), the largest professional organization in the world. IEEE has a standardization group that develops standards in the area of electrical engineering and computing. IEEE's 802 standard for local area networks is the key standard for LANs.

Who's Who in the Internet Standards World

The worldwide Internet has its own standardization mechanisms, very different from those of ITU-T and ISO. The difference can be crudely summed up by saying that the people who come to ITU or ISO standardization meetings wear suits. The people who come to Internet standardization meetings wear either jeans or military uniforms.

When the ARPANET was set up, DoD created an informal committee to oversee it. In 1983, the committee was renamed the IAB (Internet Activities Board) and given a slighter broader mission, namely, to keep the researchers involved with the ARPANET and Internet pointed more-or-less in the same direction, an activity not unlike herding cats. The meaning of the acronym "IAB" was later changed to Internet Architecture Board.

ARPANET

Let us now switch gears from LANs to WANs. In the mid-1960s, at the height of the Cold War, the DoD wanted a command and control network that could survive a nuclear war. Traditional circuit-switched telephone networks were considered too vulnerable, since the loss of one line or switch would certainly terminate all conversations using them and might even partition the network. To solve this problem, DoD turned to its research arm, ARPA (later DARPA, now ARPA again), the (periodically Defense) Advanced Research Projects Agency.

ARPA was created in response to the Soviet Union's launching Sputnik in 1957 and had the mission of advancing technology that might be useful to the military. ARPA had no scientists or laboratories, in fact, it had nothing more than an office and a small (by Pentagon standards) budget. It did its work by issuing grants and contracts to universities and companies whose ideas looked promising to it.

Several early grants went to universities for investigating the then-radical idea of packet switching, something that had been suggested by Paul Baram in a series of RAND Corporation reports published in the early 1960s. After some discussions with various experts, ARPA decided that the network the DoD needed should be a packet-switched network, consisting of a subnet and host computers.

The subnet would consist of minicomputers called IMPs (Interface Message Processors) connected by transmission lines. For high reliability, each IMP would be connected to at least two other IMPs. The subnet was to be a datagram subnet, so if some lines and IMPs were destroyed, messages could be automatically rerouted along alternative paths.

Each node of the network was to consist of an IMP and a host, in the same room, connected by a short wire. A host could send messages of up to 8063 bits to its IMP, which would then break these up into packets of at most 1008 bits and forward them independently towards the destination. Each packet was received in its entirety before being forwarded, so the subnet was the first electronic store-and-forward packet-switching network.

In addition to helping the fledging ARPANET grow, ARPA also funded research on satellite networks and mobile packet radio networks. In one famous demonstration, a truck driving around in California used the packet radio network to send messages to SRI, which were then forwarded over the ARPANET to the East Coast, where they were shipped to University College in London over the satellite network. This allowed a researcher in the truck to use a computer in London while driving around in California.

This experiment also demonstrated that the existing ARPANET protocols were not suitable for running over multiple networks. This observation led to more research on protocols, culminating with the invention of the TCP/IP was specifically designed to handle communication over internetworks, something becoming increasingly important as more and more networks were being hooked up to the ARPANET.

Lesson - 3

PHYSICAL LAYER

Theoretical Basics for Data Communication

Normally information are transmitted from one place to another place on wires carrying some physical property such as voltage or current. By representing the value of this voltage or current as a single-valued function of time, f(t), the behaviour of the function can be modeled and analyze it mathematically.

Fourier Analysis

$$G(t) = \tfrac{1}{2} C + \sum_{n=1}^{\alpha} a_n \sin (2\pi nft) + \sum_{n=1}^{\alpha} b_n \cos (2\pi nft)$$

Where $f = 1/T$ is the fundamental frequency and a_n and b_n are the sine and cosine amplitudes of the n^{th} harmonics. Such decomposition is called a Fourier series.

The a_n amplitudes can be computed for any given g(t) by multiplying both sides by sin (2 kπft) and integrating from O to T.

$$\int_0^T \sin (2\,k\pi ft) \sin (2\,\pi nft)\,dt = \begin{cases} 0 \text{ for } k \neq n \\ T/2 \text{ for } k = n \end{cases}$$

$$a_n = 2/T \int_0^T g(t) \sin(2\pi nft)dt, \quad b_n = 2/T \int_0^T g(t) \cos(2\pi nft)dt, \quad c = 2/T \int_0^T g(t)dt$$

Bandwidth Limited Signals

The Fourier analysis of this signal yields the coefficients:

$$a_n = 1/\pi n\,[\cos (n\pi/4) - \cos (\beta n\,\pi/4) + \cos(6h\pi/4) - \cos (7n\pi/4)]$$

$$b_n = 1/\pi n\,[\sin (3n\pi/4) - \sin(n\pi/4) + \sin(7n\pi/4) - \sin (6n\pi/4)]$$

$$c = 3/8$$

The time T required to transmit the character depends on both the encoding method and the signaling speed. The number of changes per second is measured in baud.

Maximum Data Rate of a Channel

Nyquist proved that if an arbitrary signal has been run through a low - pass filter of bandwidth H, the filtered signal can be completely reconstructed by making only 2H(exact) samples per second - Nyquist's theorem states:

Maximum data rate = $2 H \log_2 V$ bits/sec.

The amount of thermal noise present is measured by the ratio of the signal power to the noise power, called the signal to noise ratio.

Shannon's major result is that the maximum data rate of a noisy channel whose bandwidth is 'H' H_z, and whose signal-to-noise ratio is S/N, is given by

Maximum number of bits/sec = $H \log_2 (1 + S/N)$

Transmission Media

A raw bit from one machine to another machine is transmitted through the transmission media in the physical layer. The transmission media is divided into two parts namely guided media and unguided media. The copper wire and fiber optics are examples for guided media. The radio and lasers are the examples for unguided media.

Magnetic Media

For the data communication, the most common way is to copy the data into floppy disk or magnetic tapes, copy them back onto the destination machine, which can be read, whenever necessary. Although it takes the sophistication to be used for complex communications, it has the advantage of being cost effective.

Twisted Pair

A twisted pair consists of two insulated wires, which are twisted in a helical form, in order to reduce electrical interference. A common example is the telephone system. Repeaters are needed for transmission over long distances. Twisted pairs can be used for both analog and digital transmission.

Twisted pair comes in various categories: Category 3 twisted pair consist of two insulated wires twisted together. Four such pair are grouped together.

Category 5 is similar to category 3, but has more twists, which results in less cross talk and better quality.

Baseband Coaxial Cable

The next media is base band coaxial cable. Coaxial cable has better shielding than twisted pair. It consists of a copper wire as the core, surrounded by an insulating material. The insulator is surrounded by a cylindrical conductor, which is encased by a protective plastic sheath.

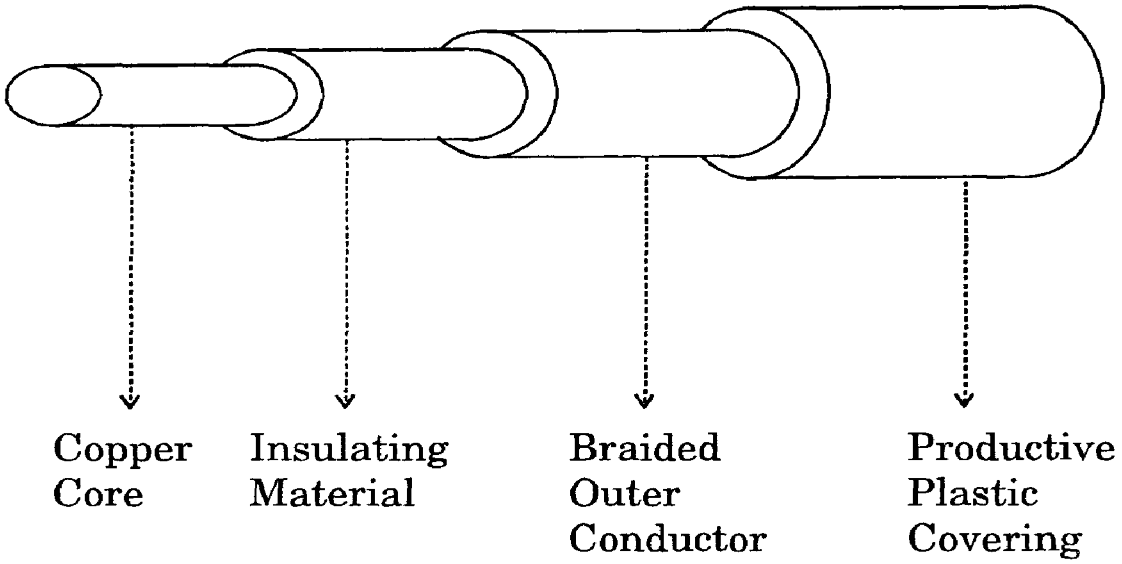


Fig. 3.1

Broadband Coaxial Cable:

We can use the broad band cable for the analog transmission. Broad band systems are divided into multiple channels, each of which can be used for analog television, co-quality audio, which can be mixed on the cable.

Broad band systems are divided into two types: Single cable & Dual cable. Dual cable systems have two identical cables. During the data transmission, the computer outputs the data to the head - end of the cable tree. The head-end transfers the signal to the cable for transmission back down the tree.

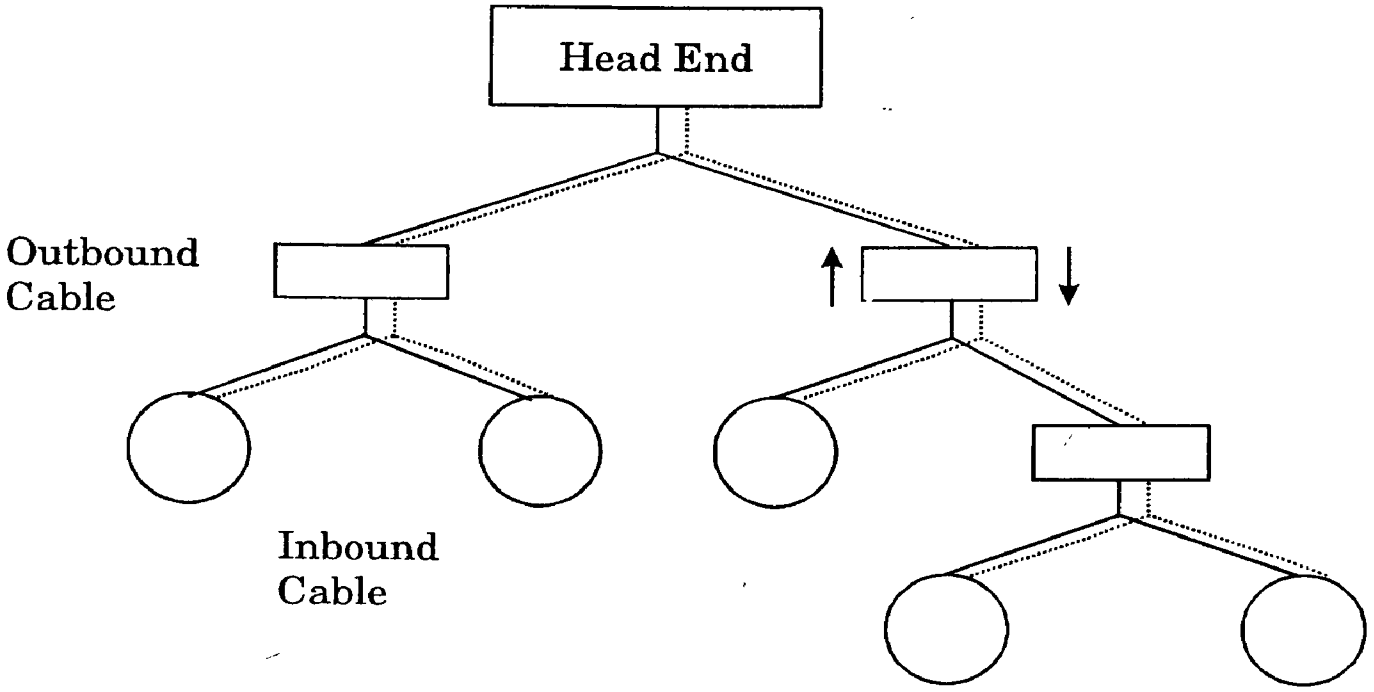
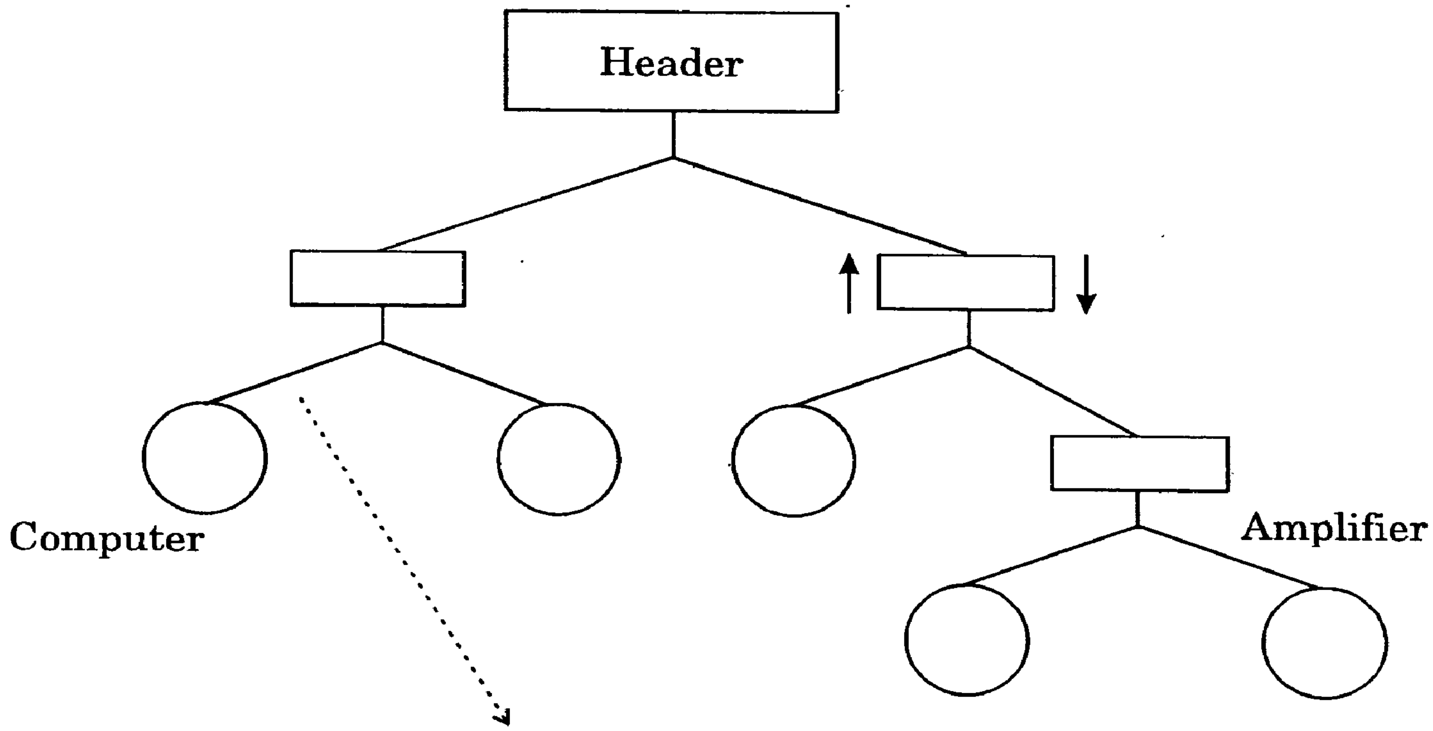


Fig. 3.2

The other type of broad band coaxial cable is the single cable. In the subsplit system, frequencies from 5 to 30Mhz are used for inbound traffic, and frequencies from 40 to 300Mhz are used for outbound traffic.

In the mid-split system, the inbound band is 5 to 116 Hz and the outbound band is 168Mhz to 300Mhz.



Single Cable Low frequencies for inbound traffic,
High frequencies for outbound traffic.

Fig. 3.3

Fiber Optics

Fiber optics is one of the most common guided media. An optical transmission system has three components: the light source, the transmission medium, and the detector. A pulse of light indicates a 1 bit and the absence of light indicates a zero bit. The transmission medium is an ultra - thin fiber of glass. The detector generates an electrical pulse when light falls on it. By attaching a light source to one end of an optical fiber and a detector to the other, we have unidirectional data transmission system.

Transmission of Light Through Fiber

Light pulses sent down a fiber spread out in length as they propagate. This spreading is called dispersion. One way to keep these spreads - out pulses from overlapping is to increase the distance between them, but reducing the signaling rate can only do this. By making the pulses in a special shape related to the reciprocal of the hyperbolic cosine, all the dispersion effects council out, and it may be possible to send pulses for thousands of kilometers without appreciable shape distortion. These pulses are called solutions.

Fiber Cables

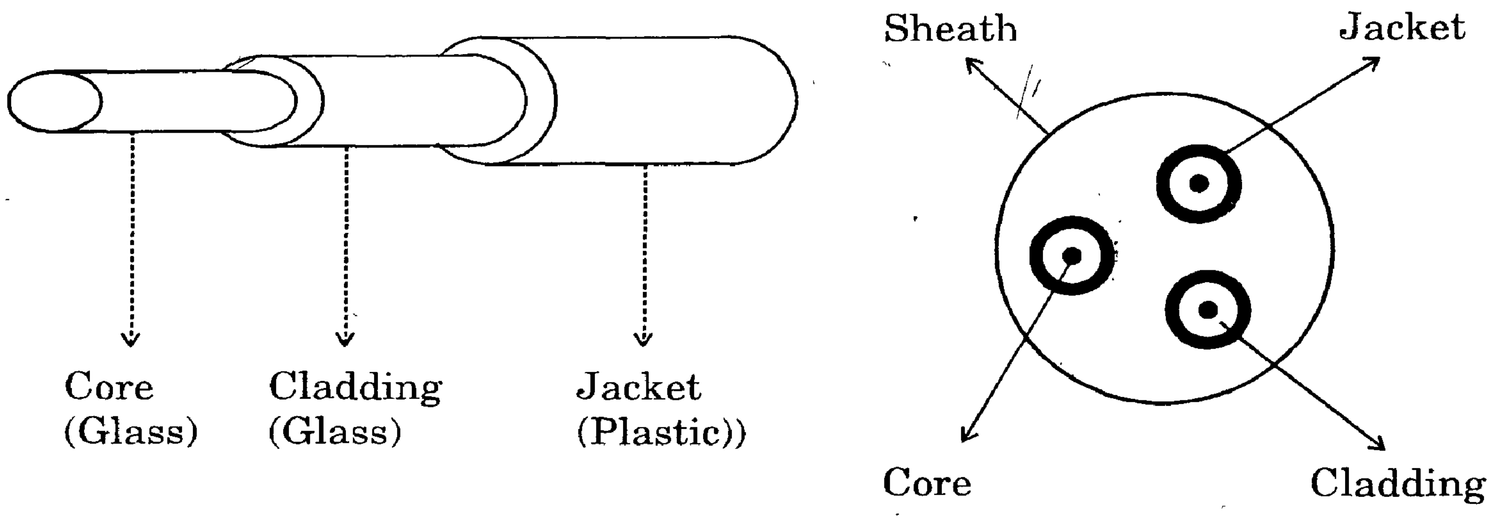


Fig. 3.4

At the centre is the glass core through which the light propagates. In multimate fibers, the core is 50 microns in diameter. The core is surrounded by a glass cladding with a lower index of refraction than the core, to keep all the light in the core, next comes a thin plastic jacket to protect the cladding. Fibers are typically grouped together in bundles protected by an outer sheath.

Fibers can be connected in three different ways:

(i) They can terminate in connectors and be plugged into fiber sockets.

(ii) They can be spliced mechanically. Mechanical splices just lay the two carefully cut ends next to each other in a special leave and clamp them in place.

(iii) Two pieces of fiber can be fused (melted) to form a solid connection.

For all three kinds of splices, reflections can occur at the point of the splice, and the reflected energy can interfere with the signal.

Comparison of Semiconductor Diodes and LEDs as Light Sources

Item	Led	Semiconductor Laser
Data rate	Low	High
Mode	Multimode	Multimode or single mode
Distance	Short	Long
Lifetime	Long life	Short life
Temperature Sensitivity	Minor	Substantial
Cost	Low Cost	Expensive

Fiber Optic Networks

In fiber optic networks, two types of interfaces are used. A passive interface consists of two taps fused onto the main fiber. One tap has an LED or laser diode at the end of it (for transmitting), and the other has a photodiode (for receiving). The top itself is passive.

The other type of interface is the repeater. The incoming light is converted to an electrical signal, regenerated to full strength if it has been weakened, and retransmitted as light.

If an active repeater fails, the ring is broken and the network goes down on the other hand, since the signal is regenerated at each interface, the individual computer-to-computer links can be kilometer long, with virtually no limit on the total size of the ring.

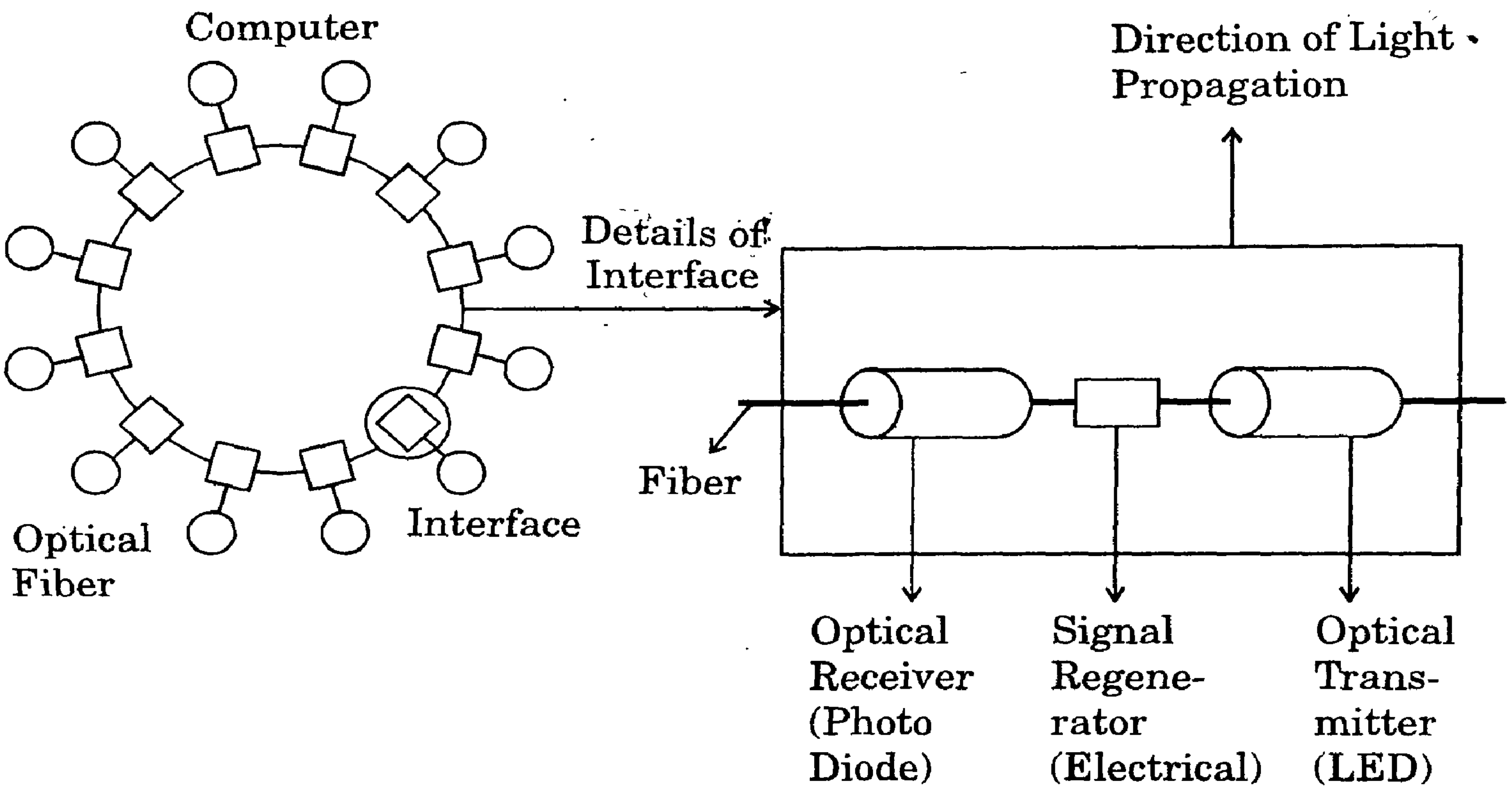

Computer
Direction of Light Propagation
Details of Interface
Fiber
Optical Fiber
Interface
Optical Receiver (Photo Diode)
Signal Regenerator (Electrical)
Optical Transmitter (LED)

Fig. 3.5

Comparison of Fiber Optics and Copper Wire

i) Fiber can handle much bandwidths than copper.

ii) Fiber requires lower installation cost.

iii) Fiber is thin and is of lightweight.

iv) Fibers do not leak light and are difficult to top.

v) Fiber is an unfamiliar technology requiring skills most engineers do not have.

vi) Fiber interface costs more than electrical interface.

Wireless Transmission

Electrical Spectrum

When electrons, more, they create electromagnetic waves that can propagate through free space (even in a vacuum). The number of oscillations per second of an electromagnetic wave is called its frequency, f, and is measured in Hz. The distance between two consecutive maxima is called the wavelength, λ. The speed of light is represented by C. The relationship between them is given by

$$\lambda f = C$$

If we solve for 'f' and differentiate w.r.t. 'λ' we get

$$df/d\lambda = c/\lambda^2$$

If we look at absolute values, we get

$$\Delta f = c\Delta\lambda/\lambda^2$$

Narrow frequency band are commonly used to get the best reception. However, in some cases, the transmitter hops from frequency to frequency in a regular pattern or the transmissions are intentionally spread over a wide frequency band. This technique is called spread spectrum. The other true spread spectrum is called direct sequence spread spectrum.

Radio Transmission

Radio waves generation are more easy. It can travel long distances, and penetrate buildings easily, so that they are widely used for communication, both indoors and outdoors. Radio waves are Omni directional, i.e., they travel in all directions from the source, so that the transmitter and receiver do not have to be carefully aligned physically.

The properties of radio waves are frequency dependent. At low frequencies, radio waves pass through obstacles well, but the power falls off sharply with distance from the source. At high frequencies, radio waves tend to travel in straight lines.

Microwave Transmission

Since the microwaves travel in a straight line, if the towers are too far apart, the earth will get in the way consequently, repeaters are needed periodically.

Unlike radio waves at lower frequencies, microwaves do not pass through buildings well. Some waves may take slightly longer to arrive than direct waves. This effect is called multipath fading.

Microwave communication is widely used for long - distance telephone communication, cellular telephones, television distribution. Microwave is also relatively inexpensive.

Infrared and Millimeter Waves

The unguided infrared and millimeter waves are used for short-range communication. The remote controls used on televisions, VCRs and stereos use infrared communication. They are relatively directional, cheap and easy to build, but they do not pass through solid objects. The infrared waves are used for indoor LANs, but they cannot be used for outdoors.

Light Wave Transmission

For many years the unguided optical signaling has been in use. An application is to connect the LANs in two buildings via lasers mounted on their rooftops. Coherent optical signaling using lasers is inherently unidirectional, so each building needs its own laser and its own photo detector. This scheme offers very high bandwidth and very low cost, It is also relatively easy to install. A disadvantage is that the laser beams cannot penetrate rain or thick fog, but they normally work well on sunny days.

Telephone System

Structure of the Telephone System

At present, the telephone system is organized as a highly redundant, multilevel hierarchy. Each telephone has two copper wires coming out of it that go directly to the telephone company's nearest end office (also called a local central office). The two-wire connections between each subscriber's telephone and the end office are known as the local loop.

Each end office has a number of outgoing lines to one or more nearby switching centres, called toll offices (Tandem offices). These lines are called toll connecting trunks. The toll, primary sectional and regional exchanges communicate with each other via high bandwidth inter toll trunks (also called inter office trunks).

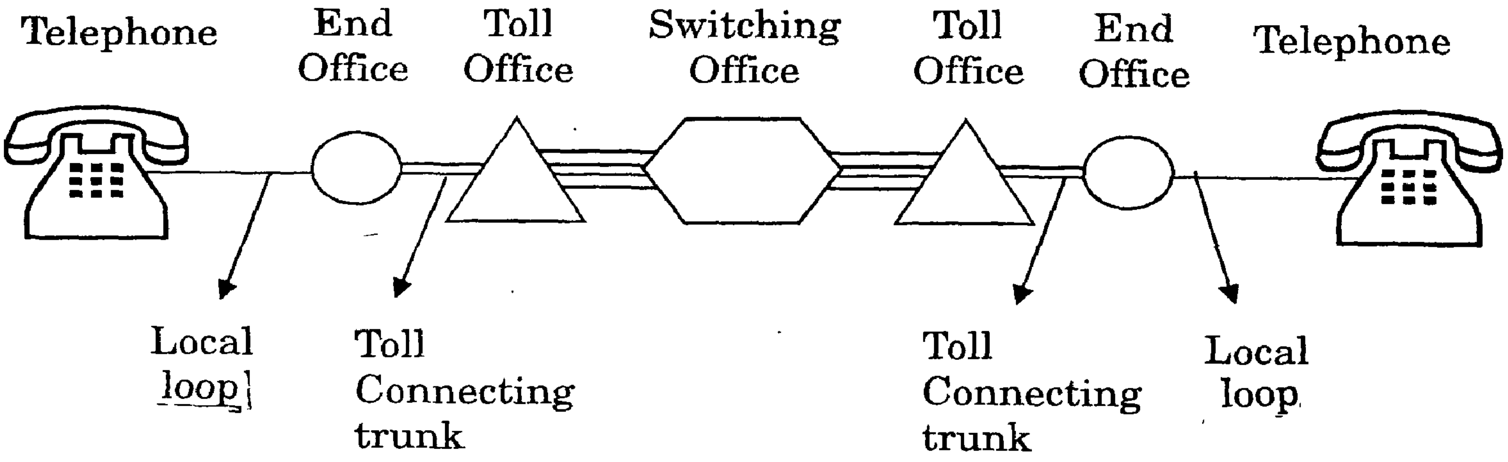


Fig. 3.6

The Telephone System consists of three major components:

1. Local loops (twisted pairs, analog signaling)
2. Trunks (fiber optics or microwave, mostly digital)
3. Switching offices.

Local Loop

The telephone system was based entirely on analog signaling. While the long distance trunks are now barley digital in the more advanced countries, the local loops are still analog. When a computer wishes to send digital data over a dial-up line, the data must first be converted to analog form, then to digital form and back to analog form.

Transmission Impairments

Transmission times suffer from three major problems:

- Attenuation is the loss of energy as the signal propagates outward.
- Delay distortion is caused by the fact that different Fourier components travel at different speeds.
- Noise is an unwanted energy from sources other than the transmitter.

The three types of communication are:

- *Simplex Transmission :* Transmission is in only one direction.
- *Half Duplex Communication :* Communication can go either a way but only one at a time.
- *Full Duplex Communication :* Communication is in both the directions.

Modem

Due to the problems just discussed, especially the fact that both attenuation and propagation speed are frequency dependent, it is undesirable to have a wide range of frequencies in the signal. Unfortunately, square waves, as in digital data, have a wide spectrum and thus are subject to strong attenuation and delay distortion. These effects make baseband (DC) signaling unsuitable except at slow speeds and over short distances.

To get around the problems associated with DC signaling, especially telephone lines, AC signaling is used. A continuous tone in the 1000-2000Hz range, called a sine wave carrier is introduced. Its amplitude, frequency, or phase can be modulated to transmit intermission. In amplitude modulation, two different voltage levels are used to represent 0 and 1 respectively. In frequency modulation, also known as frequency shift keying, two (or more) different tones are used. In the simplest form of phase modulation, the carrier wave is systematically shifted 45, 135, 225 or 315 degrees at uniformly spaced intervals. Each phase shift transmits 2 bits of information. Figure (1.1) illustrates the three terms of modulation. A device that accepts a serial stream of bits as input and produces a modulated carrier as output (or vice versa) is called a, modem (for modulator - demodulator). The modem inserted between the (digital) computer and the (analog) telephone system.

To go to higher and higher speeds, it is not possible to just keep increasing the sampling rate. Thus research on faster modems focused on getting more bits per sample (i.e., per band).

Most advanced modems use a combination of modulation techniques to transmit multiple bits per band. In figure (a) we see dots at 0,90,180, and 270, degrees, with two amplitude levels per phase shift. Amplitude is indicated by the distance from the origin. In figure (b) we see a different modulation scheme, in which 16 different combinations of amplitude and phase shift are used. Thus figure (a) has eight valid combinations and can be used to transmit 3 bits per band. In contrast figure (b) has 16 valid combinations and can thus be used to transmit 4 bits per baud. The scheme of figure (b) when used to transmit 9600 bps over a 2400 - baud line is called QAM (Quadrature Amplitude Modulation).

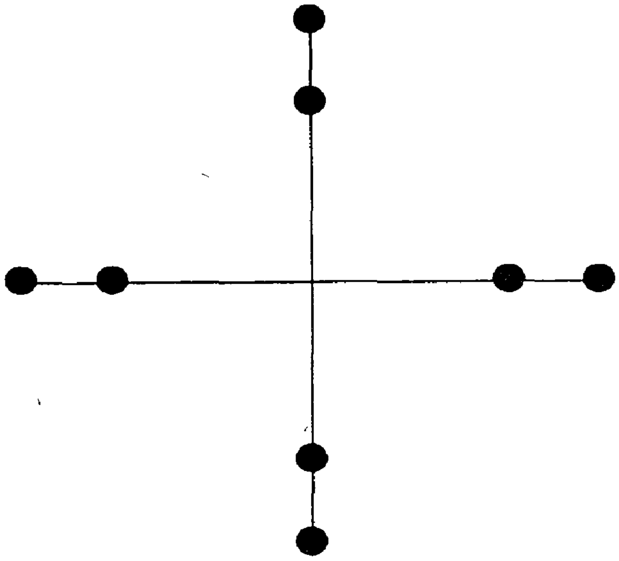


Fig. (a)
3 bits/band modulation

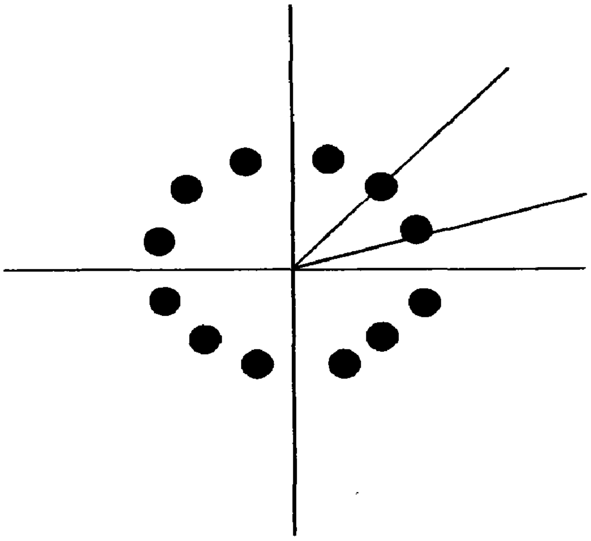


Fig. (b)
4 bits/band modulation

Diagrams such as those of figure, we show the legal combinations of amplitude and phase are called constellation patterns. Each high - speed modem standard hȧs its own constellation pattern and can talk only to other modems that use the same one although most modems can emulate all the slower. ones. The ITV V.32 9600 bps modem standard uses the constellation pattern of figure (b) for example. The next step above 9600 bps is 14,400 bps. It is called V.32 bis. This speed is achieved by transmitting 6 bits per sample at 2400 baud. Its constellation pattern has 64 points. Fax modems use this speed to transmit pages that have been scanned in as bit maps. After V.32 bis comes V.34 which runs at 28,800 bps.

Many modems now have compression and error correction built into the modems. The big advantage of this approach is that these features improve the effective data rate without requiring any changes to existing software. One popular compression scheme is MNP 5, which use run - length encoding to squeeze out runs of identical bytes. Fax modems also use run - length encoding, since runs of Os(blank paper) are very common. Another scheme is V.42 bis, which uses a Ziv - lempel compression algorithm also used in compress and other programs (Ziv and lempel, 1977).

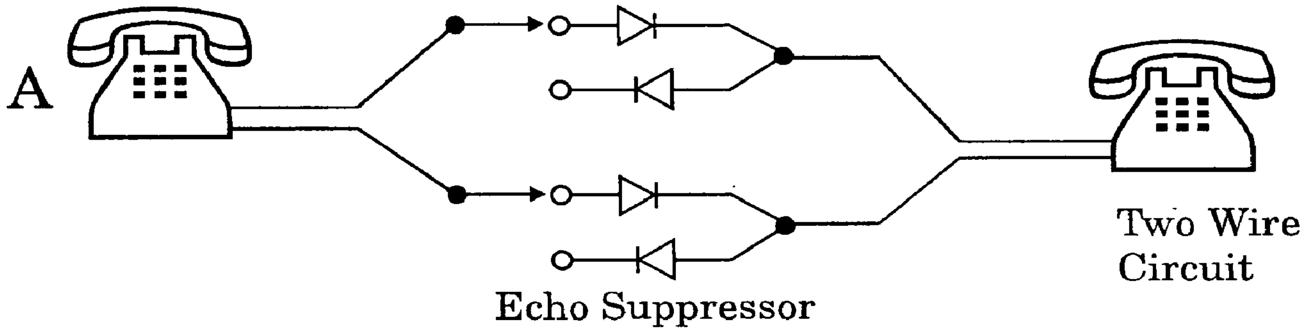


Fig. (c) A talking to B

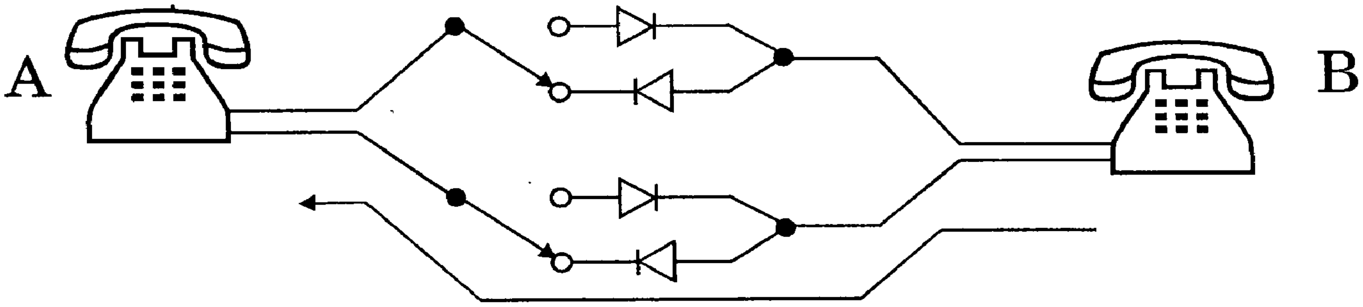


Fig. (d) *B* talking to A

Even when modems are used, another problem can occur on telephone lines: echos on a long line when the signal gets to the final destination, some of the energy may be reflected back, analogous to acoustic echoes in the mountains.

An echo suppressor is a device that detects human speech coming from one end of the connection and suppresses all signals going the other way.

When the first person stops talking and the second begins, the echo suppressor switches directions. A good echo suppressor can reverse in 2 to 5 msec. While it is functioning, however, information can only travel in one direction; echoes cannot get back to the sender. Figure (c) shows the state of the echo suppressors while A is talking to B. Figure (d) shows the state after B has started talking.

RS - 232 - C AND RS - 449

The interface between the computer or terminal and the modem is an example of a physical layer protocol. It must specify in detail the mechanical, electrical, functional and procedural interface. We will now look closely at two well known physical layer standards: RS - 232 - C and it successor, RS - 449.

Let us start with RS - 232 - C, the third revision of the original RS - 232 standard. The standard was drawn up by the Electronic Industries Association, a trade organization of electronics manufacturers, and is properly referred as EIA RS - 232 - C. The international version is given in CCITT recommendation V.24, which is similar but offers slightly some of the rarely used circuits. In the standards, the terminal or computer is officially called a DTE (Data Terminal Equipment) and modem is officially called a DCE (Data circuit - Terminating Equipment).

The mechanical specification is for a 25 - pin connector 47.04, .13mm wide with all the other dimensions equally well specified. The top row has pins numbered 1 to 13 (left to right). The bottom row has pins numbered 14 to 25 (left to right).

The electrical specification for RS - 232 - C is that a voltage more negative than - 3 volts is a binary 1 and a voltage more positive than +4 volts is a binary 0. Data rates upto 29 Kbps are permitted, and cables up to 15 meters.

The functional specification tells which circuits are connected to each of the 25 pins and what they mean.

Fig. (e) shows 9 pins that are nearly always implemented. The remaining are frequently omitted. When the terminal or computer is powered up, it asserts (i.e., sets to a logical 1) Data Terminal ready (pin 20). When the modem is powered up, it asserts Data set ready (pin 6). When the modem detects a carrier on the telephone line, it asserts carrier Detect (pin 8). Request to send (pin 4) indicates that the terminal wants to send data. Clear to send (pin 5) means that the modem is prepared to accept data. Data are transmitted on the Transmit circuit (pin 2) and received on the Receive circuit (pin 3).

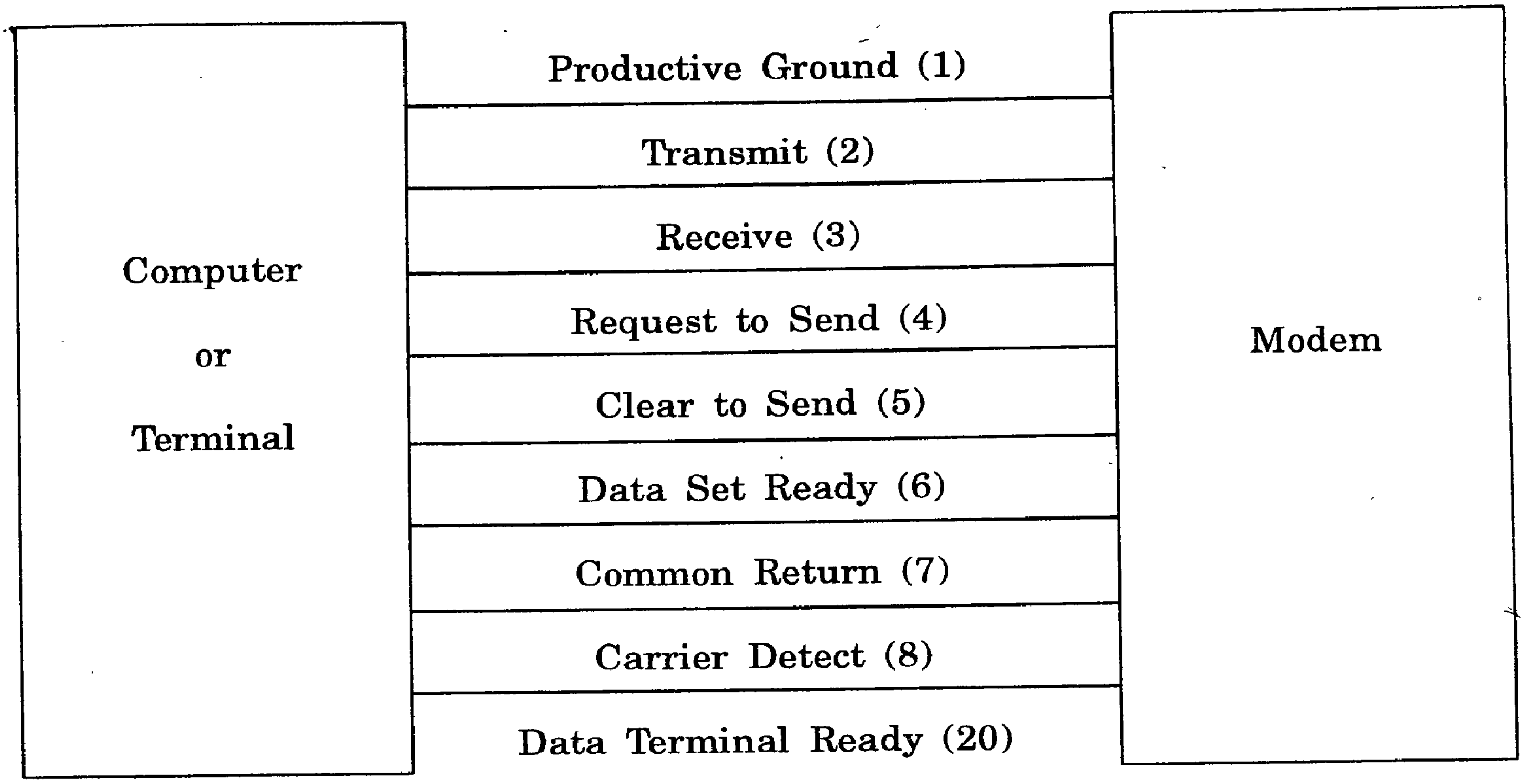


Fig. (e)
Some of the principal RS-232-C Circuits.
The Pin numbers are given in parentheses.

The procedural specification is the protocol, that is, the legal sequence of events, the protocol is based on action - reaction pairs.

It commonly occurs that two computers must be connected using RS-232-C. Since neither one is modem, there is an interface problem. This problem is solved by connecting them with a device called a null modem, which connects the transmit line of one machine to the receive line of the other. A null modem looks like a short cable.

RS-232-C has been around for years. Gradually, the limitation of the data rate to be not more than 20 kbps and 15 meter maximum cable length have become increasingly annonying. EIA had a long debate about whether to try to have a new standard that was compatible with the old one (but technically not very advanced) or a new and incompatible one that would meet all needs for years to come.

The new standard, called RS-449, is actually three standards in one. The mechanical functional and procedural interfaces are given in RS-449, but the electrical interface is given by two different standards. The first of these RS-423-A is similar to RS-232-C in that all its circuits share a common ground. This technique is called unbalanced transmission. The second electrical standard, RS-442-A, in contrast, use balanced transmission, in which each of the main circuits requires two wires, with no common ground. As a result, RS-442-A can be used at speeds upto 2 Mbps over 60 - meter cables.

Trunks and Multiplexing

The multiplexing schemes can be divided into two categories:

FDM (Frequency Division Multiplexing) and TDM (Time Division Multiplexing). In FDM, the fequency spectrum is divided among the logical channels, with each user having exclusive possession of some frequency band. In TDM, the users take turns, each one periodically getting the entire bandwidth for a little burst of time.

Frequency Division Multiplexing

In FDM, first the voice channels are raised in frequency, each by a different amount, Then they can be combined, because no two channels now occupy the same portion of the spectrum.

The basic unit is called the group. Five groups can be multiplexed to form a super group. The next unit is the master group which is five super groups or ten super groups.

Wave Length Division Multiplexing

For fiber optic channels, a variation of frequency division multiplexing is used. It is called WDM (Wavelength Division Multiplexing).

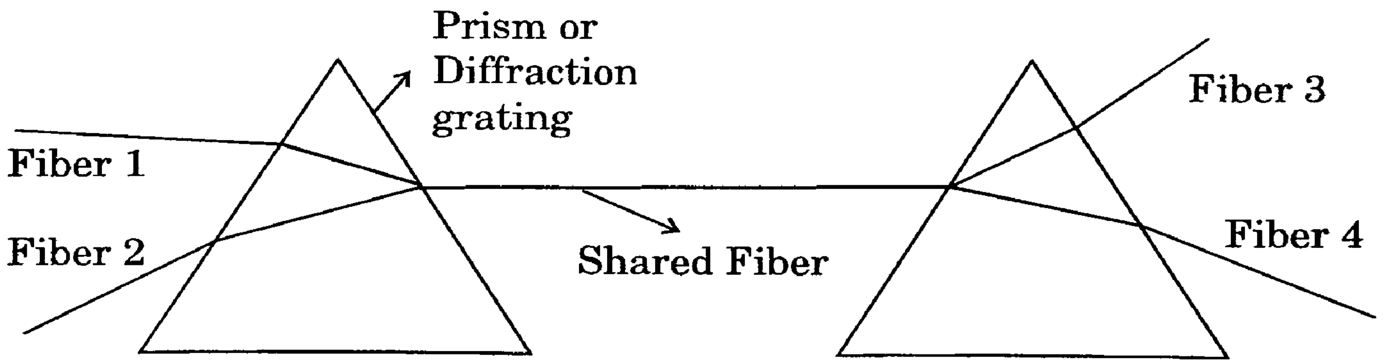


Fig. 3.7

Two fibers come together at a prism (different grating), each with its energy in a different band. The two beams are passed through the prism or grating, and combined onto a single fiber for transmission to a distant destination, where they are split again.

SONET/SDH

The SONET design had four major goals.

(i) SONET had to make it possible for different carriers to internet work.

(ii) Some means was needed to unify the U.S., European, and Japanese digital systems, all of which were based on 64 - kbps PCM channels, but all which combined them indifferent ways.

(iii) SONET had to provide a way to multiplex digital channels together.

(iv) SONET had to provide support for operations, administration, and maintenance (OAM).

A SONET system consists of switches, multiplexes, repeaters, all connected by a fiber. A fiber going directly from any device to any other device, is called a section. A run between two multiplexes is called a line. The connection between the source and destination is called a path.

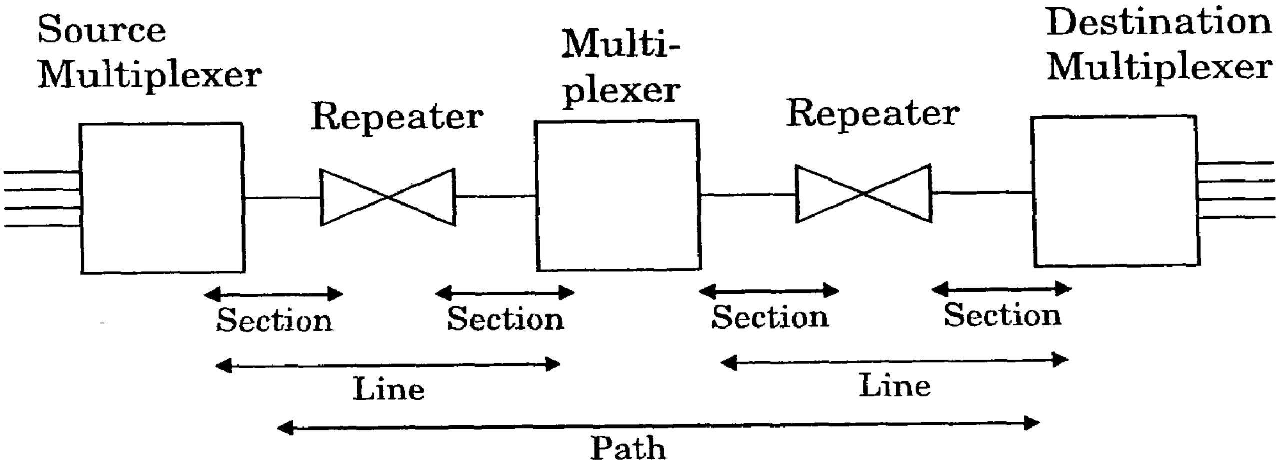


Fig. 3.8

The SONET physical layer is divided into four categories:

(i) Photonic sublayer (ii) Section sublayer
(iii) Line sublayer (iv) Path sublayer

Switching

Two different switching techniques used inside the telephone system are circuit switching and packet switching.

Circuit Switching

When the computer places a telephone call, the switching equipment within the telephone system seeks out a physical "copper" path all the way from the telephone to the receiver's telephone. This technique is called circuit switching.

An alternative switching strategy is message switching when this form of switching is used, no physical copper path is established in advance between sender and receiver. Instead when the sender has a block of data to be sent, it is stored in the first switching office and then forwarded later, one hop at a time.

Packet switching networks place a tight upper limit on block size, allowing packets to be buffered in router's main memory instead of on disk. By making sure that no user can monopolize any transmission line very long, packet - switching network are well suited to handle interactive traffic.

Crossbar Switches

The simplest kind of switch is the crossbar switch (or) crosspoint switch. In a switch 'n' input lines and 'n' output lines, the crossbar switch has n”^{2} intersections, called gross points, where an input and an output line may be connected by a semiconductor switch.

Space Division Switches

By splitting the crossbar switch into small chunks and interconnecting them, it is possible to build multistage switches with many fewer crosspoints. These are called space division switches.

Let us now compute the number of crosspoints needed for a three - stage switch. In the first stage, there are N/n crossbars, each with nk crosspoints, for a total of NK. In the second stage, there are K crossbars, each with (N/n^2) crosspoints. The third stage is the same as the fist. Adding up the three stages, wet get

Number of crosspoint = $2KN + K(N/n^2)$

Time Division Switches

In time division switching, the 'n' input lines are scanned in sequence to build up an input frame with 'n' slots. Each slot has 'k' bits.

The heart of the time division switch is the time slot interchanger, which works as follows:

When an input frame is ready to be processed, each slot is written into a RAM buffer inside the interchange. The slots are written in order, so buffer work 'I' contains slot 'I'.

Lesson - 4

MEDIUM ACCESS SUBLAYER

Networks can be divided into two categories: those using point-to-point connections and those using broadcast channels. This chapter deals with broadcast networks and their protocols. Broadcast channels are sometimes referred to as multi-access channels or random access channels.

The protocols used to determine who goes next on a multi-access channel belong to a sublayer of the data link layer called the MAC (Medium Access Control) sublayer. The MAC sublayer is especially important in LANs. Technically, the MAC sublayer is the bottom part of the data link layer.

Local Area Network (LAN)

LAN are privately-owned networks within a single building or campus of upto a few kilometers in size. They are widely used to connect personal computers and workstations in company offices and factories to share resources.

LANs are distinguished from other kinds of networks by three characteristics :

i) their size
ii) their transmission technology
iii) their topology

Metropolitan Area Network (MAN)

A man is broadly a bigger version of a LAN and normally uses similar technology. It might cover a group of nearby corporate offices or a city and might be either private or public. A man can support both data and voice, and might even be related to the local cable television network.

Static and Dynamic Allocation

The central theme of this chapter is to allocate a single broadcast channel among competing users. Broadcast networks can be divided into static and dynamic, depending on how the channel is allocated. We will first look at static and dynamic themes in general.

Static Channel Allocation in LANs and MANs

The traditional way of allocating a single channel, such as a telephone trunk, among multiple competing users is frequency division multiplexing(FDM).

If there are N users, the bandwidth is divided into N equal sized portions, each user being assigned one portion. Since each user has a private frequency band, there is no interference between users. When there is only a small and fixed number of users, each of which has a heavy (buffered) load of traffic (e.g. carriers' switching offices), FDM is a simple and efficient allocation mechanism.

Dynamic Channel Allocation in LANS and MANs

Before we get into the first of the many channel allocation methods to be discussed in this chapter, it is worth while carefully formulating the allocation problem. Underlying all the work done in this area are five key assumptions, described below.

1) Station Model

The model consists of N independent stations (computers, telephones, etc) each with a program or user that generates frames for transmission. The probability of a frame being generated in an interval of length is λt, where λ is a constant (the arrival rate of new frames). Once a frame has been generated, the station is blocked and does nothing until the frame has been successfully transmitted.

2) Single Channel Assumption

A single channel is available for all communication. All stations can transmit on it and can receive from it.

3) Collision Assumption

If two frames are transmitted simultaneously, they overlap in time and the resulting signal is garbled. This event is called collision. All stations can detect collisions. A collided frame must be transmitted again later. There are no errors other than these generated by ollisions.

4a) Cotinuous Time

Frame transmission can begin at any instant. There is no master clock dividing time into discrete intervals.

4b) Slotted Time

Time is divided into discrete intervals (Slots). Frame transmissions always begin at the start of the slot. A slot may contain 0,1 or more frames, corresponding to an idle slot, a successful transmission, or a collision, respectively.

ALOHA

In the 1970's Nermal Abramgan and hisk colleagues at the University of Hawaii devised a new and elegant method to solve the channel allocation problem. Abremgan's work, called the ALOHA system, used ground based ration broadcasting, the basic is applicable to any system in which uncoordinated users are competing for the use of a single shared channel. We will discuss two versions of ALOHA here: pure and slotted. Pure ALOHA does not require global time synchronization; slotted ALOHA does.

Pure ALOHA

Pure ALOHA is easy to implement: every station just sends whenever it wants to. The trouble is that the channel efficiency is only about 18 percent. Generally, such a low utilization factor is unacceptable for setellites that cost tens of millions of dollars each. If the frame was destroyed, the sender just waits a random amount of time and sends it again. The waiting time must be random or the same frames will collide ever and ever. Systems in which multiple users share a common channel in a way that can lead to conflicts are widely known as contention systems

A sketch of frame generation in an ALOHA system is as follows:

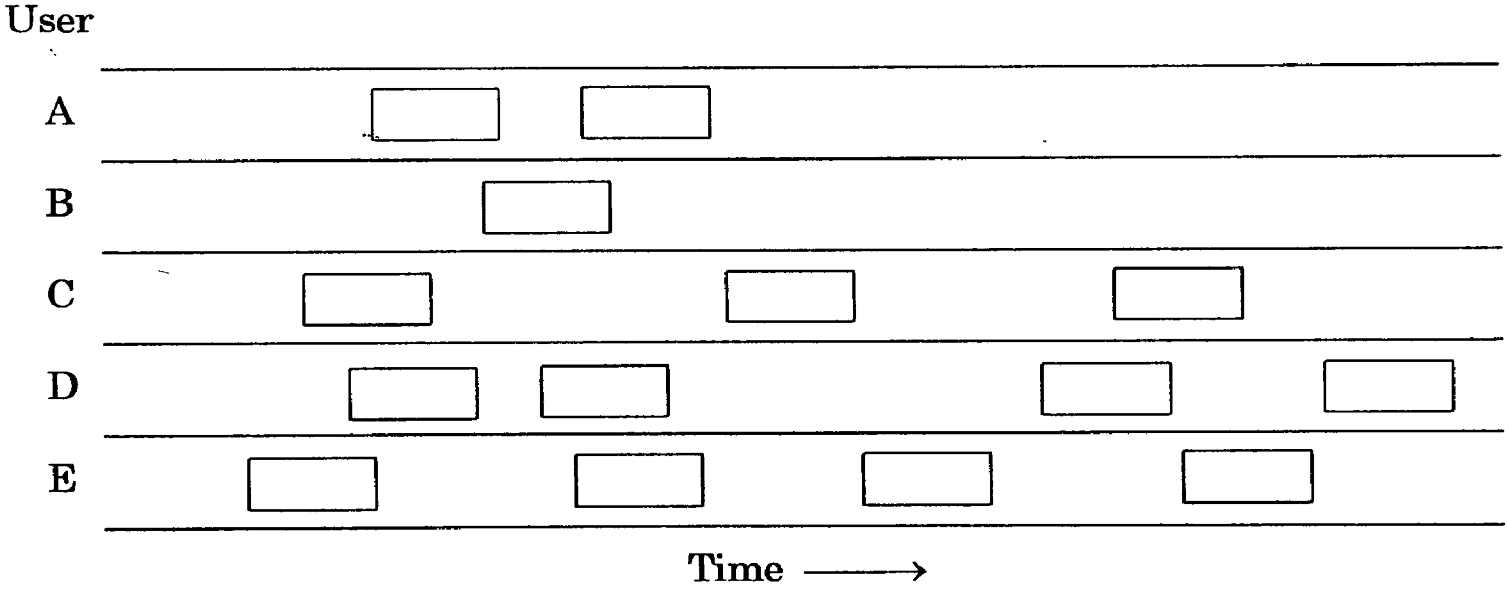


Fig. 4.1

In pure ALOHA, frames are transmitted at completely arbitrary times.

We have made the frames all the same length because the throughput of ALOHA systems is maximized by having a uniform frame size rather than allowing variable length frames.

Whenever two frames try to occupy the channel at the same time, there will be a collision and both will garbled. If the first bit•of a new frame overlaps with just the last bit of a frame almost finished, both frames will be totally destroyed and both will have to be retransmitted later. The checksum cannot distinguish between a total loss and a near miss. Bad is bad.

Slotted ALOHA

In 1972, Roberst published a method for doubling the capacity of an ALOHA system. His proposal was to divide time into discrete intervals, each interval corresponding to one frame. This approach requires the users to agree of slot boundaries. One way to achieve synchronization would be to have one special station emit a pip at the start of each interval, like a clock.

In Robert's method, which has come to be known as slotted ALOHA a computer is not permitted to send whenever a carrige return is typed. Thus the continuous pure ALOHA is turned into discrete one. Using slotted ALOHA doubles the efficiency but introduces the problem of how to synchronize all the stations. So they all know when each time slot begins.

LAN Protocols

Now we turn to the question of what layers are required for the proper operation of the Lan. For the sake of clarity, we examine the question in the context of the OSI reference model. Two characteristics of Lans are important in this context. First, data are transmitted in address frames. Second, there is no intermediate switching, hence no routing required (repeaters are used in rings and may be used in band bus LANs, but do not involves switching or routing).

These two characteristics essentially determine the answer to the question: What OSI layers are needed? Layer 1, certainly physical connnection is required. Layer 2, is also needed. Data transmitted across the LAN must be organized into frames and control must be exercised. But, What about Layer3?. The answer is yes and no. if we look at the functions performed by Layers 3, the answer would seem to be on. First there is routing. The other functions-addressing, sequencing, Flow control, error control and so on-are also performed by Layer 2.

The difference is that layer 2 performs these functions across a single link, whereas layer 3 may perform them across the sequence of links required to traverse the network.

At the highest level are the functions associated with accepting transmissions frame and delivering receptions to attached stations. These functions include:

Provide one or more service access points (SAP). A SAP is a logical interface between two adjacent layers.

- On transmission, assemble data into a frame with address and error detection fields.

- On reception, deassemple frame, perform address recognition and error detection.

- Manage communication between the link.

These are the functions typically associated with layer 2, the data link layer. The first function and related function are grouped into a logical link control (LLC) layer by IEEE 802. The last three functions are treated as a separate layer called Medium Access Control (MAC). This is done for the following reasons:

- The logic required to manage access to multiple-destination link is not found in traditional layer 2 link control.

- For the same Logical Link Control, several MAC options may be provided.

Finally, at the lowest layer, are the functions generally associated with the physical layer. These include:

- Encoding/Decoding of singnals

- Preamble generation/removal (for synchronization)

- Bit transmission or reception.

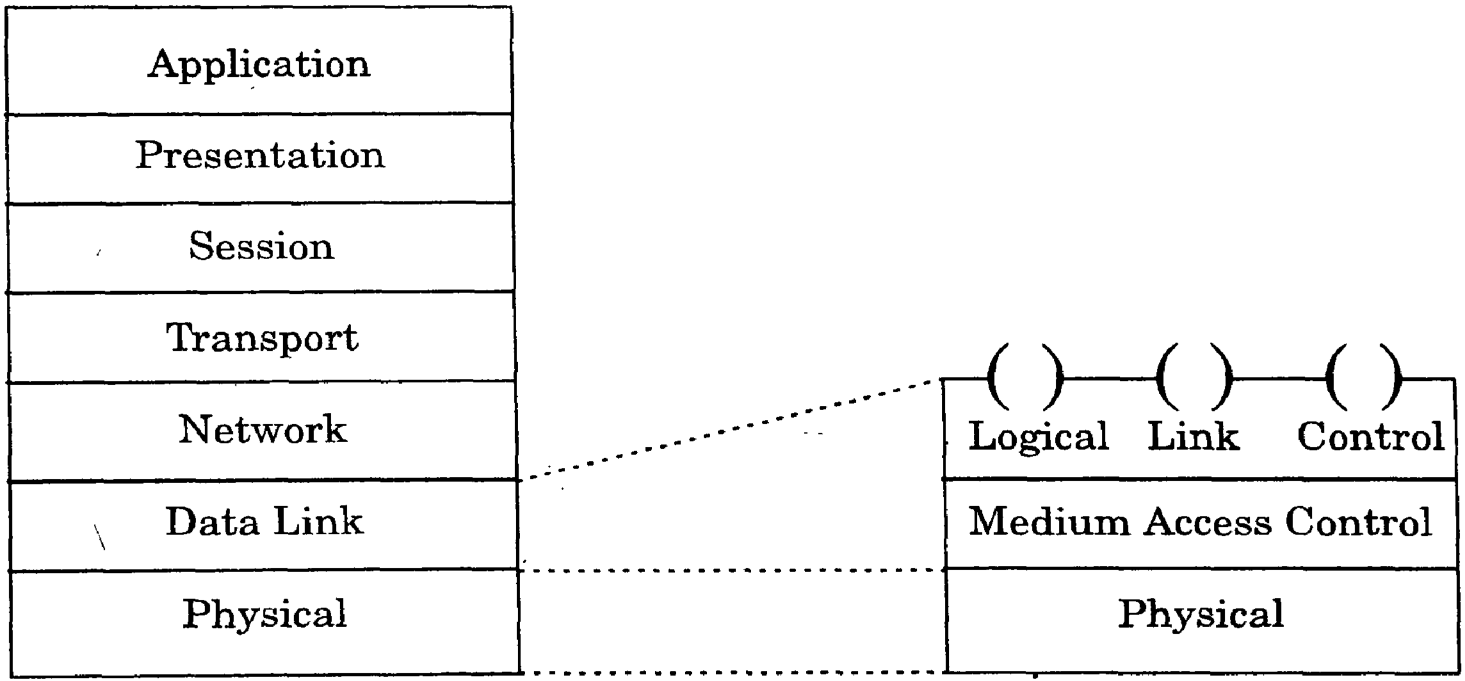


Fig. 4.2
Local Network Communication architecture compared to OSI

Data Link Control for Local Networks

As with all data link control standards, LLC is concerned with the transmission of a frame of data between two stations, with one intermediate switching node.

It differs from traditional link layers in three ways:

- It must suport the multiaccess nature of the link
- It is relived of some details of link access by the MAC layer
- It must provide some Layer 3 functions.

The following figure helps to clarify the requirements for the link layer, we consider two stations or systems that communicate via a LAN (bus or ring). Higher layer (the equivalent of transport and above) provide end-to-end services between the stations. Below the link layer a Mac layer provides the necessary logic for gaining access to the network for frame transmission and reception.

At a minimum, the link layer should perform those functions normally associated with that layer.

- *Error Control* : End-to-end error control and acknowledgement. The link layer should guarantee error free tansmission across the LAN.
- *Flow Control* : End-to-end flow control.

These functions are provided in much the same way as for High-level Data Link Control (HDLC) and other point-to-point link protocols by the use of sequence numbers.

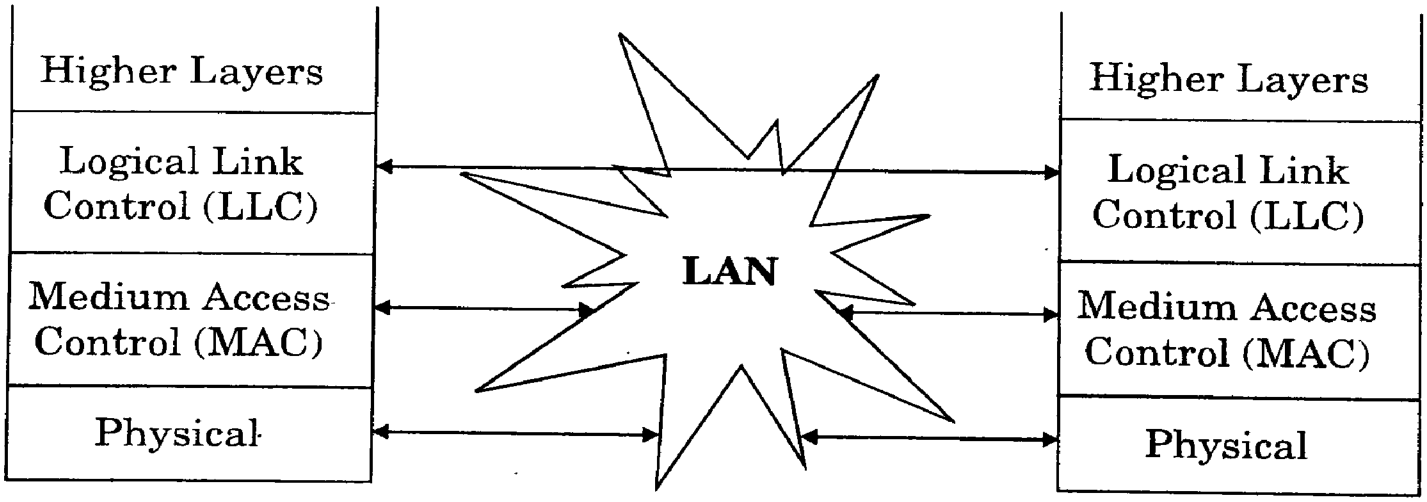

Higher Layers
Logical Link Control (LLC)
Medium Access Control (MAC)
Physical
LAN
Higher Layers
Logical Link Control (LLC)
Medium Access Control (MAC)
Physical

Fig. 4.3
LAN Communication Architecture

Addressing

The preceding discussion referred to both station and LLC addresses consider the following figure which shows the overall format as data transmitted using the LLC and Mac protocols. User data is passed down to LLC which appends a header. This header contains control information that is used to manage the protocol between the local LLC protocol entity. The combination of user data LLC and header is referred to as an LLC protocol data unit (PDU) After LLC has prepared a PDU, the PDU is then passed as data down to the MAC entity. The MAC entity appends both a header and trailer, to manage the MAC protocol. The result is a MAC-level PDU. To avoid confusion with an LLC-level PDU, we will refer to the MAC-level object as a frame; this is the item used in the standards.

The MAC header must contain a destinat~~io~~n address that uniquely identifies a station on the Local Network. This is needed as each station on the local network, will read the destination address field to determine id it should capture the MAC frame when a MAC frame is captured. The MAC entity strips off. The MAC header and trailer passes the LLC PDU up to the LLC entity. The link header must contain a destination SAP address so that LLC can determine to whom the data is to be delivered. Hence, two levels of addressing are needed:

- *MAC address :* Identifies a station on the local network

- *LLC address :* Identifies an LLC user

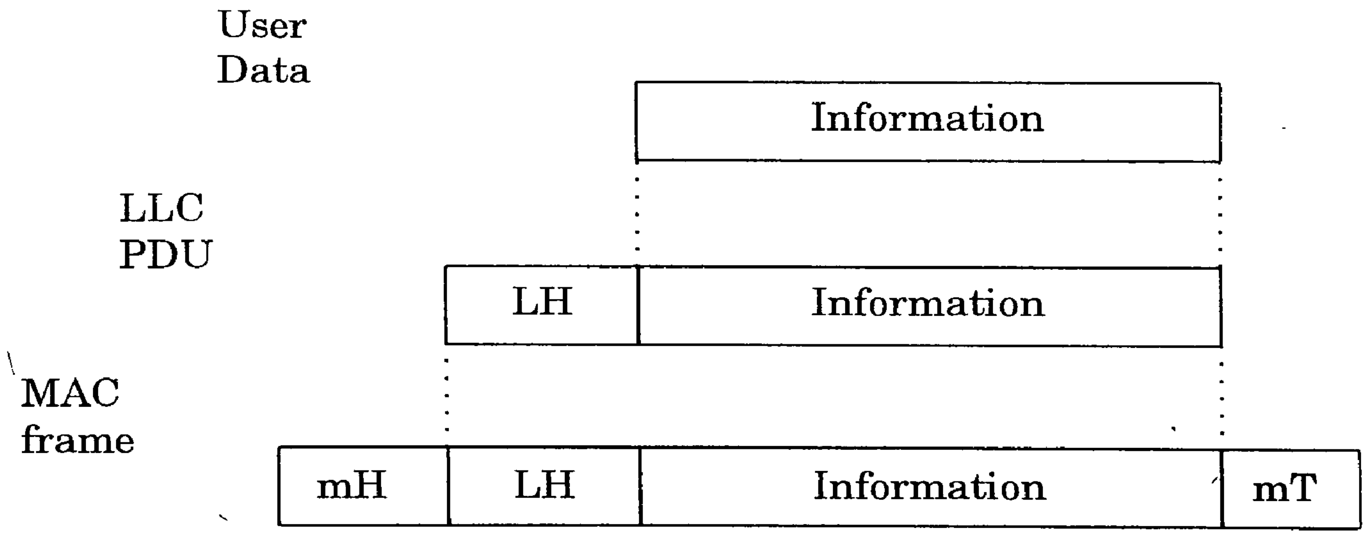


Fig. 4.4
Local Network Protocol Data Units

IEEE Standard 802 for LAN

The key to the development of the LAN market is the availability of a low cost interface. The costs to connect equipment to a Lan must be much less than the cost of the equipment alone. This requirement, plus the complexity of the LAN protocols dictate a VLSI solution. However, chip manufacturers will be reductant to commit the necessary resources unless there is a high-volume market. A LAN standared would assure that volume and also enable equipment of a variety of manufacturers to inter communicate. This is the rationale of the IEEE computer society in February of 1980 to prepare local area network standards. In 1985, the 802 committee issued a set of four standards, which were subsequently adopted in 1985 by the American National Standards Institute (ANSI) as American National Standards [IEEE 85a-d]. These standards were subsequently revised and reissued as international standards by the International Organization for Standardization (ISO) in 1987, with the designation, ISO 8802 [ISO87a-d].

Logical Link Control(LLC)	**IEEE 802.2** Unacknowledged Connectionless Service Connection - Mode Service Acknowledged Connectionless Service			
Medium Access Control (MAC)	IEEE 802.3 CSMA/CD Medium Access Control	Token - Bus Medium Access Control	Token - Ring Medium Access Control	Token - Ring Medium Access Control
Physical Medium	Baseband Coaxial 10 mbps (2 versions) Unshielded Twisted pair 1, 10 mbps Broadband Coaxial 10 mbps	Broadband Coaxial 1,5,10 mbps carrierband 1,5,10 mbps Optical Fiber 5,10,20 mbps	Shielded Twisted Pair 1,4 mbps	Optical Fiber 100 mbps
	Bus Topology		Ring Topology	

Fig. 4.5

The work of the IEEE 80 committee is currently organized the following subcommittes:

802.1	:	High Level Interface
802.2	:	Logical Link Control
802.3	:	CSMA/CD Networks
802.4	:	Token Bus Networks
802.5	:	Token Ring Networks
802.6	:	Metropolitan Area Networks
802.7	:	Broadband Technical Advisory Group
802.8	:	Fiber optic Technical Advisory Group
802.9	:	Integrated Voice and Data Lan Working Group
802.10	:	LAN security working Group.

The high level interface committee deals with issues related to network management for local networks. The discussion in this chapter on architecture and addressing is based on the work of this subcommittee.

Work has been completed on LLC, CSMA/CD, token bus and token ring for an initial set of standards. Work on new options and features continues in each subcommittee.

The work on metropolitan area networks (MAN) is just begging to make progress. The subcommittee is attempting to develop a small number of reasonable alternatives for further study. Since FDDI standards satisfy many of the requirements for a MAN, there has been less enthusiasm and less progress within 802.6 than might have otherwise been expected.

The purpose of 802.7 and 802.8 is to provide technical guidance to the other subcommittees on broadband and optical fiber technology, respectively. The broad band technical advisory group is producing a recommended practice document for broadband cabling systems. The fiber optic technical advisory group is investigating the use of optical fiber as an alternative transmission medium for 802.3, 802.4 and 802.5. It is also considering installation recommendations and a tutorial on fiber optic standards and related information.

The integrated voice and data (IVD) LAN working group was started in 1986. It is developing an architecture and an interface standard for desktop devices to 802 LANs and to integrated services Digital Networks (ISDNs), utilizing twisted-pair wiring to carry both voice and data.

The LAN security-working group was formed in 1988. It will address such issues as secure data exchange, encryption key management, security aspects of network management, and the application of the OSI security architecture to LANs.

Lesson – 5

DATA LINK LAYER

The task of the data link layer is to provide a raw transmission facility and transform it into a line that is free of transmission errors to the network layer. This is accomplished by dividing the input data up into data frames, transmit the frames sequentially and process the acknowledgement frames sent back by the receiver. As physical layer accept only raw bit of data it does not consider about structure of transmission. So to recognizing structure of stream of data, data link layer create and recognize them by frame boundaries. It is accomplished by attaching special bit patterns at the beginning and end of the frame.

The data link layer has to solve the problems caused by damaged, lost and duplicate frames. Any noise added with data during transmission can destroy a frame completely. In this case, the data link layer software cause the source machine to retransmit the frame. Multiple transmission of the source frame introduces the possibility of duplicate frames, which could be sent, only if the acknowledgement frame from the receiver back to the sender were lost.

Data Link Layer Design Issues

The data link layer has a number of functions. Some of them are

(i) Determining how the bits of the physical layer are grouped into frames

(ii) Providing a well-defined service interface to the network layer.

(iii) Dealing with transmission errors

(iv) Regulating the flow of frames so that slow receivers are not swamped by fast senders using traffic regulation mechanism.

(v) In broadcast networks, it provides control access to the shared channel.

Services provided to the Network Layer

The data link layer has to provide services to the network layer, by transferring data from the network layer on the source machine to the network layer on the destination. It has to transfer the frames from source to destination, so that they can be delivered to the network layer. It accepts packets from network layer in source machine and convert into frames and transmit frames to physical link layer in the same machine.

Fig.(1) shows the virtual transmission of bits from the source to the destination is considered as transmitting from network to datalink layer and to source to destination. Fig.(2) shows the actual transmission of data from the source to the destination.

Fig. 1

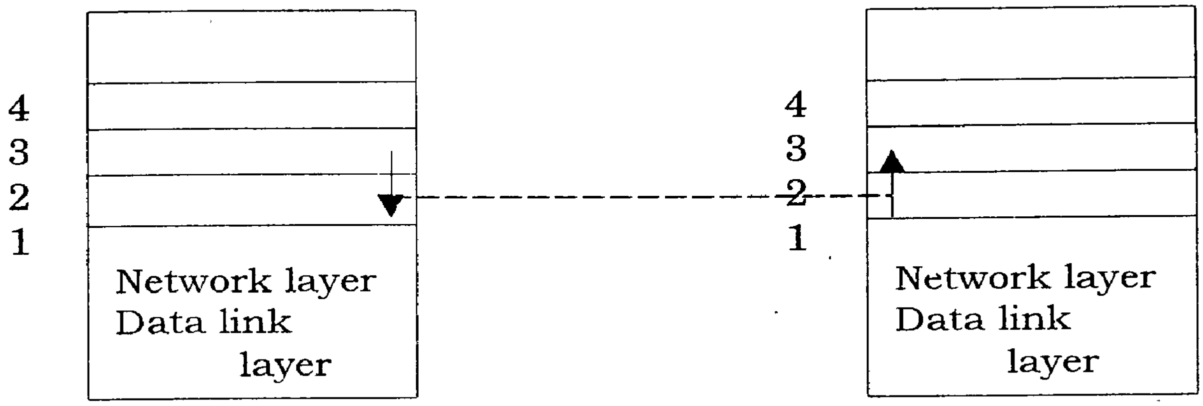
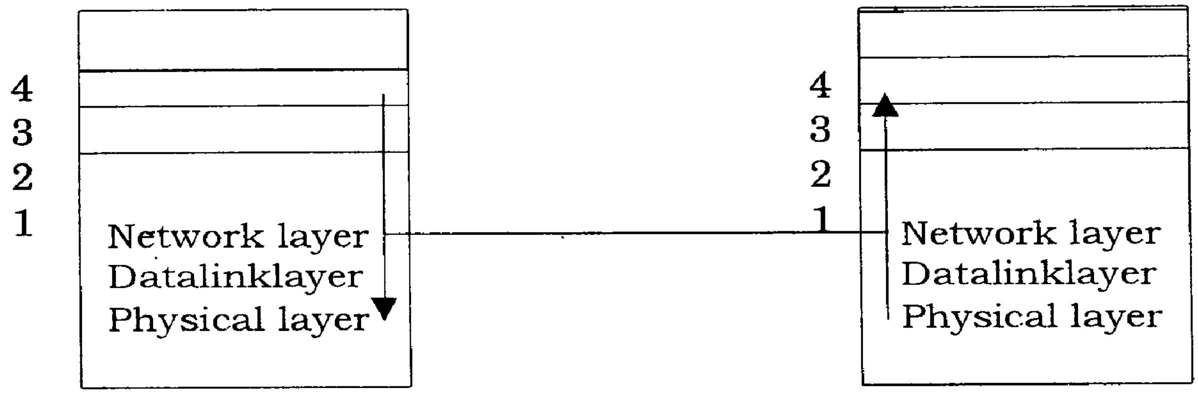


Fig - 2



Three reasonable possibilities provided for this are

(i) Unacknowledged connectionless service

(ii) Acknowledged connectionless service

(iii) Acknowledged connection-oriented service.

(i) Unacknowledged Connectionless Service

Unacknowledged connectionless services send independent frames to the destination machine without making external connections between source and destination machine, and also send frames to destination machine without receiving acknowledge from them. If a frame is lost due to noise during transmission, no attempt is made to recover frames. This type of service is suited, when the error rate is very low so that recovery is left to higher layers. This service is suited only for real time traffic such as speech, LAN etc.

(ii) Acknowledged Connectionless Service

Acknowledged connectionless service is offered, when each frame sent is individually acknowledged which makes sender know whether the frame has arrived destination or not. In case, if the frame is not arrived within the stipulated time, it can be sent again. No connection is established during transmission. This type of service is useful over unreliable channels such as wireless system.

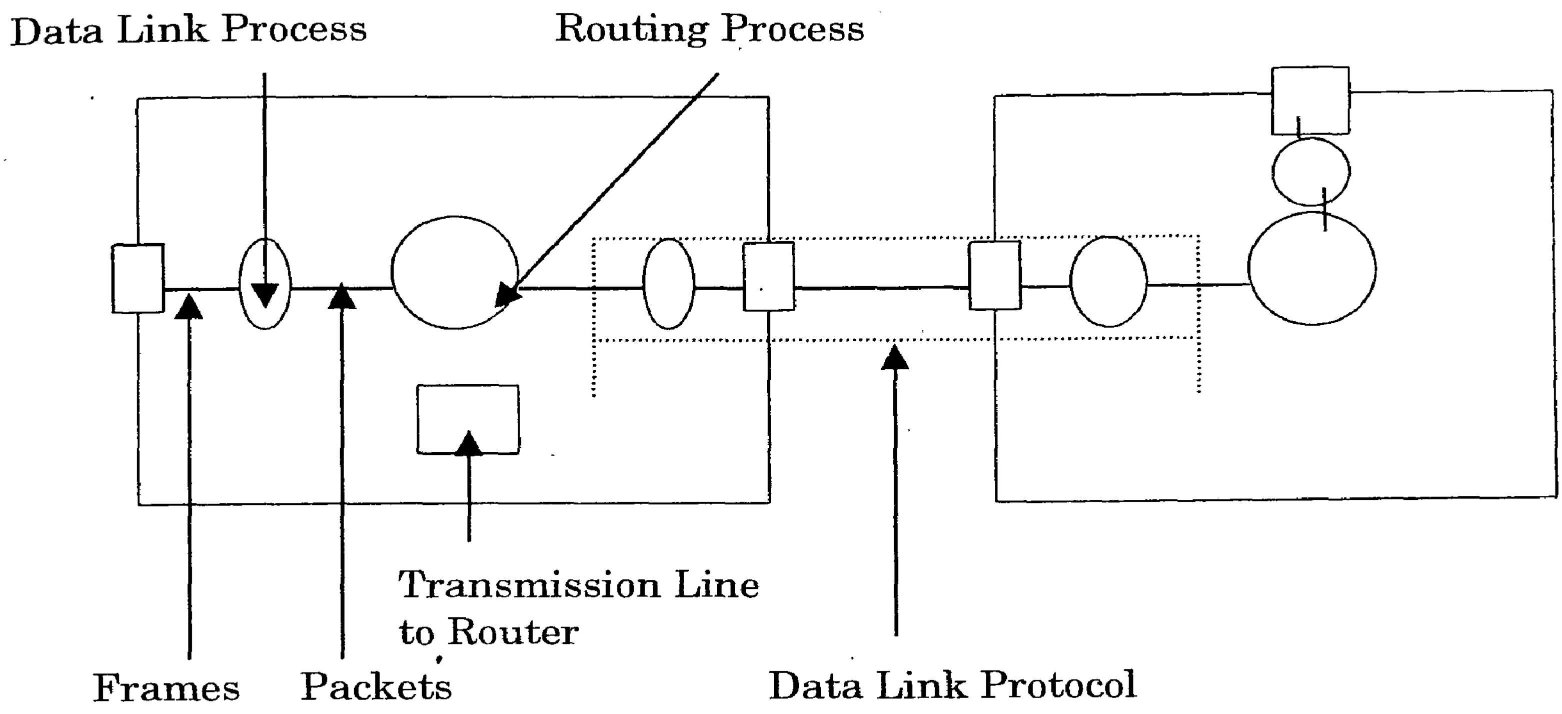
The transport layer sends a message and wait for it to be acknowledged. If the acknowledgement is not arrived at source machine before the timer goes off, the sender can send the entire message again.

(iii) Connection Oriented Service

The most sophisticated service is connection-oriented service. Here a external connection is established between the source and destination and each frame sent through the connection is numbered. Then the data link layer guaranteed that each frame sent, is received by destination in the right order. This provides the network layer processes with the equivalent of a reliable bit stream.

Three different phases are involved in connection oriented service. In the first phase, a connection is established by having both sides initialize variables and counters to keep track of frames. In the second phase one or more frames are transmitted. In the third phase the connection is released freeing up the variables, buffers and other resources used to maintain the connection.

For example, consider a WAN subnet consisting of routers connected by point to point based telephone lines. When a frame arrives at the router, the hardware verifies the checksum and sends it to the Data Link Layer software. This software checks to see if the frame is expected. In case if it is expected it gives the packet contained in the payload field to the routing software. The routing software passes the packet to the data link layer software through an appropriate outgoing line for transmission. The flow over two routers is shown in figure below.



It is the task of the data link protocol (shown in dotted lines) to make unreliable communication lines look perfect. It is an important property for wireless links. One copy handles all the lines, with different tables and data structure for each one.

Framing

The main function of the physical layer is to accept a raw bit stream and deliver it to the destination. This bit stream is not error free. The number of bits received may be less than, equal to (or) greater than the number of bits transmitted. The data link layer has to detect and correct the error.

The data link layer breaks the bit stream into a number of frames and computes the checksum for each frame. When a frame arrives at the destination, the checksum is recomputed. If the newly computed checksum is different from the one contained in the frame, then the data link layer knows that an error has occurred and use any error correction techniques to correct it.

Breaking the bit stream into a number of frames is a difficult task. One way to achieve this is to insert time gaps between frames. It is also possible that these gaps might be squeezed out (or) other gaps might be inserted during transmission since network rarely makes any guarantees about timing. Second way is to use start and stop identification at each frame.

The four method devised to mark the start and end of each frame are

a) Character count
b) Starting and ending characters, with character stuffing

c) Starting and ending flags, with bit stuffing
d) Physical layer coding violations.

a) Character Count

In the character count, a field in the header specifies the number of characters in the frame. When the data link layer sees the character count, it knows how many character follows and where the end of the frame is, the main disadvantage here is that the count can be changed during transmission.

Fig. 1. Character Count

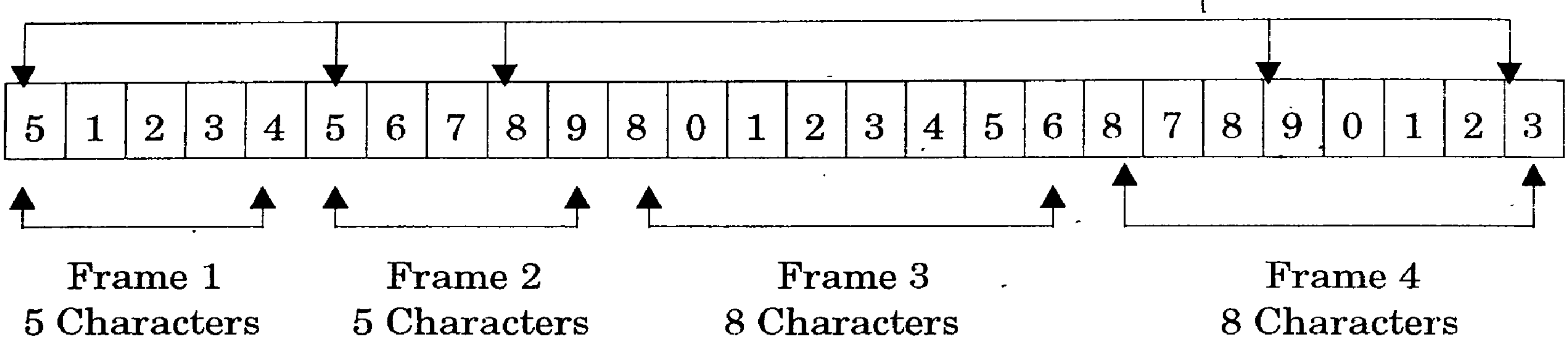
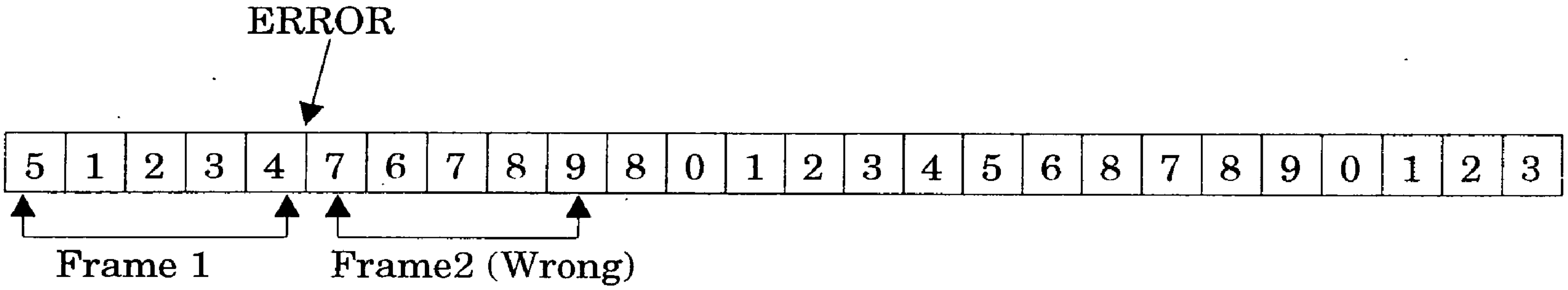


Fig. 2



In the above figure(2), if the character count of 5 in the second frame becomes 8, then the destination will be unable to locate the start of the next frame. Retransmission is very difficult since the destination does not know how many characters to skip over to get started again. For this reason, this method is rarely used.

b) Character Stuffing

Character stuffing method get around with the problem of resynchronization after an error by having each frame start with ASCII character sequence DLE STX and end with the sequence DLE ETX [DLE → Data Link Escape, STX → Start of Text, ETX → End of Text].

But the problem in this method occurs during the transmission of binary data, DLE STX (or) DLE ETX occurs in the data will interfere with the framing. This problem can be solved by inserting an ASCII DLE character just before each "accidental" DLE character in the data. At the receiving end, the Data Link layer removes the DLE before delivering it to the network layer. This technique is referred as character stuffing. Thus the presence or absence of a single DLE can differentiate a framing DLE STX (or) DLE ETX.

a) Data Send by the Network Layer

DLE	STX	A	DLE	B	DLE	ETX	

b) Data after being Character Stuffed by Data Link Layer

DLE	STX	A	DLE	DLE	B	DLE	ETX

↑
Stuffed DLE

c) *Data Passed to the Network Layer on the Receiving Side*

DLE	STX	A	DLE	B	DLE	ETX	

The above figure shows data stream before stuffing, after stuffing and after de-stuffing. A major disadvantage of using this method is that it is closely tied to 8 bit character and ASCII character code.

c) Bit Stuffing

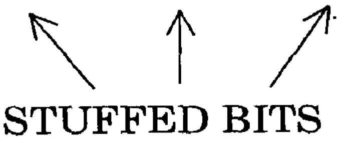
A new technique devised to allow arbitrary sized characters allows data frames to contain arbitrary number of bits and allows character codes with an arbitrary number of bits per character. Each frame begins and end with a special bit pattern, 01111110 called a flag byte.

Whenever, the sender data link layer encounters five consective ones in the data, it automatically stuffs a '0' into the outgoing bit stream. This is called bit stuffing which is similar to character stuffing. When the receiver sees five corrective incoming 1 bit followed by a 0 bit it automatically destuffs the '0' bit and extract original data on the line. With bit stuffing the boundary between two frames can be recognized by the flag pattern.

Let the ORIGINAL DATA be 011011111111111111110010

The DATA transmitted on the line be

011011111011111011111010010



The DATA after destuffing

011011111111111111111 0010

d) Physical Layer Coding Violations

The last method of framing is applicable to network in which encoding on the physical medium contains redundancy. A 1 bit is a high-low pair and 0 bit is a low-high pair. The combination of low-low and high-high are not used for data. This scheme means that every data bit has a transition in the middle.

Many DLL protocols use a combination of character count. When a frame arrives, the count field is used to locate the end of the frame. A frame is valid only if the appropriate delimiter is present at the position and checksum is correct.

Error Control

Reliable transmission is to provide sender must have a feedback of what is happening on the other end of the line to ensure for error free delivery. If the sender receives a positive acknowledgement about a frame, it knows the frame has arrived safely. Instead if it receives a negative acknowledgement, it knows the frame has not reached properly. Hence it has to be transmitted again.

In some case, due to hardware trouble frames are completely vanished. In such cases, if there is any hardware trouble during transmission, the entire system would hang forever. This can be avoided by introducing timers in the data link layer. When the sender transmits a frame, the timer also starts. The frame has to be correctly received and the acknowledgement has to be sent back to the source machine before the timer runs out.

If the frame (or) the acknowledgement is lost, the times will go off. In this case, frames has to be retransmitted. When the frames are transmitted multiple times there is a danger that the receiver will accept the same frame more than once and pass it to the network layer more than once. To rectify this it is necessary to assign a sequence to outgoing frames, so that the receiver can distinguishes retransmissions from original data.

Flow Control

Another important issue is that if a sender transmits frames faster than the receiver can accept them, there may be a loss of frames. To handle this situation, flow control is introduced to throttle the sender into sending no faster than the receiver can handle the traffic. This throttling requires a feedback mechanism so that the sender can be made aware of the situation of frames.

Error Detection and Correcúion

As a result of the physical processes, errors on some media tend to come in burst. Having errors in burst has both advantages and disadvantages. Taking into account on the advantage side, computer sends data into block of bits. For example if the size of the block is 1000bits, the error rate would be 0.001 per bit. If the errors were independent, most of the blocks would contain error. If the errors were in burst, only one or two blocks would be affected. On the disadvantage side, they are much harder to detect and correct.

a) Error-Correcting Codes

There are two basic strategies for dealing with errors. One way is to send redundant information along with the block of data so that the receiver may be able to deduce what the transmission character is. The second way is to include redundancy so that the receiver will be able to deduce that an error has occurred. But it will not spot the error exactly and requests for retransmission. The former uses error – correcting codes and the latter uses error –detecting codes.

Consider a frame with 'm' data bits and 'r' redundant (or) check bits. Let the total length be n(n=m+r). An n-bit unit containing data and check bits is referred to as n-bit codeword. Assume two codewords 10001001 and 10110001. It is possible to calculate the bit difference. To determine the bits differ just EXCLUSIVE OR of the two codewords and count the number of 1 bit in the result. In this case, 3 bits differ. The number of bit positions in which two code word differ is called the ***Hamming distance***.

If the algorithm for computing the check bits is given, it is possible to construct a complete list of the legal codewords and from this list find the minimum hamming distance of the two codewords. This distance is the Hamming distance of the complete code. The error detecting and error-correcting properties of a code depends upon the Hamming distance. To detect of error, a distance of d+1 code is needed because code with a single bit error cannot change a valid codeword into another valid codeword. Whenever the receiver finds an invalid codeword, it can tell that a transmission error has occurred. Similarly, to correct errors, a distance 2d+1 code is needed.

As an example for error-detecting code, consider a code in which a ***single parity bit*** is added to the data. The parity bit is chosen in such a way that the number of 1 bits in the codeword is either even (or) odd. When 101111 is sent in even parity by adding a bit at the end, it becomes 1011111. Consider a code with four codewords, 0000000000, 0000011111, 1111100000 & 1111111111. This code has a distance 5. If the codeword must be 0000011111 arrives the receiver knows that original code word must be 0000011111. If a triple error changes 0000000000 in 0000000111, the error will not be corrected properly.

If we design a code with 'm' message bits and 'r' check bits, that allows all single errors to be corrected. Each 2^m legal messages will have 'n' illegal codewords at a distance 1 from it. Thus 2^m legal messages require n+1 bit patterns. Since the total number of bit patterns is 2^n, we have $(n+1)2^m \leq 2^n$. Using n=m+r, it becomes $(m+r+1) \leq 2^r$. The bits of the codewords are numbered with bit 1 at the left end. The bits that are powers of 2 are check bits. The rest are filled up with the m data bits. Each check bit forces the parity of some collection of bits. A bit may be included in several parity computation.

When a codeword arrives the receiver makes the counter to zero. It then check the check bit to see if it has the correct parity. If the counter is after all the check bits have been examined, the code words is valid. In case the counter is non zero, it contains number of incorrect bits.

Use of Hamming code to correct burst errors.

Character	ASCII	Check Bits
H	1001000	00110010000
a	1100001	10111001001
m	1101101	11101010101
m	1101101	11101010101
i	1101001	01101010110
n	1101110	11111001111
g	1100111	10011000000
	0100000	10011000011
c	1100011	11111000011
o	1101111	00101011111
d	1100100	11111001100
e	1100101	00111000101

order of bits transmission

The above figure shows first bit ASCII character encoded as it, bit codewords using a Hamming code.

Hamming codes can correct only single errors. But it can also correct burst errors. A sequence of consequective codewords are arranged as a matrix. Normally the data would be transmitted one code at a time from left side. To correct burst error the data should be transmitted one column at a time from the leftmost column. When the frames arrives at the receiver, the matrix is reconstructed. If a burst error of length k occurs at most bit in each of the k code words will be affected but hamming code can correct one error per codeword. Thus the entire block can be restored.

b) Error-Detecting Codes

When the channel is simplex, retransmission cannot be requested. Error detection followed by retransmission is preferred for more efficient. For example consider a channel whose error rate is 10^{-6} per bit. Let the size of the block be 1000 bits.70 provide error connection for these block of bits, 10 check bits are needed. To detect a block with a single bit error, one parity per block is sufficient. For every 1000 blocks, an extra block has to be transmitted.

In case a single parity is added to the block & the block is garbled by means of burst error, the probability of error detection is 0.5. This can be improved by considering each block to be sent as a rectangular matrix of n bits wide and k bits high. A parity bit is completed for each column and is needed to the matrix as a last row. Then the matrix is transmitted one row at a time. At the receiver side it checks all the parity bits. If any one of them is wrong a retransmission is requested.

This method can detect a single burst of length 'n' because , bit per column will be error, the probability of n column having correct parity is 0.5. Another method that is widespread in use is the polynomial code ***(also called as cyclic redundancy code (or) CRC code)***. In this method the bit strings are representation of polynomials with coefficients as 0 and 1 only. A k bit frame is regarded as the coefficient for a polynomial with k terms ranging from x^{k-1} to x^0. Such a polynomial is said to have degree k-1.

For example 110001 has 6 bits & this represents a 6 term polynomial with coefficient 1, 1, 0, 0, 0, 1, $x^5+x^4+x^0$.

Polynomial arithmetic is done according to the rules of algebraic field theory. There are no carriers and borrows for addition as well as subtraction. Both addition and subtraction resemble XOR.

```
Eg.        1 0 0 1 1 0 1 1 (+)
           1 1 0 0 1 0 1 0
           ---------------
           0 1 0 1 0 0 0 1
           ---------------
```

When this method is employed, the sender and receiver most agree upon a generator polynomial, G(x). Both high and low order bits should be 1. To calculate the checksum with n bits, to a polynomial M(x) the frame must be longer than the generator polynomial. The idea is to append a checksum to the end of the frame so that the polynomial represented by check summed frame is divisible by G(x). If there is a remainder, there has been a transmission error.

The algorithm for calculating the checksum is given below.

(i) Let r be the degree of G(x). Append r few bits to the low-order end of the frame, so that it contains m+r bits & corresponding to the polynomial $x^r(M(x)$.

(ii) Divide the bit string corresponding to G(x) into the bit string corresponding to $x^rM(x)$ using modulo 2 division.

(iii) Subtract the remainder from the bit string corresponding to $x^rM(x)$ using modulo 2 subtraction. The result is the checksummed frame that is to be transmitted. Let it be the polynomial T(x).

Consider an example for a frame 1101011011 and $G(x) = x^4+x+1$

It should be clear that T(x) should be divisible by G(x). If a transmission error occurs, so instead of the bit string T(x), T(x)+E(x) arrives. Each 1 bit in E(x) corresponds to a bit that has been inverted. If there are k 1 bit in E(x), K-single bit errors have occurred. Upon receiving the checksummed frame, the receiver divides it by G(x), it computes [T(x)+E(x)]/G(x). If T(x)/G(x)=0, the result will be simple E(x)/G(x).

If there has been a single bit errors $E(x)=x^i+x^j$, where i>j. This can also be written as $E(x) = x^j(x^{i-j}+1)$

```
                  1100001010
10011       | 11010110110000
            | 10011
            10011
             10011
              00001
              00000
                1110   - Remainder
```

Frame – 1101011011

Generator – 10011

Message after appending 4 zero bits - 1100001010

If there are an odd number of bits in error, E(x) contains an odd number of terms. There is no polynomial with an odd number of terms that has x+1 as a factor in the module 2 system. By making x+1 a factor of G(x), we can find errors consisting of an odd number of inverted bits.

Assume that no polynomial with an odd number of terms is divisible by x+1, (x) has an odd number of terms & is divisible by x+1. Evaluate E(1)=(1+1) Q(1). Since 1+1=0 (Module 2) E(1)=0. IF E(x) has an odd number of terms substituting 1 for x will yield, as a result. Thus no polynomial with an odd number of terms is divisible by x+1.

A polynomial code with r check bits will detect all burst errors of length <r. A burst error of length K can be represented by xi (xk-1+.....1) where is determines how the right hand end of the received frame the burst is located. If the burst length is r+1, the remainder of the division by G(x) will be zero, if an only if the burst is identical to G(x).

If can also be shown that when an error longer than r+1 bits occur, several shorter burst occur, the portabibly of a bad frame getting through unnoticed is $(1/2)^r$

Three polynomials have become international standards.

CRC-12 = $x^{12}+x^{11}+x^{3}+x^{2}+x^{1}+1$

CRC –16 = $x^{16}+x^{15}+x^{2}+1$

CRC- CCITT = $x^{16}+x^{12}+x^{5}+1$

All these three contain x+1 as a prime factor, CRC-12 is usedif the length of the character is 6 bits. The other two are used for 8 bit characters. CRC-16 (or) CRC-CCITT detects all single and double errors, all errors with an odd number of bits, all burst errors of length 16, 99.9% of 17.bit error burst & 99.9% of 18. bit and longer bursts.

The computation of checksum is complicated.Peterson and Brown constructed a shift register circuit to compute and verify the checksum in hardware. It is assumed that frames to be checksum contains random bits. All analyses have been made under this assumption.

Example Data Link Protocols

Some of the widely-used data link protocols are : HDLC,SLIP, PPP etc.

HDLC – High-Level Data Link Control

The protocols used in n/w are derived from the data link protocols used in IBM'S SNA, called SDLC (Synchronous Data Link Control Protocols) ANSI modified it to become ADCCP (Advanced Data Communication Control Procedure) and ISO modified it to become HDLC (High–Level Data Link Control). CCITT then adopted and modified HDLC for its LAP (Link Access Procedure) as part of the X.25 n/w interface std but later modified it again to LAPB.

All of these protocols are bit-oriented, and all use bit stuffing for data transparency they use the frame structure as follows.

Bits	8	8	8	≥0	16	8
	01111110	Address	Control	Data	Check Sum	01111110

Field *Its Use*

Address - Used to identify one of the terminals and also used to distinguish commands from responses.

Control - Used for sequence numbers, acknowledgements, and others

Data - May contain arbitrary information.

Checksum - Minor variation on the well-known cyclic redundancy code, using CRC-CCITT as the generator polynomial.

The frame is delimited with another flag sequence (01111110) on idle point to point. Lines flag sequence are transmitted continuously. The minimum frame contains 3 fields and totals 32 bits excluding the flags on either end.

There are 3 kinds of frames. They are Information, supervisory, and unnumbered. The contents of the control field for these 3 kinds are shown in figure.

Bits	1	3	1	3	
(a)	0	Seq	P/F	Next	
(b)	1	0	Type	P/F	Next
(c)	1	1	Type	P/F	Modifier

(a) Information Frame

Field *Its use*

Seq - Frame sequence number.

Next - Used for Piggybacked acknowledgement. They use the number of first frame not received instead of Piggy backing the number of the last received correctly

P - Poll used when the computer is inviting the terminal to send data.

F - The final frame is indicated by F

The P/F bit is also used to force the other machine to send a supervisory frame immediately rather than waiting for reverse traffic onto which to piggyback the window information.

b) Supervisory Frames

They are distinguished by the Type field.

Type 0 - It is an acknowledgement frame (RECEIVE READY) used to indicate the next frame expected, and is used when there is no reverse traffic.

Type 1 - It is a negative acknowledgement frame (REJECT) and is used to indicate that transmission error has been detected.

Type 2 - RECEIVE NOT READY. It acknowledges all frames upto but not including Next, but it tells the sender to stop sending. It is used to signal certain simple problems with the receiver. When the condition has been repaired, the receiver sends a RECEIVE READY, REJECT or some control frames

Type 3 - SELECTIVE REJECT. It calls for retransmission of only the frame specified.

Next - Indicates the 1st frame in sequence not received correctly.

Unnumbered Frame

It is sometimes used for control purposes but can also be used to carry data when unreliable connectionless service is called for 5 bits are available to indicate the frame type.

The protocols provide a command called DISC (DIS Connect), which allows a machine to announce that it is going down.

SNRM(Set Normal Response Mode) is used to allow a M/C that has just come back on-line to announce its presence and force all the sequence numbers back to zero.

This mode is an asymmetric mode in which one end of the line is the master and the other side is the slave.

SABM (Set Asynchronous Balanced Mode) which resets the line and declares both parties to be equals. There are two more commands, SABME & SNRME which are the same as SABM and SNRM, except that they use 7 bit sequence numbers instead of 3 bit.

The next command is FRMR (Frame, Reject) used to indicate that a frame with a correct checksum but impossible semantics arrived.

Control frames may be cost or damaged, just like data frames. So they must be acknowledged too. A special control frame is provided for 1 ties purpose, called UA (Unnumbered Acknowledgement).

The remaining control frames deals with initialization, polling and status reporting.

The Data Link Layer in the Internet

The internet consists of individual machines (hosts and routers), and the communication infrastructure that connects them. Within a single building, LANS are widely used for interconnection, but most of the wide area infrastructure is built up from point-to-point leased lines.

In practice, point-to-point communication is primarily used in two situations. First, thousands of organizations have one or more LANs, each with some number of hosts along with a router. Often, the routers are interconnected by a backbone LAN. Typically, all connections to the outside world go through one or two routers that have point-to-point leased lines to distinct routers. It is these routers and their leased lines that make up the communication subnets on which the internet is built.

The second situation where point-to-point lines play a major role in the internet is the millions of individuals who have home connections to the internet using modems and dial-up telephone lines. Usually, what happens is that the user's home PC calls up an "Internet Provider", which includes commercial companies like America Online, CompuServe, and the Microsoft Network, but also many universities and companies that provide home Internet connectivity to their students and employees. Sometimes the home PC just functions as a character – oriented terminal logged into the Internet service provider's time sharing system. In this mode, the user can type commands and run programs, but the graphical internet services, such as the world wide web are not available. This way of working is called having a "Shell account".

Alternatively, the home PC can call an Internet service provider's router and then act like a full-blown internet host. Their method of operation is no different than having a leased line between the PC and the router, except that the connection is terminated when the user ends the session. With this approach, all Internet services, including the graphical ones, become available.

For both the router – router leased line connection and the dial-up host – router connection some point-to-point data link protocol is required on the line for framing, error control, and the other data link layer functions. Two such protocols are widely used in the Internet, SLIP and PPP.

SLIP - Serial Line IP-II-7

SLIP is the older of the two protocols. It was devised by Rick Adams in 1984 to connect sun workstations to the Internet over a dial-up line using a modem. The protocol, which is described in RFC 1055, is very simple. The workstation just sends raw if packets over the line, with a special flag byte (0x10) at the end for framing. If the flag byte occurs inside the IP packet, a form of character stuffing is used, and the two byte sequence (0 x DB, 0 x DC) is sent in its place. If o x DB occurs inside the IP packet, it too is stuffed. Some SLIP implementations attach a flag byte to both the front and back of each IP packet sent.

More recent versions of SLIP do some TCP and IP header compression. What they do is take advantage of the fact that consecutive packets often have many header fields in common. These are compressed by omitting those fields that are the same as the corresponding fields in the previous IP packet. Further more, the fields that do differ are not sent in their entirely but as increments to the previous value. These optimizations are described in RFC 1144.

Although it is still widely used, SLIP has some serious problems first, it does not do any error detection or correction, so it is up to higher layers to detect and recover from lost, damaged, or merged frames.

Second, SLIP supports only IP. With the growth of the Internet to encompass networks that do not use IP as their native language (e.g Novell LANS), this restriction is becoming increasingly serious.

Third, each side must know the other's IP address in advance; neither address can be dynamically assigned during setup. Given the current shortage of IP addresses, this limitation is a major issue as it is impossible to give each home internet user a unique IP address.

Fourth, SLIP does not provide any form of authentication, so neither party knows whom it is really talking to. With leased lines, this is not an issue, but with dial-up lines it is.

Fifth, SLIP is not an approved internet standard, so many different (and incompatible) versions exist. This situations does not make internetworking easier.

PPP – Point to Point Protocol

PPP handles error detection, supports multiple protocols, allows IP address to be negotiated at connection time, permits authentication and has many other improvements over SLIP.

PPP provides 3 things :

1. A framing method that unambiguously delineator the end of one frame and the start of the next one. The frame format also handles error detection.

2. A link control protocol for bringing lines up, testing them negotiating options, and bringing them down again gracefully when they are no longer needed. This protocol is called LCP (Link Control Protocol).

3. A way to negotiate n/w layer options in a way that is independent of the n/w layer protocol to be used. The method is to have a different NCP (N/W control protocol) for each n/w layer supported.

To make a Home PC a temporary Internet Host

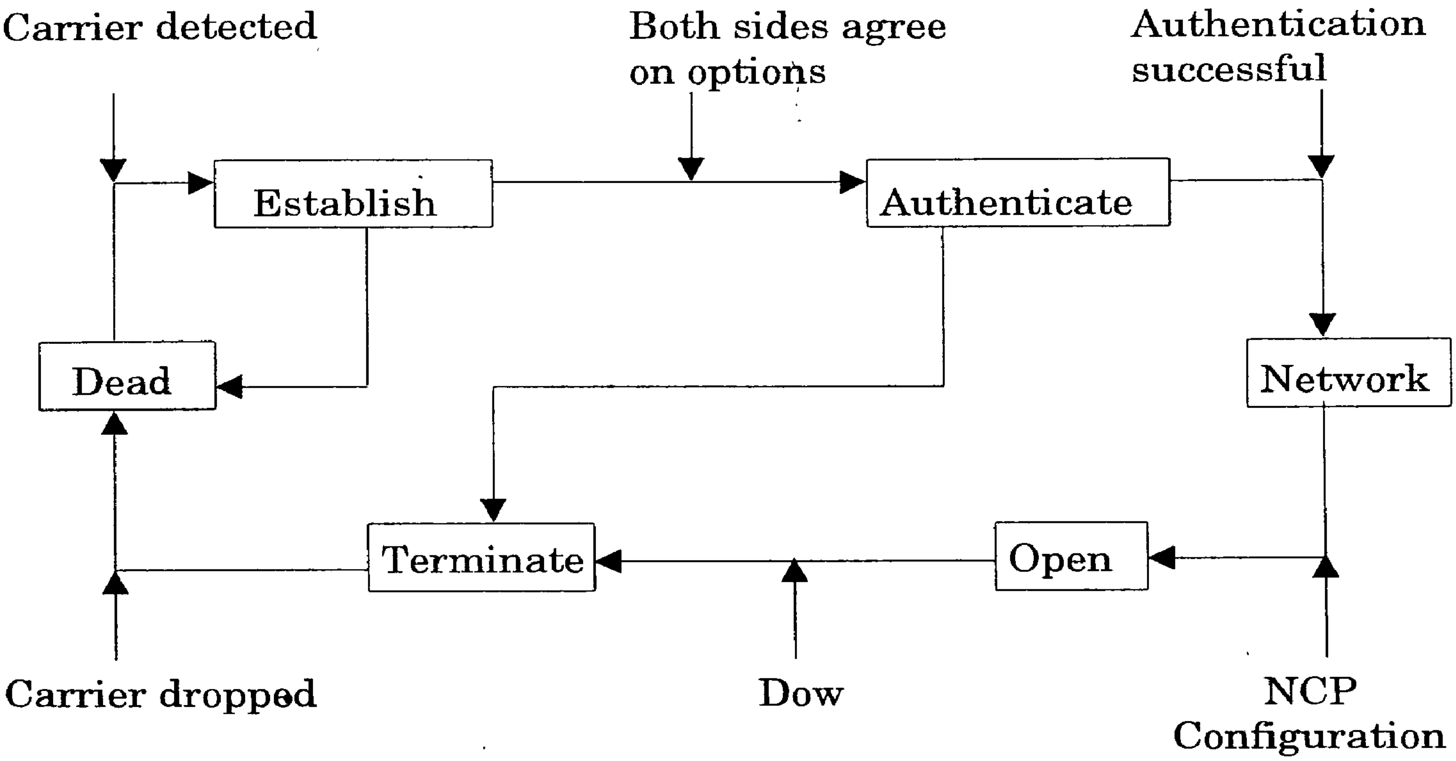
The PC first calls the Internet service provider's router via a modem. After the router's modem has answered the phone and established a physical connection, the PC sends the router a series of LCP packets in the payloads field of one or more PPP frames. These packets and their responses, select the PPP parameters to be used.

Once these have been agreed upon, a series of NCP packets are sent to configure the n/w layer. The PC needs an IP address to run a TCP/IP protocol stack. Each Internet provider gets a block of IP addresses and then dynamically assigns one to each newly attached PC during the login session. If a provider owns n IP addresses, it can have up to n machines logged in simultaneously. The NCP for IP is used to do the IP address assignment.

Now the PC is an Internet host and can send and receive IP packets. When the user is finished, the NCP is used to tear down the n/w layer connection and free up the IP address. Then LCP is used to shut down the data link layer connection. Finally the computer tells the modem to hang up the phone releasing the physical layer connection.

Bringing a Line Up and Down

When the line is DEAD, no physical layer carrier is present and no physical layer connection exists. After that, the line moves to ESTABLISHED. At that point LCP option negotiation begins, if it is successful, it leads to AUTHENTICATE.



When the NETWORK phase is entered, the appropriate NCP protocol is involved to configure the n/w layer. If the configuration is successful, OPEN is reached and data transport can take place. When data transport is finished. The line moves into the TERMINATE phase, and from these, back to DEAD when the carrier is dropped.

During the established (ESTABLISH) phase, LCP is used to negotiate data link protocol options. The LCP protocol provides a way for the initiating process to make a proposal and for the responding process to accept or reject it. It also provides a way for the 2 processes to test the line quality.

Eleven types of LCP packets are defined in RFC 1661. These are listed below.

Name	Direction	Description
Configure – request	I → R	List of proposed options and values.
Configure – ack	I ← R	All options are accepted.
Configure – nak	I ← R	Some options are not accepted.
Configure - reject	I ← R	Some options are not accepted.
Terminate- request	I ← R	Request to shut the uni down.
Terminate – ack	I ← R	OK, line shut down.
code – reject	I ← R	Unknown request received.
Protocol –reject	I ← R	Unknown protocol requested.
Echo – request	I → R	Please end this frame back.
Echo – reply	I ← R	Here is the frame back.
Discard – request	I → R	Just discard this frame. (for listing).

I - Initiator R -> Responder.

The terminate codes are used to shut down the line. The code-reject and protocol reject codes are used by the responder to indicate that it got something that it does not understand. Echo-types-used to test lines. Discard - request - Used for debugging.

The Data Link Layer in ATM

The ATM physical layer covers roughly the OSI physical and data link layers, with the physical medium dependent sublayer being functionally like the OSI physical layer and the transmission convergence (TC) sublayer having datalink functionality there are no physical layer characteristics specific to ATM. Instead, ATM cells are carried by SONET, FODI, and other transmission systems.

When an application program produces a message to be sent, that message works its way down the ATM protocol stack, having headers and tails added and undergoing segmentation into cells. Eventually, the cells reach the TC sub-layer for Transmission.

Cell Transmission

The first step is header checksumming, each cells contains a 5 byte header consisting of 4 bytes of virtual circuit and control information followed by a 1 byte checksum. Although, the contents of the header are not relevant to the TC sub layer, the checksum only covers the first four header bytes not the pay load field. It consists of the remainder after the 32 header bits have been divided by the polynomial $x^8+x^{2+n+1.}$ To this the constant 01010101 is added to provide robustness in the face of headers containing mostly 0 bits. The decision to checksum only the header was made to reduce the probability of cells being delivered incorrectly due to a header error, but to avoid paying the price of check summing the much larger payload field. It is up to higher layers to perform this function, if they so desire. For many real -time applications such as voice and video, losing a few bits once in a while is acceptable. Because it covers only the header the 8 bit checksum field is called the HEC (Header Error Control).

A factor that played a major role in this check summing scheme is the fact that ATM was designed for use over fiber, and fiber is highly reliable. The HEC scheme corrects all single bit errors and detects many multibit errors as well.

For applications that need reliable transmission in the data link layer, Shacham and Meckenney (1990) have developed a scheme in which a sequence of consecutive cells are EXCLUSIVE ORED together. The result an entire cell is appended to the sequence. If one cell is lost on badly garbled. it can be reconstructed from the available information.

Once the HEC has been generated and inserted into the cell header, the cell is ready for transmission. Transmission media come in two categories; asynchronous and synchronous. When an asynchronous medium is used a cell can be sent when ever it is ready to go no timing restrictions exist.

With a synchronous medium cells must be transmitted according to a predefined timing pattern. If no data cell is available when needed the TC sublayer must invent one. These are called "idle cells".

Another kind of non data cell is the OAM (operation and maintenance) cell . OAM cells are also used by the ATM switches for exchanging control and other information necessary for keeping the system running .OAM cells also have some other special functions.

On the receiver's side, idle cells are processed in the TC sublayer, but OAM cells are given to the ATM layer. OAM cells are distinguished from data cells by having the first three header bytes be all zeroes, something not allowed for data cells . The fourth byte describes the nature of the OAM cell.

Another important task of the TC sublayer is generating the framing information for the underlying transmission system, if any. For example, an ATM video camera might just produce a sequence of cells, on the wire, but it might also produce SONET frames with the ATM cells embedded inside the SONET payload. In the latter case, the TC sublayer would generate the SONET framing and pack the ATM cells inside, not entirely a trivial business since a SONET payload does not hold an integral number of 53-byte cells.

Although the telephone companies clearly intend to use SONET as the underlying transmission system for ATM, mappings from ATM onto the payload fields of other systems have also been defined, and new ones are being worked on. In particular, mappings onto T1, T3 and FDD1 also exist.

Cell Reception

On output, the job of the TC sublayer is to take a resequence of cells, add a HEC to each one, convert the result to a bit stream, and match the bit stream to the speed of the underlying physical transmission system by inserting OAM cells as filter. On input, the TC sublayer does not exactly the reverse. It takes an incoming bit stream, locates the cell boundaries, verifier the headers (discarding cells with invalid headers), processes the OAM cells, and passes the data cells upto the ATM layer.

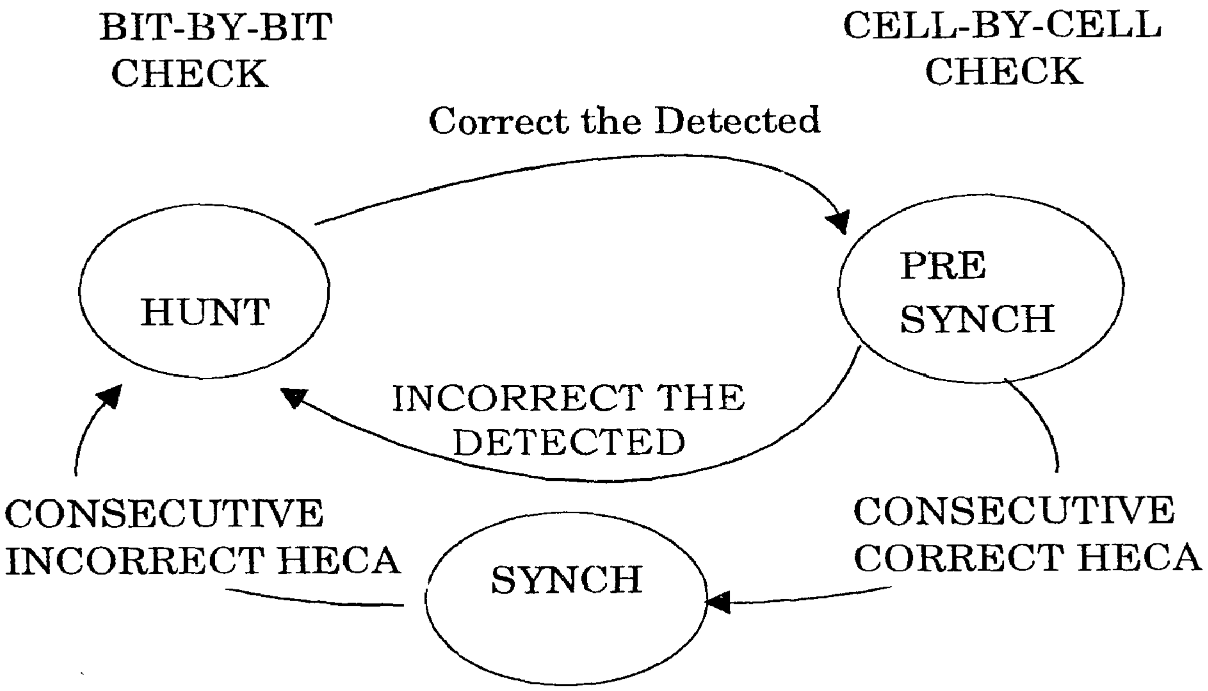
The hardest part is locating the cell boundaries in the incoming bit stream. At the bit level, a cell is just a sequence of 53x8=424 bits. No 01111110 flag bytes are present to mark the start and end of a cell, as they are in HDLC.

In some case, the underlying physical layer provides help. With SONET, for example, cells can be aligned with the synchronous payload envelope, so the SPE pointer in the SONET header points to the start of the first full cell.

The trick is use the HEC. As the bits come in, the TC sublayer maintains a 40 bit shift register, with bits entering on the left and exiting on the right. The TC sublayer then inspects the 40 bits, to see if it is potentially a valid cell header. If it is, the rightmost 8 bits will be valid HEC over the leftmost 32 bits. If this condition does not hold, the buffer does not hold a valid cell, in which case all the bits in the buffer are shifted right one bit, causing one bit to fall off the end, and a new input bit is inserted at the left end. This process is repeated until a valid HEC is located. At that point, the cell boundary is known because the shift register contains a valid header.

The trouble with this heuristic is that the HEC is only 8 bits wide. For any given shift register, even one containing random bit, the probability of finding a valid HEC is 1/256, a moderately large value. Used by itself, this procedure would incorrectly detect cell headers far too often.

To improve the accuracy of the recognition, algorithm, the finite state machine as in the fig is used. These states are used : HUNT, PRESYNCH and SYNCH. In the HUNT state, the TC sublayer is shifting bits into the shift register one at a time looking for a valid HEC. As soon as one is found, the finite state machine switches to PRESYNCH state, meaning that it has tentatively located a cell boundary. It now shifts in the next 424 bits (53 bytes) without examining them. If its guess about the cell boundary was correct, the shift register should now contain another valid cell header, So it once again runs the HEC algorithm. If the HEC is incorrect, the TC goes back to the Hunt(HUNT) state and continues to search bit-by-bit for a header whose HEC is correct.



On the other hand if the second HEC is also correct, the TC may be onto something, so it shifts in another 424 bits and tries again. It continues inspecting header in this fashion until it has found & correct header in a row, at which time it assumes that is synchronized and moves into the SYNCH state to start normal operation. In addition to synchronizing after losing synchronization (or at startup,) the TC sublayer needs a heuristic to determine when it has lost synchronization, for example after a bit has been inserted or deleted from the bit stream. It would be unwise to give up if just one HEC was incorrect, since most errors are bit inversions, not insertions or deletions. The wisest course here is just to discard the cell with the bad header and hope the next one is good. However, if HECS in al row are bad, the TC sub layer has to conclude that it has Lost synchronization and must return to the HUNT state.

Although unlikely it is conceivable that a malicious user could try to spoof the TC sub layer by inserting a data pattern into the payload field of many consecutive cells that imitates the HEC algorithm. Then if synchronization were ever lost, it might be regained in the wrong place. To make this trick much harder, the payload bits are scrambled on transmission and descrumbled on reception.

Before leaving the TC sub layer one comment is in order. The mechanism chosen for cell delineation requires the TC sub layer to understand and use the header of the ATM layer above it. Having one layer make use of the header of a high layer, is in complete violation of the basic rules of protocol engineering. The idea of having layered protocols is to make each layer be independent of the one above it.

It should be possible, for example to change the header format of the ATM layer without affecting the TC sub layer. However, due to the way cell delineation is accomplished, making such a change is not possible.

Lesson – 6

NETWORK LAYER

The network layer is used to get packets from the source all the way to the destination. Getting to the destination may require many hops at the intermediate routers along the way. The network layer is a lower layer that deals with end to end transmission.

To achieve its goals the network layer must

1. Know about the topology of the communication subnet (ie., set of all routers) and choose appropriate paths through it.

2. Take care to choose routers to avoid overloading some of the communication lines and routers while leaving others idle.

When the source and destination are in different networks, it is up to the network layer to deal with these differences and solve the problem that result from them.

Network Layer Design Issues

The network layer design involves two steps

1. The service provided to the transport layer

2. The internal design of the subnet

Services provided to the Transport Layer

The network layer provides services to the transport layer at the network layer/transport layer interface. This interface is a frequently used interface between the carrier and the customer. Its job is to deliver the packets given to it by its customers.

Main goals required to design the network layer services are given below

1. The services should be independent of the subnet Technology.

2. The transport layer should be shielded from the number, type and topology of the subnets present.

3. The network addresses made available to the transport layer should use a uniform numbering plan, even across LANS and WANS.

Now the question raised is whether the network layer should provide connection oriented service or connection less service. Persons from internet community explain that the network service should be connectionless with primitives SEND PACKET and RECEIVE PACKET. In particular no packet ordering and flow control should be done. Each packet must carry the full destination address, because each packet sent is carried independently of its predecessors.

1. A network layer process on the sending side must set up a connection to its peer on the receiving side before sending data. This connection, which is given a special identifier is then used until all the data have been sent at which time it is explicitly released.

2. When a connection is set up, the two processes can enter into a negotiation about the parameters quality and cost of the service to be provided.

3. Communication is in both directions and packets are delivered in sequence.

4. Flow control is provided automatically to present a fast sender from dumping packets into the pipe at a higher rate than the receiver can take them out, they leading to overflow.

Other properties such as guaranteed delivery, explicit confirmation of delivery and high priority packets are optioned.

In the connection oriented service, it is in the network layer (subnet); in the connection less service it is in the transport layer (hosts). Supporters of connection less service say that user computing power become cheap. Some applications, such as digitized voice and real time data collection may regard speedy delivery as much more important than accuracy delivery. Supporters of connection - oriented service say that most users are not interested in running complex transport layer protocols in their machines. Reliable and trouble free service can be provided by network connections.

Different combinations of connection available in network layer are

1. Reliable connection owned

2. Reliable connectionless

3. Unreliable connection oriented and

4. Unreliable connectionless

In theory all the four combination exist, but the dominant combinatory are reliable connection oriented and unreliable connectionless.

The internet has a connectionless network layer and ATM networks have a connection oriented network layer. In ATM, the source host first establishes an ATM network layer connection to the destination host and then sends independent packets (IP) over it as shown on the figure.

E-mail	FTP	......
TCP		
IP		
ATM		
Data ink		
Physical		

But this approach is inefficient because certain guarantees that packets are not always delivered in order, but the TCP code still contains the full mechanism for managing and recording - out - of - order packets.

Internal Organization of the Network Layer

The sub net can be organized by two ways.

1. Using connections

2. Using connectionless.

In the internal option of a sub net, a connection is usually called a virtual circuit. The independent packets of the connectionless organization are called datagrams.

Virtual Circuit

In the case of virtual circuit, it is not necessary to choose a new route for every packet. Instead when a connection is established, a route from the source machine to the destination machine is chosen as part of the connection setup and remembered. It is like a telephone system, when a connection is released, the virtual circuit is also terminated.

Suppose the packets flowing over a given virtual circuit always take the same route through the sub net, each router must remember where to forward packets. So every router must contain a table with one entry for open virtual circuit passing through it. Each packet must contain a virtual circuit number field in its header, in addition to sequence numbers, check sums, when a packet arrives at a router, the router knows on which line it arrived and its virtual circuit number. The virtual circuits have only the logical significance. If it has global significance then there is a possibility of ambiguity. The two virtual circuits having the same global virtual circuit number passes through some intermediate router leading to ambiguities.

The virtual circuits can use either full duplex or simplex. In full duplex if they have been programmed to choose the lowest number not already in use on the link, they will pick the same number when a data packet arrives later, the receiving router has no way of heeling whether it is a forward packet or reverse packet. If circuits are simplex then there is no ambiguity.

Datagrams

Here each packet is routed independently of its predecessors. The routers do not have a table with one entry for each open virtual circuit. Instead line to use for each possible destination router. When the virtual circuits are used internally, then these tables are needed to find the run for the setup packet.

Each data gram contains full destination address. When a packet comes in, the router looks up the out going line to use and sends the packet on its way.

Comparison of Virtual Circuit and Datagram Subnets

Issue	*Datagram subnet*	*Virtual Circuit subnet*
Circuit	not needed	Required
Addressing	Each packet contains the full source and destination address	Each packet contains a short virtual circuit number
State information	sub net does not hold State information	Each Virtual circuit requires sub net table space
Effects of router failure	None, except for packets lost during crash	All virtual circuits that passed through the failed router are terminated
Congestion control	Difficult	Each if enough buffers can be allocated in advance for each virtual circuit.

There are several trades off exist between virtual circuits and data grams. One is between the router memory space and bandwidth. In Data gram, each packet has full destination address. If the packets tenc to be fairly short, in such a case it washes its band width.

Another trade off is set up time and passing time, virtual circuit requires a set up phase, which takes time and consumes resources. To find out where the packet goes, the router just uses the circuit number. In a data gram sub net, a more complicated procedure is required to determine where the packet goes. If a router crashes and losses its memory then all the virtual circuits passing through it will have to be aborted. In data gram if the router goes down means only those users whose packets were quacked up in the router at that time will suffer. The loss of communication is factual to virtual circuits using it but can be easily compensated for if data grams are used.

There is a possibility of occurring a connectionless service on the top of a virtual circuit sub net. Example is running IP over an ATM sub net.

Routing Algorithms

It is used for deciding which output line an incoming packet should be transmitted on. If the subject uses the virtual circuit means the decisions are made only when a new virtual circuit is being setup. If the subject uses the data gram then this decision must be made for every arriving data packet. Routing algorithm is divided into two

(i) Non-adaptive algorithms
(ii) Adaptive algorithms

Non-Adaptive Algorithms

The algorithms not based on their routing decisions or the estimates of the current traffic and topology are known as non-adaptive algorithms.

The route which is used to get from source to destination is computed in advance and downloaded to the routers when the network is booted. This procedure is known as STATIC ROUTING.

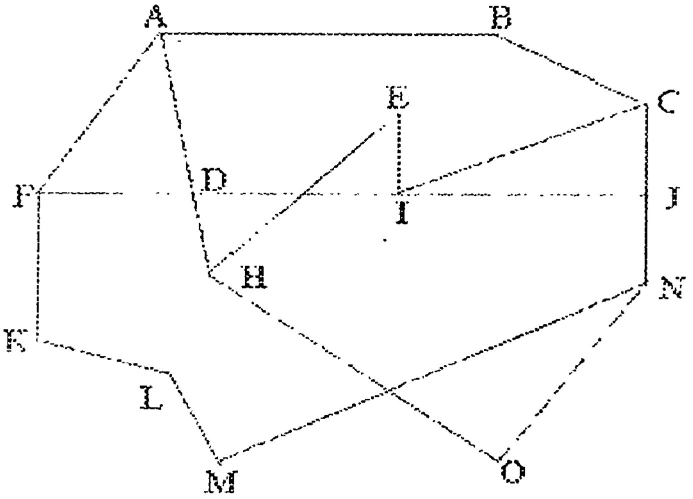
Adaptive Algorithms

The algorithms which will change their routing decisions to reflect the changes in the topology. Adaptive algorithms differ in which they get information, when they change the routes and what metric is used for optimization. (eg. distance, number of hops or estimated transit time)

The Optimality Principle

It may be helpful to note that one can make a general statement about optimal routes without regards to network topology or traffic after getting the specific algorithm. This statement is known as the optimality principle.

EX SUB NET



From the subnet the router j is on the optional path from router I to router K. The optional path from j to k also falls along the same route.

Sink Tree

The set of optimal routes from all sources to a given destination from a tree rooted at the destination. This tree is called Sink tree.

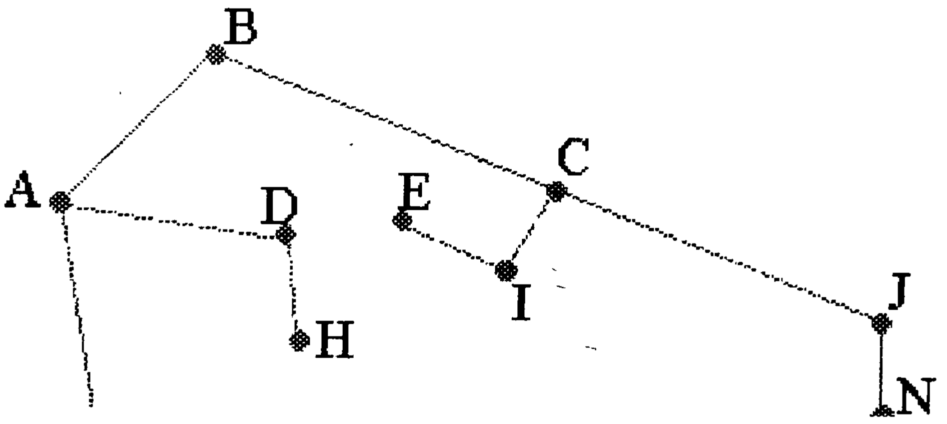


Fig. Sink Tree for B

A sink tree is indeed a tree, it does not contain any loops. So each packet will be finite and has bounded number of hops.

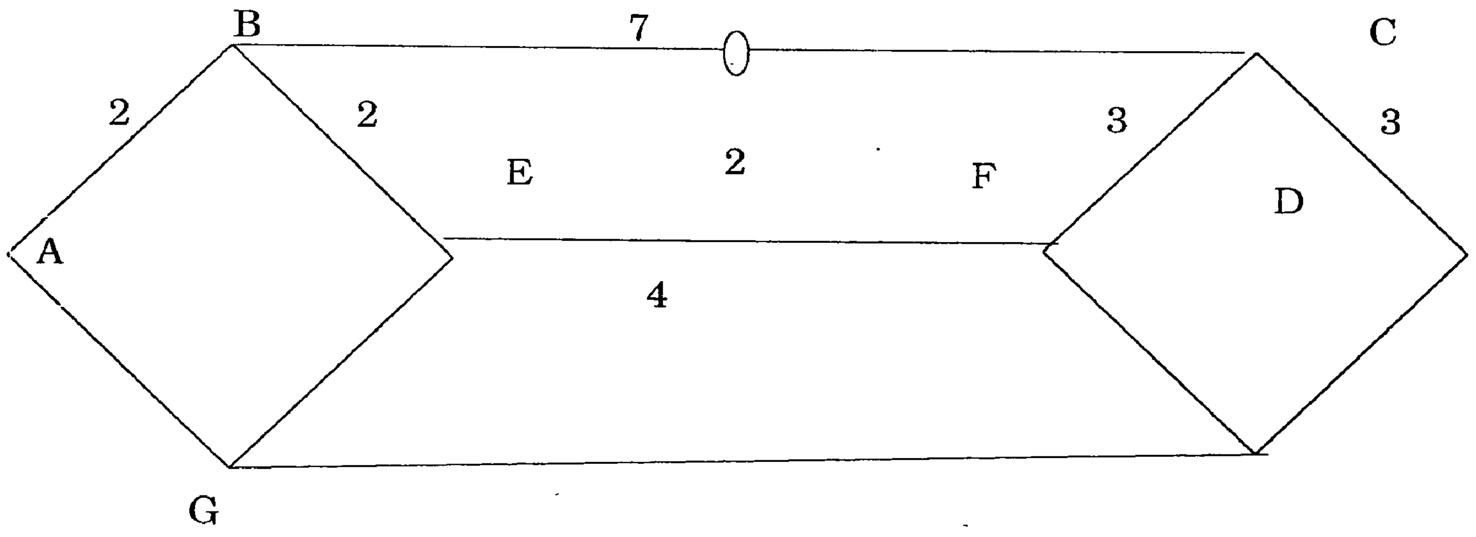
Links and routers can go down and come back during operations.

The optimality principle and the sink tree provide a bench mark against which other routing algorithms can be measured

Shortest Path Algorithm

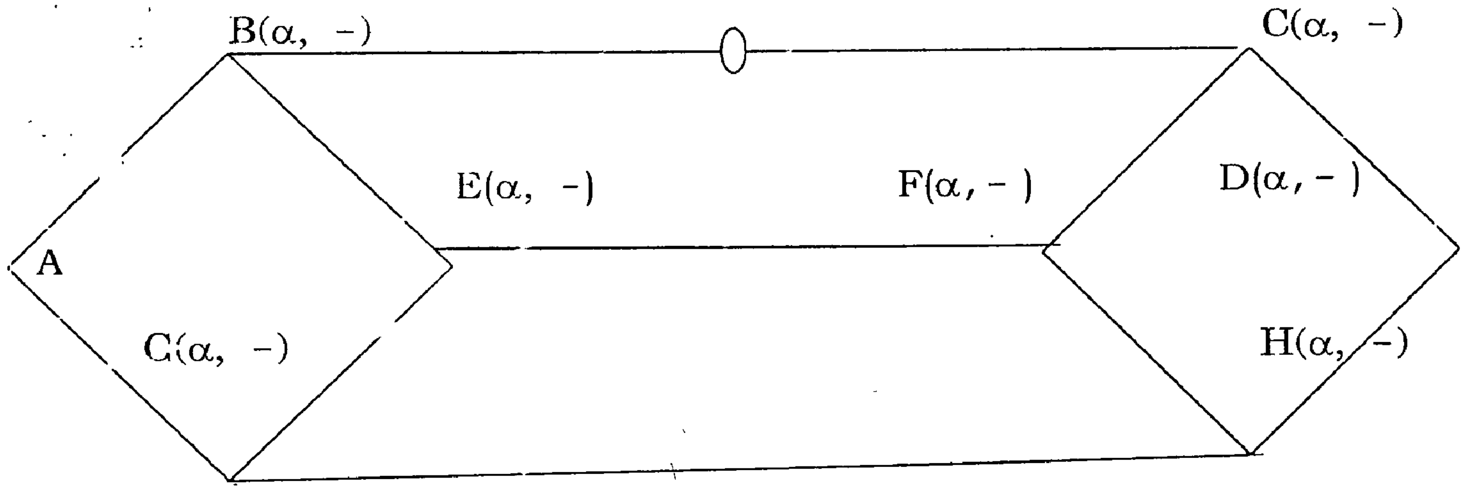
To choose the route between the given pair of routers, the algorithm just finds the shortest path between them on the graph.

One way of measuring the path length is the number of hops. Another metric is the geographic distance in kilometers. Consider the following figure

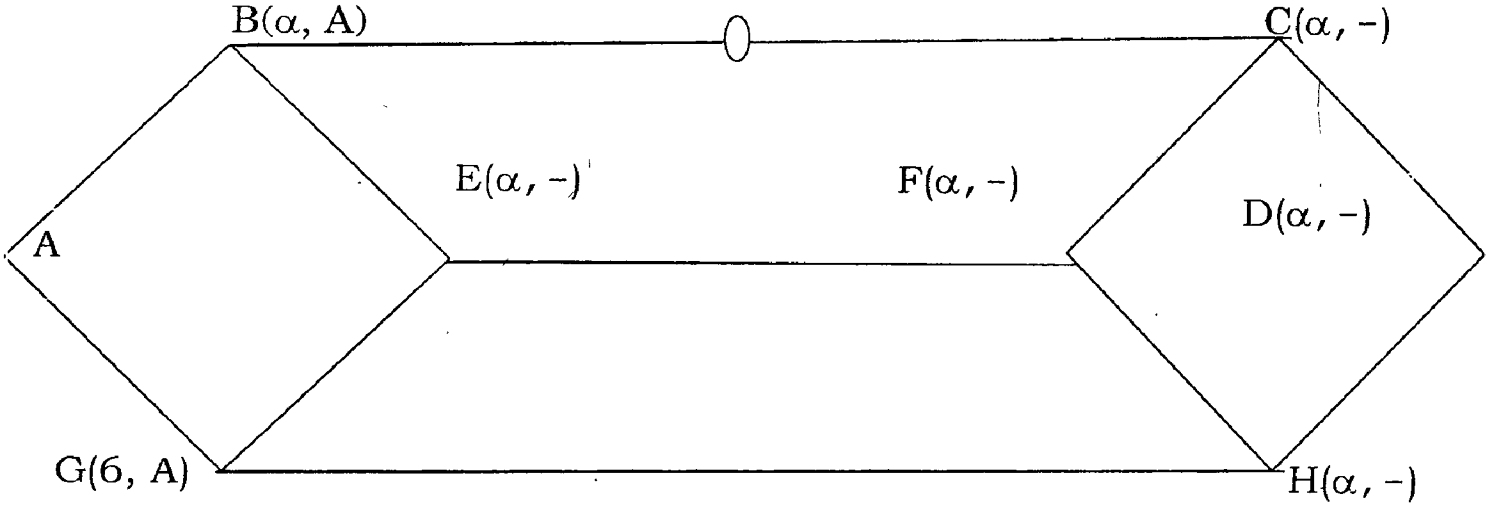


According to the first way ABC and ABE are equally long. In the second way ABC is clearly much longer than ABE.

Several algorithms are used to find the shortest path between two nodes of a graph. According to Dijkstra algorithm each node is labeled with its distance from the source code. Initially no paths are known. So all the nodes are labelled as infinity.



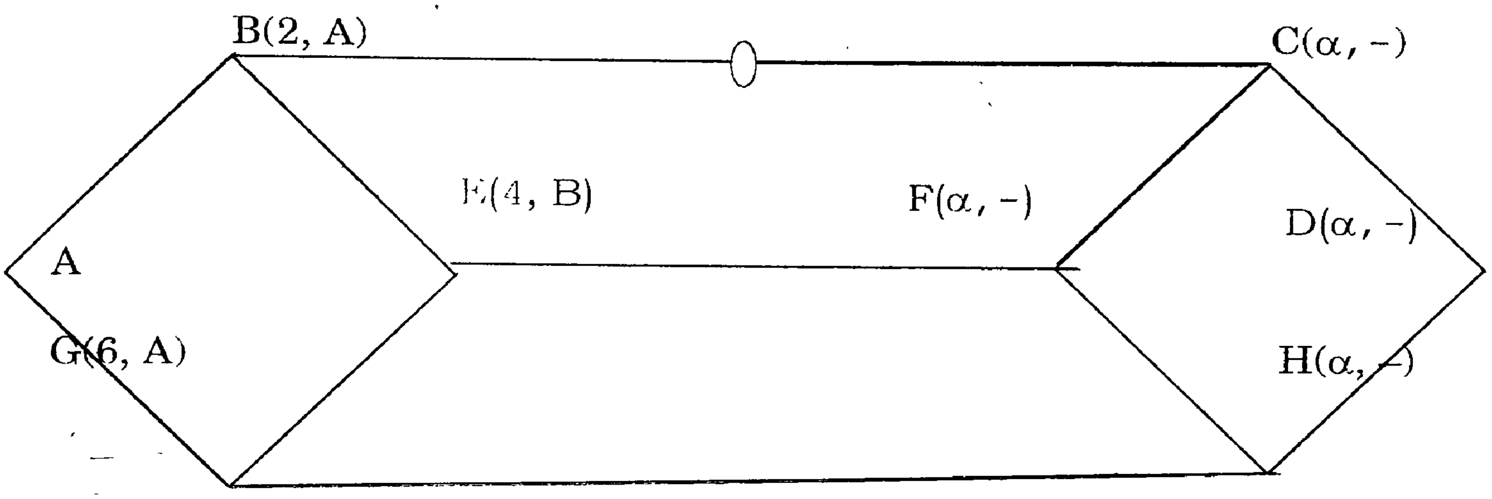
As the algorithm proceeds and paths are found the label may change. Start with the A node. A has two paths, one is to B and another is to G. So now we have two choices AB and AG. The path length of AB is 2 and AG is 6. Here minimum path length is in AB.



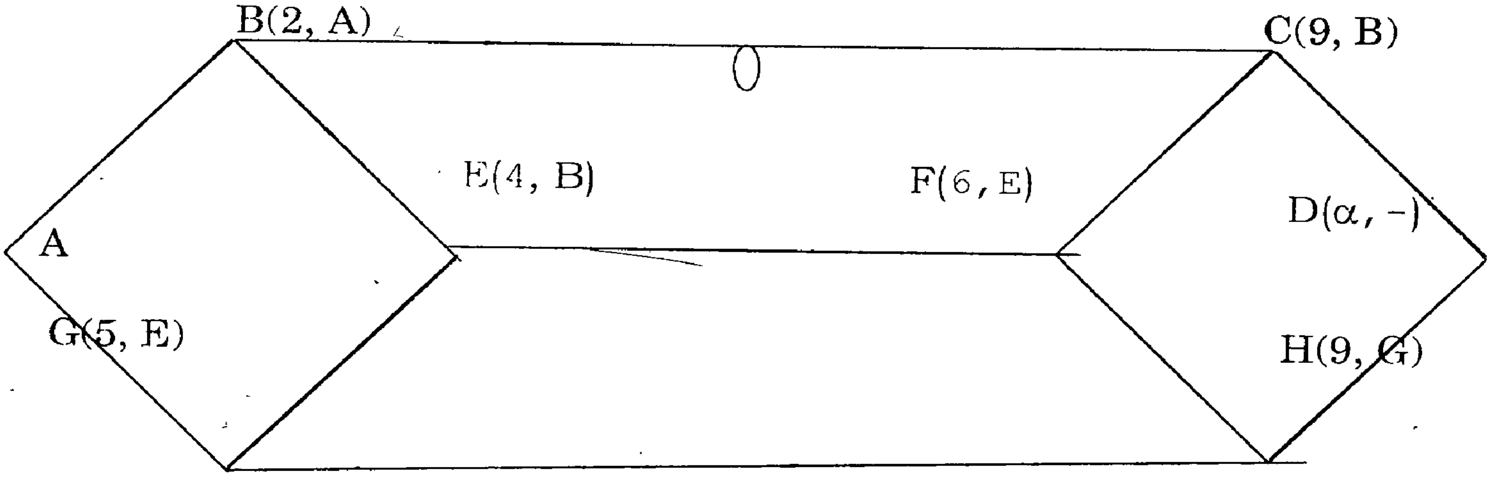
So label B as (2,A). A label may be either tentative or permanent. Initially all the labels are tentative. After finding the shortest path, label is made permanent and never changed there after.

Now start from B, Check all the adjacent node from B.

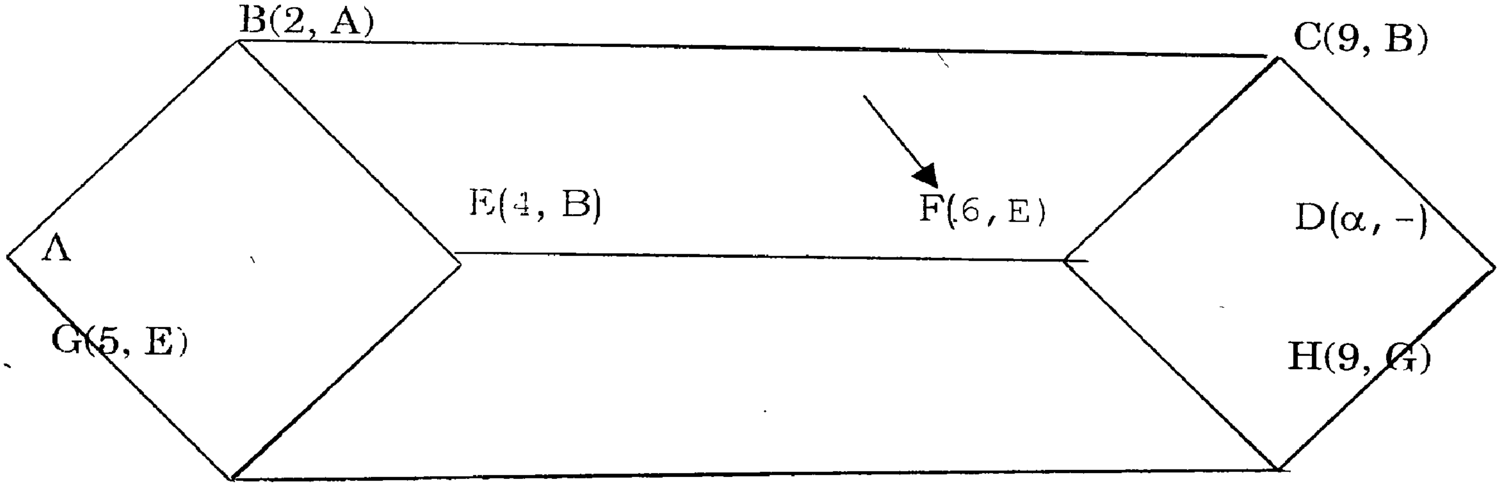
BC and BE are the adjacent nodes. The path length of BC is 9 i.e the length of B from A is 2, B to C is 7.Similarly BE is 4. So choose BE.



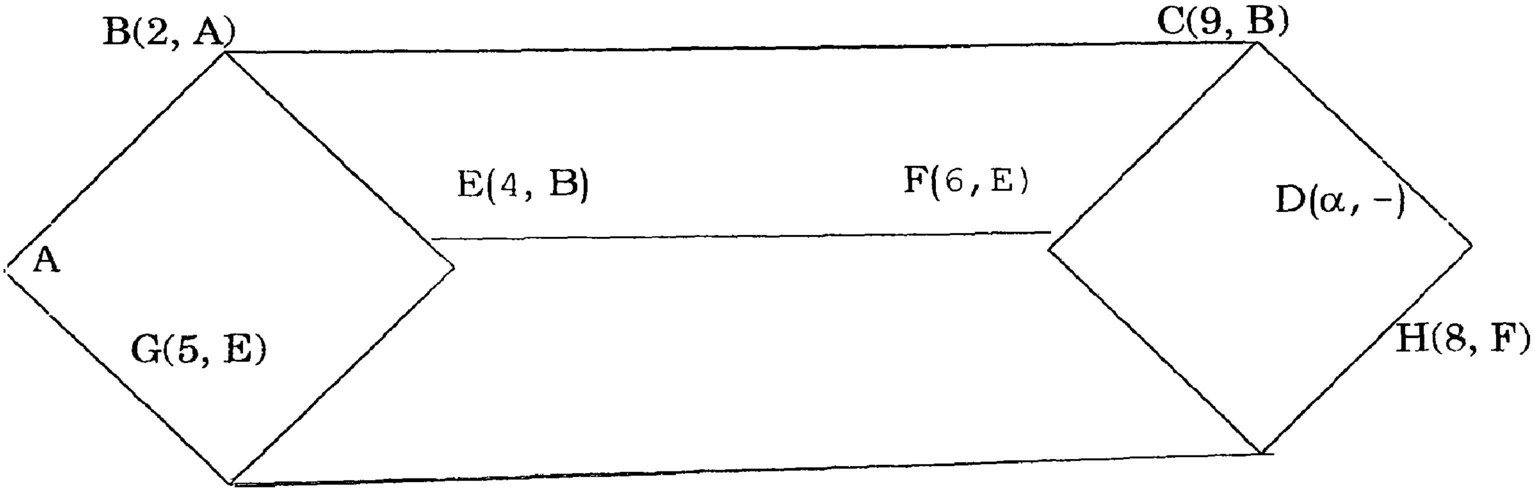
From E, we have two options EG and EF. EG has the pathlength 5 where as EF has 6. So choose EG.



Continue the same procedure for G. The resultant is GH with path length 9. Instead of choosing G, if we select f means there are two possibilities. One is FC and another is FH. C has a length 9 where as H has 8. So we can choose F as a current node.



Now apply the Dijkstra algorithm to this graph.



Here all the nodes are labeled by its predecessor, not its successor. When copying the final path into the output variable, the path is thus reversed. So the answer is produced in the correct order.

INTERNET WORKING

When two or more networks are connected together, it forms an internet. Having different networks invariably means having different protocols. A variety of different networks will always be around, for the following reasons.

1. The installed base of different networks is large and growing like TCP/IP, SNA, PEC net, Novell NCP/IPX or Apple Talk etc.

2. As computers and networks get cheaper the place where decisions get made moves downwards. Classification of purchase based on finance backward enhances it further.

3. Different networks have radically different technology. As new hardware developments occur, new software should be created.

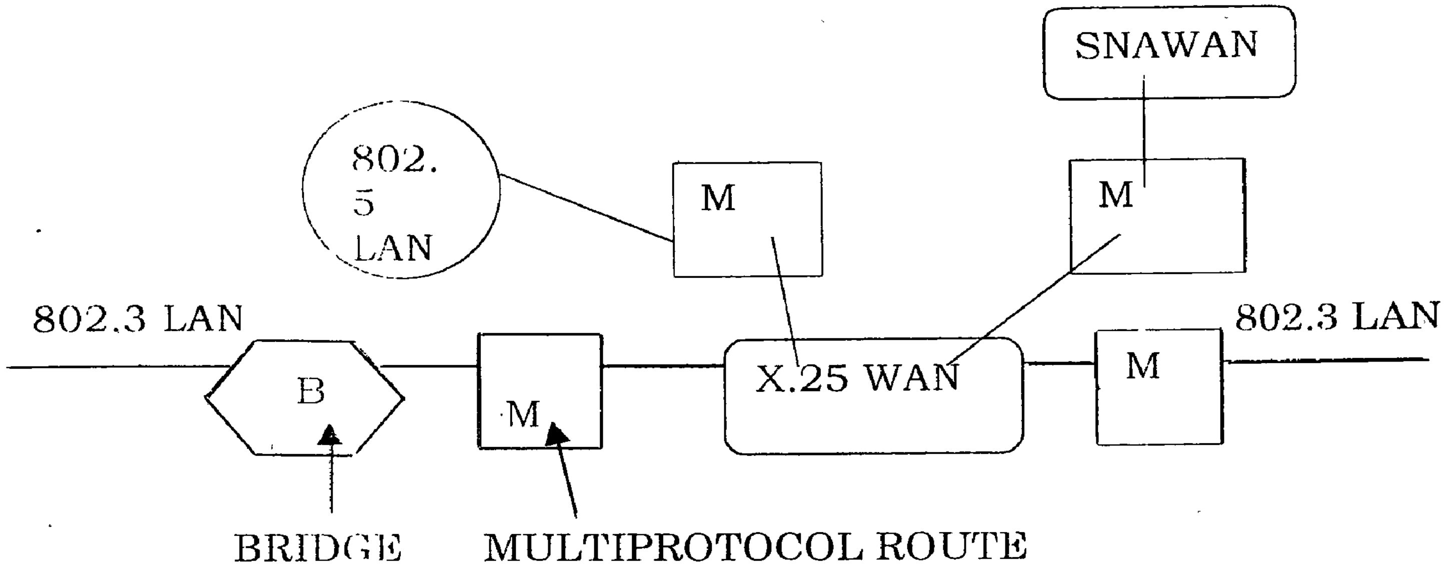
Possible Connections with Possible Examples

1. **LAN-LAN** : A computer scientist downloading a file to engineering.

2. **LAN-WAN** : A computer scientist sending mail to a distant physicist.

3. **WAN-WAN** : Two poets exchanging sonnets.

4. **LAN-LAN-LAN** : Engineers at different universities communicating.



Network Inter Connection

Layer 1 :	Repeaters copy individual bits between cable segments
Layer 2 :	Bridges store and forward data link frame between LANS.
Layer 3 :	Multi protocol routes forward packets between dissimilar networks.
Layer 4 :	Transport gateways connect byte streams in the transport layer.
Above 4 :	Application gateways allow interworking above layer 4.

Gateway is advice that connects two or more dissimilar networks.

Repeaters are low level devices that just amplify or regenerate weak signals. They are needed to provide current to drive long cables.

Bridges store and forward devices. It accepts the entire frame, verifies and then sends to the physical layer.

It can make mirror changes to the frame before forwarding it, such as adding or deleting some fields from the frame header.

Multi Protocol Routes : These are conceptually similar to bridges except that they are found in the network layer.

Transport Gateways : They make connections between two networks at the transport layer.

Application Gateways : It connects two parts of an application in the application layer.

The gateway is effectively ripped apart in the middle and the two parts are connected with a wire. Each of the halves is called a half gateway.

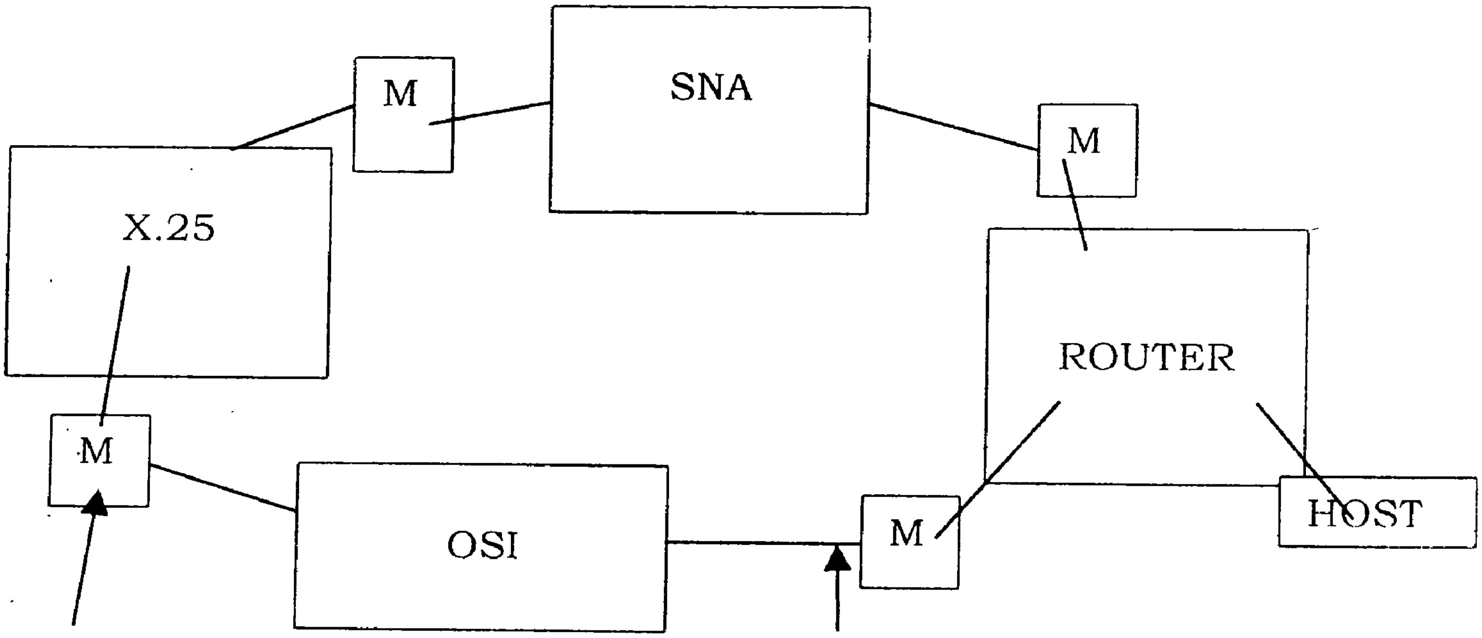
Pure Bridge : Key property is that it examines data link layer frame headers and does not inspect or modify the network layer packets inside the frames.

How Networks Differ : It can differ in many ways like:-

1) The conversions adaptations required when dealing with connection- oriented to connectionless and vice-versa.
2) Maximum packet size, different qualities of service.
3) Error flow, congestion control, security mechanisms, parameter settings and accounting rules.

Concatenated Virtual Circuits

Two Styles of Inter Working are Common : a connection oriented concentration of virtual circuit subnets and a datagram internet style. In this a connection to the host in a distant network is set up in a way similar to the way connections are normally established. The subnet then builds a virtual circuit to the routes nearest to the destination work. It then constructs an external gateway.



Multi protocol Router End to End Concatenated virtual circuits

All data packets must traverse the same sequence of gateways and thus arrive in order. The essential feature of this approach is that a sequence of virtual circuits is setup from the source through one or more gateways to the destination. Concatenated virtual circuits are common in transport layer.

Connectionless Interworking

In this the only service the network layer offers to the transport layer is the ability to inject datagrams into the subnet and hope for the best. It does not requires all packets belonging to one connection.

A decision is made separately for each packet, possibly depending on the traffic at the moment the packet is sent. there is no guarantee that the packets arrive at the destination in order assuming that they arrive at all.

A Connectionless Internet

One serious problem in addressing is sending an IP packet to a host on an adjoining OSI host. OSI hosts do not have 32 bit Internet address.

Advantages

(i) Buffers can be reserved in advance
(ii) Sequencing can be guaranteed.
(iii) Short headers can be used and
(iv) Troubles caused by delayed duplicate packets can be avoided

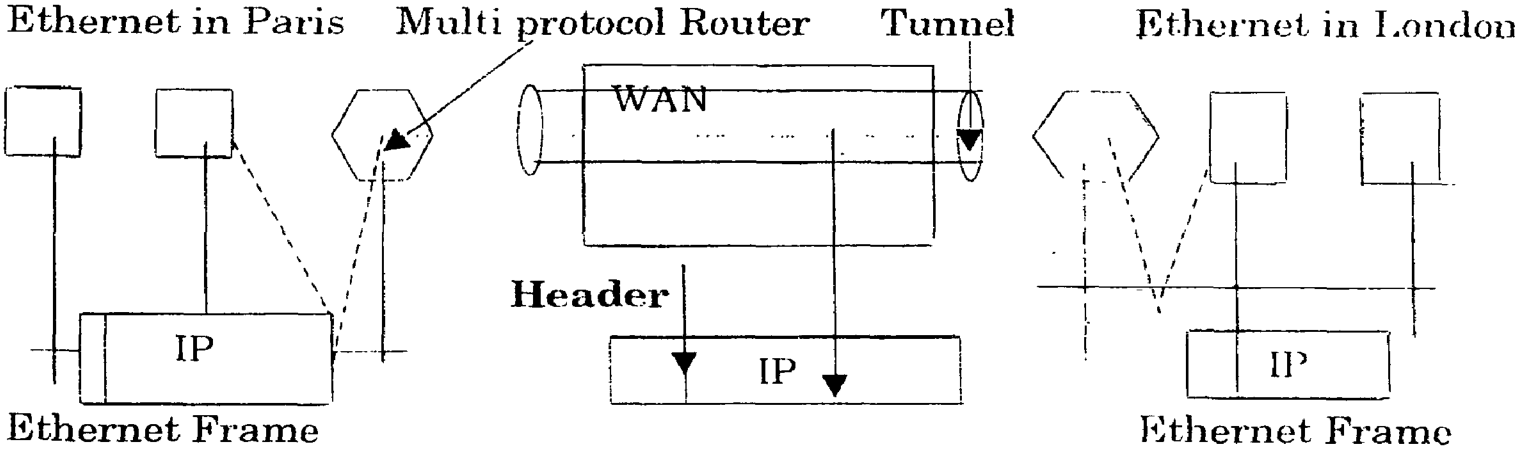
Disadvantages

(i) Table space required in the routes for each open connection
(ii) Vulteineberality to routes failures
(iii) No alternate routing to avoid congested areas.

Tunneling

The problems dealing with different network interworks can be overcome by tunneling. The WAN can be seen as a big tunnel extending from one multi protocol routes to the others. The IP packet just travels from one end of the tunnel to the other, snug in its nice box. It does not have to lay about dealing with the LAN at all. Only the multi protocol routes has to understand IP and LAN packets.

Eg : Tunneling a car from France to England



Internet Work Routing

Routing through an internet work is similar to routing within a single subnet, but with some added compilations. The following is an internet work in which five networks are connected by multi protocol routers.

Drawing a graph is complicated by the fact that every multi protocol router can directly access (is send packets to) every other router connected to any network to which it is connected. B can directly access A and C via network 2 and also D via network

Within each network an interior gateway protocol is used, but between the networks, an extension gateway protocol is used ("gateway" is an olden term for routes).

An Inter Network

Each network in an inter network is independent of all the others and it is referred to as an *Autonomus System (AS)*.

Working : A typical internet packet starts out of its LAN addressed to the local mulitprotocol routes. After it gets there the network layer code decides which protocol routes to forward the Packet by using its own routing tables. If it cannot be forwarded directly, then it is tuned. This process is separated till the packet reaches the destination network.

Difference between internet work routing and inter network routing is that the previous one requires crossing international boundaries.

Fragmentation

Factors that limits (or) imposes some maximum size on its packets are

(i) *Hardware* (eg. the width of a TPM transmission Blot)
(ii) *Operating System* (eg. all buffers are 512 bytes
(iii) *Protocol* (eg. the number of bits in the packet length field
(iv) *Compliance with* some (inter) national standard.
(v) Desire to reduce error induced retransmission to some level.
(vi) Desire to Prevent one packet from occupying channel too long.

Maximum payloads range from 48 bytes (HTM cells) to 65,515 bytes (IP packets), although the payload size in higher layer is often larger.

Working : The solution to the problem of transferring large packet, is to allow gateways to break packets up into fragments, sending each fragment as a separate internet packet. Breaking into small fragments is easier than the reverse process.

In the AIM world, fragmentation is called segmentation. It is simple and has some problems, like

(1) The exit gateway must know when it has received all the pieces so that either a count field on an " end of packet" bit must be included in each packet.
(2) All the packets must exit via the same gateway.
(3) The other strategy is to refrain from combining fragments at any intermediate gateways. Once a packet has been fragmented, each fragment is treated as though it were an original packet. All fragments are passed through the exit gateway. Recombination's access only at the destination host.

Disadvantages

(i) It requires every host to be able, to do reassemble.
(ii) Each fragment must have header when a Packet is fragmented, the fragments must be numbered in such a way that the original data stream can be reconstructed. If packet 0 must be split up, the pieces are called 0.0.0,0.01,0.0.2,.......0.1.2 etc.

However if even one network loses or discards packets, there is a need for end-to-end transmission. with unfortunate effects for the numbering system.

Another method is to define an elementary fragment size small enough such that the elementary fragment can pass through every network. The internet header must provide the original packet number and the number of the (first) elementary fragment contained in the packet.

This method requires two sequence fields in the internet header the original packet number and the fragment number. The limit here is to have the elementary fragment to a single bit or byte, with the fragment number being the bit or byte offset within the original packet.

Firewalls

The ability to connect any computer, to any other computer is a mixed blessing. But there may be changes of information leaking out and information leaking in. Also, Viruses. Worms and other digital pests can breach security. destroy valuable data.

Firewalls are just a modern adaptation of security standby. The firewall in this configuration has two components.

(a) Two routers that has packet filtering, and
(b) Application gateway.

Each Packet Filter

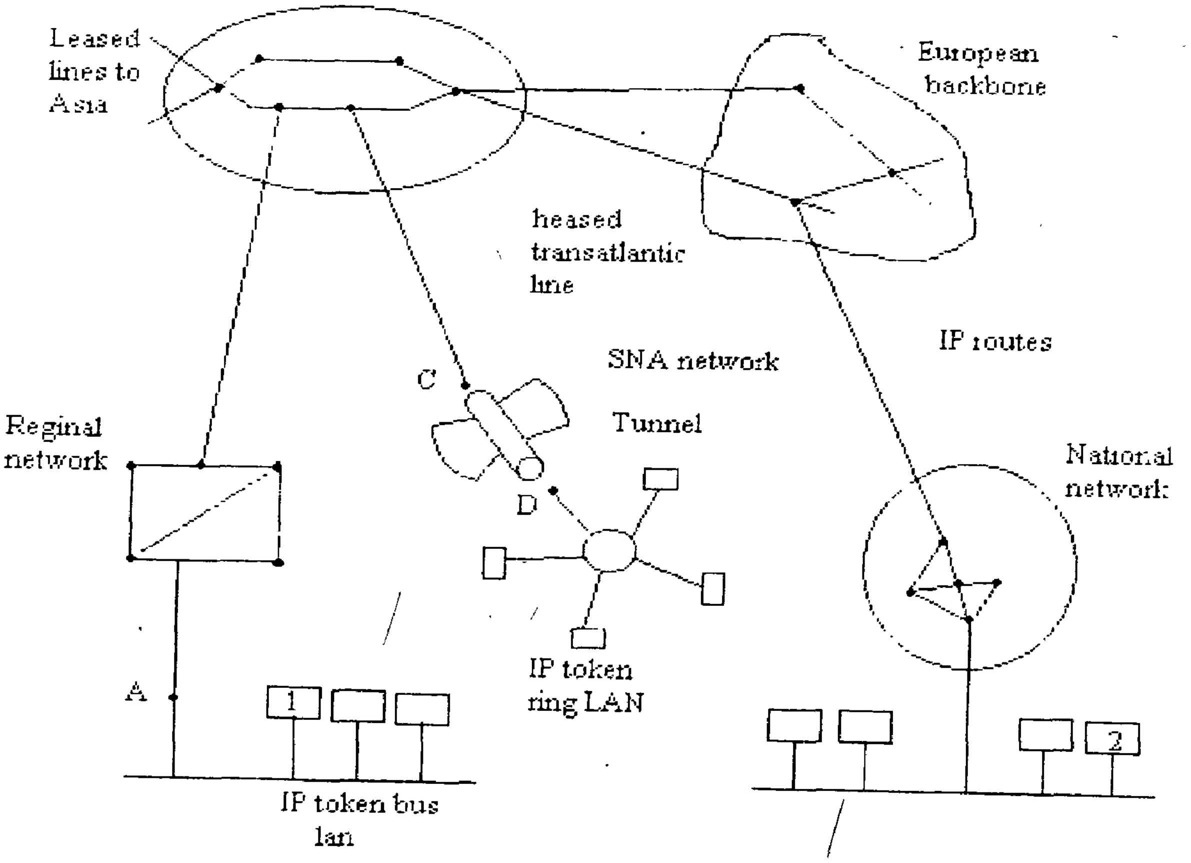
It is a standard router equipped with some extra functionality. The extra functionality allows every incoming or outgoing packet to be inspected. Packets meeting some direction are forwarded normally. The point of putting the two packet filters on different LANs is to ensure that no packet gets in or out without having to pass through the application gateway.

For some important services, such as FTP (File Transfer Protocol) port numbers are assigned dynamically. The second half of the fire-wall mechanism is the *application gateway*. The gateway operates at the application level. A mail gateway, can be set up to examine each message going in or coming out.

Network Layer in the Internet

In network layer,

- The internet can be viewed as a collection of *subnetworks* or *Autonomous systems* (Ases) that are connected together.



- No real structure, but major backbones exist. Attached to the backbones are regional networks, and these regional networks are the LANs at many universities, companies and Internet Service providers.

- The glue that holds the Internet together is the network layer protocol IP(Internet Protocol).

Communication in the Internet

- The transport layer takes data streams and breaks them up into datagrams.

- Each datagram is transmitted through the Internet, possibly fragmented into smaller units.

- When all the pieces finally get to the destination machine, they are reassembled by the network layer into the original datagrams.

- This datagram is then handed to the transport layer, which insist it into the receiving process input stream.

The IP Protocol

The IP datagram consists of

(1) Header part
(2) Text part

The HP Header

Version Field → keeps track of which version of the protocol the datagram belongs to.

IHL → is provided to tell how long the header is, in 32-bit words. The minimum value is 5, which applies when no options are present.

Type of Service → allows the host tell the subnet what kind of service it wants.

Total Length → includes everything in the datagram - both header and data. The maximum length is 65,535 bytes.

Identification Field → is needed to allow the destination host to determine to which datagram a newly arrived fragment belongs to. All the fragments of a datagram contain the same identification value.

Fragment Offset → tells where in the current datagram this fragment belongs.

Time to Line → it is a counter used to limit packet life times. It is supposed to allow counting time in seconds allowing maximum lifetime of 255 seconds.

Protocol Field → tells network layer which transport process to give it to.

Header Checksum → to verify header, useful for detecting errors generated by bad memory words inside the routes.

Source Address and Destination Address → indicate the network number and host number.

Option Field

Five options are defined, but not all routers support all of them.

Option	*Description*
Security →	Specifies how the datagram secret is
Strict source routing →	Gives complete path to be followed
Loose source routing →	Gives a list of routes not to be
Record route →	Makes each route append its address.
Timestamp →	Makes each routes append its address and timestamp

IP Addresses

Every host and routes on the Internet has an IP address which encodes its network number and host number. The combination is unique. No two machines have the same IP address. All IP addresses are 32 bits long and are used in the source and destination address of IP packets.

SUBNETS

- As we have known that all the host in a network must have the same network number.
- The above property causes problems as networks grow.
- If the number of distinct local networks grow, managing them can become a serious headache.
- Moving a machine from one LAN to other cause serious problem.

The above are the problems. The solution to the above problems is to allow a network to be split into several parts for internal use but still act like a single network to the outside world. In the internal literature these parts are known as the subnets.

Internet Control Message

The operation of the internet is monitored closely by the routers. When something unexpected occurs, the event is reported by the ICMP, which is also used to test the internet.

About a dozen types of ICMP messages are defined. Each ICMP message type is encapsulated in a IP packet.

The Principal ICMP Message Types

Message type	*Description*
Destination unreachable	Packet could not be delivered.
Time exceeded	Time to leave field bit 0.
Parameter problem	Invated headed field
Sassee quench	Choke pocket
Redirect	Teach a router, flow geography.
Echo report	Alert, a machine it is alive
Echo reply	Yes, I am alive
Timestand report	Same as echo but with timestamp
Timestand reply	Same as Echo reply, with timestamp.

The Address Resolution Protocol

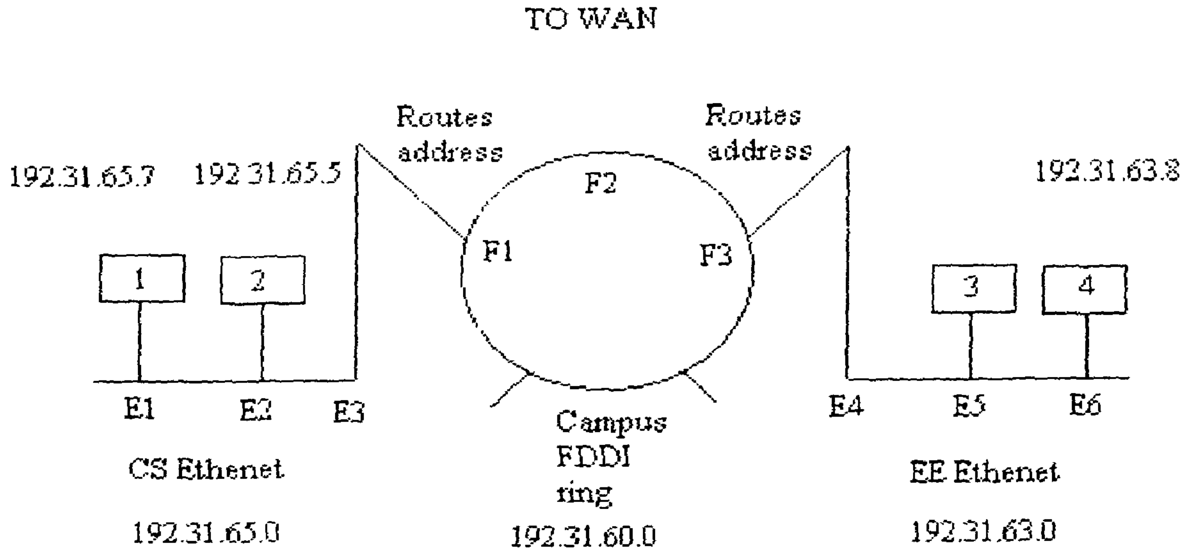
- Although every machine on the Internet has one (or more) IP addresses, these cannot actually be used for sending packets because the data link layer hardware does not understand Internet addresses.

- Most of the hosts are attached to a LAN by an *interface board* that only understands LAN addresses. The boards send and receive data based on 48-bit Ethernet addresses.

The question is

How to IP addresses get mapped onto data link layer addresses, such as Ethernet?

Working Diagrammatic Representation



Host 1 Sending Packet to Host 2

- ✱ First finds IP address to host 2, this is performed by Domain name system. DNS return IP address for host 2.

- ✱ The upper layer of the software on host 1 now builds a packet with host 2 address in destination address field and gives it to the IP software to transmit.

- ✱ The IP software can look at the address and see that it's destination is on its own network, but it needs a way to find the destination's Ethernet address.

The Protocol ARP (Address Resolution Protocol) is also used for this, but it sometimes fails to support fully. So other solution is with the use of proxy ARP.

The second solution is to have host 1 immediately see that the destination is on a remote network and just send all the traffic to default Ethernet address.

The Reverse Address Resolution Protocol

RARP solves problem of which ethernet address corresponds to a given IP address. Sometimes Reverse problem has to be solved.

Issue : Given an ethernet address what is the IP address?

This problem occurs when routing a diskless workstation. But how does it lead it to IP address.

The solution is to use RARP (Reverse Address Resolution Protocol). This protocol allows a newly routed workstation to broadcast its ethernet address.

The RARP server sees this request, look up ethernet address in its configuration files and sends to the corresponding IP address.

The Interior Gateway Routing Protocol : OSPF

Internet is made up of a large number of autonomous systems. Each autonomous system is operated by a different organization and can use its own routing algorithm inside. A routing algorithm within is called an interior gateway protocol. An algorithm for routing between Ases is called an exterior gateway protocol.

The distance vector protocol (Internet interior gateway protocol) worked well in small systems but failed when the autonomous systems (AS) that are bigger. It suffered from count to infinity problem and generally slow convergence. So it was replaced by a link state protocol. Open shortest path First (OSPF) will become the main Interior Gateway Protocol in the near future because many vendors are supporting it.

The requirements in designing the new protocol:

(1) The algorithm should be published in open literature.

(2) It should support a variety of distant machines.

(3) The algorithm should be dynamic (ie) adapted to changes in the topology.

(4) The real-time traffic should be routed the one way and the other traffic, a different way.

(5) Load balancing should be done (ie) splitting the load to multiple line rather than sending all the packets over the best route.

(6) Support for hierarchical systems was needed.

(7) The fun-loving students should be prevented from spooling routers by sending them false routing information. So proper security should be given.

(8) Proper provision was needed for dealing with routers that were connected to the Internet via a tunnel.

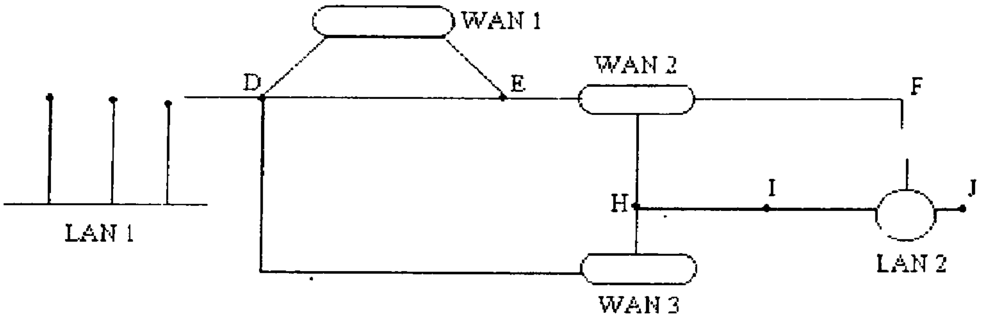
OSPF supports 3 kinds of connections between 2 networks.

(1) Point -to - point between exactly two networks.

(2) Multi-access Networks with broad casting (eg) LAN'S

(3) Multi-access Networks without broad casting. (eg) packet switched Wan's.

A multi-access network is one, that can have multiple routers on it, each of which can directly communicate with all the others.



A Graph Representation of an AS

The Exterior Gateway Routing Protocol - BGP

Within a single As, the recommended routing protocol on the Internet is OSPF (although it is certainly not the only one in use). Between Ases, a different protocol, BGP (Border Gateway Protocol), is used. A different protocol is needed between Ases because the goals of an interior gateway, protocol and an exterior gateway protocol are not the same. All an interior gateway protocol has to do is move packets as efficiently as possible from the source to the destination. It does not have to worry about politics.

Exterior gateway protocol routers have to worry about politics a great deal. For example, a corporate As might want the ability to send packets to any Internet site and receive packets from any Internet site. However, it might be unwilling to carry transit packets originating in a foreign As and ending in a different foreign As, even if its own As was on the shortest path between the two foreign Ases ("That's their problem, not ours") on the other hand, it might be willing to carry transit traffic for its neighbours, or even for specific other Ases that paid it for this service. Telephone companies, for example, might be happy to act as a carrier for their customers, but not for others. Exterior gateway protocols, in general and BGP in particular have been designed to allow many kinds of routing policies to be enforced in the inter as traffic.

Typical policies involve political security or economic consideration. A new examples of routing constraints are,

1. No transit traffic through certain Ases.

2. Never put Iraq on a route starting at the Pentagon.

3. Do not use the United States to get from British Columbia to Ontario.

4. Only transit Albania if there is no alternative to the destination.

5. Traffic starting or ending at IBM should not transit Microsoft.

Policies are manually configured into each BGP router. They are not part of the protocol itself.

From the point of view of a BGP router the world consists of other BGP routers and lines connecting them. Two BGP routers are considered connected if they share a common network. Given BGP's special interest in transit traffic, networks are grouped into one of three categories. The first category is the "Stub Networks" which have only one connection to the BGP graph. These cannot be used for transit traffic because there is no one on the other side. Then come the "multi connected networks". These could be used for transit traffic, except that they refuse. Finally, there are the "transit networks", such as backbones which are willing to handle third party packets possibly with some restrictions.

Pairs of BGP routers communicate with each other by establishing TCP connections. Operating this way provides reliable communications and hides all the details of the network being passed through.

Internet Multicasting

Normal IP communication is between one sender and one receiver. However, for some applications it is useful for a process to be able to send to a large number of receivers simultaneously. Examples are updating replicated, distributed databases, transmitting stock quotes to multiple brokers, and handling digital conference ((i.e.) multiparty) telephone calls.

IP supports multicasting. Twenty-eight bits are available for identifying groups, so over 250 million groups can exist at the same time. When a process sends a packet, a best-efforts attempt is made to deliver it to all the members of the group addressed, but no guarantees are given some members may not get the packet.

Two kinds of group addresses are supported: Permanent addresses and temporary ones. A permanent group is always there and does not have to be set up. Each permanent group has a permanent group address. Some examples of permanent group addresses are:

224.0.0.1.1	All systems on a LAN
224.0.0.1.2	All routers on a LAN
224.0.0.5.1	All OSPF routers on a LAN
224.0.0.5.2	All designated OSPF routers on a LAN.

Temporary groups must be created before they can be used. A process can ask its host to join a specific group. It can also ask its host to leave the group. When the last process on a host leaves a group, that group is no longer present on the host. Each host keeps track of which groups, its processes currently belong to.

Multicasting is implemented by a special multicast routers.

These query and response packets use a protocol called IGMP (Internet Group Management Protocol), which is vaguely analogous to ICMP. It has only two kinds of packets : query and response, each with a simple fixed format containing some control information in the first word of the payload field and address in the second word. It is described in RFC 1112.

Multicast routing is done using spanning trees. Each multicast router exchanges information with its neighbours using a modified distance vector protocol. Various optimization are used to prone the tree to eliminate routers and networks not interested in particular, groups. The protocol makes heavy use of tunneling to avoid bothering nodes not in a spanning tree.

Mobile IP

Mobile IP makes working far from home easier. Every IP address contains 3 fields. class, the network number and the host number. Routes all the world have routing table telling which line to use to get to the network. Routers can also use complete IP addresses for routing instead of just the class and N/W. The working group have formulated number of goals that are as follows:

1. Each mobile host must be able to use its home IP address anywhere.

2. Software changes to the fixed host was not permitted.

3. Changes to the router software and tables were not persisted.

4. Most packets are mobile hosts, should not make detours on the way.

5. No overhead should be incurred when a mobile host is at home.

Gratuitous ARP

At the time, the mobile host moves the router probably has its Ethernet address catched. To replace that Ethernet address with some home agents, a trick called Gratuitous ARP, it is a special unsolicited message to the router.

When a mobile host arrives somewhere it can just listen some broad cast called advertisement. The IETF soul for mobile host solves a number of problems. Cryptographic Authentication protocols are used for security purpose. Two levels of mobility are available.

CIDR - Clember Inter Domain Routing

For most organization a network with 16 million addresses is too big and a class C network with 256 address is too small. A clear B network with 65,536 is just right. In network this situation is known as three Bears pbm (as in Goldilock and the three Bears).

The main disadvantage of IP is its running out of addresses.

Solution to the IP address problem can be overcome by CIDR. It is now being implemented and which will give the internet a bit of extra breathing room.

The allocation was as follows

Addresses 194.0.0.0 to 195-255. 255-255 are for Europe

Addresses 198.0.0.0 to 199.255.255.255 are for N.America

Address 200.0.0.0 to 201.255.258.255 are for Central and S.America

Addresses 202.0.0.0 to 203.255.255.255 are for Asia and Pacific

The world was partitioned into four zones and each one given a portion of the clear in C address space.

Advantages

It is now possible that any routes outside of Europe that gets addresses to 194.xx.yy.33 (or) 194.xx.yy.33 can just send it to its standard European gateway.

IPVF

IETF started work on a new Version on IP, would solve a variety of other problems and be more flexible and efficient. Its major goals were to

1. Reduce the size of routing table.
2. Provide better security
3. Aid multicasting
4. Simplification of Protocol.
5. Allow the protocol to evolve in the futures.
6. Make it possible for a host to roam without changing its address.
7. Support billions of hosts, even with inefficient address space allocation.
8. Pay more attention to type of real time data service.

SIPP - Simple Internet Protocol Plus

This was selected and given the designation IPV6. It meets goals, maintain good features of IP, discards bad ones and add new ones.

Features

1. It has longer addresses than IPV4, they are 16 bytes long.
2. Simplification of headers, it contains only seven fields.
3. Better supports for options. This feature speeds up packet processing time.
4. It represents a big advancement in security.
5. More attention is paid to type of service than in the past.

The Main IPV6 Header

The version field is always 6 for IPV6. The priority field is used to distinguish between packets whose sources can be flow controlled and those that cannot. The distinction of values allows routers to deal with packets better in the event of congestion.

The Flow table will be used to allow a source and distinction to set up a pseudo connection with particular properties and requirements. When a packet with a non zero flow label shows up, all the routers can look it up in internal tables to see what kind of special treatment it requires.

Each flow is designated by the source address, destination address and flow number. It is expected that flow numbers will be chosen randomly rather than assigned sequentially starting at 1, to make it easy for routers to ask them.

The payload length field tells how many bytes follow the 40 bytes header.

The next header field lets the cat out of the bag. The reason the header could be simplified is that there can be additional (optional) extension headers.

In addition to supporting the standard unicast and multicast addresses. IPV6 also supports a new kind of addressing any cast. Any casting is like multicasting in that the destination is a group of addresses. A new notation has been devised for writing 16-byte addresses. They are written as eight groups of four hexadecimal digits with colons between the groups, like this : 8000 : 0000 : 0000 : 0000 : 0123 : 4567 : 89AB : CDEF.

IPV4 addresses can be written as pair of colons and as an old dotted decimal number for ex. ::192.31.20.46. The ±HL field is gone because the IPV6 header has a fixed length. The protocol field was taken out because the next header field tells what follows the last IP header.

The check sum field is gone because calculating it greatly reduces performance.

Extension Header

Extension header can be supplied to provide extra information. but encoded in an efficient way. Six kinds of extension headers are defined. each one is optional.

Extension header	Description
Hop by hop options	Miscellaneous information for routers
Routing	Full or partial route to follow
Fragmentation	Management of datagram fragments.
Authentication	Verification of the sender's identity
Encrypted security payload	Information about the encrypted contents
Destination options	Additional information for the destination

The following type is 1 byte field telling which option is selected. The choices are a skip the option, discard the packet, discard the packet and send back an ICMP packet. The length is also a 1-byte field. It tells how long the value is. The value is any information required up to 255 bytes.

NEXT header	0	194	0
Jumbo payload length			

Jumbo Grams

Datagrams using the following type of header extension are called Jumbograms. Sizes less than 651536 are not permitted and will result in the first route discarding the packet sending back an ICMP error message.

NEXT header	0	Number of addresses	Next address
Bit MAP.			
1 - 24 Address			

The extension header is for routing. The Fragment header deals with fragmentation. The header holds the datagram identifier, fragment number and a bit telling whether more fragments will follow.

The Network Layer in ATM Networks

ATM - Asynchronous Transfer Mode.

In network system the lowest layer that goes from source to destination and thus involves routing and switching is the network layer. The ATM layer deals with moving cells from source to destination and hence involves routing and switching and hence functionally the ATM layer performs the work expected of the network layer.

The more advantageous part of ATM layer is it resembles layer of x.25 which everyone agrees is a network layer protocol. It has the characteristics of a network layer protocol end_to_end virtual circuits, switching and routing.

ATM layer is connection oriented. The basic element of ATM layer is the virtual circuit. It is nothing but a connection from one source to one destination, although multicast connections are permitted. Virtual circuits are unidirectional. Furthermore ATM networks are often used for real-time traffic, such as audio and video. It supports two level connection hierarchies that is visible to the transport layer. A group of virtual circuits can be grouped to what is called a virtual path.

The main disadvantages of ATM layer is its lack of acknowledgement but it does periodic one guarantee that cells sent along a virtual circuit will hence arrive out of order. This chapter fully deals with ATM structure in network layer and its applications.

Cell Format

The ATM has two faces

UNI (user-network interfaces) [host - ATM]

NNI (network-network interfaces) [ATM -ATM]

The difference is in the arrangement of headers in the structure cell format.

UNI-ATM Header

NNI -ATM-Header

VPI	VCI	P T I	C L P	HEC

Connection Setup

ATM supports both permanent virtual circuits and switched virtual circuits. The following describes how switched virtual circuits are established.

The normal way to connect is

- First acquire a VC for signaling and use it
- Cells containing request are sent.
- If successful, a new Vc is opened in which connection setup request are sent and received.

They generally use 6 message types.

Message	Meaning
Set up	Sent by host to establish a circuit
Call proceeding	Call request is Attempted.
Connect	accept the call
Connect act	thanks for accepting
Release	terminate call
Release complete	Act for release

Steps involved in establishing connection.

(1) Sent set up message on a Vc.

(2) Network responds with call proceeding.

(3) The setup is acknowledged by host.

(4) The destination host responds with connect to amp.

(5) The network then sends connect acknowledgement message.

(6) Connect message propagates towards source.

(7) The source acknowledges by connect ack.

Releasing Connection

(1) Source sets up release.

(2) Destination receives release and Sends.

Release Complete to Source

(3) Connection is released when Release complete reaches source.

Routing and Switching

When virtual circuit is setup, the setup message wants its way through the network from source to destination. The ATM standard does not specify any particular routing algorithm. So the carrier is free to choose among the algorithms discussed earlier in the systems data link layer.

The general idea is to route on the VPI field, but not the VCI field except at the final hop in each direction, when cells are sent between a switch and a host. Between two switches only virtual path was to be used. Routing of individual cells is easier when all virtual circuits for a given path are always in the same bundle, basing all routing on virtual paths generally makes it easier to switch a whole group of virtual circuits, hence VPIs between interior switches has several advantages.

The VPI introduction hence makes it easier for carriers to offer closed user groups to corporate astronomers. Let us consider a circuit in real time to see Vc paths. Consider one aim is to concentrate on switch a, it has five lines from o → 4, when swtich is locked all the entries in the VPI -table structure are marked as not in use.

As each virtual path is setup, entries are made in the table as shown where the structure is given for switch among many switches in the circuits.

Source	Inline	InVpi	Dest	Outline	OutVpi	Path
NY	1	1	SF	4	1	New
SF	4	5	M	0	2	New
NY	1	1	SF	4	1	Old

The 1st call generates the (4,1) entry for VPI line the DC table because it refers to cells coming in online ’I’ with VPI ’I’ and going to SF. An entry is also made in the D table for VPI.

We can explain how cells are processed inside a switch. Suppose that a cell arrives on line 1 with VPI 3. The switch hardware or software uses the 3 as an index into the table for line 1 and see that the cell should go out on line 3 with VPI 2. It overwrites the VPI field with a2 and puts the outgoing line number 3 somewhere in the cell. At this point, it is straight forward to see how an entire bundle of virtual circuits can be rerouted as it done on the diagram of circuits and to process it differ in table entries for each routes are written with respect to incoming VPI’s and outgoing line along with VPI

Quality of Service (QOS)

Quality of service is an important issue for ATM network. The system deals with legal contract between the customer and network based on traffic offered service agreed and compliance recreant. To make it possible to have concrete traffic central the ATM standard defines a number of QOS paranets.

PCR (Peak Cell Rate)

The maximum rate at which the sender is planning to send cells.

SCR (Sustained Cell Rate)

The expected or required cell rate averaged over a long time interval.

MCR (Minimum Cell Rate)

It is the minimum number of cells /sec that the customer considers acceptable

CVDT (Cell Variation Delay Tolerance)

The parametre tells how much variation will be present in cell transmissions.

LR (Cell Loss Ratio)

It measures the fraction of transmitted cells that are not delivered at all or elivered so late.

TD (Cell Transfer Delay)

It is the average transit line from source to destination.

CDV (Cell Delay Variation)

Measures how uniformly the cells are delivered. It is thus the routing of virtual circuits are done.

Service Categories

There are several customers requesting services that are available. The service categories are given below:

CBR (Constant Bit Rate)

They are essentially to make a smooth transition between the current telephone system and future B-ISDN (i,e) Bits are put at one end come off at other end.

ABR (Available Bit Rate)

They are designed for burst traffic whose full bandwidth range is known roughly. It avoids making a long-term commitment to a fixed bandwidth and provides a rate of feedback to the sender.

UBR (Unspecified Bit Rate)

This is well suited to send IP packets, since IP also makes no about delivery. If congestion occurs, UBR cells will be discarded, with no feedback and no expectation.

CER (Cell Error Ratio)

It is the fraction of cells that are delivered with one (or) more bits wrong.

SCEBR (Severaly-errored Cell Block Ratio)

It is the fraction of N-cell blocks of which M (or) more cells are error

CMR (Cell Missinesertion Rate)

It is the number of cells/sec that are there delivered to the wrong destination.

Traffic Shaping and Policing

The mechanism for using and enforcing the quality of service parameters is based on the generic cell rate algorithm (GCRA). It has two parameters one is the PCR - arrived rate and the (CDVT) - event of variation that are tolerable. The reciprocal of PCR, T=1/PCR is the minimum cell inter arrival time. A sender is always permitted to spare consecutive cells more widely than T.

The problem arise with senders that tend to jump the gun (i,e) the cell that arrives a little early (or later than t, + T-L)

If the cell arrives more than L micro sec early it is declared as non-conforming. The treatment of nonconforming cells is up to the carrier. Some carrier may simply discard them others may keep them, but set the CLP bit, to mark them as low priority to allow switches to drop non conforming cells first in the congestion.

Congestion Control

ATM networks do not automatically meet the performance requirements set forth in the traffic contract. For eg. Congestion at intermediate switches is always a potential problem. ATM networks deal with both long-term congestion, caused by more traffic and short-term congestion, caused by business in the traffic.

Admission Control

In a network it is usually adequate to wait for congestion to occur and then react to it by telling the source of the packets to slow down. In high speed network, this approach often works poorly, because in the interval between sending the notification and notification arriving at the source, thousands of additional packets may arrive. A major tool is admission control where the host wants a new virtual circuit, it must decide the traffic to be offered and the service expected.

Rate-based Congestion Control

With CBR and VBR traffic it is generally not possible for sender to slow down, even in the event of congestion due to the inherent real-time (or) semi-realtime nature of the information source.

However, with ABR traffic it is possible and reasonable for the network to signal one or more senders and ask them to slow down temporarily until the network can recover.

The next way is to mark the packet boundaries by a bit in last cell to relieve the congestion. The next argument is to send a RM cell (resource management) cell after every K cell. This cell travels along the same path as the datacell, but it is treated specially by the switches along the way when it gets to the destination it is examined, updated and sent back to sender.

Flow Based Routing

The static algorithm which uses both topology and local for routing is called flow based routing.

The following techniques must be known

1. The subnet topology.
2. The traffic(f_{ij})
3. Line capacity matrix (C_{ij}) (bps)
4. Routing algorithm must be chosen.

E.g: ***Full Duplex Subnet***

[illegible] Here the access give capacities (C_{ij}) in each direction measured in kbps.

The Traffic in Packets /Sec and Routing Matrix

	A	B	C	D	E	F
A		9 AB	4 ABC	1 ABFD	7 AE	4 AEF
B	9 BA		8 BC	3 BFD	2 BFE	4 BF
C	4 CBA	8 CB		3 CD	3 CE	2 CEF
D	1 DFBA	3 DFB	3 DC		3 DCE	4 DF
E	7 EA	2EFB	3 EC	3 ECD		5 EF
F	4 FEA	4 FB	2 FEC	4 FD	5 FE	

Here the matrix has as entry for each source destination pair. The entry for source i to destination j shows route to be used for ij traffic and no. of packets/sec to be sent from source to its destination j.

Eg: · 3 packets/sec go from B to D using route BFD is 3 packets /sec to each of 2 lines (BF line and FD line).

Analysis Of The Above Sub Net Using A Mean Packet Size Of 800 Bits:

i	line	λi (packets /sec)	Ci (kbps)	μci(packet /sec	Ti(mscs)	wt
1	AB	14	20	25	91	0.171
2	BC	12	20	25	77	0.146
3	CD	6	10	12.5	154	0.073
4	AE	11	20	25	71	0.134
5	EF	13	50	62.5	20	0.159
6	FD	8	10	12.5	222	0.098
7	BF	10	20	25	67	0.122
8	EC	8	20	25	59	0.098

λi = Total traffic in each bound line. Eg; Here all traffic is symmetric in xy traffic is identified to yx traffic ∇*dy

1) The next column gives mean delay for each line derived by queuing theory formula,

$$T = \frac{1}{\mu c - \lambda}$$

where $1/\mu$ = mean packet size in bits
C = capacity in bps.
λ = mean flow in packets/sec.

Eg.:	Let capacity μc	= 25 packets/sec.
	Actual flow λ	= 14 packets/sec.
	Mean delay	= 91 msec.

bits λ=0, mean delay is still 40msec. Since capacity is 25 packets/sec.

Distance Vector Routing

Distance vector routing algorithms operate by having each route maintain a table giving the best known distance to each destination and which the line to use to get there.

eg. A subnet

(a) 1/p from A, I, H, K and new routing table

A	I	H	K	New estimated delay from J line	
0	24	20	21	8	A
12	36	31	28	20	A
25	18	19	36	28	I
40	27	8	24	20	H

(b) 1/p from A,I,H,K and new routing table

'A' claims to have 12 msec. delay to B, 25 msec. delay to C, 40 msec. delayed to 10 etc., Suppose J has measured or estimated its delay to neighbours A, I, H and K as 8, 10, 12 and 6 msec. ie., spetunets.

<u>**Eg :**</u> J computers its new route to route G. A-8 msec. A to G=18 sec. J has delay of 26 msec. to G. Similarly it computes delay to G nil I, H and K as 41(31+10), 18 (6+12) and 37(31+6) msec. respectively.

The best of these values = 18 (route to use is via it)

Routing Algorithms

Link State Routing

The idea behind link state routing is simple and can be stated as five parts. Each router must

1. Discover its neighbour and learn their network address.
2. Measure the delay or cost to each of its neighbour.
3. Construct a packet telling all it has just learned.
4. Send their packet to all other routers.
5. Compute the shortest path to every other router.

The complete technology and all delays are experimentally measured and distributed to every router. Then Dijkstra's algorithm can be used to find shortest path to every other router.

Learning about the Neighbours

When a router is spotted, its first task is to learn who its neighbours are. It accomplishes their goal by sending a special HELLO packet to each point-to-point line. The router in the other hand is expected to send back a reply telling whom it is. These names must be globally unique.

Measuring Line Cort

The link state routing algorithm requires each router to know, or atleast have a reasonable estimate of the delay to each of its neighbours. The most direct way to determine this delay is to send a special ECM packet over the line that the other side is required to send back immediately. By measuring the sound-trip time and dividing it by two, the sending router can get a reasonable estimate of the delay.

Routing for Mobile Hosts

The mobile hosts introduce a new complication to route a packet to a mobile.

Here we have a WAN continuity of routers and hosts. Users who never move are said to be stationary. They are connected to the network by copper wires or fiber optics.

All users are assumed to have a permanent home location that never changes. Users also have a permanent home address.

The routing goal in systems with mobile users is to make it possible to send packets to mobile users using their home address and have the packets efficiently reach them, whenever they may be.

The world is divided up into small units called areas. Each area has one or more foreign agents. In addition, each area has a home agent, which keeps track of its users.

When a new user enters on area, either by connecting to it, or just wondering into the all, his computer must register itself. The registration procedure works like

1) Periodically, each foreign agent broadcast a packet announcing its existence and address. The mobile host may wait for one of these massages.

2) The mobile host contacts with the foreign agent giving its home address.

3) The foreign agent contacts the mobile host's home agents.

4) The home agent confirms to security information.

5) When the foreign agents gets the acknowledgement from the home agent, it make an entry in its tables and informs the mobile host that it is now registered.

Packet Routing for Mobile Hosts

1) Packet is sent to the mobile host's home address.

2) Packet is termed to the foreign agent.

3) Sender is given foreign agent's address.

4) Subsequent packets are tunneled to the foreign agents.

Broadcast Routing

Hosts need to send messages to many or all other hosts.

Ex : Weather reports, stock market update sending a packet to all destination simultaneously is called broadcasting.

One broadcasting method that requires no special features from the subnet is for the source to simply send a distinct packet to each destination. Not only is the method wasteful of bandwidth.

The problem with flooding as a broadcast technique is the same problem it has as a point to point routing algorithm. It generates too many packets and consumes too much bandwidths.

Multidestination Routing

If this method is used, each packet contains either a list of destination or a bit map indicating the desired destination. When a packet arrives at a router, the router checks all the destination to determine the set of O/P (output) lines that will be needed.

The router generates a new copy of the packet for each O/P line to be used and includes in each packet only those destination that are to use the line.

In effect the destination set is partitioned among the O/P lines. After a sufficient number of hops, each packet will carry only one destination and can be treated as a normal packet. Multi destination routing is like separately addressed packets except that when several packets follow the same route, one of them pays full fare and the rest ride free.

A spanning tree is a subset of the subnet that includes all the routers but contain no loops. This method makes excellent use of bandwidth, generating the absolute minimum number of packets necessary to do the job.

When a broadcast packet arrives at a router, the router checks to see if the packet arrived on the line that is normally used for sending packets to the source of the broadcast. If so, there is an excellent chance that the broadcast packet itself follows the best route from the router and is therefore the first copy to arrive at the router.

Multicast Routing

Sending a message to such a widely separated processes work together in groups is called multicasting and its routing algorithm is called multicast routing.

To do multicasting, group management is required some way is needed to create and destroy groups and for processes to join and leave groups.

To do multicast routing, each router has a spanning tree covering all other routers in the subnet.

When a process sends a multicast packet to a group, the first router examines the spanning tree and prunes. It removes all lines that do not lead to hosts that are members of the groups.

Multicast packets are forwarded only along the spanning tree. The spanning tree can be pruned by starting at the end of each path and working towards the root, removing all routers that do not belong to the group in question.

Disadvantages of this Algorithm

A network has n groups, each with an average of m members. For each group, m pruned spanning tree must be stored for a total of mn trees.

Lesson - 7

TRANSPORT LAYER

Without the transport layer, the whole concept of layered protocols would make little sense. The transport layer is not just another layer, its task is to provide reliable data transport from the source machine to the destination machine, independent of the physical network or networks currently in use. Let us discuss in detail, the services, design, protocols and performance.

Services provided to the Upper Layers

The ultimate goal of the transport layer is to provide efficient reliable and cost-effective service to its users, normally processes in the application layer. The transport layer makes use of the services provided by the network layer. The hardware & software within the transport layer that does the work is called the transport entity. The transport entity can be in the operating system Kernel, in a separate user process, in a library package bound into network applications or on the network interface card. In some cases, the carries may even provide reliable transport service, in which case the transport entity lives on special interface machines at the edge of the subnet to which hosts connect the logical relationship of the network, transport and application layer. This is given below:

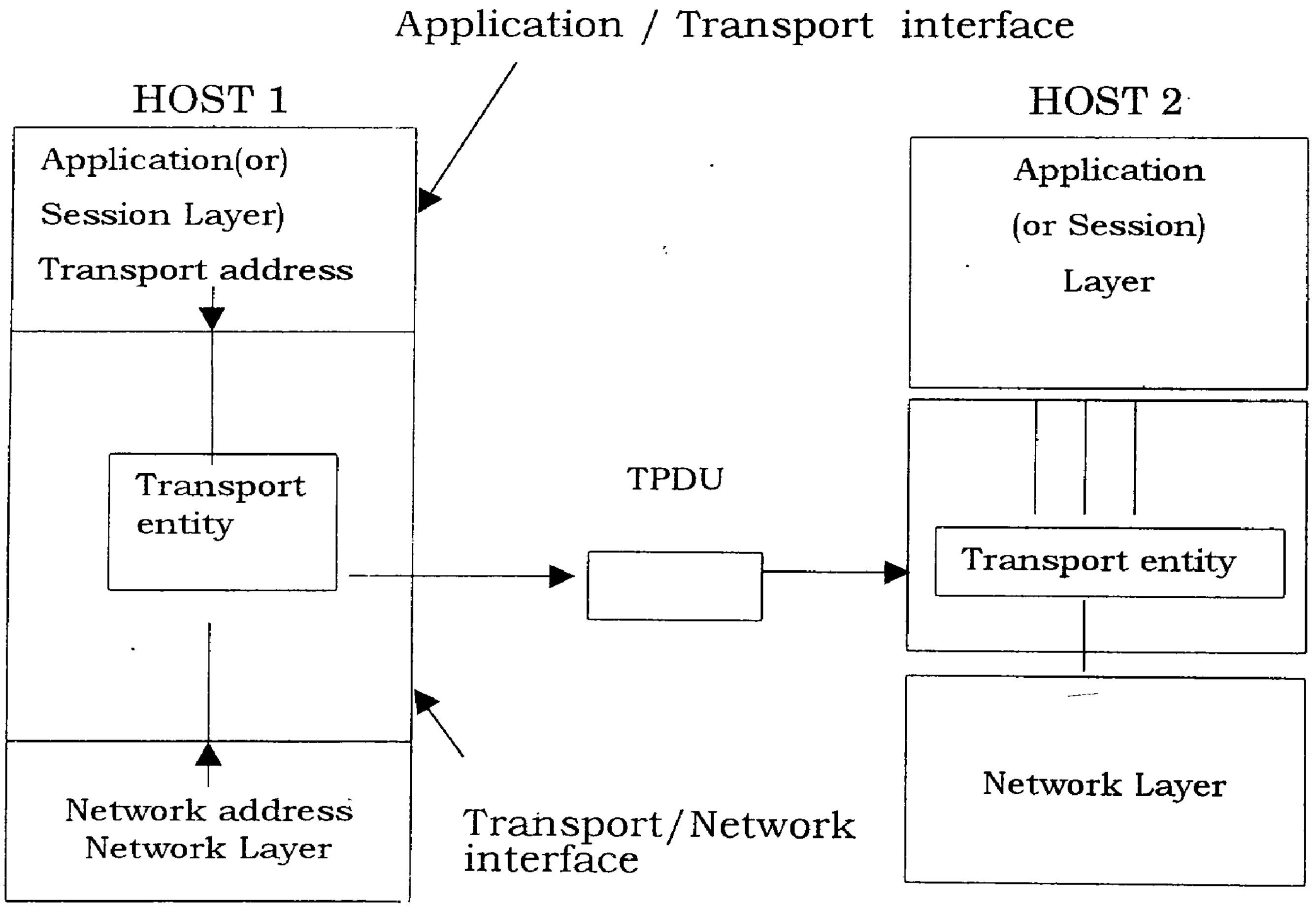


Fig. 7.1 Network, Transport of Application Layer

Two types of network services: connection - oriented and connectionless. The connection-oriented transport service is similar to the connection-oriented network service in many ways. In both cases, connections have three phases :

- establishment
- data transfer and
- release

Addressing and flow controls are also similar in both layers. The connectionless transport service is also very similar to the connectionless network services.

If a transport entity is informed halfway through a long transmission that its network connection has been abruptly terminated, with no indication of what has happened to the data currently in transit, it can set up a new network connection to the remote transport entity. Using this new network connection, it can send a query to its peer asking which data arrived and which did not, and then pick up from where it left off.

The existence of the transport layer makes it possible for the transport service to be more reliable than the underlying network service. Lost packets and managed data can be detected and compensated for by the transport layer. The transport service primitives can be designed to be independent of the network service primitives which may vary considerably from network to network. Eg: connectionless LAN service may be quite different than connection oriented WAN service.

The Transport Layer

The basic function of the transport layer is to split into smaller units if need be, has these to the network layer and ensure that the faces all arrive correctly at the end from the session layer. The transport connection required by the session layer is created by the transport layer. The transport layer multiplex several transport connections on to the same network connection to reduce the cost. In all cases the transport layer is required to make the multiplexing transparent to the session layer.

The transport layer also determines what type of service to provide the session layer and ultimately, the uses of the network. The most popular type of transport connection is an error-free point-to -point channel that deliver images (or) bytes in the order in which they were sent.

When the connection is established the type of service is determined. The transport layer is a true end-to end layer from source to destination. A program on the source machine carries on a conversation with similar program on the message header and control messages. In lower layers the protocols are between each machine and its immediate neighbours and not by the ultimate source and destination machine Which may be separated by many routes.

In addition to multiplexing several message stream on to one channel, the transport layer must take care of establishing and delaying connections across the network. There must also be a mechanism to regulate the flow of information, so that fast host cannot over run a slow one. Such a mechanism is called flow control and plays a key role in the transport layer. Flow control between route is distinct from flow control between hosts.

The Session Layer

A session allows ordinary data transport as does the transport layer, but it also provides enhanced services useful in some application. It is used to transfer a file between two machines. The service of the session layer is to manage dialogue control.

A related session service is token management. For some protocols, it is central that both sides do not attempt the same operation at the same time. To mange these activities the session layer provides token that can be exchanged. Only the side holding the token can perform the critical operation. To eliminate the time consuming problem the senior layer provides a way to insist check points into the data stream, so that after a crash only the data transferred after the last check point have to be repeated.

X. 25 Networks

It was developed during the 1970s by CCITT to provide an interface between Public Packet Switched networks and their customers. The physical layer protocol called x.21 specify the physical, electrical and procedural interface between the host and the network.

The data link layer standard has a number of variations. They all are designed to deal with transmission errors on the telephone lines between the users equipment and the public network. The network layer protocol deals with addressing, flow control delivery conformation interrupts and related issues. The packets are delivered reliably and in order. Most x.25 networks work at speeds up to 64 kbps, which makes them absolute for many purposes. Nevertheless, they are still widespread, so readers should be aware of their resistance x.25 in connection oriented and supports both switched virtual circuits and permanent ones.

A switched virtual circuit is created where one computer sends a packet to network asking to make a call to a remote computer. Once established packets can be sent over the connection, always arriving in order. X.25 provides flow control to make sure a fast sender cannot swamp a slow or busy receiver.

A standard protocol has been defined between the terminal and the PAD, called x.23, another standard protocol exists between the PAD and the network called x.29. Together, these three recommendations are often called triplex.

Frame Relay

Frame relay is the service for people who want an absolute bare-bones connection - oriented way to move bits from A to B at reasonable speed and low cost. The complex protocols were required to mask errors.

The situation has changed radically , leased telephone lines are now fast, digital and reliable, and computers are fast and inexpensive. This suggest the use of simple protocol with most of the work being done alone by the users computer rather than by the network. Frame relay can be thought of as a virtual based line. The customer leases a permanent virtual circuit between two points and can then send frames of up to 1600 bytes between them.

Frame relay provides a minimal service, primarily a way to determine the start and end of each frame, and detection of transmission errors. If a bad frame is received, the frame relay service empty discards it.

Quality of Service

The primary function of transport layer is enhancing the QOS (quality of service). If the network service is poor, the transport layer has to bridge the gap between what the transport users want and what the network layer provides. The transport service may allow the user to specify preferred, acceptable and minimum values for various service parameters at the time a connection is set up. The transport layer examines the parameters, which applies connectionless transport.

Typical transport layer quality service parameters

Connection Establishment delay
Connection establishment failure
Probability
Throughout
Transits delay
Residual error ratio
Protection
Priority
Resistance

The Connection Establishment Delay

The time taken by the transport connection that is being requested and confirmation being received by the user of the transport service. If the delay is shorter, the service is better.

The Connection Establishment Failure Probability

It is the chance of a connection not being established within the maximum establishment delay time.

Throughput : It is the parameter that measures the number of bytes of user data transferred per second, measured over some time interval. It is measured separately for each direction.

Transit Delay : Transit delay measures the time between the transport of the message and receiving of the message at the destination machine.

Residual Error Ratio : It is the measure of the messages that has been lost (or) garbled when it is totally sent.

Protection : Protection parameter provides a way for the transport user to specify interest in having the transport layer provide protection against unauthorized third parties to trap the data.

Priority : It is the parameter that provides a way for a transport user to indicate that some of its connection are important and in the event of congestion, the high priority connections are serviced first and then the low-priority ones.

Resilence : It is the parameter that gives the probability of the transport layer after terminating a connection due to internal problems or congestion by itself.

Conclusion : In some cases the transport layer does not achieve the desired goal but it can achieve a lower but still acceptable rate. Then it is sent to a remote machine, to check whether it can establish a connection. The remote machine finally inform whether the connection can be established (or) rejected according to the rate. This process is called option negotiation.

Transport Service Primitives

The transport service primitives allow transport users to access the transport service. There are important differences between network service and transport service. The transport service in contrast is reliable.

The second difference between the network service and transport service is to whom the services are intended for transport service might be like the five primitives listed below.

PRIMITIVE	*TPDUSENT*	*MEANING*
LISTEN	(NONE)	Block until some process tries to connect
CONNECT	CONNECTION REQ.	Actively attempt to establish a connection
SEND	DATA	Send information
RECEIVE	(NONE)	Block until a DATA TPDU arrives.
DISCONNECT	DISCONNECTION REQ.	This side wants to release the connection.

This transport service is truly bare bones but it gives the essential flavour of what a connection-oriented transport interface has to do. And allows the program to establish, use and release connection, which is sufficient for many application.

To see how the primitives are used, consider an application with a server and a number of remote clients.

To start with, the server executes a LISTEN primitive typically by calling a library procedure that makes the system call, to block the server until a client turns up. When a client wants to talk to the server it executes a CONNECT primitive. The transport entity carries out this primitive by blocking the caller and sending a packet to the server. Encapsulated in payload of this packet is the transport layer message for the server transport entity.

A quick note on a terminology is now in order. For lack of better term we will reluctantly use the somewhat ungainly acronym TPDU (Transport Protocol Data Unit) for messages sent from and to-to transport entity. These TPDU are contained in packets. In turn packets are contained in frames. When frame arrives, the data like layer crosses the frame header and takes the contents to the network entity. The network entity processes the packet header and panes the content of the packet payload up to the transport entity. This is of next form. The following figure shows the nesting.

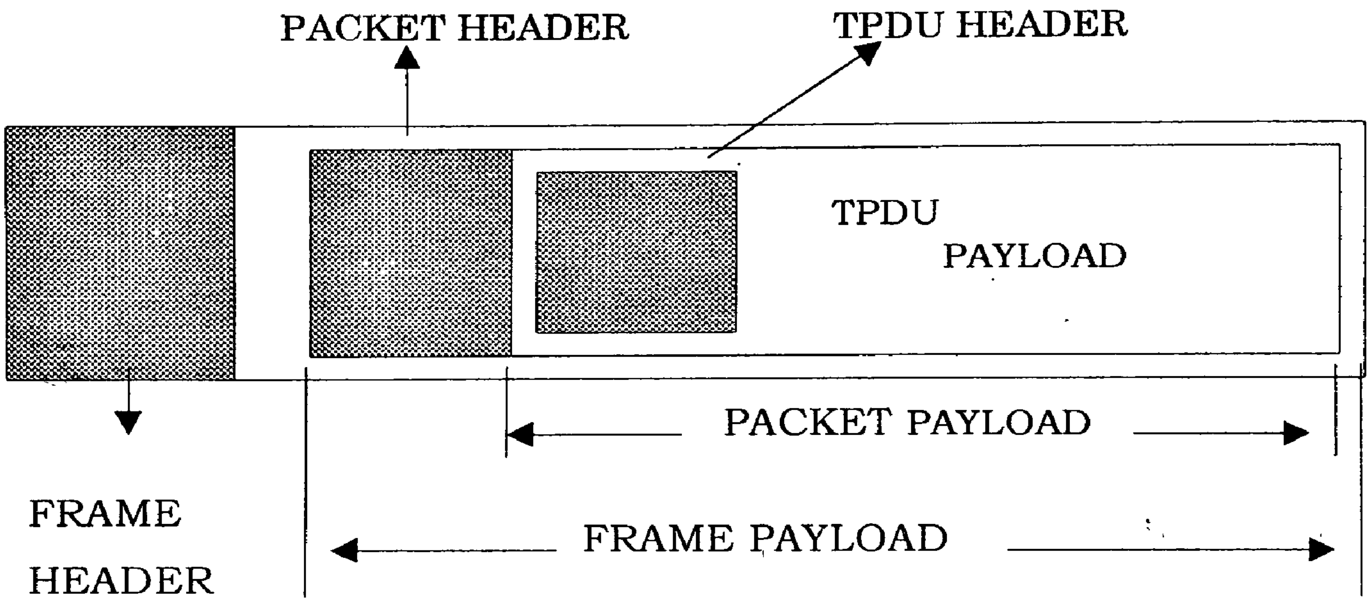


Fig 7.2

Hence the clients CONNECT call causes a connection request TPDU to be sent to the server. When it arrives the transport entity checks the server is blocked on a LISTEN. If unblock then the server sent CONNECTION ACCEPTED TPDU back to the client, thus the connection is established.

Data can be exchanged usually with the SEND and RECEIVE primitive as long as both sides can keep track of sending and receiving.

At the network layer even a simple unit direction data exchange is more complicated. Every data packet control TPDUS and send are also acknowledged implicitly (or) explicitly. These are managed by the transport entities.

To the transport users, a connection is a reliable "bit-pipe". One user stuffs bits in and they appear at another end. This ability to hide complexity is the reason that layered protocols are such a powerful tool.

When connection is no longer needed it must be realized to free up table space. "Disconnection " as two variance Asymmetric and Symmetric.

In Asymmetric variant either transport user can issue a DISCONNECT primitive, on getting the DISCONNECT TPDU, the connection is released.

In the symmetric variant each direction is closed separately independently of the other one. When one side does a DISCONNECT, it has no more data to send. But it is still willing to accept data. To this model when both the sides have done a DISCONNECT, the connection is released.

Berkely Sockets

The socket primitives which is used in Berkeley UNIX for TCP is discussed above.

Primitive	*Meaning*
SOCKET	Create a new communication end point
BIND	Attach a local address to a socket
LISTEN	Announce willingness to accept connections; give queue size
ACCEPT	Block the caller until a connection attempt arrives
CONNECT	Actively.attempt to establish a connection
SEND	send some data over the connection
RECEIVE	receive some data from the connection
CLOSE	release the connection

The SOCKET primitive creates a new end point and allocates table space for it within the transport entity. A successful SOCKET call returns an ordinary file descriptor for use in succeeding calls the same way an OPEN call does. Newly created sockets do not have address and assigned using the BIND primitive. When the socket is bound with an address, remote clients will be connected.

LISTEN call allocates space to queue incoming calls for the case that several subnets try to connect at the same time.

ACCEPT primitive is executed by the server when the block is waiting for an incoming connection. When TPDU asking for a connection arrives the transport entity creates a new socket with the same properties as the original one and returns a file descriptor. The CONNECT Primitive blocks the caller and actively starts the connection process. When it completes the client process is unblocked and the connection is established. Both sides can now use SEND and RECEIVE to transmit and receive data over the full-duplex connection. The connection is released when the socket is symmetric (ie) both sides have executed a CLOSE primitive.

Elements of Transport Protocols

Transport protocol, which is between two transport entity implements the transport service. Transport protocol and data link protocol resembles each other in some ways. The main difference between the transport protocol and data link protocol is the physical channel in the data link, which is replaced by the entire subnet in the transport protocol. This difference has many important implications for the protocols.

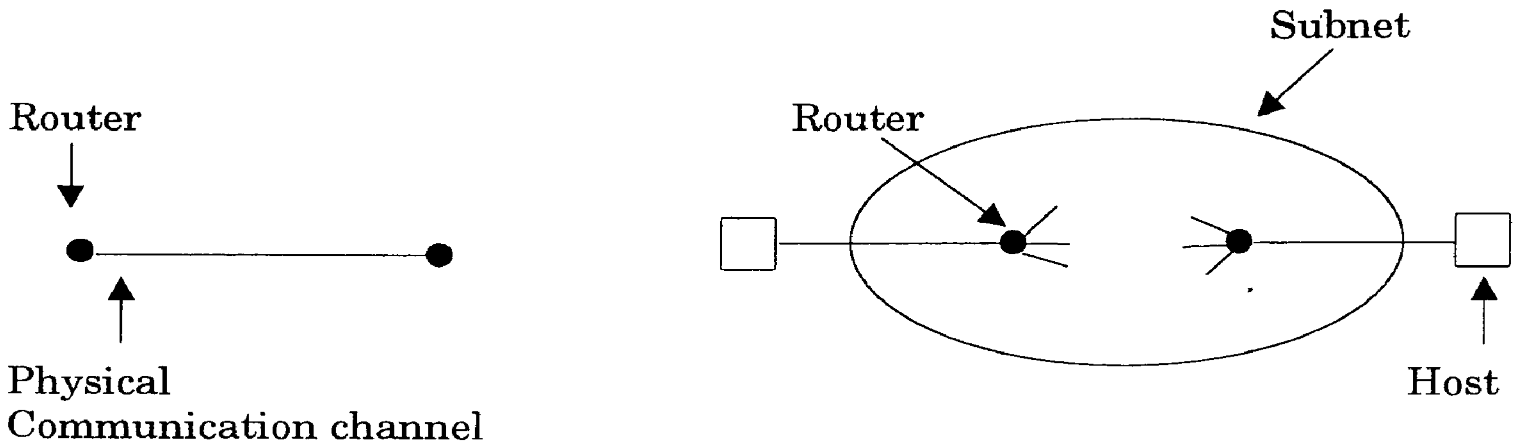


Fig 7.3

Data link layer, it is not necessary for a router to specify which router it wants to talk ,as each outgoing line uniquely specifies a particular router. In the transport layer explicit addressing of destination is required. The arranging difference between the data link layer and the transport layer is the potential existence of storage capacity in the subnet.

A final difference between the data link and transport protocol is the presence of large and dynamically varying number of connections in the transport layer but buffering and flow control are needed in both layers.

Addressing

Usually when an application process wishes to set up a connection to a remote application process we define, transport address to which connection requests. In ATM networks, they are AAL-SAPS. The neutral trim TSAP (Transport service Access Point) is used. The analogous end points in the network layer are called NSAPS.

TSAPS, NSAPs, and Connections

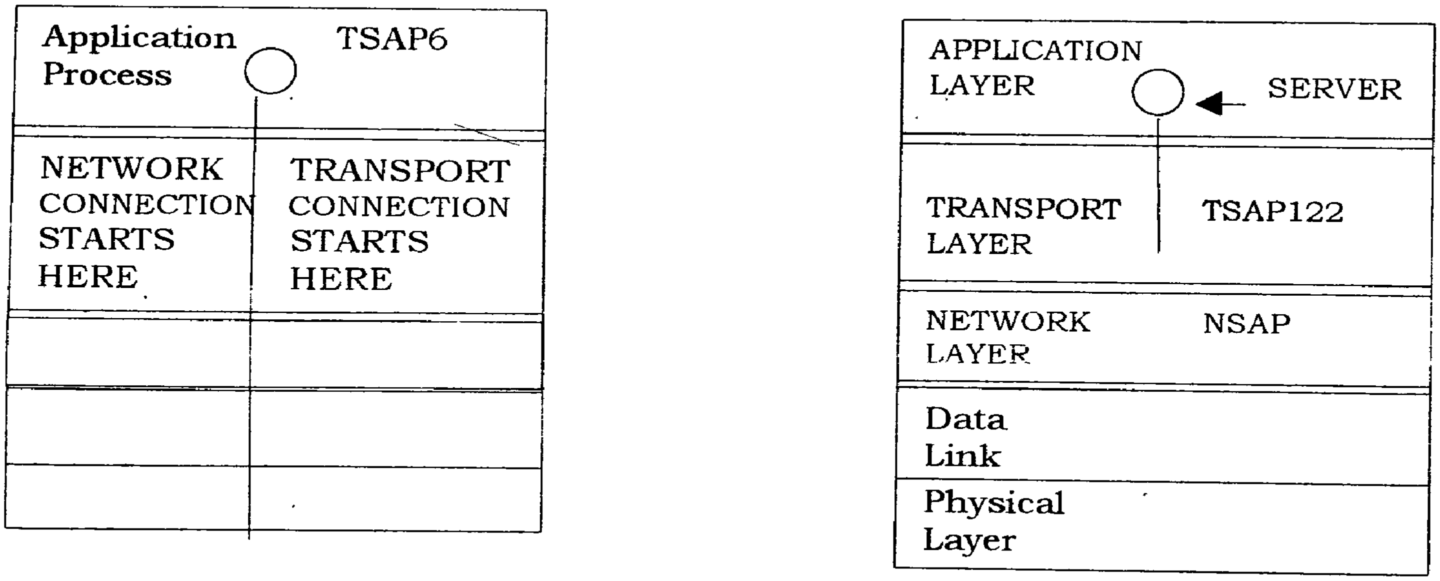


Fig. 7.4

The figure shows the reaction between the NSAP, TSAP, network connection and transport connection for a connection–oriented subnet.

Brief description about the connection are a time of day server process on host 2 attaches itself to TSAP 122 to wait for an incoming call. The attachment of the process to a TSAP is outside the networking model and depends entirely on the local operating system.

An application process on host wants to find out the time –of –day, so it issues a CONNECT request specifying TSAP 6 as the source and TSAP 122 as destination. The host1 transport entity can talk to the transport enmity on host 2 by network connection, which is set up by host, which selects a network address on its machine. The transport connection is established once the TSAP accepts the new connection in host 2.

The transport connection goes from TSAP to TSAP, whereas the network connection only goes part way, from NSAP to NSAP. Furthermore, if there are potentially many server processes most of which are rarely used, it is wasteful to have each of them active and listening to a stable TSAP address all day long.

Establishing a Connection

It would be sufficient for one transport entity to send a connection request for establishing a connection. The problem occurs where network can lose, store and duplicate packets.

Imagine a subnet that gets acknowledgements in time and each packet in time outs. Suppose it uses datagram inside and every packet follows a different route and some of the packets may get stick and take a long time to arrive, that's why they are stored in the subnet and popped out later. By establishing connection in a reliable way we can avoid the nightmares.

To cruse of the problem is the existence of delayed duplicates. One way is to use thrown away transport address. In this approach each time a transport address, which is new is generated, when the connection is released, the address is discarded.

Another possibility is to give each connection a connection identifier. Thus whenever a request came in it could be checked and it updates a table listing absolute connection when released. It is basic fault as it requires each transport entity to maintain certain amount of history information indefinitely. If the machine crashes ,the connection identifier already in use will not be known.

Instead we can allow packets to live forever within the subnet and bill - off aged packets. If we can ensure that no packet lives longer than some time the problem becomes more manageable. The lifetime is restricted by the following techniques.

1) ***Restricted subnet design :*** If prevents packets from locking, combined with some way of bounding conjunction delay over the longest possible path.

2) ***Putting a hop-counter in each packet :*** It has a hop-count incremented each time the packet is forwarded. The data link protocol destroy the packet which exceeds certain hop-counter value.

3) ***Time sampling each packet :*** It requires each packet to bear the time it was created to discard any packet- older then some agreed upon time. It requires the clocks to be synchronized. It is achieved extended to the network, for example by listening to "WWW" or some station that broadcast the precise time periodically.

With packet life time bounded it is possible to establish connection safely. Once such method described by TOM LINSON and it was further refined by sunshine and Dallas. Variance of it is widely used in practice.

TOM LINSON proposed equipping each host with a time of day clock, which need not be synchronized to avoid the losing of all memory, when the device crashes. Each clock takes the form of a binary counter that increments itself at uniform intervals. The clock is assumed to continue running even if the host goes down. The basic idea is to avoid identically numbered TPDUs outstand at the same time. When a connection is set up the low order K-bits of the clock are used as the initial sequence number. The sequence space should be so large that when the NEW TPDUs comes, the old ones to the same sequence number are gone. This linear relation between time and initial sequence number is as follows.

Once both transport entities have agreed on the initial sequence number any sliding window protocol used for data control. Actually the sequence number curve is not linear. Problem occurs when a host crashes. Its transport entity does not know where it was in the sequence ways. One solution is to make the transport entities idle 40 seconds after a recovery to let all old TPDUS die-off. But in a compels inter-network the T may be large.

To prevent the problems due to crashing, we must prevent sequence nos. from being used before their potential used as initial sequence numbers. Before the TPDU has been sent it must read the clock and check that it is not in the hidden region. The protocol can get into trouble in two different ways.

If a host sent too much data too fast on a newly opened connection the curve may raise more steeper (ie) the maximum data rate is 1-TPDU per clock rates-tick. It means, the transport entity must wait until the clock ticks before opening a new connection after a crash restart, test the same number being used twice. To get into trouble by sending too fast is not the only way. It is clear that at any data rate less than the clock rate the curve will eventually run into the forbidden region from the left. If the slope is greater then the event will be delayed longer.

The clock based method solves the delayed duplicate problem for data TPDUs but it is useful, to establish a connection. Since TPDUs may also be delayed thus the potential problem in getting both sides to agree on the initial sequence number.

To solve this TOM LISON introduced 3 -way hand shakes. This establishment protocol does not require both sides to begin sending with the same sequence numbers so it can be used with synchronization methods other than the global clock method.

The normal set up procedure when host initiates is shown in the following steps.

- Host 1 chooses a sequence number x, and sends a CONNECTION REQUEST TPDU containing it to host 2.
- Host 2 replies with a CONNECTION ACCEPTED TPDU acknowledging a and announcing its own initial sequence number.
- Finally host 1 Acknowledges host 2's choice of an initial sequence number in the first data TPDU that it sends.

Old Duplicate Connection request appearing out of now Here

The three way hand shakes work in the presence of delayed duplicate control TPDUs. The first TPDU is a delayed duplicate CONNECTION REQUEST from an old connection. This TPDU arrives at host 2 without host 1 s knowledge. Host 2 reacts to this TPDU by sending host 1 a connection accepted TPDU, in effect asking for verification that host was indeed trying to set up a new connection. When host 1 rejects host 2's attempt to establish, host 2 realizes that it was touched by a delayed duplicate and abandons the connection. In this way a delayed duplicate does on damage. The worst case is when both a delayed CONNECTION Request and an acknowledgement to a CONNECTION ACCEPTED are floating around in the sub net.

Duplicate Connection Request and Duplicate

In this case the host 2 gets a delayed CONNECTION REQUEST and replies to it. At this point it is crucial to release that host 2 has proposed using Y as the initial sequence numbers for host 2 to host., traffic knowing full well that no TPDUs containing sequence number Y or acknowledgment to y are still in existence when the second delayed TPDU arrives at host 2 the fact that 2 has been acknowledged rather than Y tells host 2 that this, two is an old duplicate. The important thing to realize there is that no combination of old CONNECTION REQUEST, CONNECTION ACCEPTED (or) other TPDUs that can cause the protocol to fail and have a connection set up by accident when no one wants it.

Releasing a Connection

Releasing a connection is easier than establishing one. There are two types of release.

(i) ***Asymmetric Release :*** Asymmetric release is the way the telephone system works, where one party hangs up and the connection is broken. Asymmetric release is abrupt and may result in data loss.

After the connection is established , host 1 sends a TPDU that arrives properly at host 2.

The host 1 sends other TPDU. The host issues a DISCONNECT before the second TPDU arrives. The result is that the connection is released and data are lost.

(ii) ***Symmetric :*** Symmetric release treats the connection as two separate unidirectional connections and requires each one to be released separately. A more sophisticated release protocol is required to avoid data loss. In symmetric release the scenarios of release uses a three -way- handshakes. The protocol is not infallible, it is usually adequate.

Normal Case of Three-way Hand Shake

The normal case in which one of the user sends a DR (Disconnection request) TPDU in order to initiate the connection release. When it arrives, the recipient sends back a DR TPDU too and starts a timer, just in case its DR is lost when this DR arrives the original sender sends back an ACK TPDU and releases the connection. Finally, when the ACK TPDU arrives the receiver also releases the connection.

Releasing a connection means that the transport entity removes the information about the connection from its table of open connection and signals the connection owner some how. This action is different from a transport user using a DISCONNECT primitive.

Final ACK Lost

If the final ACK. TPDU is lost, the situation is saved by the timer. When the timer expires the connection is released any way. The user initiating the disconnection will not receive the expected response, will time out and will start all over again in the case of the second DR being lost.

Response Lost

In this case we assume that the second time no TPDUs are lost and all TPDUs are delivered correctly and on time.

Response Lost and Subsequent DR Lost

This is same as the response lost except that now we assume all the repeated attempts to retransmit the DR also fail due to lost TPDUs. After N times the sender just gives up and release the connection. Meanwhile the receiver times out and also exist, while this protocol usually suffices in theory it can fail if the initial DR and N retransmissions are all lost. The sends will give up and release the connection, while the other side knows nothing at all about the attempts to disconnect and is fully active. This situation results in a half-open connection.

One way to kill the half-open connections is to have a rule saying that if no TPDUs have arrived for a certain number of seconds the connection is automatically disconnected. If one side ever disconnects, the other side will detect the lack of activity and also disconnect. If the Automatic disconnect rule is used and two many dummy TPDUs in a row are lost on an otherwise idle connection first one side, then the other side will automatically disconnect.

Low Control and Buffering

The transport layer and the data link layer have the same flow control problem. The basic similarity is that in both layers a sliding window or other scheme is needed on each connection to keep a fast transmitter from overrunning a slow receiver. The host has many numerous connections whereas the router usually has relatively few lines, which is the main difference between them. This difference makes it impractical to implement the data link buffering strategy in the transport layer.

In the data link layer the sending side must buffer outgoing frames because they might have to be retransmitted. If the subnet provides datagram service, the sending transport entity must also buffer and for same reason. If the receiver knows that the sender buffer all TPDUs until they are acknowledged, the receiver may or may not dedicate specific buffers to specific connections as it sees fit. When TPDU comes in, an attempt is made to dynamically acquire a new buffer. If one is available, the TPDU is accepted otherwise it is discarded. The network service is unreliable, the sender must buffer all. TPDUs sent just as in data link layer. Buffering is must, since receiver cannot guarantee that every incoming TPDU will be accepted.

Buffer size depends on TPDUs size . If there is wide variation in TPDU size from a few character typed at a terminal to a 1000 of character from file transfers a tool of fixed–sized, buffer presents problems. Buffer size is wasted when TPDU is less in size. If TPDU is larger than buffer size will be needed for long TPDUs with attendant complexity.

The optimum trade–off between source buffering and destination buffering depends on the type of traffic carried by the connection. For low- band width busty traffic, such as that produced by an interactive terminal , it is better not to dedicate any buffer, but rather to acquire them dynamically at both ends. For low bandwidth busy traffic, it is better to buffer at the sender and for high bandwidth smooth traffic it is better to buffer at the receiver.

Buffer allocation are adjusted according to the connection that are opened and closed and as the traffic pattern changes. Sending host should request buffer space of the transport protocol. Buffers should be allocated for connection or collectively for all the connection running between the two host.

Potential problems with buffer allocation schemes of any kind can arise in datagram network if control TPDUs can get lost. Each host should periodically send control TPDUs giving the acknowledgement and buffer status on each connection. That way the dead lock will be broken sooner or later. Until now we have tacitly assumed that the only limit imposed on the senders data rate is the amount of buffer space available in the receiver. The problem is that if the dramatically it may become feasible to equip hosts with so much memory that lack of buffers is rarely if ever a problem. When buffer space no longer limits the maximum flow, another and the bottleneck will affect the carrying capacity of the sub net. The carrying capacity can be determined by simply counting the number of TPDUs acknowledged during some time period and then dividing by the time period. During the measurement, the sender should send as fast as it can to make sure that the network is carrying its capacity and not the low input rate as the factor limiting the acknowledgement rate.

Multiplexing

Multiplexing several conversations on to connections virtual circuits and physical link plays a role in several layers of the network architecture. In the transport layer the need for multiplexing can arise in a number of ways.

Upward Multiplexing

The price structure that heavily penalizes installation for having many virtual circuits open for long period of time is to make multiplexing of different transport connection onto the same network connection. This form of multiplexing is called upward multiplexing.

Four distinct transport connection, all use the same network connection to remote host in ATM virtual Circuit which is an example of upward multiplexing. When upward multiplexing is used with ATM, we have the same situation of having to identify the connection using a field in the transport header, even though ATM provides more than 4000 virtual circuit numbers per virtual path expressly for that purpose.

Down Ward Multiplexing

Downward multiplexing is a module operation, where the transport layer open multiple network connections and distribute the traffic among them on a round - robin basis. Downward multiplexing can also be used to increase the performance if multiple lines are available. Eg. 4095 virtual circuit

Crash Recovery

If hosts and routers are subject to crashes, recovery from these crashes becomes an issue and if the transport entity is entirely within the hosts, recovery from network and router crashes is straightforward.

A more troublesome problem is the recoverment from host crashes. In particular, it may be desirable for clients to be able to continue working when servers crash and quickly reboot. The transport layer on the server simply passes the incoming TPDUs to the transport user, one by one. Part way through the transmission, the server crashes, when it comes backup, its tables are reinitialized, so it no longer knows precisely where it was.

The server might send a broadcast TPDU to all other hosts, announcing that it had just crashed and requesting that its clients inform it of the status of all open connection to recover its previous status. Client can be in one of the 2 forms

- TPDU outstanding , SI
- TPDU outstanding, SO

Based on only the first, the client decides whether or not to retransmit the most recent TPDU.

The client should retransmit only if it has an unacknowledged TPDU outstanding when it learns of the crash. Sending an acknowledgement and writing a TPDU onto the output stream cannot be done simultaneously. If the crash occurs after the acknowledgement has been sent but before the write has been done, the client will receive the acknowledgement and thus be in state so when the crash recovery announcement arrives. Thus the client without retransmitting leads to a missing TPDU. If the write has been done but the crash occurs before the acknowledgement can be sent, then the client will be in the state & SI and thus retransmit leading to an undetected duplicate TPDU in the out-put stream to the server application process. There are always protocols, which fails without recovering. The server can be programmed in one of two ways. Acknowledge first (or) write first. The client can be programmed in one of four ways. Always retransmit the last TPDU, never retransmit the last TPDU, retransmit only in the state so (or) retransmit only in state SI.

This gives eight combination, but for each combination there is some set of events that makes the protocol fails.

Three different events are possible at the server. Sending an acknowledgment (A) writing to the output process (W) and crashing(C).

The three events can occur in six different orderings. AC (W), AWC, C(AW), C(WA), WAC, WC(A). Parenthesis are used to indicate that neither A row W may follow C.

The eight combinations of client and server strategy and the valid event sequence for each one is shown below when each strategy there is some sequence of events that causes the protocol to fail.

Making the protocol more elaborate does not heap. Even if the client and server exchange several TPDUs before the server attempts to write, so that the client knows exactly what is about to happen the client has no way of knowing whether a crash occurred just before (or) just after the write. A truly end -to -end acknowledgement, whose receipt means that the work has actually been done and back there of means that it has not, is probably impossible to achieve.

A Simple Transport Protocal

Concrete service of transport protocol consists of five primitives.

Connect, Listen Disconect, Send and Receive

These five primitives corresponds exactly with a library procedure that executes the primitive.

```
Connum    = LISTEN (local)
Connum    = CONNECT (local, remote)
Status    = SEND (connum,buffer, bytes)
Status    = RECEIVE (connum, buffer,bytes)
Status    = DISCONNECT (connum)
```

- **LISTEN :** This primitive announces the caller's willingness to accept connection requests directed at the indicated.
- TSAP. The user of the primitive is blocked until an attempt is made to connect to it. There is no timeout.

- **CONNECT :** It takes two parameters local TSAP and remote TSAP, and tries to establish a transport connection between the two. If it succeeds it return in connum a nonnegative number used to identify the connection. If it fails, the reason for failure is but in connumar a negative number. The reason for the failure is that one of the transport addresses is currently in use. There are some other reasons for failure such as remote host down illegal local address, and illegal remote address.

- **SEND :** This primitive transmits the contents of the buffer as a message on the indicated transport connection possibly in several units if it is too big.

- **RECEIVE :** It indicates the caller's desire to accept data. The size of the incoming message is placed in bytes.

- **DISCONNECT :** The primitive terminates a transport connection. The parameter connum tells which one.

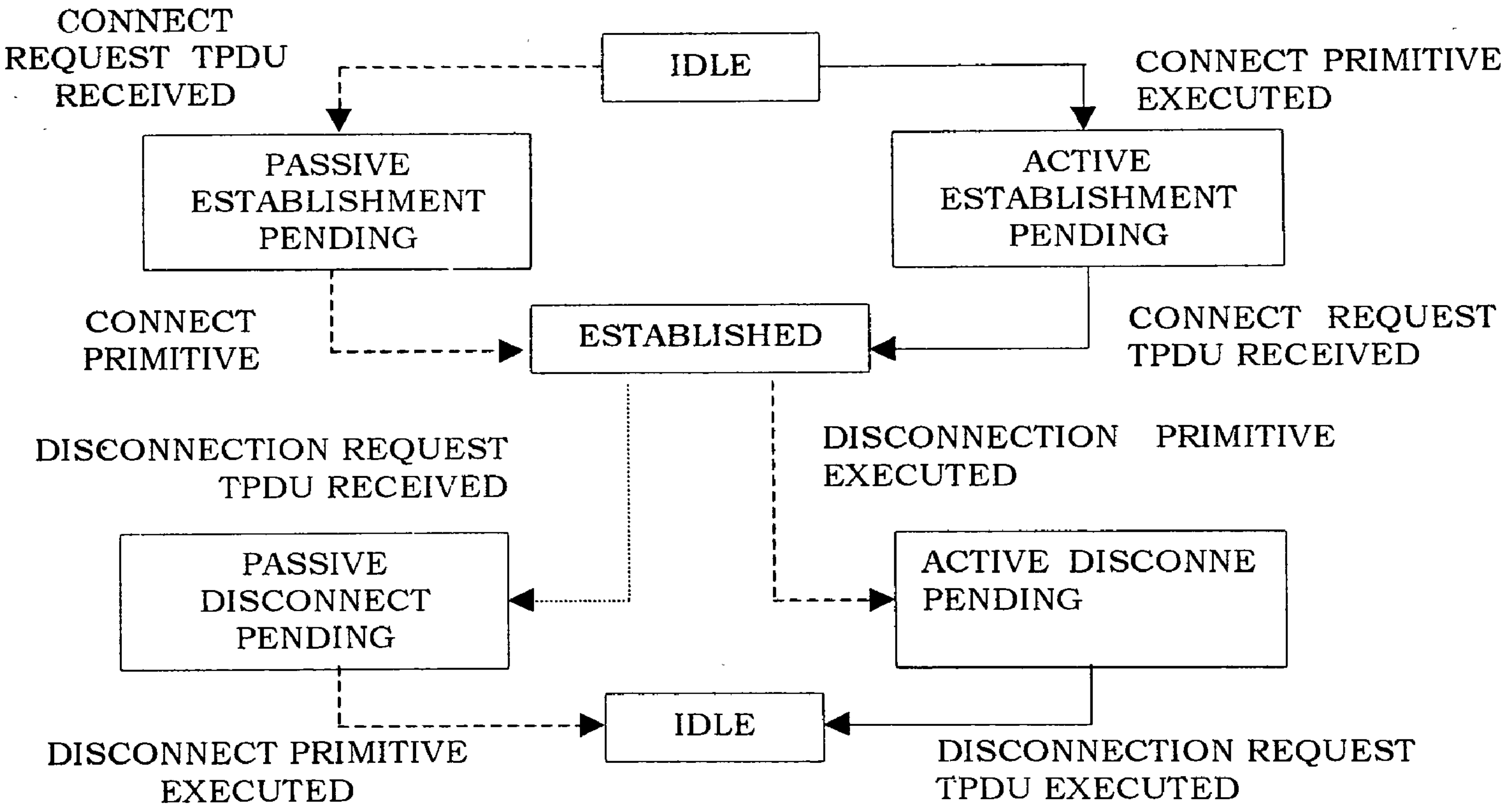

CONNECT REQUEST TPDU RECEIVED
IDLE
CONNECT PRIMITIVE EXECUTED
PASSIVE ESTABLISHMENT PENDING
ACTIVE ESTABLISHMENT PENDING
CONNECT PRIMITIVE
ESTABLISHED
CONNECT REQUEST TPDU RECEIVED
DISCONNECTION REQUEST TPDU RECEIVED
DISCONNECTION PRIMITIVE EXECUTED
PASSIVE DISCONNECT PENDING
ACTIVE DISCONNE PENDING
IDLE
DISCONNECT PRIMITIVE EXECUTED
DISCONNECTION REQUEST TPDU EXECUTED

Fig. 7.5 A state diagram for a simple connection management scheme

LESSON - 8

PRESENTATION LAYER
DNS - DOMAIN NAME SYSTEM

Programs rarely refer to hosts and other resources by their binary network addresses. Instead of binary numbers, they use ASCII strings. The network can understand only binary address, so some mechanism is needed to convert the ASCII strings to network addresses.

In the ARPANET, there was simply a file, host.txt that listed all the hosts and their IP (Internet Protocol) addresses. For a network of a few hundred time sharing machines, this type of approach works well. When thousands of workstations were connected to the net, this approach does not suit well, since the size of the file would become too large. Even host name conflicts would occur. To solve these problems DNS was invented.

The essence of DNS is the invention of a hierarchical, domain-based naming scheme and a distributed database system for implementing the naming scheme. It is primarily used for mapping host names and email destinations to IP addresses.

To map a name onto an IP address, an application program calls a library procedure called the resolver, passing it the name as a parameter. The resolver sends a UDP (Uniform Datagram Packet) packet to a local resolver, which then looks up the name and returns the IP address to the resolver ,which then returns it to the caller. The program can then establish a TCP connection with the destination or send it UDP packets.

DNS Name Space

The internet is divided into several hundred top-level domains, where each domain covers many hosts. Each domain is partitioned into sub domains and these are also further partitioned. All these domains can be represented by a tree.

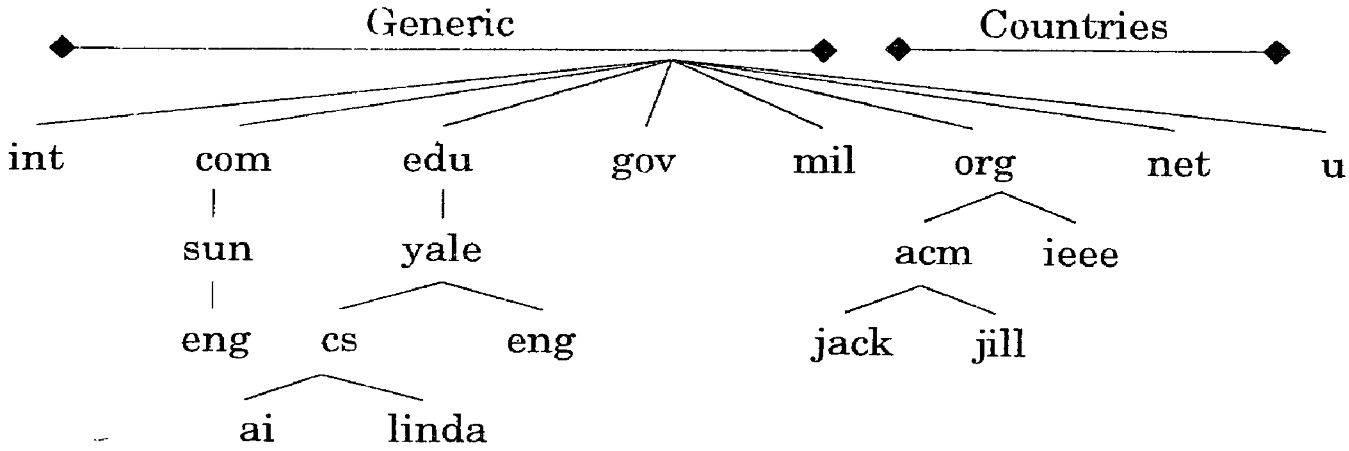


Figure 8.1

The leaves of the tree represent domains that have no sub domains. A leaf domain may contain a single host or thousand of hosts.

The top-level domains come in two flavors

(i) Generic (ii) Countries

The generic domains are com (commercial), edu (educational institutions), gov (the U.S. federal government), int (certain international organizations). mil (the U.S. armed forces), net (network providers), and org (non profit organizations). The country domain includes one entry for every country.

Each domain is named by the path upward from it to the root. The components are separated by periods (or) dots. Domain names can be either absolute or relative. An absolute domain names end with a period where as a relative domain name does not. In both cases, a named domain refers to a specific node in the tree and all the nodes under it.

Domain names are case insensitive. Component names can be up to 63 characters long and full path names must not exceed 255 characters.

To create a new domain, permission is required of the domain in which it will be included. For example, if a VLSI group is started at Anna and wants to be known as Vlsi.CS.anna.edu, it needs permission from whomever manager CS.anna.edu. Once a new domain has been created and registered, it can create sub domains without getting permission from anybody higher up the tree.

Resource Record

Every domain, whether it is a single host or top-level domains, can have a set of resources records associated with it. The real function of DNS is to map domain names onto resource records.

A resource record is a five-tuple. The format is as follows :

Domain_name Time_to_live Class Type Value

The Domain_name tells the domain to which this record applies. Many records exist for each domain and each copy of the database holds information about multiple domains. When a query is made about a domain, all the matching records of the class requested are returned.

The Time_to_live field gives an indication of how stable the record is. Information with high stability is assigned a large Value. Information that is highly volatile is assigned a small value.

The Type field tells what kind of record this is. The most important types are listed below.

Type	*Meaning*	*Value*
SOA	Start of Authority	Parameters for this zone
A	IP address of a host	32 Bit integer
MX	Mail exchange	Priority, domain willing to accept email
NS	Name Server	Name of a server for this domain
CNAME	Canonical name	Domain name
PTR	Pointer	Alias for an IP address
HINFO	Host description	CPU and OS in ASCII
TXT	Text	Uninterested ASCII text

The Principal DNS Resource Record Types

An SOA record provides the name of the primary sources of information about the name servers zone, the email address of its administrator, a unique serial number, various flags and time outs.

A (Address) record is an important record type. It holds a 32 bit IP address for some host. Interest host must have at least one IP address, so other machines can communicate with it some hosts have two or more network connections, in which case they will have one type A resource record per network connection.

The next important record type is the MX record. It specifies the name of the host prepared to accept email for the specified domain. The main use of this type of record is to allow a machine that is not on the Internet to receive email from internet site.

The NS records specify name servers. Every DNS database normally has domains, so email can be sent to distant parts of the naming tree.

CNAME records allow aliases to be created. When a person wants to send a message to someone whose login name is Jones in CS department at M.I.T would guess that *jones@CS.mit.edu* will work. Actually this address will not work because the domain for M.I.T's Computer science department is lcs.mit.edu. As a service to people who do not know this, M.I.T would create a CNAME entry to point people and programs in the right direction.

PTR points to another name. PTR is a regular DNS data type whose interpretation depends on the context in which it is found. It is used to associate a name with an IP address to allow look ups of the IP address and return the name of the corresponding machine.

HINFO *record allows people to find out what kind of machine and operating* system a domain corresponds to.

TXT record types allow domains to identify themselves in arbitrary ways.

The fourth field of every resource record is the Class. For Internet information, it is always IN. For non-internet information, other codes can be used.

Final field is the Value. This field can be a number, a domain name or an ASCII string. The semantics depend on the record type.

The below figure shows a portion of a possible DNS Database for cs.vu.nl. The database contains several types of resource records.

Authoritative data for *cs.vu.nl*

cs.vu.nl	86400	IN	SOA	Stal boss (952771, 7200, 7200, 2419200, 86)
cs.vu.nl	86400	IN	TXT	"Facultect Wiskunde en Informatica"
cs.vu.nl	86400	IN	TXT	"Vriji Universitiet Amsterdan".
cs.vu.nl	86400	IN	MX	1 zephys.es.vu.nl.
cs.vu.nl	86400	IN	MX	2 top.cs.vu.nl
flits.cs.vu.nl	86400	IN	HINFO	Sun unix
flits.cs.vu.nl	86400	IN	A	130.37.16.112
flits.cs.vu.nl	86400	IN	A	192.31.231.165
flits.cs.vu.nl	86400	IN	MX	1 flits.cs.vu.nl
flits.cs.vu.nl	86400	IN	MX	2 Zephys.cs.vu.nl
flits.cs.vu.nl	86400	IN	MX	3 top.cs.vu.nl
flits.cs.vu.nl	86400	IN	CNAME	star.cs.vu.nl
flits.cs.vu.nl	86400	IN	CNAME	zephys.cs.vu.nl
rowboat		IN	A	130.37.56.201
		IN	MS	1 rowboat
		IN	MS	2 zephyr
		IN	MS	Sun unix
little-sister		IN	A	130.37.62.23
		IN	HINFO	Mac Macos
Laserjet		IN	A	192.31.231.216
		IN	HINFO	"HP Laserjet IIISI" Proprietary.

The first noncomment line gives some basic information about the domain. The next two lines give textual information about where the domain is located. Then come two entries giving the first and second places to try to deliver email sent to person@cs.vu.nl. The Zephyr (a specific machine) should be tried first. If this fails, the *top* should be tried next.

Flits is a Sun workstation running UNIX and giving both of its IP address. Three choices are given of handling email sent to flits.cs.vu.nl. First choice is the flits itself, the zephyr and top are the second and third choices. Next is the alias www.cs.vu.nl, so that this address can be used without designating a specific machine. Creating this alias allows to change its www server without invalidating the address people use to get to it.

The next four lines contain a typical entry for a workstation. The information provided contains the IP address, the primary and secondary mail drops, and information about the machine. Then comes an entry for a non-unix system that is not capable of receiving mail itself, followed by an entry for a laser printer.

Name Servers

A single name server could contain the entire DNS database and respond to all queries about it. In practice, this server would be so overloaded as to be useless. If it ever went down, the entire Internet would be crippled.

To avoid this problem, the DNS name space is divided into non overlapping zones. One possible way to divide up the name space is shown in the below figure.

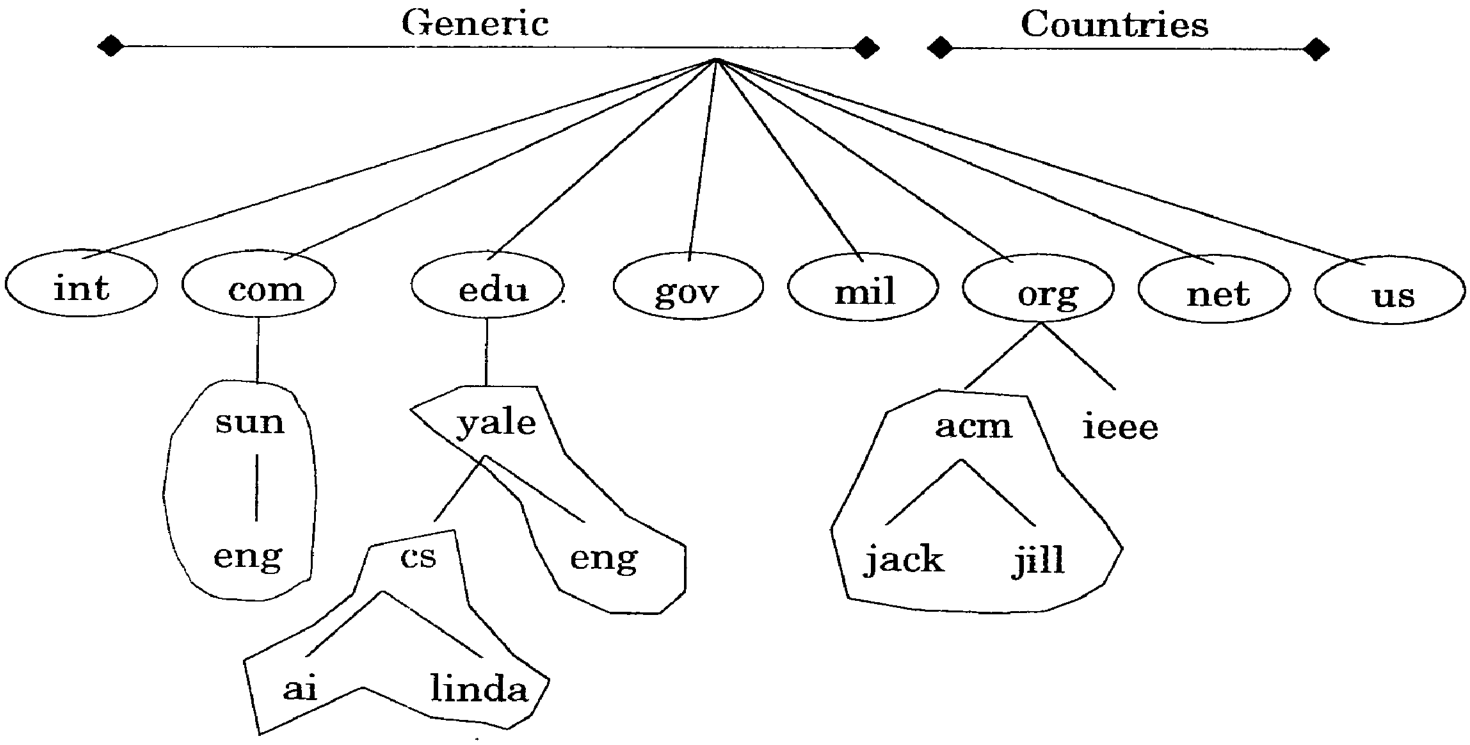


Figure 8.2

Each zone contains some part of the tree and also contain name servers holding the authoritative information of that zone. A zone will have one primary name server, which gets its information from a file on its disk and one or more secondary name servers, which get their information from the primary name server. Some servers for a zone can be located outside the zone to improve reliability.

It is the decision of the zone's administrator to place the zone boundaries within the zone. When a resolver has a query about a domain name, it passes the query to one of the local name servers. If the domain being sought falls under the jurisdiction of the name server, it returns the authoritative resource records.

If the domain is remote and no information about the requested domain is available locally, the name server sends a query message to the top level name server for the domain requested.

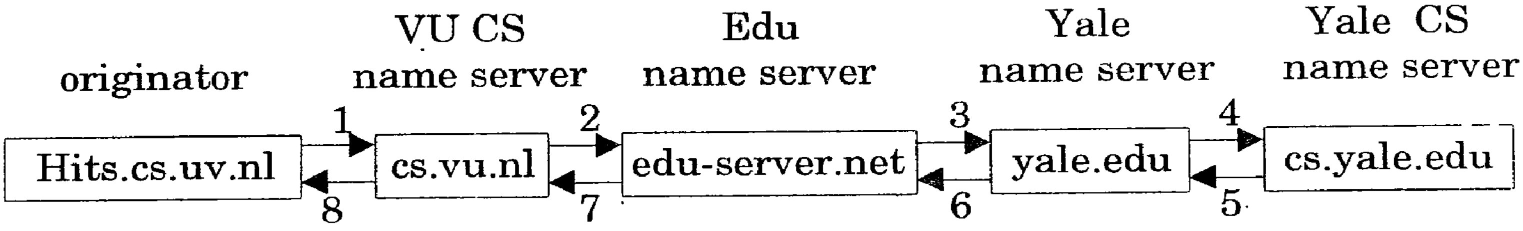


Figure 8.3

How a resolver looks up a remote name in 8 steps.

In step 1 it sends a query to the local name server, cs.vu.nl. This query contains the domain name sought, the type (A) and the class (IN).

Suppose the local name server has never had a query for this domain before and knows nothing about it. It may ask a few other nearby name servers, but if none of them know, it sends a UDP packet to the server for edu given in its database edu-server.net. This server knows the address of linda.cs.anna.edu and probably does not know cs.anna.edu so it forwards the request to the name server for anna.edu. In turn, this forwards the request to cs.anna.edu, which must have the authoritative resource records. Since each request is from a client to a server, the resource record requested works its way back in steps 5 through 8.

Once these records get back to the cs.vu.nl server, they will be entered into a cache. But this information is not authoritative since changes in cs.anna.edu will not be propagated to all the caches. For this reason, cache entries should not live too long. That is why, the Time_to_live field is included in each resource records.It tells remote name servers how long to cache records. If a certain machine had the same IP address for years, it may be safe to cache that information for 1 day. For more volatile information, it might be safer to purge the records after a few seconds or a minute.

The query mentioned here is recursive query, since each server does not have the requested information goes and finds it some where then reports back. An alternative form is also possible. When a query cannot be satisfied the query fails but the name of the next server along the line to try is returned. This gives the client more control over the search process. Some servers do not implement recursive queries and always return the name of the next server to try.

ELECTRONIC MAIL

Introduction

1. Electronic mail on E-mail has been around for over two decades (one decade – ten years)

2. The first email systems consisted of FTPs (File Transfer Protocols) with the recipient address in the first line.

   Limitations of the above mail system with file transfer approach were,

   1. Sending a message to a group of people was inconvenient.
      eg. Managers try to send memos to all of their subordinates.

   2. Messages had no internal structure, making computer processing difficult eg. If a forwarded message was included in the body of another message, extracting the forwarded part from the received message was difficult.

   3. No acknowledgements for the messages sent by the sender.

   4. If someone was away for several weeks and wanted all incoming email to be handed, it was not easy.

   5. The user interfaces was poor, with the transmission system requires the users to edit a file, first and invoke it leaving the editor.

   6. It was not possible to create and send messages containing a mixture of tent, drawing, facsimile and voice.

Architecture and Services

Email systems consists of two subsystems.

1. The user agents : It allows the people to read and send email.

2. The message transfer agents : It moves the messages from the source to the destination.

User agents are local programs that provide a command – based, menu based or graphical method for interacting with the email system. The message transfer agents are typically system demons that run in the background and move email through the system.

Email system supports five basic functions

1. Composition
2. Transfer
3. Reporting
4. Displaying
5. Disposition

Composition

1. It refers to the process of creating messages and answers.

2. Any text editor can be used for the body of the message and the system provides assistance with addressing and numerous header fields are attached to each message.

Transfer

It refers to moving messages from the originator to the recipient.

Reporting

It refers to reporting the originator the confirmation of delivery.

Displaying

Incoming messages are needed for people to read their email. Conversion is required if the message is a postscript file or digitized voice.

Disposition

It deals with the act of the recipient with the message after receiving it.

They may throw it away before reading, after reading, saving it, retrieve and reread saved messages, forward them, or process them in other ways.

Mailboxes are used to store incoming mails & mailing list of email addresses. When a message is sent to the mailing list, identical copies are delivered to everyone on list.

Registered email is to allow the originator to know that his message has arrived. Automatic notification of undeliverable email is also desired.

Other advanced features are carbon copies, high priority email, secret email, etc.

Envelope encapsulates the message which consists of address, priority and security level. It consists of two parts namely the header and the body.

The User Agent

A user agent is normally a program that accepts a variety of commands for composing, receiving and replying to messages and for manipulating mailboxes.

Sending Email

While sending the destination address many user agents expect DNS addresses of the form mailbox @ location. X.400 addresses look different than DNS addresses.

e.g.

/C=US/SP=MASSACHUSETIS/L=CAMBRIDGE/PA=360 MEMORIAL

DR./CN=KENSMITH/

This specifies a country, state, locality, personal address and a common name. This is less convenient than DNS names. Most email systems have always allowed users to have alias files. Mailing lists are used to send the same message to a list of people.

e.g. if a group of bird watchers have a mailing list called birders installed on *meadowlark.arizona.edu,* then any message sent to *birders@meadowlark.arizona.edu* will be routed to the University of Arizona and expanded there into individual messages to all the mailing list members, wherever in the world they may be.

Reading Email

An example display of the contents of a mailbox.

#	*Flags*	*Bytes*	*Sender*	*Subject*
1	KA	6348	john	Comments on material you sent me
2	K	1030	peter	changes to minix
3	KF	4519	armstrong	Request for information

The first field is the message number. The second field, flags, K refers that the message is not new, A - says that it has already been answered

F – meaning that the message has been forwarded to source one else. The third field tells, how long the message is, and the fourth one tells, who sent the message.

After the headers have been displayed, the user can perform any of the commands available.

The # sign means that the numbers of a message is expected

Command	*Parameter*	*Description*
H	#	Display header(s) on the Screen
C		Display current header only
T	#	Type message(s) on the Screen
S	Address	Send a message
F	#	Forward message(s)
A	#	Answer message(s)
D	#	Delete message(s)
U	#	undelete previously deleted message(s)
M	#	move message(s) to another mailbox
K	#	keep message(s) after exiting
R	mailbox	Read a new mailbox

Command	*Parameter*	*Description*
R		Go to the next message and display it
B		Backup to the previous message and display it
G	#	Go to a specific message but do not display it
E		Exit the mail system and update the mailbox

Typical mail handling commands

h – displays one or more headers in the format of the previous figure
c – prints the requested message's header
t – types the requested message or messages

t3 to type message 3, t 4-6, to type messages 4 through 6, and t a to type them all.

The next group of three commands deals with sending messages rather than receiving them.

s – sends a message by calling an appropriate editor

f – forwards a message from the mailbox, prompting for an address to send it to

a – extracts the source address

The next group of commands is for manipulating mailboxes.

Message Formats

RFC 822 messages consist of a primitive envelope, some number of header fields, a blank line, and then the message body.

The principal heard fields related to message transport are

Header	*Meaning*
To	Email address(es) of primary recipients(s)
CC	Email address(es) of secondary recipient(s)
Bcc	Email address(es) for blind carbon copies
From	Person or people who created the message
Sender	Person or people who created the message
Received	Line added by each transfer agent along the route
Return-Path	Can be used to identify a path back to the sender

In addition to the principal header fields, the most common ones are

Header	*Meaning*
Date	The date and time the message was sent
Reply-To	Email address to which replies should be sent
Message-ID	Unique number for referencing this message late
In-Reply-To	Message-de of the message to which this in a reply
References	Other relevant message-ids
Keywords	User chosen keywords
Subject	Short summary of the message for the one line display

MIME – Multipurpose Internet Mail Extensions

MIME adds structure to the message body and defines encoding rules for non-ASCII messages.

It defines five new message headers

Header	*Meaning*
MIME – Version	Identifies the MIME version
Content – Description	Human readable string telling what is in the message
Content - ID	Unique identifier
Content-Transfer-Encoding	Now the body is wrapped for transmission
Content – Type	Nature of the message

The content description : This is needed so the recipient will know whether it is worth decoding and reading the message.

The Content – id : Identifies the content

The content – Transfer-Encoding : Tells how the body is wrapped for transmission through a network that may object to most character other than letters, numbers, and punctuation marks.

Content-type : Video/mpeg

The subtype must be given explicitly in the header; no defaults are provided.

The text plain combination is for ordinary messages that can be displayed as received, with no encoding and no further processing.

SMTP (Simple Mail Transfer Protocol)

It is a simple ASCII protocol. Client sends the message and the server acknowledges it. The line sent by the client are marked c:, those sent by the server are marked as S; The first command from the client is HELO, a four characters abbreviation for HELLO. Since, the message is sent to only one recipient, only one RCPT command is used. A few problems can arise in SMTP, one relates to message length. Another relates to timeouts, if one of them gives up, while the other is still busy. To get around these problems, extended SMTP, has been defined. Instead of HELO, EHLO message is used. Email Gateways Email SMTP works when both the sender and receiver are on the Internet and can support TCP connections between sender and receiver. Many machines that are not on the Internet still want to send and receive email from Internet sites. Another problem occurs when the sender speaks only RFC 822 and the receiver speaks only X.400. Direct communication is impossible due to this problem. Both of these problems are solved using email gateways.

The procedure is that HOST 1 should establish a TCP connection to the gateway and then use. SMTP to transfer a message(s) there the demon on the gateway then put the message in a buffer of messages destined for host 2. Later a

TPH connection is established with host2 and the message is transferred using the OSI equivalent of SMTP. The gateways work is to extract the incoming messages from the quace and deposit them in another. It looks easy, but it is not. The first problem is that Internet addresses and x.400 addresses are totally different. An elaborate mapping is to be done. The second problem is that envelope or header fields that one present in one system and the other does not have this concept at all.

Final Delivery

All users who work on machines are not capable of sending and receiving emails.

A Simple protocol used for fetching email from a remote mailbox is Pop3 (Post Office Protocol). It has commands for the user to log in, log out fetch messages, and delete messages. The point of POP3 into fetch email from the remote mailbox and store it on the user's local machine to be read later. A more sophisticated delivery protocol is IMAP (Interactive Mail Access Protocol). The basic idea is for the email server to maintain a central repository that can be accessed from any machine. It does not copy email to the users personal machine, because the user may have several.

A third delivery protocol is DMSP (Distributed Mail System Protocol). It allows users to download email from the server to a workstation, PC or laptop and then disconnect. A valuable tool for many email users is the ability to set up filters. These are rules that are checked when email comes in or when the user agent is started. Each rule specifies a condition and an action.

Another delivery feature often provided is the ability to (temporarily) forward incoming email to a different address. Another feature of final delivery is the ability to install a vacation daemon. This is a program that examines each incoming message and sends the sender replies such as Hi.I'm on vacation. I'll be back on the 24th August. Have a nice day.

Email Privacy

Many people would like to be able to send email that can be read by the intended recipient and no one else : not their boss, not hackers, not even the Government. This desire has stimulated several people and groups to apply the cryptographic principles, to produce secure email.

PGP - Pretty Good Privacy

It is the brainchild of Phil Zimmermann. It provides privacy, authentication, digital signatures, and compression all in easy to use form. It intentionally uses existing cryptographic algorithms rather than inventing new ones. It is based on

RSA, IDEA and MD5. These algorithms have withstood expensive peer review. It supports text compression, secrecy and digital signatures and also provides extensive key management facilities.

Consider the following example : Alice wants to send a signed plaintext message P, to Bob in a secure way. Both Alice and Bob have private (Dx) and public (Ex) RSA keys. Each one knows the others public key. Alice starts out by involving PGP program on her computer. PGP hashes her message, P using MD5 and then encrypts the result using her private RSA key DA. When Bob eventually gets the message, he can decrypt the hash with alice's public key and verity that the hash is correct. Even if someone else could acquire the hash at this stage, and decrypt it with alice's known public key, it is infeasible to produce another message with the same MD5 hash.

The encrypted hash and the original message are now concatenated into a single message,P1 and compressed using the ZIP program. When Bob gets the message, he reverses the base 64 encoding and decrypts the IDEA key using his private RSA key. Using this key, he decrypts the message to get P1.Z.After decompressing it, Bob separates the plaintext from the encrypted hash and decrypts the hash using alice's public key. If the plaintext hash agrees with his own MD5 computation, he knows that p is the correct message and that it came from Alice.

PEM - Privacy Enhanced Mail

It is an official internet Standard. Messages sent using PEM are first converted to a canonical form so they all have the same conventions about white space and the use of carriage returns and line feeds. A message hash is computed using MD2 or MD5. Then it is encrypted using DES. This is encoded with base 64 coding and transmitted to the recipient. Each message is encrypted with a one time key that is enclosed along with the message. The key can be protected either with RSA or with triple DES using EDE.

Lesson - 9

APPLICATION LAYER

Design Issues of Application Layer

1. File Transfer, Access and Management:

File Transfer and remote file access are two of the most common applications in any Computer Network. There may be a necessity among people to transfer files. For this, one has to have the original file on a machine and have copies on other machines as needed. Remote file access is similar to file transfer, except that only piece of files are read or written, rather than entire files.

The key idea behind most modern file servers is that of a virtual key store, an abstract files server either free standing or running as a process on a timeshared computer. The virtual file store presents a standardized interface to its clients and provides a set of standardized operations that the clients can execute. Transfers to and from the virtual file store use standardized protocols (set of rules). If the real file server has a different internal structure that of the virtual file store, it will need some application layer software to hide the truth from the clients and make only the virtual file store interface visible.

2. Electronic Mail

Electronic Mail can be viewed as just a special case of file transfer. In this case the ultimate senders and receivers are always people, not machines. The difference between electronic mail and general purpose file transfer is that mail messages are lightly structured documents.

Many companies are interested in offering Electronic Mail as a standard service to companies and individual subscribers. To prevent world wide chaos, in 1984 CCITT(a committee responsible for issuing recommendations in the field of data communication) defined a series of protocols for what it calls MITs (Message Handling System) in its X•400 series of recommendations. ISO tried to incorporate them into the OSI application layer under the name MOTIS (Message Oriented Text Interchange System).

3. Virtual terminals

Nearly all terminals accept certain character sequences, called escape sequences for cursor motion. Entering and leaving reverse video mode, inserting and deleting characters and lines and so on. Each manufacturer has it own escape sequences, which may differ from those of every other manufacturer. So, it is difficult for anyone to write a screen editor that works with all keyboards and displays.

The OSI (Open System Interconnections) approach to solve this problem is to design a virtual terminal, which is really a data structure that represents the state of the real terminal in an abstract form. This data structure can be manipulated by both keyboard and the computer, with the current state of the data structure being reflected on the display.

4. Other Applications

Numerous other applications are there. Remote job entry allows a user working on one computer to submit a job for execution on another computer. Typically it is the user of a personal computer submitting a batch job for execution on a large mainframe somewhere else. In many cases, the program, data files, and job control statements must all be collected, possibly from different machines, bundled together and submitted as a unit. Finally, the output must be directed to the appropriate destinations.

Telematics is the collective name for public information services for home and office use. Teletex is a simple system in which a small amount of information can be sent to large number of people using television broadcasting. Videotext is an interactive service in which users can access large public databases and perform simple transactions like making railway reservations.

Virtual terminals

As already mentioned virtual terminal is really an abstract data structure that represents the abstract state of the real terminal.

They fall into three broad classes :

1. Scroll Mode
2. Page mode
3. Form mode

Scroll Mode Terminals

Scroll mode terminals do not have built in microprocessors or any local editing capability. When a key is hit it is sent over the line. When a key comes in over the line, it is just displayed. As new lines are displayed, the old ones just scroll upward. Some CRT (Cathode Ray Tube) terminals are of this type.

Despite their simplicity, they differ in many ways, for example, character set, line length, overprinting, carriage return (Enter key), line feed, horizontal tab, vertical tab, backspace, form feed, and break are handled. Some terminals also have potential timing problems.

Because scroll mode terminals do not have any processing power, they cannot communicate with the network using any of the network's standard protocol. To solve this problem, a "black box" was inserted between the network and the terminal. The black box speaks RS-232 to the terminal and some standard protocol to the network. It is generally called a PAD (Packet Assembler/Disassembler)

Page mode terminals

Page mode terminals are typically CRT terminals that can display 25 lines of 80 characters each. The computer can move the cursor around the screen to modify selected portions of the display. These terminals have all the same problems as scroll mode terminals, plus a few more such as page length, cursor addressing and the presence or absence of blinking.

When an editor starts up, it inquires about the terminal type and then reads in the entry for the terminal from a database called term cap (terminal capabilities). This entry gives the escape sequence required for each virtual command. As long as the software restricts itself to issuing virtual terminal commands, it will run on any terminal having a term cap entry.

Form mode terminals

Form mode terminals are terminals with built in microprocessors. In applications such as banking and airline reservations, the computer can download a form to the terminal, with some of the fields being read only and containing information, with the other fields to be filled in by keyboard input. The microprocessor can allow local editing and other facilities. These terminals are often called form mode terminals. When the form has been filled up, the microprocessor can run a quick syntax check on it. If the form is syntactically correct, the modified portion can be uploaded across the network back to the computer.

There are 2 types of virtual terminal models :

1. Synchronous model
2. Asynchronous model

In synchronous model, there is a single abstract data structure representing the screen image. Two identical copies of it is maintained, one by the virtual terminal software running near the terminal's microprocessor and the other by the virtual terminal software running near the application program on a distant host computer.

The person at the terminal can modify the microprocessor's copy of the abstract data structure by typing on the keyboard. These changes are visible in the screen because the screen display is driven by the local data structure. The microprocessor then sends virtual terminal commands to the distant host over the network using virtual terminal protocol. These protocol data units(PDU) cause the remote copy of the data structure to be brought up to date with the local one. The modified data structure can be read by the application program using appropriate commands.

Similarly, the application program can modify its copy of the data structure, which causes PDU's to be sent to the microprocessor to update the terminals data structure and hence its display.

To prevent from both sides trying to modify the data structure simultaneously a token is used to control update access and only the token holder may update the data structure. The token may be requested and passed back and forth, analogously to the tokens in the session layer, although virtual terminal token is unrelated to those tokens.

Alternately the Asynchronous model consists of two independent dialogs, rather than a single dialog. Each end of the connection has a data structure for input and a second one for output. In this, each copy of the data structure has a single reader and a single writer, so simultaneous write conflicts cannot occur.

Many virtual terminal protocols allow each side to send any legal sequence of commands that achieves the same net effect as the commands that were actually typed in. This optimization is called net effecting. The issue of how long the microprocessor should collect key stokes before transmitting them of their net effect is called delivery control. The more it collects, the larger the chance that some optimization can be done.

Another issue that virtual terminals must address is attention handling. All systems have some key that the terminal user can hit to interrupt the current command or program. Sometimes, it is the break hit; sometimes it is DEL. For killing an operation the user might hit the DEL key and kill that operation. But, the DEL must be acknowledged in such a way that the terminal can tell which characters were output before the DEL and which were output after it. If this is acknowledged then the terminal will be able to tell when to stop discarding.

Lesson - 10

MULTIMEDIA

Multimedia is just two or more media. It generally contains two media : text and graphics. Nevertheless, when most people refer to multimedia, they generally mean the combination of two or more continuous media, media that have to be played during some well-defined time interval, usually with some clear interaction. In practice, the two media are normally audio and video, that is, sound plus moving pictures.

Audio

When an acoustic wave strikes a microphone, the microphone generates an electrical signal, representing the sound amplitude as a function of time. The representation, processing, storage and transmission of such audio signals are a major part of the study of multimedia systems. The frequency range of the human ear runs from 20Hz to 20,000Hz. The ratio of two sounds with amplitudes B and C is conventionally expressed in dB according to the formula, dB = 20 log 10 (B/C).

If we define the lower limit of audibility for a 1Hz sine wave as 0 dB, an ordinary conversation is about 50 dB and the pain threshold is about 120 dB, a dynamic range of a factor of 1 million. To avoid any confusion, A and B above are amplitudes. If we were to use the power level, which is proportional to the square of the amplitude, the coefficient of the logarithm would be 10, not 20. Jitter of only a few milliseconds during a multimedia transmission affects the perceived sound quality more than it affects the perceived image quality.

The error introduced by the finite number of bits per sample is called the quantisation noise. Pulse code modulation, as used within the telephone system, uses 7-bit or 8 bit samples 8000 times per second. Audio CDs are digital with a sampling rate of 44,100 samples /sec, enough to capture frequencies up to 22,050Hz. The samples are 16 bits each, and are linear over the range of amplitudes. Using only 16bits/sample introduce some quantisation noise.

Digitised sound can be easily processed by computers in software. Many musical instruments have digital interface now. When digital interface first came out, each one 'had its own interface, but after a while, a standard. MIDI (Music Instrument Digital Interface), was developed. Each MIDI message consists of status byte followed by zero or more date bytes. A MIDI message conveys one musically significant event. Typical events are a key being pressed, a slider being moved, or a foot pedal being released. The status byte indicates the event and the data byte give parameters, such as which key was depressed and with what velocity it was moved.

Every instrument has a MIDI code assigned to it. For example, a grand piano is 0, a marimba is 12, and a violine is 40. The number of "instruments" defined is 127. The heart of every MIDI system is a synthesizer that accepts messages and generates music from them. The advantage of transmitting music using MIDI compared to sending a digitized waveform is the enormous reduction in bandwidth, often by a factor of 1000. Human speech tends to be in the 600Hz to 6000Hz range.

Vowels are produced when the vocal tract is unobstructed. Consonants are produced when vocal tract is partially blocked. These sounds are less regular than vowels. Some speech generation and transmission systems make use of models of the vocal system to reduce speech to a few parameters, rather than just sampling the speech waveform.

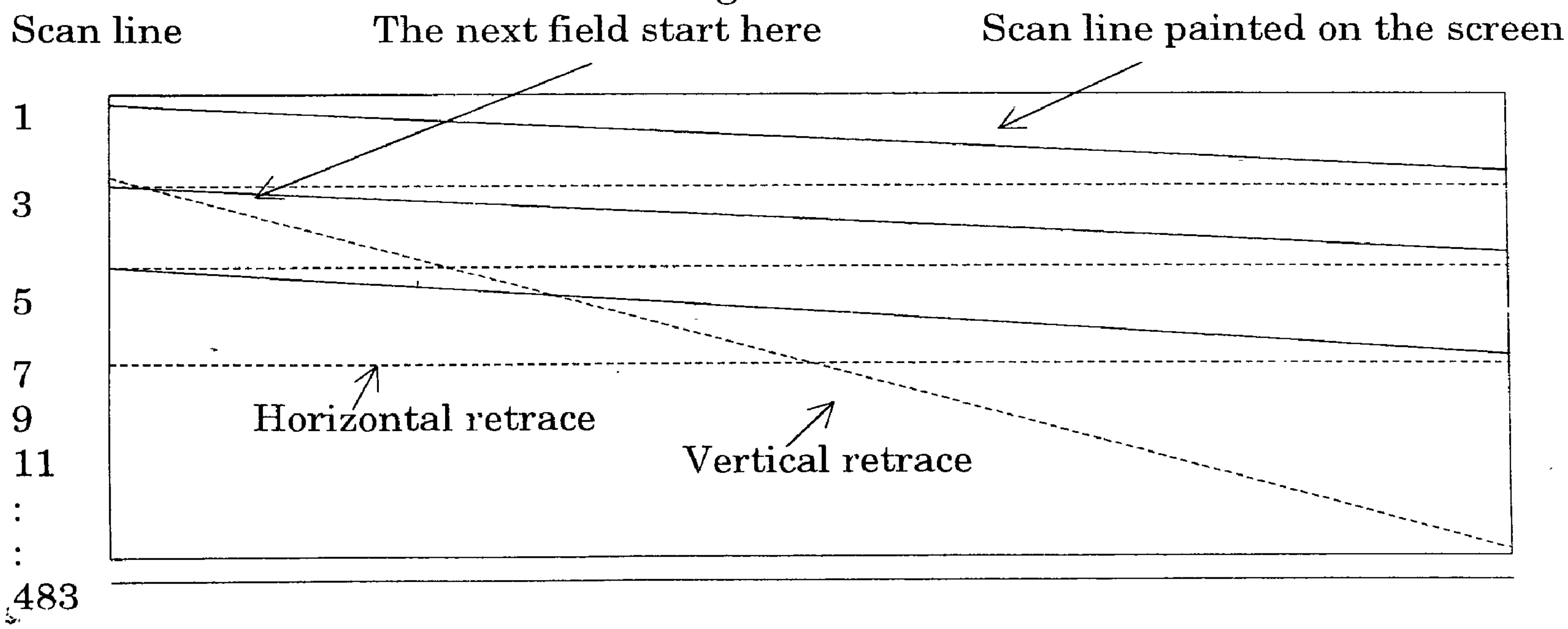
Video

The human eye has the property that when an image is flashed on the retina, it is retained for some number of milliseconds before decaying. If a sequence of images is flashed at 50 or more images/sec, the eye does not notice that it is looking at discrete images. All video systems exploit this principle to produce moving pictures.

Analog System

To understand Video systems, it is best to start with simple black-and-white television. To represent the 2-D image in front of it as 1-D voltage as a function of time, the camera scans an electron beam rapidly across the image and slowly down it, recording the light intensity as it goes. At the end of the scan, called a frame, the beam retraces. This intensity as a function of time is broad cast, and receivers repeat the scanning process to reconstruct the image. The scanning pattern used by both the camera and the receiver is shown in

Figure 10.1.



The exact scanning parameter's vary from country to country. The system used in North and South America and Japan has 525 scan lines, a horizontal to vertical aspect ratio of 4:3 and 30 frames/sec. The European system has 625 scan line, the same aspect ratio of 4:3 and 25 frames/sec.

In both systems the top few and bottom few lines are not displayed. The beam is turned off during the vertical retrace, so many stations use this interval to broadcast Tele Text. While 25 frames/sec. is enough to capture smooth motion, at that frame rate many people will perceive the image to flicker. Instead of displaying the scan lines in order, first all the odd scan lines are displayed, then the even ones are displayed. Each of these half frames is called a field. Experiments have shown that although people notice flicker at 25 frames/sec, they do not notice it at 50 fields/sec. This technique is called interlacing. Non interlaced television or video is said to be progressive.

Color video uses the same scanning pattern as monochrome except that instead of displaying the image with one moving beam, three beams moving in unison are used. One beam is used for each of the three additive primary colors: red, green and blue (RGB). This technique works because any color can be constructed from a linear superposition of red, green and blue with appropriate intensities. For transmission on a single channel, the three color signals must be combined into a single composite signal.

The first color system was standardized in the united states by the National Television Standards Committee (NTSC), the SECAM (Sequential Couleur Avec Memoire) is used in France and Eastern Europe, and PAL (Phase Alternating line) used in east of Europe.

To allow color transmissions to be viewed on black and white receivers all three systems linearly combine the RGB signals into a luminance signal, and two chrominance signals, which are broadcasted in narrowbands at higher frequencies.

Digital Systems

The simplest representation of digital video is a sequence of frames, each consisting of rectangular grid of picture elements, or pixels. Each pixel can be single bit, to represent either black or white. The next step up is to use 8 bits per pixel to represent 256 gray levels. This scheme gives high-quality black and white video. For color video, good systems use 8 bits for each of the RGB colors, although nearly all systems mix these into composite video for transmission. While using 24 bits per pixel limits the number of color to about 16 million, the human eye cannot even distinguish this many colors, let alone more. Digital color images are produced using three scanning beams, one per color.

To produce smooth motion, digital video, like analog video, must display at least 25 frames/sec. Since good quality computer monitors often rescan the screen from the images stored in memory at 75 times/second or more, interlaced scanning is not needed and consequently is not normally used. Just repainting the same frame three times in a row is enough to eliminate flicker.

Smoothness of motion is determined by the number of different images per second. A movie with 20 different frames per second, each of which is painted four times in a row, will not flicker, but the motion will appear jerky.

A better solution is to transmit 25 frames/sec and have the computer store each one and paint it twice. Broadcast television does not use this strategy because T.V. sets do not have memory, and in any event, analog signals cannot be stored in RAM without first converting them to digital form. As a consequence, interlacing is needed for broadcast television but not for digital video.

Data Compression

All compression systems require two algorithms: one for compressing the data at the source, and another for decompressing it at the destination. These algorithms are referred to as the encoding and decoding algorithms. For e.g.: A movie will only be encoded once but will be decoded thousands of times. It is acceptable for the encoding algorithm to be slow and require expensive hardware provided that the decoding algorithm is fast and does not require expensive hardware.

For real time multimedia, such as video conferencing, encoding must happen on the fly in the real time. When the decoded output is not exactly equal to the original input, the system is said to be lossy. If the input and output are identical, the system is lossless. Lossy systems are important because accepting a small amount of information loss can give a huge pay off in terms of the compression ratio possible.

Entrophy Encoding

Entropy Encoding just manipulates bit streams without regard to what the bits mean. It is a general, lossless, fully reversible technique, applicable to all data. It can be illustrated by three examples.

The first example of entropy encoding is run length encoding. In many kinds of data, strings of repeated symbols are common. These can be replaced by a special marker not otherwise allowed in the data, followed by the symbol comprising the run, followed by how many times it occurred. If the special marker itself occurs in the data, it is duplicated. For example, consider the following string of decimal digits.

3150000000000000845871111111111111111635467400000000000000000000. 65 If we now introduce A as the marker and use two digit numbers for the repetition count, we an encode the above digit string as 315S01284587S00316354674A02265.

Here run-length encoding has cut the data string in half. Our second example of entropy encoding is statistical encoding. By this we mean using a short code to represent common symbols and long ones to represent infrequent ones. Morse code uses this principle with E being '.' and Q being '-----' and so on.

The third example of entropy encoding is CLUT (Color Look up Table) encoding. Consider an image using RGB encoding with 3 bytes/pixel. In theory, the image might contain as 2 power 22 different colour values. A factor of almost three compression can be achieved by building a 768-byte table listing the RGB values of the 256 colors actually used, and then representing each pixel by the index of its RGB value in the table.

Source Encoding

Our first example is differential encoding, in which a sequence of values are encoded by representing each one as the difference from the previous value. It is lossy because the signal might jump so much between two consecutive values that the difference does not fit in the field provided for expressing differences, so at least one incorrect value will be recorded and some information lost.

The second example of source encoding consists of transformations. By transforming signals from one domain to another, compression may become much easier. Transformations are also applicable to 2-D image data. Suppose that the 4 x 4 matrix of fig 10.2(a) represents the gray scale values of a mono chrome image. We can transform these data by subtracting the value in the upper-left-hand corner from all elements except itself as shown in fig 10.2(b). This transformation might be useful if variable-length encoding is used. For example, values between -7 and +7 could be encoded with 4 bit numbers and values between 0 and 255 could b encoded as a special 4 bit code (-8) followed by an 8 bit number.

Pixel Value

160	160	161	160
161	165	166	158
160	167	165	161
159	160	160	160

(a)

4 pixels

160	0	1	0
1	5	6	-2
0	7	5	1
-1	0	1	0

(b)

Figure. 10.2

The third example of source encoding is vector quantization, which is also directly applicable to image data. Here, the image is divided up into fixed size rectangles.

In addition to the image itself, we also need a table of rectangles. This table is called the code books. Each rectangle is transmitted by looking it up in the code book and just sending the index instead of the rectangle. If the code book is created dynamically, it must be transmitted, too. If a small number of rectangles dominate the image, large savings in bandwidth are possible here.

The JPEG Standard

The JPEG (Joint Photographic Experts Group) standard for compressing continuous-tone still pictures. JPEG has four modes and many options.

Step 1 of encoding an image with JPEG is block preparation. For the sake of specificity, let us assume that the JPEG input is a 640x480 RGB image with 24 bits/pixel. Since using luminance and chrominance gives better compression. We first compute the luminance, Y, and the two chrominance I and Q according to the following formulas

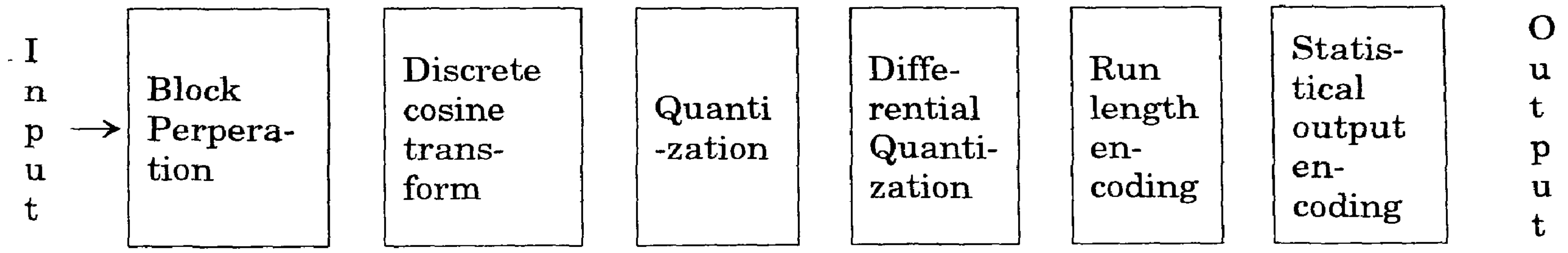

Input
Block Perpera-tion
Discrete cosine trans-form
Quanti-zation
Diffe-rential Quanti-zation
Run length en-coding
Statis-tical output en-coding
Output

Figure 10.3. The operation of JPEG in lossy sequential mode

$$
\begin{aligned}
Y &= 0.30R + 0.59G + 0.11B \\
I &= 0.60R - 0.28G - 0.32B \\
Q &= 0.21R - 0.52G + 0.31B
\end{aligned}
$$

Step 2 of JPEG is to apply a discrete cosine transformation of each of the 7200 blocks separately. The output of each DCT is an 8x8 motive of DCT coefficients. DCT element (0,0) is the average value of the blocks. In theory, a DCT is lossless, but in practice using floating-point numbers and transcendental functions always introduces some round off error that results in a little information loss.

Step 3, called quantization, in which the less important DCT coefficients are wiped out. This (lossy) transformation is done by dividing each of the cofficients in the 8x8 DCT matrix by a weight taken from a table. If all the weight are 1, the transformation does nothing. However, if the weights increase sharply from the origin, higher spatial frequencies are dropped quickly.

Step 4, reduces the (0,0) value of each block by replacing it with the amount it differs from the corresponding element in the previous blocks. The (0,0) values are referred to as the DC components; the other values are the AC components.

Step 5, linearizes the 64 elements and applies run-length encoding to the list. Scanning the block from left to right and then top to bottom will not concentrate the zero's together, so a zig-zag scanning pattern is used.

Step 6, Huffman encodes the numbers for storage or transmission. If often produces 20:1compression or better it is widely used.

The MPEG standard

The first standard to be finalized was MPEG-1. MPEG-1 can be transmitted over twisted pair transmission lines for modest distances. MPEG-1 is also used for storing movies on CD-ROM in CD-I and CD-Video format. The next standard in the MPEG family was MPEG-2, which was originally designed for compressing broad cast quality video into 4 to 6 mbps.

MPEG-4 is for medium-resolution video conferencing with low frame rates. This will permit video conferences to be held over a single N-ISDN B channel.

MPEG-1 has three parts : audio, video, and system, which integrates the other two as shown in fig. The audio and video encoders work independently which raises the issue of how the two streams get synchronized at the receiver. This problem is solved by having a 90KHz system clock that outputs the current time value to both encoders. These values are 33 bits, to allow films to run for 24 hours without wrapping around. These time stamps are included in the encoded output and propagated all the way to the receiver, which can use them to synchronize the audio and video streams.

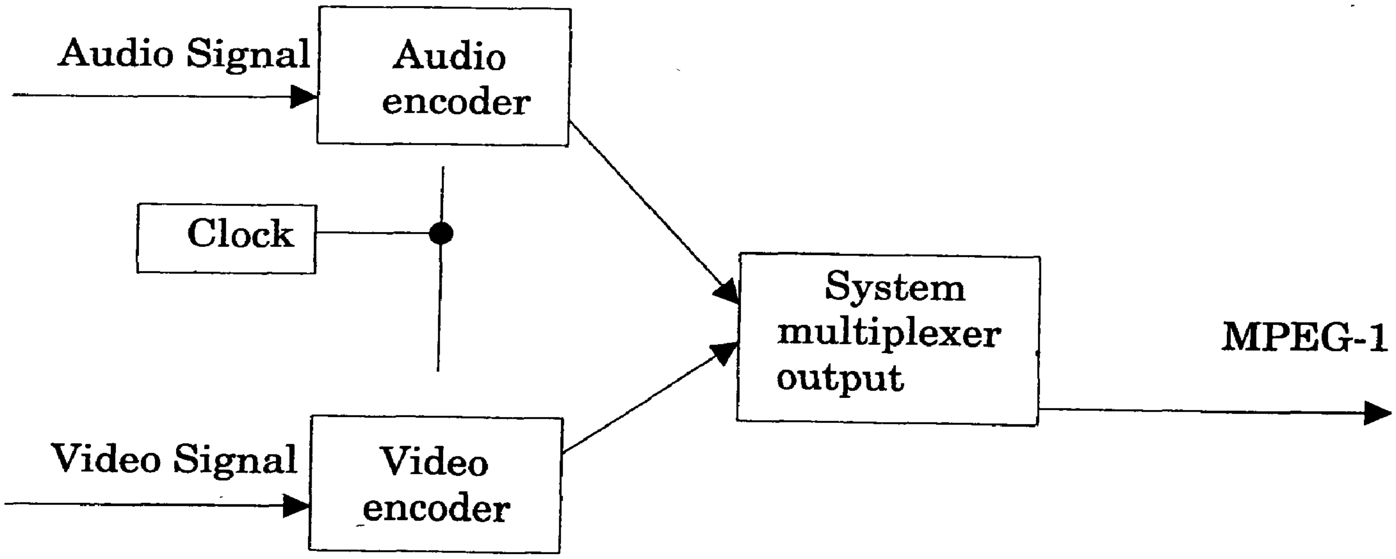


Figure 10.4 Synchronization of the Audio and Video streams in MPEG-1

MPEG audio compression is organized as three layers. Layer 1 is the basic scheme used in the DCC digital tape system. Layer 2 adds advanced bit allocation to the basic scheme. Layer 3 adds hybrid filters, non uniform quantization, Huffman coding, and other advance techniques.

MPEG-1 video compression has two kinds of redundancies existing in movies, Spatial and temporal. MPEG-1 uses both. Spatial redundancy can be utilized by simply coding each frame separately with JPEG.

MPEG-1 output consists of four kinds of frames :

1. I (Intracoded) frames : self-contained JPEG-encoded still pictures.
2. P (predictive) frames : Block-by-block difference with the last frame.
3. B (Bi-directional) frames : Differences with the last and next frame.
4. D (DC-coded) frames : Block averages used for fast forward.

I frames are just still pictures coded using JPEG, also using full-resolution luminance and half-resolution chrominance along each axis. P-frames in contrast, code inter frame differences. They are based on the idea of macro blocks,which cover 16x16 pixels in luminance space and 8x8 pixels in chrominance space. A macro block is encoded by searching the previous frame for it or something only slightly different from it.

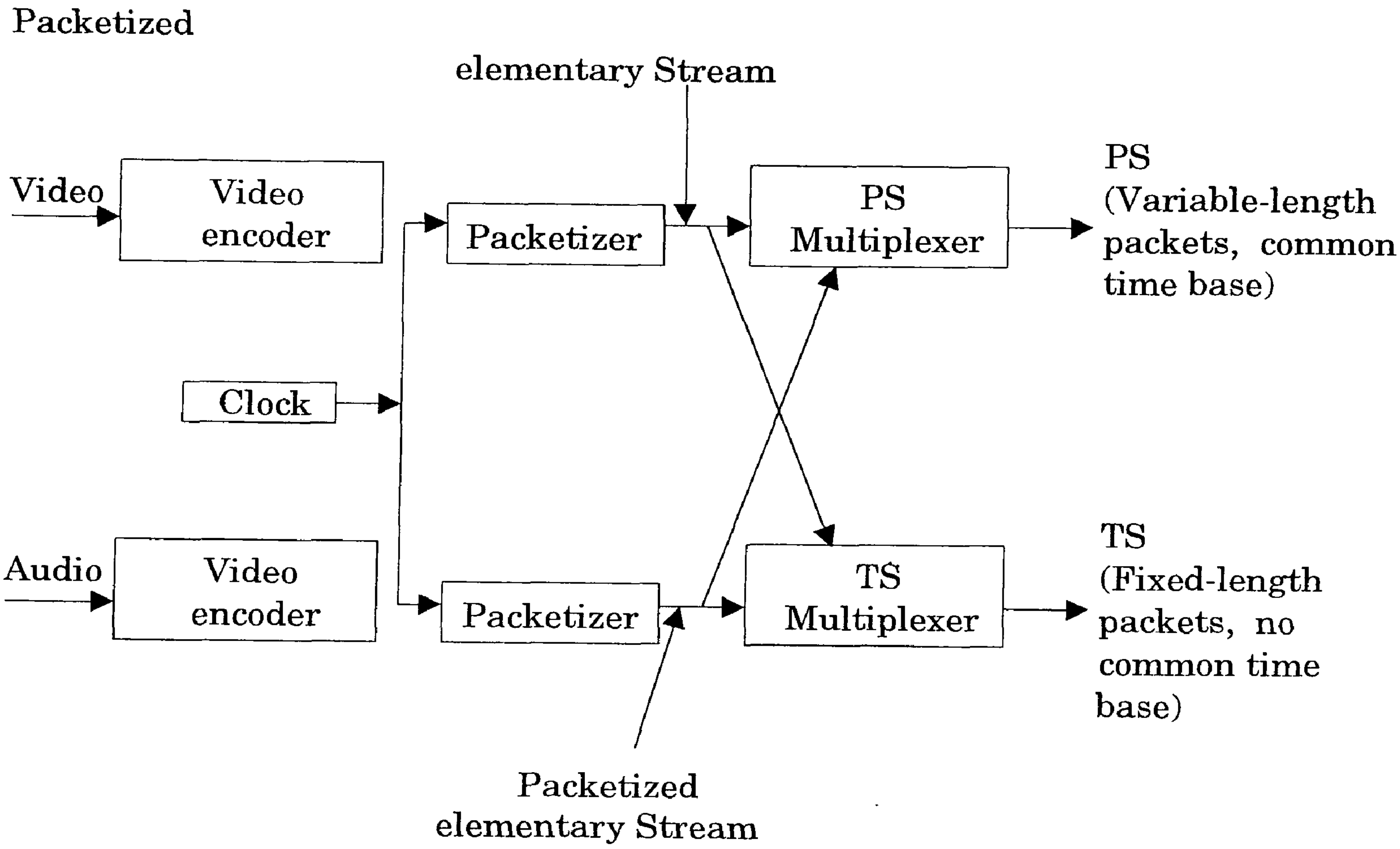
If a macro block is found, it is encoded by taking the difference with its value in the previous frame. These difference matrices are then subject to discrete cosine transformation, quanitzation, run-length encoding and Huffman encoding, just as with JPEG. The value for macro block in the output stream is then the motion vector, followed by the Huffman encoded list of numbers. If the macro block is not located in the previous frame, the current value is encoded with JPEG, just as in an I-frame.

B-frames are similar to P-frames, except that they allow the reference macro block to be in either a previous frame, or in a succeeding frame. This additional freedom allows improved motion compensation, and is also useful when objects pass in front of, or behind, other objects. Although B-frames give the best compression, not all implementations support them.

D-frames are only used to make it possible to display a low-resolution image when doing a rewind or fast forward. The D-frames are used to produce low-resolution images. Each D-frames entry is just the average value of one block, with no further encoding, making it easy to display in real time. This facility is important to allow people to scan through a video at high speed in search of a particular scene.

MPEG-2 does not support D-frames. In addition to having four resolution levels, MPEG-2 also supports 5 profiles. The main profile is for general purpose use, and probably most chips will be optimized for the main profile and the main resolution level. The simple profile is similar to the main one, except that it excludes the use of B-frames, to make software encoding and decoding easier.

The compressed data rate for each combination of resolution and profile is different. Each of the streams is first packetized with timestamps. A simple two stream example is shown in fig.



Packetized
elementary Stream
Video
Video encoder
Packetizer
PS Multiplexer
PS (Variable-length packets, common time base)
Clock
Audio
Video encoder
Packetizer
TS Multiplexer
TS (Fixed-length packets, no common time base)
Packetized elementary Stream

Fig.10.5 Multiplexing of two streams in MPEG-2

The output of each packetizer is PES (Packetized Elementary Stream). Each PES packet has about 30 header fields and flags, including lengths, stream identifiers, encryption control, copyright status, time stamps and a CRC. Two types of streams are defined. The MPEG-2 program stream is similar to the MPEG-1 systems stream. The other MPEG-2 stream is the transport stream. All the encoding schemes we have discussed are based on the model of lossy encoding followed by lossless transmission.

Video on Demand

Video on Demand is where the user selects any one of a large number of available videos and takes it home to view. Unlike video, television viewers do not expect to put programs on pause.

If video on demand is seen more as advanced television, then it may be sufficient to have the video provider start each popular video, say, every 10 minutes, and run these non-stop. Although pause/resume is not possible here, a viewer returning to the living room after a short break can switch to another channel showing the same video but 10 minutes behind. Some material will be repeated but nothing will be missed. This scheme is called 'near video on demand'. If offers the potential for much lower lost.

Video Servers

To have video on demand, we need video servers capable of storing and outputting a large number of movies simultaneously. The total number of movies ever made is estimated at 65,000. When compressed in MPEG-2, a normal movie occupies roughly 4GB of storage, so 65,000 of them would require something like 260 terabytes.

The cheapest way to store large volumes of information is on magnetic tape. A DAT type can store 8GB at a cost of about 5 dollars/gigabyte. The problem with these systems is the access time, the transfer rate, and the limited number of tape drives.

When there are N movies available, the fraction of all requests being for the Kth most popular one is approximately C/K. Here C is computed to normalize the sum to 1, namely

$$C=1/(1+1/2+1/3+1/4+1/5+\dots+1/N)$$

Thus the most popular movie is seven times as popular as the number seven movie. This result is known as Zipf's law.

The performance increases as one moves up the hierarchy in the fig 10.6.

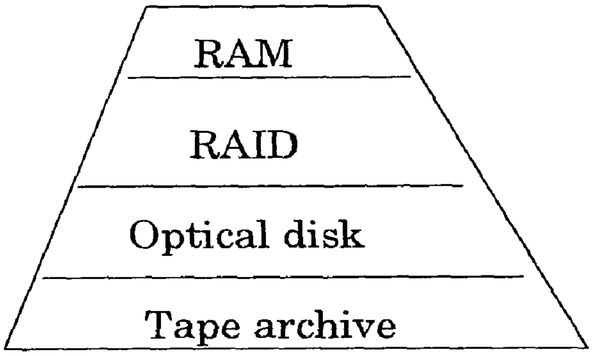


Figure 10.6. A Video Server storage hierarchy

An alternative to tape is optical storage. Current CD-ROMs hold only 650 MB, but the next generation will hold about 4GB, to make them suitable for distributing MPEG-2 movies. Magnetic disks have short access times (10msec), high transfer rates (10MB/sec), and substantial capacities (10GB) which make them well suited to holding movies that are actually being transmitted. The main drawback is the high cost for storing movies that are rarely accessed.

RAM is the fastest storage medium, but also the most expensive. It is best suited to movies for which different parts are being sent to different destinations at the same time.

Two possible ways of organizing disk storage are the disk farm and the disk array. With the disk farm, each drive holds a few entire movies. For performance and reliability reasons, each movie should be present on at least two drives may be more. The other storage organization is the disk array or RAID (Redundant Array of Inexpensive Disks), in which each movie is spread out over multiple drives.

Disks are scheduled using the elevator algorithm, which starts the arm moving inward and keeps going until it hits the innermost cylinder, processing all requests it hits in cylinder order. When it gets as far as it can, the arm reverses and starts moving outward, again processing all pending requests along the way in order.

The Distribution Networks

The distribution network is the set of switches and lines between the source and destination. The four main local distribution schemes for video on demands go by the acronyms ADSL, FTTC, FTTH and HFC.

ADSL (Asymmetric Digital Subscriber Line) is of the idea that virtually every house in united states, Europe, and Japan already has a copper twisted pair going into it. If these wires could be used for video on demand, the telephone companies could clean up.

In FTTC (Fiber To The Curb) the telephone company runs optical fibre from the end office into each residential neighbourhood, terminating in a device called an ONU (Optical Network Unit).

In FTTH (Fibre To The Home) scheme, everyone can have an OC-1, OC-3, or even higher carrier if that is required. A completely different approach is HFC (Hybrid Fibre Coax), which is preferred solution currently being installed by cable TV providers. The current 300 to 450 MHZ Coax cables will be replaced by 750MHZ coax cables.

Set- Top Boxes

PC's usually have small screens, are located in studies or dens rather than in living rooms and are traditionally used by one person at a time. They also emit significantly less time than television sets.

The local network operator rents or sells the user a set-top box to which the network and television set are connected. This approach has the advantage that every one has a television but not everyone has a PC, and many of the PCs that people do have are old, peculiar or other wise unsuited to MPEG decoding.

MBone-Multicast Backbone

MBone can be thought of as Internet radio and television. Unlike Video on demand, where the emphasis is on calling up and viewing recompressed movies stored on server, MBone is used for broadcasting live audio and video in digital form all over the world via the Internet.

Network Security

Substitution Ciphers

In this each letter or group of letters is replaced by another letter or group of letters to disguise it.

Plain text : abcdefghijklmnopqrstuvwxyz

Cipher text : QWERTYUIOPASDFGHJKLZXCVBNM

This system is called a monoalphabetic situation.For the key above, the plain text "attack" becomes "QZZQEA".By using the most common 2 letter combinations (or) diagrams and 3 letter combinations called Trigrams we can break the cipher text.

Another method is guessing a word or a phrase. If the cipher text for e.g. Pertains to an accounting firm, the word financial would surely occur more than once. Thus, tracking such words will break the cipher.

The next method is polyalphabetic cipher (or) vignere ciphers. It consists of a square matrix with 26 alphabets. Then a unique key is used to encryption.

Other methods are Random number generation, porta's cipher, etc.

Codes

Codes are used for larger cipher units. A cipher encrypts fixed size unit of plaintext where as a code encrypts variable length linguistic unit. Codes are of 2 types namely one part code and two part code. An one part code uses same code for encryption and decryption while a 2 part code uses separate codes for encryption and decryption.

Codes and ciphers are combined to form superencipherment.

Transposition Ciphers

These ciphers reorder the letters rather than disguising them. For e.g., consider "Megabuck" as a key. The purpose of key is to number the columns. To break a Transposition cipher, a person should be aware that he is dealing with a transposition cipher. The method of Transpositioning is as follows,

MEGABUCK
7 4 5 1 2 8 3 6
P l e a s e t r
a n s f e r o n
e m i l l i o n
d o l l a r s t
o m yswi s s
b a n k a c co
u n t s i x t w
o t w o a b cd

Plain text : pleasetransferonemilliondollarsto
myswissbankaccountsixtwotwo

Cipher text : AFLLSKSOSELAWAIATOOSSCTCLNMOMANT
ESILYNTWRNNTSOWDPAEDOBUOERIRICXB

Again, phrase guessing may be helpful breaking these ciphers.

Symmetric Cryptography

This type of cryptography uses a same key called "secret key" for both encryption and decryption. The Data Encryption standard (DES) is a popular algorithm for symmetric cryptography. The currently used algorithms use long codes and short keys.

The circuits used for ciphers are as follows :

1. P-Box

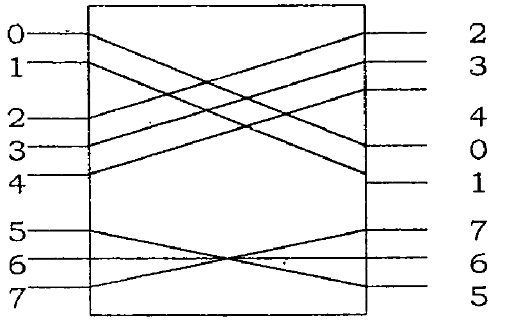


Figure 10.7

P-Box is called as Permutation box and it produces 8 bit Transposition. In the above figure from top to bottom, the plain text of the form 0 1 2 3 4 5 6 7 becomes 2 3 4 0 1 7 6 5 after 2 permutated transposition.

2. S-Box :

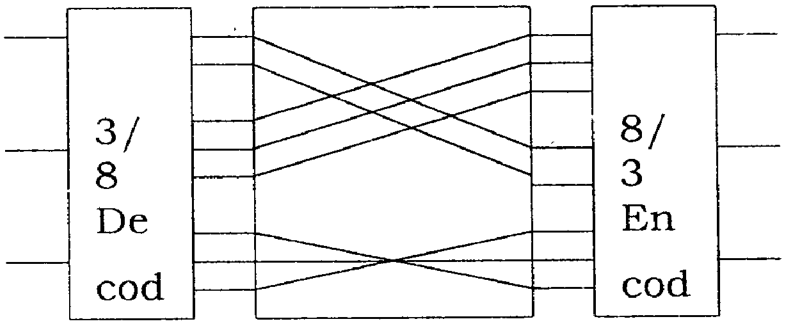


Figure 10.8

Here, a 3 bit plain text is given as input and 3 bit cipher text is got at the output. The 3 bit input selects one of the 8 lines and sets it to 1. The second state is P-Box. The final stage encodes input to binary again.

The power of S-Box and P-Box is revealed when they are cascaded resulting in product cipher.

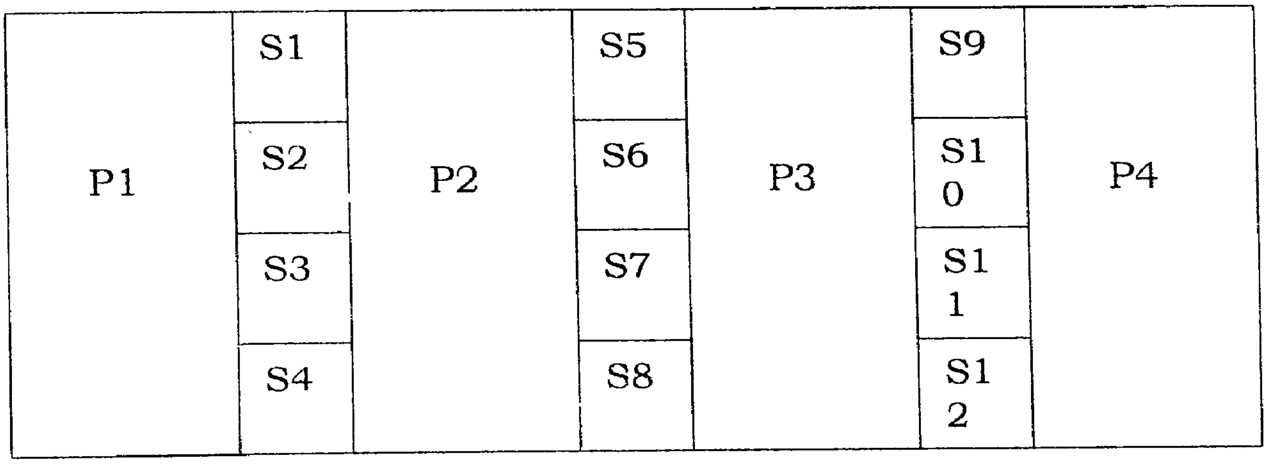


Figure 10.9

Here,

1, 2 input lines are transposed at first stage. Theoretically, a second stage (S-Box) would require 212 = 4096 wires. Instead of doing that, the input is broken into group of 3 bits. Thus, as no. of stages increases, output becomes non clear to that of input.

(DES) Data Encryption Standard is the popular algorithm for secret key cryptography.

DES

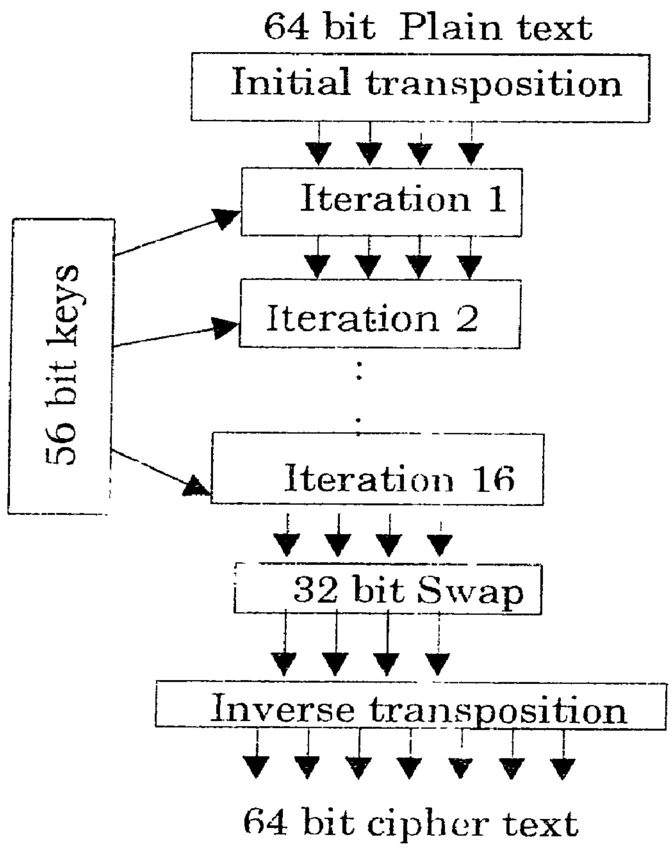


Figure (a)

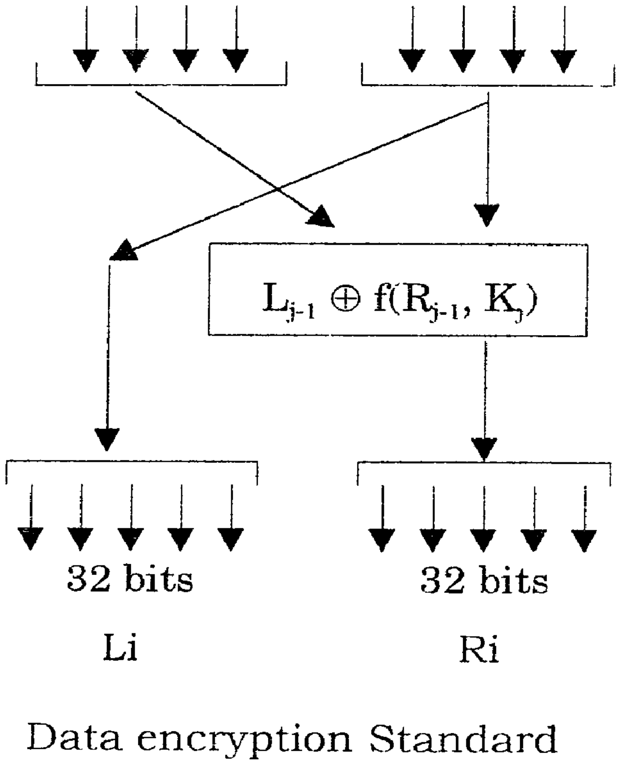


Figure (b)

The plain text is encrypted in blocks of 64 bits yielding 64 bit cipher text. The algorithm, which is parameterized by a 56 bit key, has 19 dishnet stages. The first stage is a key independent transposition on 64 bit plain text. The last stage is exactly the inverse of Transposition. The stage prior to last exchanges left most 32 bits with rightmost 32 bits. The remaining 16 stages are functionally identical.

The operation of intermediate stages is shown in Fig (b). Each stage takes 2 32 bit inputs and produces 2 32 bit outputs. The left output is a copy of right input. The right output is bitwise XOR of left input and a function of right input and key ki for that stage.

The function has 4 steps,

1. A 48 bit number 'E' is constructed by expanding 32 bit Ri-1 by Fixed Transposition and duplication rule.
2. E and ki are xor'ed. The output is partitioned into r groups of 6 bits each and fed into S-Box. The S-Box produces 4 instead of 6 outputs.
3. Each of 64 inputs to S-Box is mapped on to a 4 bit output. The result is a list of 8 4 bit nos.
4. Finally, these 32 bits are passed into P-Box.

Each iteration uses a different key. At first, 56 bit Transposition is applied to the key. For each iteration, it is partitioned into 2 28 bit units, each of which is rotated left by a no. of bits defending on iteration no. ki is derived from this rotated key by applying yet another 56 bit Transportation to it.

Stream Encryption

When stream cipher is used, both sender & receiver operate DES chips in encryption mode. Each DES chip 64 bit input register, which operates as a shift register. A plain text is XOR ed with 8 bits of output register 01. The character thus created is both transmitted to receiver and shifted into input register, pushing I8 off end.

At receiving end, incoming character is XOR ed with 01 and shifted to I1.

Public key cryptography

When there exists 2 keys, one for encryption and one for Decryption, such a for of cryptography is called as Asymmetric (or) Public key cryptography.

Public key cryptography has 3 requirements namely,

1. D (E(P)) = P where
   D Decryption Algorithm
   E Encryption Algorithm
   P Plain Text

2. It is exceedingly difficult to deduce D from E
3. E cannot be broken by a chosen plaintext attack.

Consider 2 parties A and B who need to have a secure Transaction. Both A's Encryption keys. EA and B's Encryption key EB. Now, A takes message P, computer EB(P) and sends it to B. B then decrypts it applying secret key DB. Thus DB(EB(P)) = P.

The MIT Algorithm

The MIT Algorithm can be explained as follows :

1. Choose 2 large primes, p1 and q1, each greater than 10^{100}.
2. Compute n = p1 x q1 and x = (p1-1) (q1-1)
3. Choose a no. relatively prime to Z and call it d.
4. Find e such that e x d = 1 mod Z

Using these parameters in advance, we are ready for encryption.

To encrypt a message P, compute C = pe (mod n). To decrypt C, compute P=cd (mod n). To perform encryption, e and n are needed. To perform decryption, d and n are needed.

For eg.,

p1 = 3	n1 = 33
q1 = 11	z1 = 20

Now, d1 = 7

7e = 1(mod 20)   e = 3

$c = p^3$ (mod 33)
$p1 = c^7$ (mod 33)

Such a level of encryption requires 10^{13} years for decryption.

Authentication and Digital Signatures

Authentication is a process of revealing the identity of sender and Digital signatures. Help in doing so, authentication is required because

1. The receiver can verify identity of sender

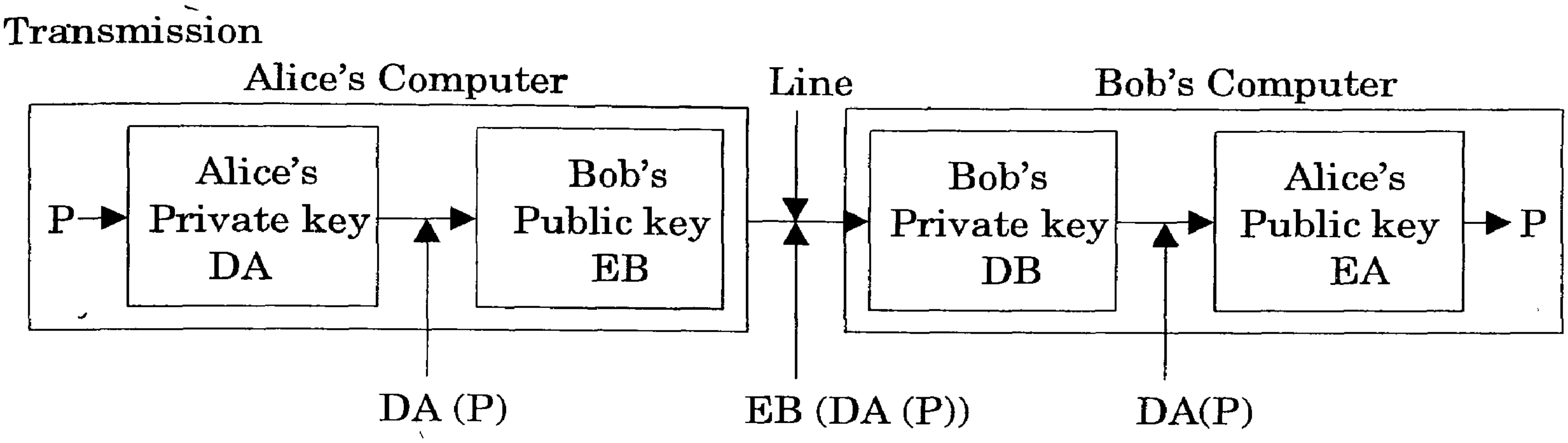
2. The sender cannot repudiate message later

Digital Signatures with Public key cryptography

For public key cryptography with digital signatures, the following properties are to be satisfied,

(i) E (D(P)) = P

(ii) D (E(P)) = P



Digital Signatures using Public-key Cryptography

Figure (10.11a) *Figure (10.11b)*

Digital Signatures using Public-key Cryptography

Fig 10.11b explains signatures. Here, the plain text p is encrypted first by A's secret key (DA) and then by B's public key (EB). In the receiving end, B applies his secret key (DB) to get DA(P) and then applies A's public key which results in the plain text p.

The above method reveals both senders and receivers identity and hence its secure.

One Way checksum

When, the sender and receiver are interested more in Authentication rather than security, this method is highly useful. Here, a check sum function CK is used and checksum value for plain text 'p1' is calculated (CK(P1)). Now, A1 applies has private key DA1 yielding DA1 (CK(P1)). Now, A1 transmits P1, DA1(CK(P1)) to B1.

B1 applies EA1 to yield CK(P1). B1 holds P1, CK(P1), DA1(CK(P1)). B1 now applies CK to P1 to check whether received CK(P1) and computed CK(P1) match one another. If so, message is correct. Else, he concludes that message has been tampered with and initiates retransmission.

Thus, such methods enhance the overall Network security.

SNMP - Simple Network Management Protocol

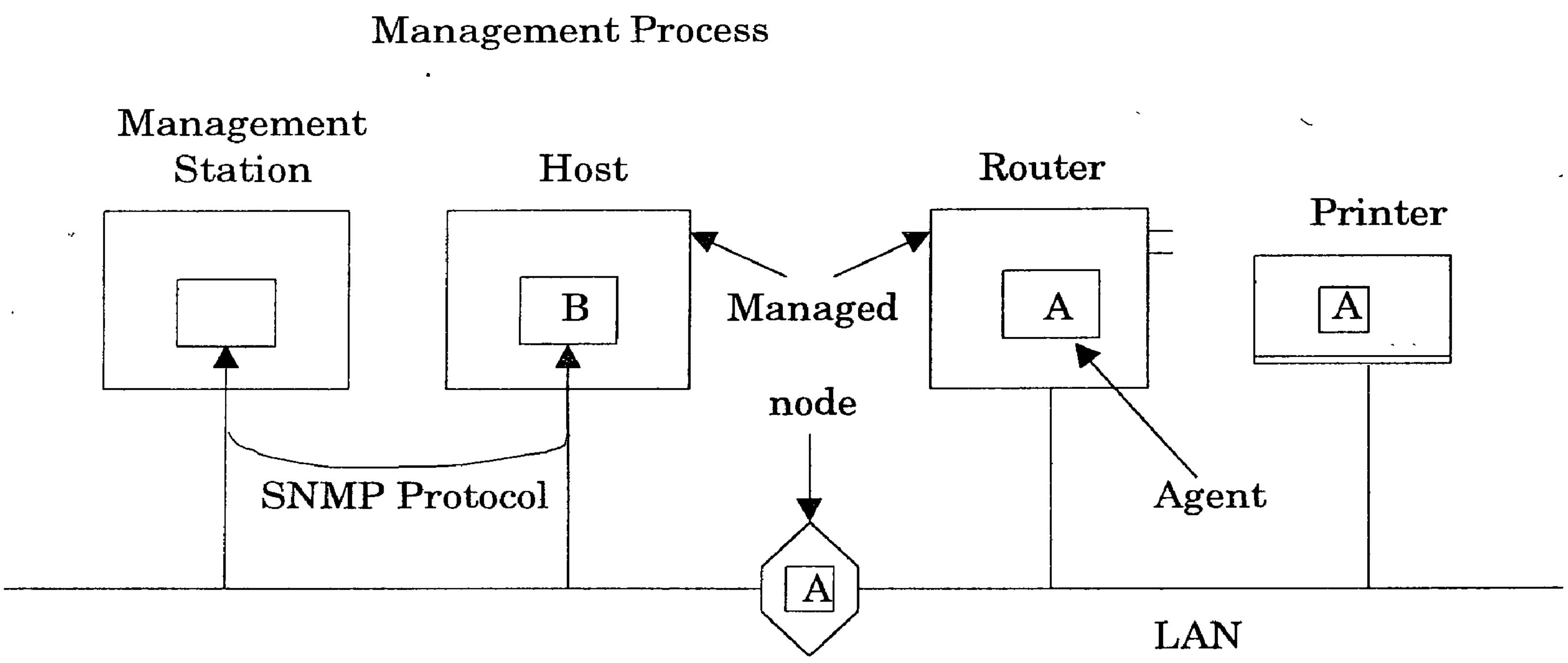
SNMP provides a systematic way of monitoring and managing a computer network. This framework and protocol were widely implemented in commercial products and became the defaults standards for network management.

SNMP Model

The SNMP model of a managed network consists of four components.

1. Managed Nodes
2. Management Stations
3. Management Information
4. Management Protocol

These pieces are illustrated as below.



Componenets of SNMP Model

Figure 10.12

The managed modes can be hosts, routes, bridges, printers or any other devices capable of communicating status information to the outside world. To be managed directly by SNMP, a node must be capable of running an SNMP management process, called an SNMP agent.

Network Management is done from start, which are, in fact, general purpose computers running special management software. The management stations contain one or more processes that communicate with the agents over the network, issuing commands and getting responses. In this design all the intelligence is in the management stations, in order to keep the agents as simple as possible and minimize their impact on the devices they are running on.

In order to allow a management station to talk to all these diverse components, the nature of the information maintained by all the devices must be rigidly specified. Therefore, SNMP describes the exact information each kind of agent has to maintain and the format it has to supply it.

Each device in the network, maintains one or more variables that describe its state. In the SNMP literature, these variables are called objects. The collection of all possible objects in a network is given in a data structure called the MIB (Management Information Base). The management station interacts with the agents using the SNMP protocol. This protocol allows the management station to query the state of an agent's local objects and change them if necessary.

ASN 1- Abstract Syntax Notation 1

The heart of the SNMP model is the set of objects managed by the agents and read and written by the management station. To make multivendor communication possible, it is essential that these objects be defined in a standard and Vendor-neutral way. Further more, a standard way is needed to encode them for transfer over a network.

For this reason, a standard object definition language, along with encoding rules is needed. The one used by SNMP is taken from OSI and called ASN 1. The one alleged strength of ASN 1 is now really a weakness, because the encoding rules are optimized to minimize the number of bits on the wire, all the cost of wasting CPU time at both ends encoding and decoding them. Nevertheless, for better or worse SNMP is drenched in ASN 1. So anyone wishing to truly understood SNMP must become fluent in ASN 1.

The ASN 1 abstract syntax is essentially a primitive data declaration language. It allows the user to define primitive objects and then combine them into more complex ones. A series of declarations in ASN 1 is functionally similar to the declarations found in the header files associated with many 'C' programs.

SNMP has some lexical conventions. These are not entirely the same as pure ASN 1 uses, however built-in datatypes are written in uppercase. User-define types begin with an uppercase letter but must contain at least one character other than an uppercase letter. Identifiers may contain upper and lowercase letters, digits and hyphens, but must begin with a lowercase letter. White space is not significant.

The mechanism that is used to define a standard tree, and place every object in every standard at a unique location in the tree.

The top level of the tree lists all the important standards organizations in the world, namely ISO and CCITT plus the combination of the two. From the ISO node, four are defined, one of which is for identified organization which is ISO's concession that may be some other folks are vaguely involved with standards too.

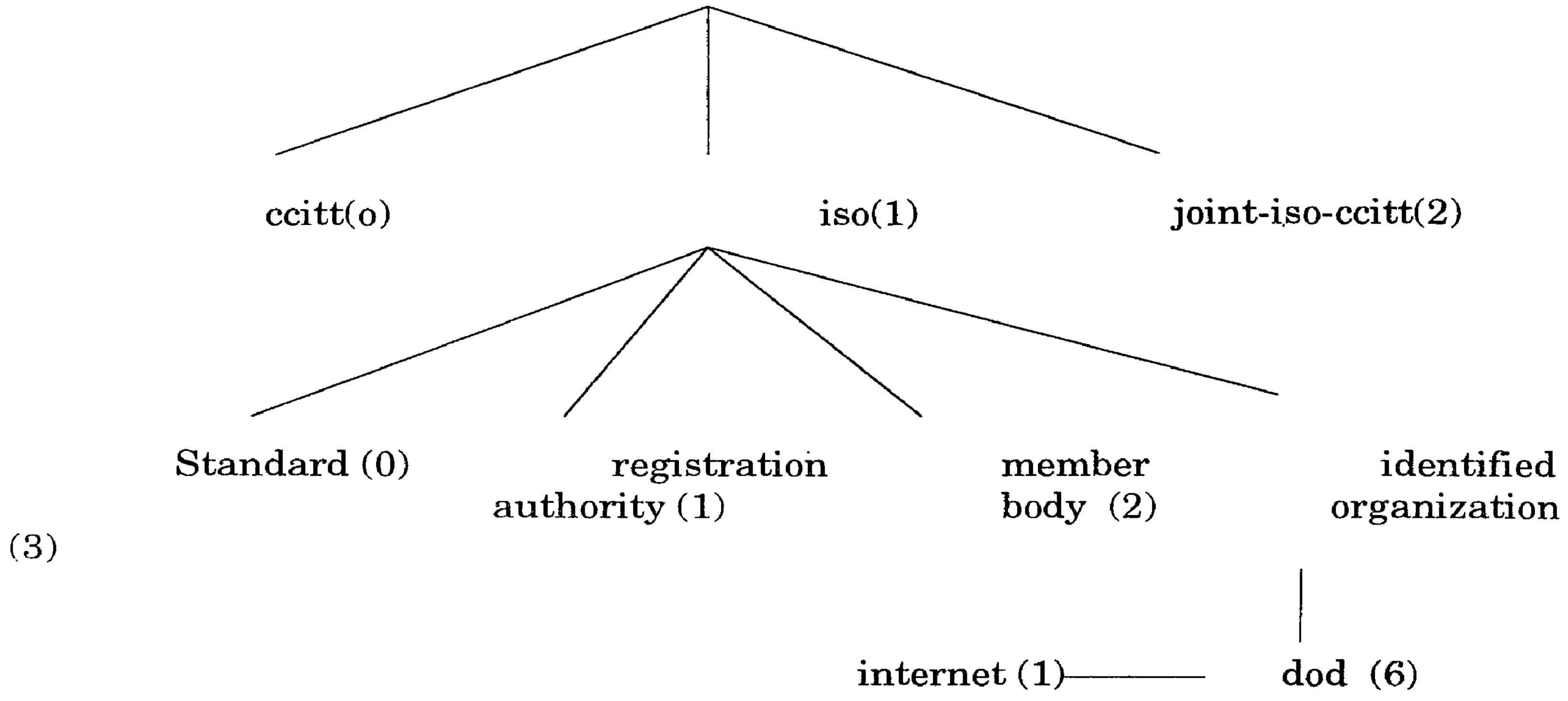


Figure 10.13

Every arc in the fig. has both a label and number, so nodes can be identified by a list of arcs, using label or numbers. Thus all SNMP MIB objects are identified by a label of the form.

ISO identified organization (3) clod (6) internet (1) mgmt (2)
mib 1 (1)

or alternatively [136121 ....]

ASN 1 Transfer Syntax

An ASN 1 transfer syntax defines how values of ASN 1 types are unambiguously converted to a sequence of bytes of transmission. The transfer syntax used by ASN 1 is called BER (Basic Encoding Rules).

The rules are recursive, so the encoding of a structured object is just the concatenation of the encodings of the component objects. In this way all object encodings can be reduced to a well defined sequence of encoded primitive objects. The encoding of these objects, in turn is defined by BER.

The guiding principle behind the basic encoding rules is that every value transmitted both primitive and constructed ones, consists of up to four fields.

1. The identifier
2. The length of the data field in bytes
3. The data field
4. The ends of content fley if the data length is unknown.

The last one is permitted by ASN 1, but specifically forbidden by SNMP, so we will assume the data length is always known.

The encoding of the data field depends on the type of data present. Integers are encoded in two's complement. A positive integer below 128 requires 1 byte, a positive integer below 32,768 requires 2 bytes and so forth. The most significant bit is transmitted first.

Bit strings are encoded as themselves. The only problem is how to indicate the length. The length filed tells how many bytes the value has, not how many bits. The solution chosen is to transmit 1 byte before the actual bit string telling how many bits of the final byte are unused. Thus the encoding of the 9 bits string '010011111' would be 07,4F,80.

Octet Strings are easy. The bytes of the string are transmitted in standard big Indian style, left to right.

The null value is indicated by setting the length field to O.No, numerical value is actually transmitted.

An OBJECT IDENTIFIER is encoded as the sequence of integers it represents.

An example showing encoding of some values is given in following figure. The values encoded are INTEGER-49, the OCTET STRING '110', 'XY', the only possible value 'NULL', the OBJECT IDENTIFIER for the internet {1,3,6,1} and a Guage 32 value of 14.

SMI – Structure of Management Information

It is this sub-super-set of ASN 1 which goes by the ungainly name of SMI (Structure of Management Information), that is really used to define the SNMP data structures.

At the lowest level SNMP variables are defined as individual objects. Related objects are collected together into groups and groups are assembled into modules. For example, group exist for IP objects and TCP objects.

All MIB modules start with an invocation of the MODULE-IDENTITY macro. Its parameters provide the name and address of the implementer, the revision history and other administrative information. Typically, this call is followed by an invocation of the OBJECT-IDENTIFY macro, which tells where the module fits in the naming tree.

Later on some or more invocations of the OBJECT – TYPE macro, which name the actual variables being managed and specify their properties, Grouping variables into groups is done by convention, there are no BEGIN – GROUP and END-GROUP statements in ASN 1 or SMI.

A Simple example of an OBJECT TYPE declaration is given in following fig. The variable is called lost packets and might be useful in a router or other device dealing with packets. The value after the ::: sign places it in the tree, lost packets OBJECT TYPE

```
    SYNTAX Counter 32             ---    Use a 32 bit Counter
    MAX – ACCESS read only        ---    the management station may not
                                         change it
    STATUS Current
    DESCRIPTION                   ---    this variable is not obsolete (yet)

    "The number of packets lost since the boot"
:: = {experimental 20}
```

MIB - Management Information Base

The collection of objects managed by SNMP is defined in the MIB for convenience, these objects are grouped into ten categories, which correspond to the ten nodes under mib-2 in naming tree. The ten categories are intended to provide a basis of what a management station should understand. The ten categories are summarized as follows:

GROUP	OBJECTS	DESCRIPTION
System	7	Name, Location and description of the equipment
Interfaces	23	Net work interfaces and their measured traffic
AT	3	Address Translation
IP	42	IP packet statistics
ICMP	26	Statistics about ICMP messages received

TCP	19	TCP algorithms, parameters and statistics
UDP	6	UDP traffic statistics
EGP	20	Exterior gateway protocol traffic statistics
Transmission	0	Reserved for media specific MIBs
SNMP	29	SNMP traffic Statistics

SNMP PROTOCOL

The normal way that SNMP is used is that the management station sends a request to an agent asking it for information or commanding it to update its state in a certain way. Ideally, the agent just replies with the requested information or confirms that it has updated its state as requested. Data are sent using the ASN.1 transfer syntax. However, various errors can also be reported, such as No such variable.

SNMP defines seven messages that can be sent. The six messages from an initiator are listed in the following table, and the seventh message is the response message. The first format names the variables it wants explicitly. The second one asks for the next variable, allowing a manager to step through the entire MIB alphabetically. The third is for large transfer such as tables.

MESSAGE	DESCRIPTION
Get - Request	Requests the value of one or more variables
Get-next-request	Requests the variable following this one.
Get-back-request	Fetches a large table
Set-request	Updates one or more variables
Inform-request	Manager-to-manager message describing local MIB

Then comes a message allows the manager to update an agent's variables, to the extent that the object specification permits such update of course. Next is an informational request that allows one manager to tell another one which variables it is managing finally comes the message sent from an agent to a manager when trap has spring.

USENET News

The most popular application of computer networking is the world wide system of newsgroups called net news. Net news is referred to as USE-NET News. It is used to separate UNIX to UNIX physical network that once carried the traffic using a program called UUCP.

USENET and INTERNET are not the same. Internet sites do not get news but some other sites get news without being on the Internet.

The User View of USENET

A newsgroup is a world wide discussion on specific topic. Those people who are interested can subscribe to this newsgroup. Subscribers should use a special kind of a user agent called a new stealer to read all the articles posted to a news group. Over all news group is somewhat like a mailing list, but internally it is implemented differently.

The number of groups is so large that they are arranged in a hierarchy to make them manageable. The most common USENET hierarchies

Name	*Topic covered*
Comp	Computers, Computer Science & Industry
Sci	The physical sciences and engineering.
Humanities	Literature and the Humanities.
Rec	Recreational sports music etc.

The comp groups were the original USENET groups. These groups are popular among computer professionals. The science and Humanities groups are populated by scientists, scholars and amateurs with interest in physics, chemistry etc.

Each categories is broken into subcategories recursively. For example rec. sport is about sports rec. sport. basket ball.

Numerous News Readers exists. When the newsreader is started, it checks a file to see which news group the user subscribes to. It then typically displays a one line summary of each as yet unread article in the first newsgroup and wait for the user to select one or more for reading.

Functions of News Reader

News reader allow user to subscribe and unsubscribe to news group. Changing a subscription simply means editing the local file listing which newsgroups the user is subscribed to Newsreader allow handle posting. The user composes an article and then gives a command or cricks on an icon to send the article on its way. Within a day it will reach almost every one in the world subscribing to the news group to which it posted.

The most important properties of an article is its "Cross post" nature. Cross post means sending of multiple groups with a single command.

Unfortunately some people use their new found power to communicate to a large group irresponsibly. When some one post a message saying "People like you should be slot" tempers flare and a torrent of abusive posting called a flame war.

This type of situation is solved in two ways one individual and one collective. Individual users can install a kill file, which specifies that articles with a certain subject or from a certain person are to be discarded upon arrival, prior to being displayed.

A moderated news group is one in which only one person, the moderator can post the article to the newsgroup. All postings to a moderated news group are automatically sent to the moderator who post the good one and discounts the bad one.

Since thousands of people subscribe to USENET for the first time every day the same beginners questions tend to be asked over and over. To reduce this traffic, many news (FAQ) that tries to answer the Questions that beginners have.

USENET is fuel of jargon such as BTW (By the way), ROFL (Rolling on the floor laughing) and the IMHO (In My Humble Option). The ASCII Symbols used are called smileys (or) emoticons. Although most people use their real name in posting, some people wish to remain totally anonymous, especially when posting personal ads to news group dealing with finding partners.

This desire had lead to the creation of anonymous remailers, which are servers that accept email messages, and changes the From; Sender and Reply to. As more and more people subscribe to USENET, there is a constant demand for new and more specialized groups. Consequently a procedure has been created for creating new ones.

News group creation is less formal in the all hierarchy and this is in fact that the reason alt exists. (alt-It is a chaotic, unregulated mish-mash of news groups on all topics some of which are popular and most of which are world wide). Some of the news group that are close to the legal and moral edge of what is tolerable that they would never have been accepted in a public vote. Much of the alt hierarchy is fairly conventional.

Implementation of USENET

Some smaller newsgroups are implemented using mailing lists. To post an article to one such newsgroup (ie mailing list) one sends to mailing list address, which causes copies to be sent to each address on mailing list.

USENET is not generally implemented using mailing list. Instead each site stores incoming subdirectories for comp.sci,etc. These in turn have many subdirectories.

News readers just fetch the articles from the directories. This arrangement means that each sites need only one copy of each new article, on matter how many people subscribe to its newsgroup.

A site must have a new speed from another site in USENET. The transmission line connecting pair of nodes forms the arcs of the graph. This Graph is USENET. Periodically each site that wants news can poll its news feeds asking if any new news has arrived since the previous contact. If so that news is collected and stored in the appropriate subdirectory of news.

Not every site get all news groups. There are several reasons here.

(i) Total news feed exceeds 500 HB per day and is grouping rapidly. Storing it all required large desk space.

(ii) Transmission time and cost are issues.

(iii) Not every site is interested in every topic.

News articles have the same format as RFC 822 email messages but with the addition of a few new extra headers.

The headers are

Path : Header is the list of the message traversed to get from the poster to the recipient. At each hop the forwarding machine puts its name at the front of the list. The use of exclamation marks (pronounced bang) go back to UseNet address which predate DNS.

News groups : Tells which news group the message belong. It may contain more than one group. Message correspond to multiple news groups will contain all of their names.

Follow UpTo : It is needed to tell people where tc post comments and reactions to put all of the subsequent discussion in one group.

NnTp-Posting Host : It tells which machine actually posted the article even it was composed in different machines.

The Reference : Indicate that this article is a response to an earlier article and gives the ID of that article.

Organization : Tells what company, university or agency the poster is associated with.

The Lines : Gives the length of the body.

The Subject : It ties the discussion threads together.

Summary : Used to summarize to follow up article.

NNTP - Network News Transfer Protocol

The initial algorithm just flooded articles onto every line with in USENET. The volume of traffic made this scheme impossible. This disadvantage was overcome by a protocol called Network News Transfer Protocol, which is defined in RFC 1977. NTTP has something of same flavor as SMTP with a client issuing commands in ASCII and a server issuing responses as decimal numbers coded in ASCII.

NTTP was designed for two purposes.

(i) Allow news article to propagate from one machine to another over a reliable connection.

(ii) To allow users who develop computers cannot receive news to read, read news remotely.

Two general approaches are possible.

(i) News pull, the client calls one of its news feed and ask for new news.

(ii) News push the new feed calls the news and announces that it has news.

To acquire recent articles the client must first establish a TCP connection with port 119 on one of its news feeds. Behind this port is the NTTP daemon which is neither their all the time waiting for clients or is created on the fly as needed. After the connection has been established, the client and server communicate using a set of commands.

The main ones used for moving command between daemons are.

Command Meaning

The LIST & NEWSGROUP commands allow the client to find out which group the server has. the LIST gives the complete list. News group gives those groups created after time and date specified.

Eg. comp.'s* means all news group that start with comp.os.

After the assemble of all articles in the list we an ask some particular article using ARTICLE command. The client can offer articles using I Have command and can post by using the Post command. When the client is done it terminates the session using Quit command.

Example :

Whole some.net (information provider)
(aim : to avoid controversy at all wst)

News group offered are

Soc.couples	misc.: kids

(i) Whole some.com first checks for any news to soc. couples. If there it is stored in a separate file. Each file is named by (article) number.

(ii) Having got all the news about the groups it carries, it checks for new groups and it is told that two news groups have appeared.

(iii) Next whole some.com offers feeder.com a new article posted by someone on the site.

(iv) The new push approach is similar.

A problem with new push & pull is that they use stop & wait.

The World Wide Web

The World Wide Web is an architectural frame work for accessing linked documents spread out over thousands of machines all over the Internet. The web began in 19789 at CERN the European centre for nuclear research. The initial proposal of a web for linked documents came from CERN physicist Tim Berners Lee in March 1989. The first prototype was operational 18 months later. In 1994 CERN

M.I.T. Signed an agreement setting up the world wide web consortium, an organization devoted to further developing the web standardizing protocols and encouraging interoperability between sites.

The Client Side

Web consist of a vast, world wide collection of documents usually just called pages for short. Each page may contain links to other related pages anywhere in the world. Pages that point to other pages are said to be HyperText.

Pages are viewed with a program called Browser of which Mosaic and Netscape are famous. The Browser fetches the page requested interrupts the text and formatting commands that it contains and displays the page properly formatted on the screen.

Strings of the text that are links to other pages called Hyper Links are highlighted either by underlying (or) displaying them in a special colour or both. To follow a link a user places the user on the highlighted area and enter it. (see fig).

Most Browsers have numerous buttons and features to make it easier to navigate the web. Many have a button to go to the previous page, next page or borne page.

In addition to having ordinary text and hypertext, web pages can also contain icon line drawing maps and photograph. Clicking on one of these elements causes the browser to fetch the linked page and display it, the same as clicking text.

Welcome to the University of East Podunk's WWW

- Campus Information
- Admission information.
- Campus map.
- Direction to campus.
- Academic Departments
  - Department of Animal Psychology.
  - Department of alternative studies.

The Department of Animal Psychology

- Information for prospective majors.
- Personnel
  - Faculty members
  - Graduate students
- POSITION VALUE
- Research Projects

When Hyper Text pages are mixed with other media it is called Hyper Media. Some browsers can display all kinds of hyper media but others cannot. Instead they check a configuration file. The file gives the name of program called external viewer (or) a helper application.

Many web Pages contain large images, which take a long time to load. Some page writers attempt to placate potentially bored users by displaying images in a special way. First the image quickly appears in a coarse resolution. Then details are gradually filled in. For the user seeing after few seconds albeit at low resolution, is often preferable of seeing it to build up slowly from the top scan line by scan line.

Some pages contain forms that request the user to enter information. Typical applications of these forms are searching a database; for a user supplied item ordering a product or participating in a public opinion survey. Other web pages may contain maps. Some Browsers use the local disk to cache pages that they have fetched.

To host a web Browser a machine must be directly on the **Inter net or at least** have a SLIP or PPP connection to a router or other machine that is **directly on the** Internet. This requirement exist because the way the Browser fetches a **page is to** establish a TCP connection to an arbitrary machine on the Internet, a **browser, a** browser will not work.

The Server Side

Every website has a server process listening to TCP port 80 for incoming connections from client. After a connection has been established the client sends one request and server sends one reply. Then the connection is released. The steps that occur between the user dick and the page being displayed are as follows

(i) The Browser determines the URL.
(ii) The Browser asks DNS for the IP address.
(iii) DNS reply.
(iv) The Browser makes a connection to TCP.
(v) It then sends a GET/hypertext command.
(vi) The server sends the file.
(vii) The TCP connection is released.
(viii) The Browser displays all the text in Project HTML.
(ix) The Browser fetches and displays all images in Project HTML

Many Browsers display which step they are currently executing in a status line at the bottom on the screen. When the performance is poor the user can see it if it is due to DNS not corresponding.

Since HTTP is an ASCII protocol like SMTP it is quite talk to web servers. All that is needed is a TCP connection to port 80 on the server. The simplest way to get such connection is to use the Telnet program.

Not all servers speak HTTP. In particular many older servers use the FTP, Gopher or other protocols. Since a great deal of useful information is available on FTP and Gopher servers, one of the design goals of the web was to make this information available to web users.

Instead a different solution is often used, proxy servers. A proxy server is a kind of gateway that speaks HTTP to the browser but FTP, Gopher or some other protocols to the server. It accepts HTTP requests and translates them into say. FTP request, so the Browser does not have to understand any protocol except HTTP. The proxy server can be a program running on the same machine as the browser but it can also be on a free standing machine somewhere in the network serving many Browsers.

Name	*Used For*	*Example*
http	HTML	http://www.cs.nl.nl/~ast/
ftp	FTP	ftp://ftp.cs.vu.nl/pub/minix/README
file	Local File	/usr/srzanne/prog.c
news	News group	news : comp.o.minix.
news	News article	new : AA0134223112@cs.utan edu
gopher	Gopher	gopher ://gopher.tc.umn.edu/11/Libraries
mailto	Sending Email	mailto.kin@acm.org
telnet	Remote login	telnet://www.w3.org:8

In addition to acting as a go between for unknown protocols, proxy servers have a number of other important functions such as caching. A caching proxy server conects and keep all the pages that pass through it when a user asks for a page, the proxy server to checks if it has the page. If so it can check to see if the page is still current. In the event that the page is still current it is passed to the user. Otherwise a new copy is fetched.

Hyper Text Transfer Protocol-HTTP

The standard web Transfer protocol is HTTP. HTTP is constantly evolving. Several versions are in use and others are under development. The material presented below is relatively basic and is unlikely to change in concept, but some details may be a little different in future versions.

The HTTP protocol consists of two fairly distinct items the set of requests from browsers to servers and set of responses going back the other way. All the new version of HTTP support two kinds of request. Simple request and Full request. Simple request is just like a single GET line naming the page desired with out the protocol version. Full requests are indicated by the presence of the protocol on the GET request line. Request may consist of multiple lines followed by a blank line to indicate the end of request, which is why blank line is needed. The built in HTTP request methods.

Method	*Description*
GET	Request to read a web Page
HEAD	Request to read a web Page's Header
PUT	Request to store a web Page
POST	Append to a named resource (e.g.a web Page)
DELETE	Remove the web Page.
LINK	Connects two existing resources
UNLINK	Breaks an existing connection between two resources.

The GET method requests the server to send the page. Get request is followed by an if modifier since header, the server only sends the data if it has been modified since the data supplied.

The Head method just asks for the message header, without the actual page.

The put method is the reverse of get instead of reading the page, it writes the page. This method makes it possible to build a collection of web page on remote server.

Somewhat similar to PUT is the POST method. It too bears a URL but instead of replacing the existing data, the new data is appended to it some generalized sense.

Delete does what you might except. There is no guarantee that DELETE succeeds since even if the remote HTTP server is willing to delete the page.

The LINK and UNLINK methods allow connections to be established between existing pages or other resources. The HTTP standards describe message headers and bodies in considerable detail.

Writing a Web Page in HTML

Web Pages are written in languages called HTML. It allow users to produce Web pages that include text, graphics and painters.

Uniform Resource Locator

When the web was created it was immediately apparent that having one page point to another web page required name for mechanism and locating pages. In particular there were three questions had to be answered before a selected page could be displayed.

1. What is the page called?
2. Where is the page located?
3. How can the page be accessed?

Each page is assigned a URL that effectively serves as the pages world wide name. URL'S have three parts (i) Protocol (http) (ii) DNS name of host (iii) Filename (welcome.html). Many sites have certain shortcuts for filename built in. For example ~user/might be mapped onto users www directory. Some common URL (1-1).

Name	*Used For*	*Example*
http	hyper text (HTML)	http://www.cs.nl.nl/~ast/
file	Local File	/usr/srzanne/prog.c
news	News group	news : comp.o.minix.
telnet	Remote login	telnet://www.w3.org:8

Figure 10.14 Some Common URLs

The http protocol is the web native language, the one spoken by HTTP servers.

The ftp protocol is used to access files by FTP, the Internet file transfer protocol. FTP has been around more than two decades and is well entrenched.

The news protocol allows a web user to call up a news article as though it were a web page. This means that a web Browser is simultaneously a news reader. Two formats are supported for News protocol. The first format specifies news group. The second one requires the identifier of a specific news article to be given.

The Gopher protocol is used by Gopher System. Gopher's big advantage over the web is that it works very well with ASCII terminals.

The telnet protocol is used to establish an on line connection to a remote machine. It is used the same way as the Telnet program, which is not suprising since most browsers must call the program as a helper application.

Despite all the nice properties the growing use of web has turned upon a internet weakness in the URL scheme. A URL paints to one specific host. For pages that are heavily referenced, it is desirable to have multiple copies far apart to reduce network traffic. The trouble is that URL's do not provide any way to reference a page without simultaneously telling where it is.

HTML-Hyper Text Mark Up Language

HTML is a markup language a language for describing how documents are to be formatted. The term "mark up" comes from the old days when copy editors actually marked up documents to tell the printer a human being which is fonts to use and so on.

Documents written in Mark Up Language can be contrasted to documents, produced with a WYSIWYG (What you see is what you Get) word processor such as MS WORD or Word Perfect. By embedding the mark up commands within each HTML file and standardizing them it become possible for any web browser to read and format any web page.

Disadvantages of WYSIWYG

(i) Internal markup languages are not standardized among vendors machines, and operating system.

(ii) They do not handle reformatting for different sized windows and different resolution displays.

However, work processing program can offer the option of saving documents in HTML instead of in vendor proprietary form, and some of them already do.

A proper web page consist of head and a body enclosed body <HTML> and </HTML> tags although most browsers do not complain if these tags are (complaining) missing. The head is directed <HEAD> and the </HEAD> tags and the Body is bracketed by the </BODY> tags. The commands inside the tags are directives. MOST HTML tags have this format of <SOMETHING> to begin and </SOMETHING> to end. Tags can be either in lower case (or) upper case. Thus <HEAD> and <head> have same meaning.

Some of the common HTML tags are:

Tag	*Description*
<HTML>......</HTML>	Declares web page to be written in HTML
<HEAD>......</HEAD>	Delimits the Pages Head.
<TITLE>......</TITLE>	Defines the title
<BODY>......</BODY>	Delimits the Body
<Hn>......</Hn>	Delimit a level n heading
<B>......</B>	Set in....boldface
<L>......</L>	Set in.....Italics
<UL>......</UL>	Blocks an unordered list.
<OL>......</OL>	ordered list
<MENU>......</MENU>	Brackets a menu of <LI> items.
<LI>......</LI>	Start a list item
<BR>......</BR>	Force a Break here
<P>......</P>	Start a paragraph
<HR>......</HR>	Horizontal Rule
<PRE>......</PRE>	Preformatted Text, do not reformat
<IMG SRC="...>	Load an image

Heading is generated by Hn tag where n is digit in range from 1 to 6. The tags <B> and <l>. If the Browser is not capable of displaying boldface at italics it must use some other method of rendering them.

The <UL> tag starts an unordered list. <OL> is for ordered list. <LI> items are number by the browser. <MENU> which typically produces a more compact list on the screen with no bullets and no numbers.

<DIR> can be used for making short tables. Also <DL> and </DL> can make definition list with two part entries whose part are defined by <DT> and <DD> respectively. The <BR> <P> and <HR> tags all indicate a boundary between sections of text. The <BR> tag just forces a line break. In contrast <P> starts a (photograph) paragraph, which might, for example insert a blank line and possibly some indentation.

To prevent Browsers from messing up carefully laid out the text the <PRE> and </PRE> tags were provided. Other parameters of <ING> are ALIGN which controls the alignment of text with respect to base line. ALT which provides text to use instead of the image when the user has disabled images and ISMAP a flag indicating that the image also activates hyper link. Finally we come to hyper links which use <A> and </A> tags. One feature that HTML did not link is the ability to create tables.

As a consequence a large amount of work was done to add tables to HTML. HTML tables consist of one (or) more rows each consisting of one (or) more cells.

An HTML table definition is listed in Fig and a possible redefinition shown below.

ITEM	*HTML 1-0*	*HTML 2-0*	*HTML 3-0*
Active maps & Images		x	x
Equations			x
Forms		x	x
Hyper Links	X	x	x
Images	X	x	x

<CAPTION> tag is used to provide caption. <TR> Table Row tag <TH> Table Header Tag <TD> Table Date.

Forms

As more and more commercial organizations began using the web there was a large d^mand for two way traffic. These demands led to the inclusion of forms starting in HTML 2.0. Forms contain boxes or buttons that allow users to fill in information back to pages owner. They use <INPUT> tag for this purpose.

Like all forms this one is enclosed between the <FORM> and </FORM> tags. Text not enclosed in a tag is just displayed. The VALUE parameters are used to indicate which radio button was pushed. Depending on which of the credit card options the user has chosen the variable cc will bc set to either the string "master card" or string "viscard". After the two sets of radio Buttons, we come to slipping option represented by a box of type CHECKBOX.

As an aside for very long lists from which a choice must be made, radio buttons are somewhat inconvenient. Therefore the <SELECT> & </SSELECT> tags are provided to a bracket a list of alternatives but with the semantics of radio buttons. We have now seen two of the built in types for the Input tag RADIO and CHECKBOX. In fact we have already seen a third one as well: TEXT. Because this type is the default, we did not bother to include the parameter TYPE=TEXT but we could have. Two other types are PASSWORD and TEXTAREA.

A PASSWORD box is the same as a TEXT BOX except that the characters are not displayed as they are typed.

A TEXTAREA BOX is also same as a TEXTBOX except that it can contain multiple lines.

The Browser also understands RESET button. Two more types are worth nothing. The first is the HIDDN type. This is output only it can be clicked or modified only. The last type is IMAGE, which is for active maps. The ACTION parameter specifies the URL to tell about submission and METHOD parameter tells which method to use.

The way the form's variable are sent back to the page's owner depends on the value of METHOD parameter. If the POST METHOD is used, the body of the message contains the forms variable and their messages.

The standard which is used for handling form datas are called common Gateway Interface. (CGI) CGI scripts can also produce output and do many other this as well as accepting input from forms.

JAVA

The main idea of using JAVA for interactive web pages is that a web page can paint to a small JAVA program called an APPLET. When the browser reaches it the applet is down loaded to the client machine and there is a secure way. Applets allow web pages to become interactive. They make it possible to add animation and sound web pages without having to spawn external viewers.

It is entirely possible that in the long run, the model of people buying programs, installing them and running them locally will be replaced by a model in which people click on web pages, get applets download to do work for them, possibly in conjunction, with a remote server (or) database. The JAVA system has three parts.

(i) A JAVA to byte code compiler.

(ii) A Browser that understands applets.

(iii) A Byte code Interpreter

Introduction to Java Language

Java has 8 primitive data types. Each type as a specific size, independent of local implementation. Thus unlike C where on integer may be 16, 32, or 64 Bits depending on underlying machine architecture.

Type	*Size*	*Description*
Byte	1 Bytes	A signed integer between 128-127
Short	2 Bytes	A signed 2 Byte integer
Int	4 Bytes	A signed 4 Byte integer
Float	8 Bytes	A signed 8 Byte integer
Long	4 Bytes	A 4 byte IFEE flptnol
Double	8 Bytes	An 8 Byte IFEE flat pt nol.
Boolean	1 Byte	The only values are True & Fales
Char	2 Bytes	A character in Unicode

Arithmetic variables can be combined using the usual arithmetic operators and compared using the usual relational operators. Conversion between types are permitted where they make sense. Java allows one dimensional array to be declared. For example unit [ ] table. The Java control statements are.

Statement	*Description*	*Example*
Assignment	Assign a value	n=i+j
Switch	Select a case	Switch (b) { cose 1:i++;}

For	Iteration	for (i=0;i<n-1; i++
		a[1]=b[1];

Do	Repetition	do {n=n+n}
Throw	Raise exception	throw new illegal argument Excepted
Try	Exception scoping	try {........} catch exception
Synchronized	Mutual Exclusion	Synchronized valid update

The throw statement raises an exception and the try statement defines a scope to associate exception handlers with a block of code in which an exception might occur.

The synchronized statement is new Java and has to do fact that Java programs can have multiple threads to control. To avoid race conditions this statement is used to delimit block of code that must not have more than one thread activate it at once. Such blocks of code as usually called critical regions.

Java programs can be called with arguments. Command-line processing is similar to C except that the argument away is called args instead of argc & argv[o] is first parameter not program name.

Object Orientation in Java

A Java program consists of one or more packages each of which contain some definitions. Packages can be accessed remotely over a network. So these intended for use by a wide audience must have unique name. A class definition is a template for stamping out object instances, each of which contains the same variables and same as all the object instance of its class.

Each class is based on another class. A newly defined class is said to have subclass of the class on which it is based the superclass. A subclass always intents the methods of its superclass. The property of a class automatically acquiring all the methods of its superclass is called inheritance and it is the most important property of Java. Adding new methods to superclass methods is called extending superclass.

Let us now take a look at an example of object oriented concepts presented so far.

```
Class ComplexNumber

        protected double re,in;
        public void complex (double x, double y) {re=x; in=y;}
        public double Real ( ) {return re;}
        public double Imaginary ( ) {return in.;}
        public double magnitude ( ) {return Math. sqrt (re*rer+irn*in); }
        public double angle ( ) {return Math. atam (in/re);

Class test {
        public static void main (string aigs [ ]) {
        ComplexNumber C;
        C=new ComplexNumber ( );
        C. complex  (3.0 4.0);
        system. out. println ("The magnitude of c is "+c. Magnitude ( ) );  }
```

Each object of this class contain two hidden variables re and im, both 64 Bit floating point numbers. Five methods are defined on objects belonging to class complex Number.

When this package is compiled the Java compiler produces two Binary levels one containing each of the classes and named often each classes. Typing the command.

JAVA test results in invoking the Java interpreter with class test as parameter. An extension of complex number which we will call is hairy Number. The new class automatically inherits the five methods present in super class. To make thing interesting we will define a six method, Add to in the subclass, which adds a complex number to objects, increasing its real and imaginary parts.

A Java class may define multiple methods with the same name but different parameters and different definitions. When the compiler sees the name invocation using this name, it has to use parameter type to determine which method to use. This property is called overloading (or) polymorphism.

The Application Programmer's Interface

In addition to Bare-language the Java designers have defined and implemented about 200 classes with initial release. The classes are written in Java. So, they are portable to all platforms and O.S. The 200 classes are grouped into seven packages of uneven size each of which is focused on central theme. The seven packages are.

Package	*Example Functionality*
java.lang	Classes, threads, math, strings
java.co	I/o on streams & random access files
java.net	Sockets, IP addresses
java.util	Stacks, hash tables, vectors, time & date
java.applet	Getting and displaying web pages
java.awt	Events dialog, menu, foots.
java.awt.image	Colors, image, filtering
java.awt.peer.	Access to underlying window system

Java language package contains classes that can be viewed as part of language, but are technically not. Input/output is done by loading using java.io package. Java.util contains classes and methods for common data structure such as stacks and trash tables. Java.applet package contains some of basic machinery for applets including methods for getting web pages starting from their URL's.

AWT stands for Abstract Window Toolkit and is designed to make applets portable across window system.

Security

One of the most important aspects of Java is its security properties. The first line of defense is a type safe language. Java has strong typing true arrays with bounds checking and no pointers.

The second life of defense is that before an incoming applet is executed, it is run through a byte code verifier. The byte code verifier looks for attempt to manufacture pointers, execute instruction or call methods with invalid parameters, use variables before they are initialized and so on.

The third line of defense us the class loader. Since classes can be loaded on the fly, there is a danger, than an applet could load one of its own classes to replace a critical system class, thus by pacing that class security checks.

The fourth line of defense is that some standard classes have their own security measures built in.

Despite all these measures security problems all to be expected

(i) There can be bugs in Java software that clever programmers can exploits to bypass the security.

(ii) It may be possible to prevent an applet from doing anything excepting writing to the screen many applets will need more power. So when they ask for additional privileges users may grudgingly grant them. But even in the unlikely event that applets are allowed no network access at all, they may be able to transmit information using covert channels.

To acquire the send information the applet owner can establish a connection to client's machine to read some of its public web pages or FTP some of its public files. In short Java introduces many new possibilities and opportunities into www. It allows web pages to be interactive and to contain animation and sound.

Locating Information on the Web

Programs that search the Web are sometimes called Search Engines, Spiders, crawlers, worms or knowbots. For our algorithm we will need three data structures.

(i) Need a large linear array, Un-table that contains millions of entry, ultimately one pet-webpage.

(ii) Each entry contains two pointers one to the page's URL and one to page's type. The heap is our second data structure.

(iii) The third data structure is a hash table of size entries. Any URL can be run through a hash function to produce a non negative integer less than n (fig).

Building the index requires two phases searching and indexing. The heart of a search engine is a recursive procedure process on, which takes a URL string as input.

It hashes the URL to see, if it is already present in URL table. If it is not ready its pages are fetched. Finally process extracts all hyperlink.

In practice actual search engines first collect all the hyperlinks on each page they read, remove all the ones that have already been processed & save the rest.

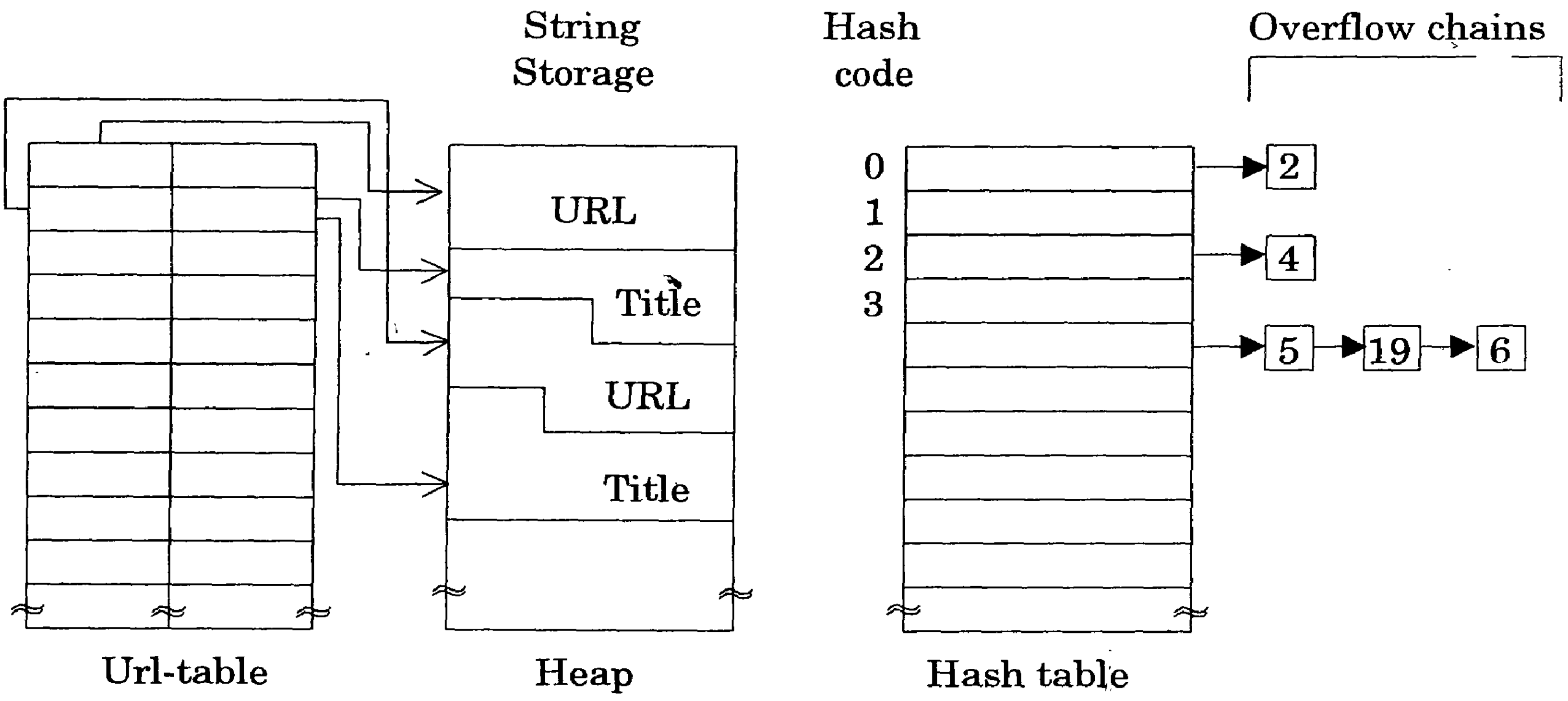


Figure 10.15 - Data structures used in a simple search engine

Indexing

Indexing procedure goes down Un-table linearly processing each entry in turn. For each entry it examines the title and select out all words. For each word selected it writes a line consisting of word followed by current un-table. The unindex will have to be stored on disk and can be used as follows. The user fills in a form listing one or more keywords and on SOBMIT button. This action causes a post request to be done on CGI scrip. The script now indexes into URL table to find all titles & URL.

The following problems have to be solved in practical system.

1. Some URL are obsolete
2. Some machines will be temporarily unreachable.
3. Not all pages may be reachable from starting URL
4. Some pages may be reachable from active maps only
5. Some documents cannot be indexed.
6. Not all documents have titles.
7. The search Engine could non out of memory.
8. The entire process might take too long.

One small request is in order. Although writing a search engine sounds easy, a buggy one wreak havoc with network by generating vast numbers of spurious requests, not only wasting bandwidth but bringing many servers to their knees due to load. If you cannot resist the temptation to write, your own search engine, proper etiquette requires restricting it to your own local DVS domain until it is totally debugged.

M.B.A. DEGREE EXAMINATION

COMPUTER NETWORKS

MODEL QUESTIONS

UNIT - I

1. Explain the uses of Computer Networks.

2. Explain the different Network Topology.

3. Explain about LAN, MAN and WAN.

4. Explain the differences between Connection oriented and Connectionless services.

5. Briefly explain about Network Software.

UNIT - II

1. Explain the Layers in OSI reference model.

2. Compare the OSI and TCP reference models.

3. Write a note on ARPANET.

4. Explain about Public Networks.

5. Explain the Computer Networks Standards.

UNIT - III

1. Explain about Bandwidth Limited Signals.

2. Compare and contrast twisted pair and fibre optics.

3. Explain the two types of Coaxial Cables.

4. Explain about Wireless Transmission.

5. Explain the structure of the Telephone Network.

UNIT - IV

1. Explain the Channel Allocation Problem.
2. Write a note on ALOHA.
3. Explain about Wireless LAN Protocols.
4. Explain the IEEE Standard 802 for LANs and MANs.
5. Explain about Fiber Optic Networks.

UNIT - V

1. Briefly discuss the Design Issues of Data Link Layer.
2. Explain the Methods for Error Detection.
3. Discuss about Error Correction.
4. Explain about character Stuffing and Bit Suffing.
5. Explain about the Elementary Data Link Protocols.

UNIT - VI

1. Discuss the Design Issues of the Network Layer.
2. Explain different Routing Algorithms.
3. Explain about Inter Networking.
4. What is Flooding?
5. Write a short note on Firewalls.

UNIT - VII

1. Explain the design issues of Transport Layer.
2. Explain about Connection Management.

3. What are the elements of Transport Protocols?

4. Explain the Internet Protocols.

5. Discuss the activities of X.25.

UNIT - VIII

1. Explain the benefits of Presentation Layer.

2. What is Encryption and Decryption?

3. Explain Sink Tree.

4. Explain Top Level Domains.

5. Write briefly about Cybher Text.

UNIT - IX

1. Explain about Network Security.

2. Explain the Components of SNMP Model.

3. Explain about DNS.

4. Write a note on Electronic Mail.

5. What is the function of USER AGENT?

UNIT - X

1. Explain about Data Compression.

2. What are the features of Multimedia?

3. Write a note on Video Servers.

4. Explain about World Wide Web.

5. Explain the use of Analog Systems in Video.

M.B.A. DEGREE EXAMINATION, APRIL 2002

Fourth Semester (Full Time)

Elective : Computer Networks

(For those who joined in July 1998 and after)

Time : Three hours Maximum : 60 marks

PART A (15 x 1 = 15 marks)
Answer ALL questions

1. Define OSI reference model.
2. What is USENET?
3. What is Ring topology?
4. Define Modem.
5. Define baseband.
6. What is LAN?
7. List different routing algorithms.
8. Define data link layer.
9. Define DTE.
10. What is SNICF?
11. What is T-connect?
12. What is remote procedure call?
13. What is EBCDIC code?
14. What is presentation layer?
15. Define fibre optics.

PART B (5 x 4 = 20 marks)
Answer any FOUR questions.

16. What are the applications of computer network?
17. Write short notes on metropolitan network.
18. Briefly discuss the design issues of data link layer.
19. Discuss the transport protocol on top of X 25.
20. Compare and contrast twisted pair cabel and fibre optics.
21. Explain polling technique.

PART C - (2 x 12 1/2 = 25 marks)
Answer any TWO questions.

22. Explain the structure of telephone network.
23. Explain different routing algorithms.
24. Discuss the design issues and connection management of transport layer.