COURSE GUIDE

CIT 423
COMPUTER NETWORKS AND COMMUNICATION

Course Team

NATIONAL OPEN UNIVERSITY OF NIGERIA

National Open University of Nigeria
Headquarters
University Village
Plot 91, Cadastral Zone
Nnamdi Azikiwe Expressway
Jabi, Abuja

Lagos Office
14/16 Ahmadu Bello Way
Victoria Island, Lagos

e-mail: centralinfo@nou.edu.ng
URL: www.nou.edu.ng

Printed 2022

ISBN:

CONTENT **PAGE**

Introduction

The aim and objective of computer networks as a course is computer literacy. Information Technology is the frontier hero of the new century, driven by ambition and full of courage, replicating itself like a virus and sweeping all before it. So in order not to be marginalized one needs to get involved as Information Technology influences virtually all the areas of human endeavour.

What You will Learn in this Course

This course will give you in brief what you need to know in Computer Science and Technology. At the end of the course you will be an expert of some sort in Computer Science and Technology.

Aim of the Course

Each module, each unit, in the course contains notes as well as set exercises. The set exercises are listed in form of a table. The table has two volumes. The last column contains the heading "what you do" while the right column is headed "comments/prompts". The left column contains the steps that you must follow. The right column serves as additional information.

Computers will be provided at study centres and will be equipped with the required programs. If you have the means, buy your PC and the stipulated software.

Each unit contains a *Tutor-Marked Assignment*, which must be done as stipulated and handed to the tutor on schedule.

Working through this Course

To complete this course you are required to read each study unit, read the textbooks and read other materials which may be provided by the National Open University of Nigeria.

Each unit contains self-assessment exercises and at certain points in the course you will be required to submit assignments for assessment purposes. At the end of the course there is final examination. Below you will find listed all the components of the course, what you have to do and how you should allocate your time to each unit in order to complete the course on time and successfully.

This course demands that you spend a lot of time to study. My advice is that you optimise the opportunity provided by the tutorial sessions where you have the opportunity of comparing your knowledge with that of your colleagues.

The Course Materials

The main components of the course are:

1. The Course Guide
2. Study Units
3. References/Further Readings
4. Assignments
5. Presentation Schedule

Study Unit

The study units in this course are as follows:

Module 1: Introduction to Computer Networks

Unit 1:	Network Classification and Reference Models
Unit 2:	Network Structures

Module 2: Network Devices and Technology

Unit 1	Network Technology
Unit 2	Network Devices–I
Unit 3	Network Devices–II
Unit 4	Integrated Service Digital Network (ISDN)
Unit5	Asynchronous Transfer Mode (ATM)
Unit 6	Data Transmission and Multiplexing
Unit 7	Medium Access Control and Data Link Layer

Module 3: Network Administration

Unit1	Network Administration: Scope, Goals, Philosophy and Standards
Unit2	Network Protocols
Unit3	Network, Transport and Application Layers

Note: each unit consists of one or two weeks work and includes introduction, objectives, reading materials, exercises, conclusion and summary, Tutor-Marked Assignment (TMAs), references and other resources. The unit directs you to work on these exercises related to required reading. In general, these exercises test you on the materials thereby assisting you to evaluate your progress and to reinforce your comprehension of the material. Together with the TMAs these exercises will help you in achieving the stated learning objectives of the individual units and of the course as a whole.

Presentation Schedule

Your course materials have important dates for early and timely completion and submission of your TMAs and attending tutorials. You should remember that you are required to submit all your assignments by the stipulated time and date. You should guide against falling behind in your work.

Assessment

There are three aspects to the assessment of the course. First is made up of self-assessment exercises, second consists of the TMA and third is the written examination/end of course examination.

You are advised to do the exercises. In tackling the assignments, you are expected to apply information, knowledge and techniques you gathered during the course. The assignments must be submitted to your facilitator for formal assessment in accordance with the deadlines stated in the presentation schedule and the assignment file. The work you submit to your tutor for assessment accounts for 30% of your total course work. At the end of the course you will need to sit for a final examination or end of course examination of about three hour duration. This examination will count for 70% of the total course mark.

Tutor-Marked Assignment

This is the continuous assessment component of your course. It accounts for 30% of the total score. You will be given four TMAs (4) to answer. Three of these must be answered before you are allowed to sit for the end of the course examination. The assignment questions for the units in the course are contained in the assignment file. You will be able to complete them through your reading the information contained in the reading materials, references and the study units. You are advised to research deeper into topics so as have a broader view of the discussions.

Endeavour to get the assignments to the facilitator on or before the deadline. If for any reason you cannot complete the work on time, contact your facilitator before the assignment is due to discuss the possibility of an extension. Extension will not be granted after the due date has passed unless on exceptional circumstances.

Final Examination and Grading

The end of course examination for Network Administration will be for about 3 hours and it has a value of 70% of the total course work. The examination will reflect the type of self-testing, practice exercise and tutor-marked assignment problems you have previously encountered. All these areas of the course will be assessed.

Use the time between finishing the last unit and sitting for the examination to revise the whole course. It might be useful to review your self tests, TMAs and the comments on them before the course examination. The end of course examination covers information from all parts of the course.

Course Marking Scheme

Facilitators/Tutors and Tutorial

There are 21 hours of tutorials provided in support of this course. You will be notified of the dates, times and venues of these tutorials as well as the name and phone numbers of the facilitator, as soon as you are allocated to a tutorial group.

Your facilitator will mark and comment on your assignments, keep a close watch on your progress and any difficulties you might face and provide assistance to you during the course. You are expected to mail TMA to your facilitator at least two working days before the schedule date. The TMAs will be marked by your tutor returned back to you as soon as possible.

Do not delay to contact your facilitator by telephone or email if you need

assistance. The following might lead to your needing your facilitator's

assistance:

You do not understand any part of the study or assigned reading
You have difficulty with the self test
You have a question or a problem with an assignment or with the grading of an assignment

Endeavour to attend tutorials. It affords you the opportunity of face to face contact with the facilitator and to ask questions which are answered instantly. You also raise problems encountered in the course of study.

SUMMARY

Computers in Society intend to make you computer literate. At the end of the course you will achieve the objective if you follow the instructions and do what you are expected to do.

We wish you a huge success.

CONTENTS **PAGE**

Module 1: Introduction to Computer Networks

Module 2: Network Devices and Technology

Module 3: Network Administration

MODULE 1 INTRODUCTION TO COMPUTER NETWORKS

UNIT 1: NETWORK CLASSIFICATION AND REFERENCE MODELS
UNIT 2: NETWORK STRUCTURE

UNIT 1: NETWORK CLASSIFICATION AND REFERENCE MODELS

1.0 INTRODUCTION

Earlier computers used to be stand alone. Different computers were used for information gathering, processing or distribution. Due to rapid technological progress, the areas of information gathering, processing and distribution are rapidly converging and differences between them are quickly disappearing. In this unit, we will learn about the different types of networks, their applications, networking models and topologies. We will also examine references, the various layers and functions of each layer.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- define and classify network
- distinguish between different types of networks
- understand what OSI model is, and TCP reference model and functions of each layer.

3.0 MAIN CONTENT

3.1 What Is A Network

In the simplest form, data transfer can take place between two devices which are directly connected by some form of communication medium. But it is not practical for two devices to be directly Point–to–Point connected. This is due to the following reasons:

(i) The devices are very far apart.
(ii) There is a set of devices, each of which may require to connect to others at various times.

Solution to this problem is to connect each device to a communication network. Computer network means interconnected set of autonomous systems that permit distributed processing of information.

In order to meet the needs of various applications, networks are available with different interconnection layouts and pLANs, method of access, protocols and media. Networks can be classified on the basis of geographical coverage.

3.1.1 Classification of Networks

- Local Area Network (LAN)
- Metropolitan Area Network (MAN)
- Wide Area Network (WAN)

3.1.2 Local Area Network (LAN)

A local area network is a relatively smaller and privately owned network with maximum span of 10km to provide local connectivity within a building or small geographical area. The LANs are distinguished from other kinds of networks by three characteristics:

(i) Size
(ii) Transmission technology, and
(iii) Topology

Accordingly, there are many LAN standards known as IEEE area standards 802 x.

3.1.3 Metropolitan Area Network (MAN)

Metropolitan Area Network is defined as less than 50km and provides regional connectivity typically within a campus or small geographical area. It is designed to extend over an entire city. It may be a single network, such as cable television network, or it may be a means of connecting a number of LANs into a large network, so that resources may be shared LAN–to–LAN as well as device to device. For example, a company can use a MAN to connect to the LANs in all of its offices throughout a city.

3.1.4 Wide Area Network (WAN)

Wide Area Network provides no limit of distance. In most WAN, the subnet consists of two distinct components. Transmission lines, also called circuits or channels, and routers. Transmission lines are used for moving bits between machines, whereas routers are used to connect two or more transmission lines

A WAN provides long distance transmission of data, voice, image and video information over large geographical areas that may comprise a country, a continent or even the whole world. In contrast to LANs (which depend on their own hardware for transmission), WANs may utilize public, leased or private communication devices usually in combination, and span own unlimited number of miles.

A WAN that is wholly owned by a single company is often referred to as an enterprise network.

3.2 Computer Network Goals/Motivation

The main goal of a computer network is to enable its users to share resources and to access these resources (i.e hard disks, high quality expensive laser printer, modems, peripheral devices, licensed software. etc.), regardless of their physical locations. Physical locations may be a few feet or even thousands of miles apart, but users exchange data and programs in the same way. In other words, distance is removed as a barrier for the above application. The computer network thus creates a global environment for its users and computers. Another goal is to provide communication services (such as E–mail) and in general, to provide robust transport network. i.e., (highway) over which application can be built.

3.3 Applications of Networks

The following is the list of some applications of computer network.

Generic application

- Resource sharing (CPU, peripherals, information and software)
- Personal communication (text+graphics+audio+video)
- Network wide information discovery and retrieval.

We are now moving from personalized computing to network computing. Therefore, its applications are increasing everyday.

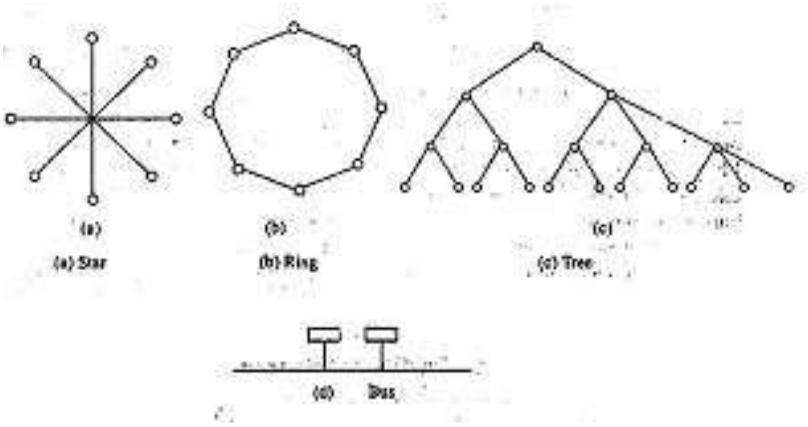
3.4 Types of Network

There are basically two types of network based on whether the network contains switching elements or not. These are Point–to–Point network and Broadcast network.

3.5.1 Point–to–Point Network or Switch Network

Point–to–Point networks consist of many connections between individual pairs of machines. To go from the to the source destination, a packet on this type of network may have to first visit one or more intermediate machine routers. When a packet is sent from one router to another intermediate router, the entire packet is stored at each intermediate router, till the output line is free and then forwarded. A subnet using this principle is called Point–to–Point or Packet switched network.

Some possible topologies for a Point–to–Point subnet are:



Star

In a star topology, each device has a dedicated Point–to–Point link only to a central controller, usually called a hub. These devices are not linked to each other. If one device wants to send data to another, it sends to the hub which then relays the data to the other connected devices. In a star, each device needs only one link and one I/O Port to connect it to any number of other devices. This factor makes it easy to install and configure. Far less cabling need to be housed and additions, moves and deletions involve only one connection between that device and the hub.

Tree

A tree topology is a variation of a star. As in a star modes in a tree are linked to a central hub that controls the traffic to the network. However, not every device plugs directly into the central hub. The majority of devices connect to a secondary hub that in turn is connected to the central hub.

The advantages and disadvantages of a tree topology are generally the same as those of stars. The addition of secondary hubs, however, brings two further advantages. First, it allows more devices to be attached to a single central hub and can, therefore, increase the distance a signal can travel between devices. Second, it isolates the network and prioritizes communication from different computers.

Ring

In a ring topology, each device has a dedicated Point–to–Point line configuration only, with the two devices on either side of it. A signal is passed along the ring in one direction from device to device, until it reaches its destination. Each device in the ring incorporates a repeater. When a device receives a signal intended for another device, its repeater regenerates the bits and passes them along.

A ring is relatively easy to install and reconfigure. Each device is linked to its immediate neighbours. However, unidirectional traffic can be a disadvantage. In a simple ring, a break in ring can disable the entire network. This weakness can be solved by using a dual ring or switch capable of closing off the break.

Bus

Bus, unlike other topologies, is a multi–point configuration. One long cable acts as a backbone to link all the devices in the network. Advantages of a bus topology include use of installation. A disadvantage includes difficult reconfiguration and fault isolation.

3.4.2 Broadcast Networks

Broadcast networks have a single communication channel that is shared by all the machines on the network. Short messages, called packets, sent by any machine are received by all the others. An address field within the packet specifies for when it is intended. Upon receiving a packet, a machine checks the address field. If the packet is intended for itself, it processes the packet; if the packet is intended for some other machine, it is just ignored.

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 very machine on the network, and this mode of operation is called broadcasting. Some broadcast systems also support transmission to a subset of the machines, something known as multicasting. One possible scheme is to reserve one bit multicasting. The remaining (n–1) address bits can hold a group number. Each machine can “subscribe” to any or all of the groups. When a packet is sent to a certain group, it is delivered to all machines subscribing to that group.

3.5 Reference Model

In this section, we will discuss two important network architectures: the OSI reference model and the TCP/IP reference model.

3.5.1 OSI (Open System Interconnection) Reference Model

The OSI model is based on a proposal developed by the International Standards Organisation as a first step towards international standardization of the protocols used in the various layers. The model is called the ISO – OSI (International Standard Organisation–Open Systems Interconnection) Reference Model because it deals with connecting open systems – that is, systems that are open for communication with other systems.

Its main objectives were to:

(i) Allow the manufacture of different systems to interconnect equipment through standard interfaces.
(ii) Allow software and hardware to integrate well and be portable on different systems.

The OSI model has seven layers shown in figure 2. The principles that were applied to arrive at the seven layers are as follows:

1. Each layer should perform a well–defined function.
2. The function of each layer should be chosen with an eye toward defining internationally standardized protocols.
3. The layer boundaries should be chosen minimize the information flow across the interfaces.

The seven layers of ISO OSI Reference Model are: (a)

Physical Layer
(b) Data Link Layer
(c) Network Layer
(d) Transport Layer
(e) Session Layer
(f) Presentation Layer
(g) Application Layer.

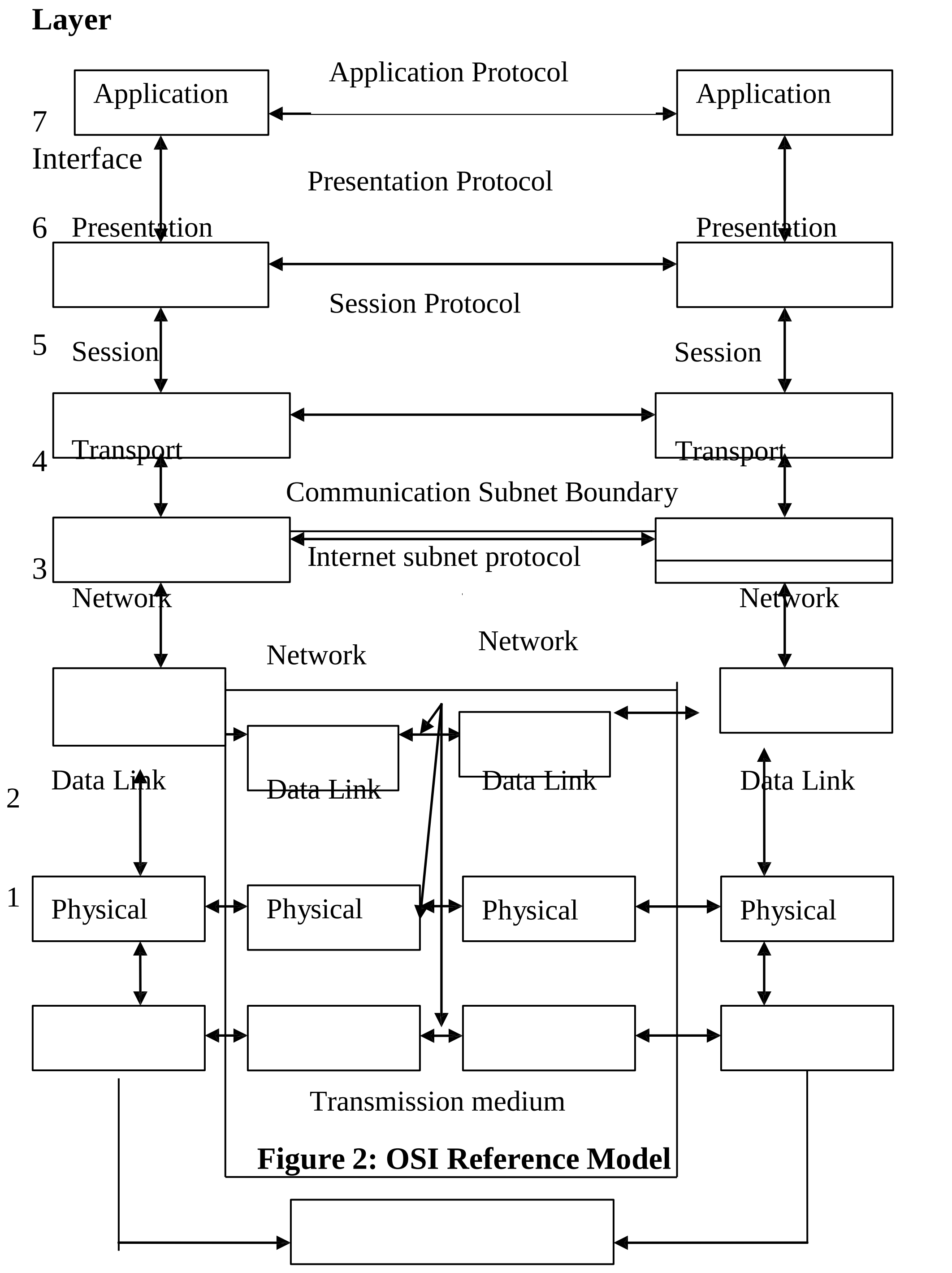


Figure 2: OSI Reference Model

3.5.1.1 The Physical Layer

Physical Layer defines electrical and mechanical specifications of cables, connectors and signaling options that physically link two nodes on a network.

3.5.1.2 The Data Link Layer

The main task of the Data Link Layer is to provide error free transmission. It accomplishes this task by having the sender break the input data up into data frames, transmit the frames sequentially, and process the acknowledgement frames sent back to the receiver.

The Data Link Layer creates and recognises frame boundaries. This can be accomplished by attaching special bit patterns to the beginning and end of the frame. If these bit patterns can accidentally occur in the data, special care must be taken to make sure these patterns are not incorrectly interpreted as frame delimiters

3.5.1.3 The Network Layer

Whereas the Data Link Layer is responsible for end to end delivery, the network layer ensures that each packet travels from its source to destination successfully and efficiently. A key design issue is determining how packets are routed from source to destination. Routes can be based on static tables that are “wired into” the network and rarely changed.

They can also be determined at the start of each conversation, for example, a terminal session. Finally, they can be highly dynamic, being determined anew for each packet, to reflect the current network load.

When a packet has to travel from one network to another to get to its destination, many problems can arise. The addressing used by the second network may be different from the first one. The second one may not accept the packet at all because it is too large. The protocols may differ, and so on. It is up to the network layer to overcome all these problems to allow heterogeneous networks to be interconnected.

3.5.1.4 The Transport Layer

The basic function of the transport layer is to accept data from the session layer, split it up into smaller units if need be, pass these to the Network Layer, and ensure the pieces all arrive correctly at the other end. Furthermore, all this must be done efficiently, and in a way that isolates the upper layers from the inevitable changes in the hardware technology.

The transport layer provides location and media independent data transfer service to session and upper layers.

3.5.1.5 The Session Layer

The main tasks of the session layer are to provide:

- Session establishment
- Session Release– Orderly or Abort
- Data Exchange
- Expedited Data Exchange.

The session layer allows users on different machines to establish sessions between them. A session allows ordinary data transport, as does the transport layer, but it also provides enhanced services useful in some applications. A session might be used to allow a user to log into a remote time sharing system or to transfer a file between two machines.

One of the services of the session layer is to manage dialogue control. Sessions can allow traffic to go in both directions at the same time, or in only one direction at a time. If traffic can only go one way at a time (analogous to a single railroad track), the session layer can help keep track of whose turn it is.

A related session service is token management. For some protocols, it is essential that both sides do not attempt the same operation at the same time. To manage these activities, the session layer provides tokens that can be exchanged. Only the side holding the token may perform the critical operation.

Another session service is synchronization. Consider the problem that might occur when trying to do a two–hour file transfer between two machines with a one hour mean time between crashes. After each transfer is aborted, the whole transfer would have to start over again and would probably fail again the next time as well. To eliminate this problem, the session layer provides a way to insert after the last checkpoint has to be repeated.

3.5.1.6 The Presentation Layer

Unlike all the lower layers which are just interested in moving bits reliably from here to there, the presentation layer is concerned with the syntax and semantics of the information transmitted.

A typical example of a presentation service is encoding data in a standard agreed upon way. Most user programs do not exchange random binary bit strings, they exchange things such as people’s names, dates, amounts of money and invoices. These items are represented as character strings, integers, floating–point number, and data structures composed of several simpler items. Different computers have different codes for representing character strings, (e.g., ASCII and Unicode), integers (e.g., one's complement and two's complement), and so on. In order to make it possible for computers with different representations to communicate, the data structure to be exchanged can be defined in an abstract way, along with a standard encoding to be used "on the wire". The presentation layer manages these abstract data structure and converts from the representation used inside the computer to the network standard representation and back.

3.5.1.7 Application Layer

Application layer supports functions that control and supervise OSI application processes such as start/maintain/stop application, allocate/de–allocate OSI resources, accounting, check point and recovering. It also supports remote job execution, file transfer protocol, message transfer and virtual terminal.

3.5.2 TCP Reference Model

The TCP/IP network architecture is a set of protocols that allow communication across multiple device networks. The architecture evolved out of research that had the original objective of transferring packets across three different packet networks: the **ARPANET** packet– switching networks, a packet radio network, and a packet satellite network. The military orientation of the research placed a premium on robustness with regards to failures in the network and on flexibility in operating over diverse networks. The environment led to a set of protocols that are highly effective in enabling communication among the many different types of computer systems and networks. Today, the internet has become the primary fabric for interconnecting the world's computers. In this section, we introduce the TCP/IP network architecture and TCP/IP is the main protocol for carrying information.

Figure 3 shows the TCP/IP network architecture, which consists of four layers. The Application Layer provides services that can be used by other applications. For example, protocols have been developed for remote login, for e–mail, for file transfer, and for network management.

The Application Layer programs are intended to run directly over the transport layer. Two basic types of services are offered in the transport layer. The first service consists of reliable connection–oriented transfer of a byte stream, which is provided by the **Transmission Control Protocol (TCP).** The second service consists of best–effort connectionless transfer of individual messages, which is provided by the **User Datagram Protocol** (UDP). This service provides no mechanisms for error recovery or flow control. UDP is used for applications that require quick but necessary or flow control. UDP is used for application that require but necessarily reliable delivery layer.

Application Layer
Transport Layer
Internet Layer
Network Interface Layer

Figure 3: TCP/IP Network Architecture

The TCP/IP model does not require strict layering. In other words, the application layer has the option or bypassing intermediate layers. For example, an application layer may run directly over the internet.

The **Internet Layer** handles the transfer of information across multiple networks through the use of gateways of routers, as shown in figure 4. The Internet Layer corresponds to the part of the OSI network layer that is concerned with the transfer of packets between machines that are connected to different networks. It must, therefore, deal with the routing of packets across these networks as well as with the control of congestion. A key aspect of the internet layer is the definition of globally unique addresses for machines that are attached to the Internet. The internet layer provides a single service, namely: best–effort connectionless packet transfer. IP packets are exchanged between routers without a connection set up; the packets are routed independently, and so they may traverse different paths. For this reason, IP packets also called **datagrams.** The connectionless approach makes the system robust; that is, if failures occur in the network, the packets are routed around the points of failure; there is no need to set up the connections. The gateways that interconnect the intermediate networks may discard packets when congestion occurs. The responsibility for recovery from these losses is passed on to the transport layer.

Finally, the Network Interface layer is concerned with the network–specific aspects of the transfer of packets. As such, it must deal with the part of the OSI network layer and data link layer. Various interfaces are available for connecting end computer systems to specific networks such as X.25, ATM, frame relay, Ethernet, and token ring.

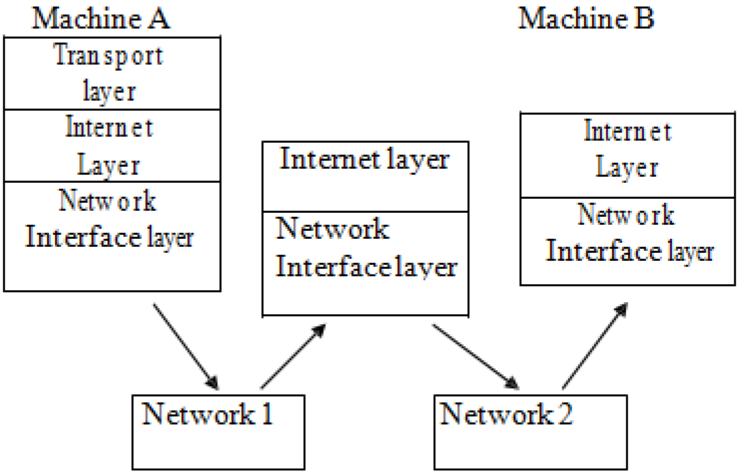


Figure 4: The Internet Layer and Network Interface Layers

The network interface layer is particularly concerned with the protocols that access the intermediate networks. At each gateway, the network access protocol encapsulates the IP packet into a packet or frame of the underlying network or link. The IP packet is recovered at the exit gateway of the given network. This gateway must then encapsulate the IP packet into a packet or frame of the type of the next network or link.

This approach provides a clear separation of the internet layer from the technology dependent network interface layer. This approach also allows the internet layer to provide a data transfer service that is transparent sense of not depending on the details of the underlying networks. The next section provides a detailed example of how IP operates over the underlying networks.

Figure 5 shows some of the protocols of the TCP/IP protocol suite. The figure shows two of the many protocols that operate over TCP, namely, HTTP and SMTP. The figure also shows DNS and Real time Protocol (RTP), which operate over UDP. The transport layer protocols TCP and UDP, on the other hand, operate over IP. Many network interfaces are defined to support IP. The salient part of figure 5 is that all higher–layer protocols access the network interfaces through IP. This feature provides the capability to operate over multiple networks. The IP protocol is complemented by additional protocols (ICMP, IGMP, ARP, and RARP) that are required to operate an internet.

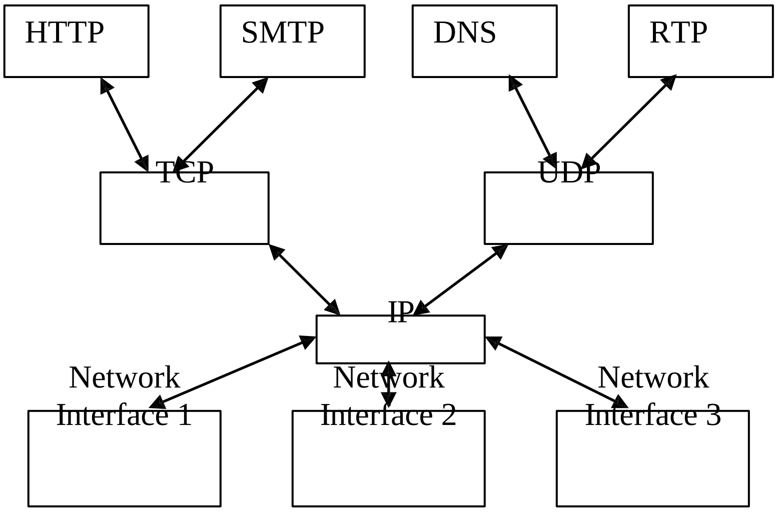


Figure 5: TCP/IP Protocol Graph

The hourglass shape of the TCP/P protocol graph underscores the features that make TCP/IP so powerful. The operation of the single IP protocol over various networks provides independence from the underlying network technologies. The communication services of TCP and UDP provide a network independent platform on which applications can be developed. By allowing multiple network technologies to coexist, the internet is able to provide ubiquitous connectivity and to achieve enormous economies of scale.

3.5.3 Difference between OSI Reference Model & TCP Reference Model

OSI Reference Model	TCP Reference Model
1. Seven layers 2. It distinguishes between service, interface and protocol. 3. First comes description of model and protocol comes next 4. Both have Network 5. supports connectionless and connection oriented communication in network layer and only connection–oriented communication in transport layer (Co$_2$ T. service is visual to the User) 6. Protocol in OSI model are better hidden and can be replaced relatively easily (No Transparency)	1. 4 layers 2. Does not clearly distinguish between service, interface and protocol 3. protocol comes first and description of model later. 4. Transport and Application layer. 5. TCP/IP has only one mode in Network layer (connectionless) but supports both modes in Transport layer. 6. Protocols in TCP/IP are not hidden and thus, cannot be easily replaced. (Transparency)

3.6 IEEE Standards for LAN

Although there are many standards, we will configure here to just three of them:

- IEEE Standard 802.3 and Ethernet
- IEEE Standard 804 Token Bus
- IEEE Standard 802.5 Token Ring

IEEE Standard 802.3 and Ethernet

1. 802.3 is a simple protocol, Station can be installed on fly without taking network down. A passive cable is used and modems are not required. Delay at low load is practically zero. A station does not have to wait for a token, they just transmit immediately. Each station has to be able to detect the signal of the weakest station even when it is transmitting itself and all of the collision detect circuiting in the transceiver is analog. Minimum valid frame is 64 bytes.

2. 802.4 Bus – It uses highly reliable cable envision equipment which is available from numerous vendors. It is more deterministic than 802.3 although repeated losses of the token at critical moments can introduce more uncertainty than its supporters like to admit–Token Bus also supports priorities.

3 Token Ring–Point–to–Point connection means that the engineering is easy and can be fully digital. Ring can be built virtually in a transmission medium from carrier pigeon to fibre optics. The standard twisted pair is cheap and simple to install like the Token bus in token ring priorities are possible.

4.0 CONCLUSION

This unit has introduced you to Computer Networks. We have classified the different types of networks, goal and motivation of Computer Networks. This unit has introduced you to the two types of network models as well as the difference between these two. The unit has also done a good job of defining various standards of LANs.

5.0 SUMMARY

A communication system that supports many udders is called a network. In a network, many computers are connected to each other by various topologies like star, ring, complete, interconnected or irregular.

Depending on the area of coverage, a network can be classified as LAN, MAN, or WAN. A network is required for better utilisation of expensive resources, sharing information, collaboration among different groups, multimedia communication and video conferencing.

Two different types of networking models OSI and TCP/IP exist. The difference between these models was discussed in detail.

6.0 TUTOR–MARKED ASSIGNMENT

i. What are the various types of networks?

ii. What is the difference between broadcasting and multicasting?

7.0 REFERENCES/FURTHER READING

UNIT 2: NETWORK STRUCTURE

1.0 INTRODUCTION

This unit provides a survey of the basic network structures or topologies. Topology can be considered as a virtual shape or structure of a network.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- name the four basic network topologies
- cite advantages and disadvantages of each type
- state the criteria necessary for an effective and efficient network.

3.0 MAIN CONTENT

3.1 PHYSICAL TOPOLOGY

The term physical topology refers to the way in which a network is laid out physically. Two or more devices connect to a link; two or more links form a topology. The topology of a network is the geometric representation of the relationship of all the links and linking devices (usually called nodes) to one another. There are four basic topologies possible: mesh, star, bus and ring (see figure 1).

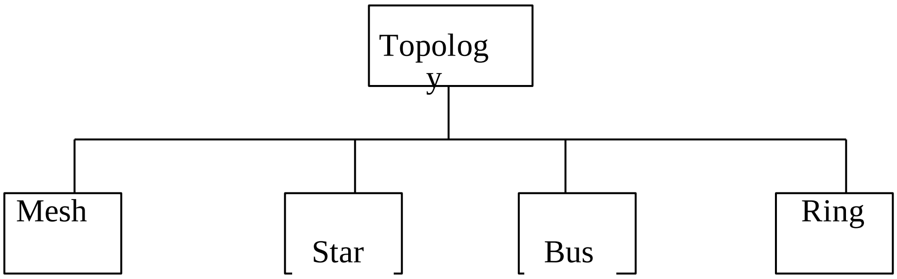


Figure 1 Categories of topology

3.1.1 Mesh

In a mesh topology, every device has a dedicated point to point link to every other device. The term dedicated means that the link carries traffic only between the two devices it connects. (see figure 2)

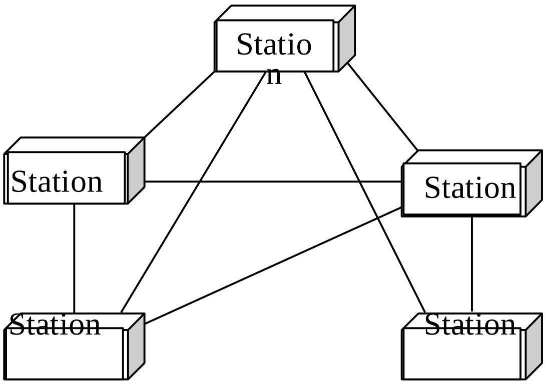


Figure 2 A fully connected mesh topology (five devices)

A mesh offers several advantages over other network topologies. First, the use of dedicated links guarantees that each connection can carry its own data load, thus eliminating the traffic problems that can occur when links must be shared by multiple devices. Second, a mesh topology is robust. If one link becomes unavailable it does not incapacitate the entire system. Third, there is the advantage of privacy or security. Whenever message travels along a dedicated line, only the intended recipient sees it. Finally, point to point links make fault identification and fault isolation easy. Traffic can be routed to avoid links with suspected problems. This enables the network manger to discover the precise location of the fault and aids in finding its cause and solution.

The main disadvantages of a mesh are related to the amount of cabling and that of I/O ports required. First, because every device must be connected to every other device, installation and reconnection are difficult. Second, the sheer bulk of the wiring can be greater than the available space (in walls, ceilings or floors) can accommodate. Finally, the hardware required to connect each link (I/O ports and cables) can be prohibitively expensive. For these reasons, a mesh topology is usually implemented in a limited fashion, for example, as a backbone connecting the main computers of a hybrid network that can include several other topologies.

One practical example of a mesh topology is the connection of a telephone regional office in which each regional office needs to be connected to every other regional office.

3.1.2 Star Topology

In a star topology, each device has a dedicated point-to-point link only to a central controller, usually called a hub. These devices are not directly linked to one another. Unlike a mesh topology, a star topology does not allow direct traffic between devices. The controller acts as an exchange. If one device wants to send data to another it sends the data to the controller, which then relay the data to the other connected devices (see figure 3)

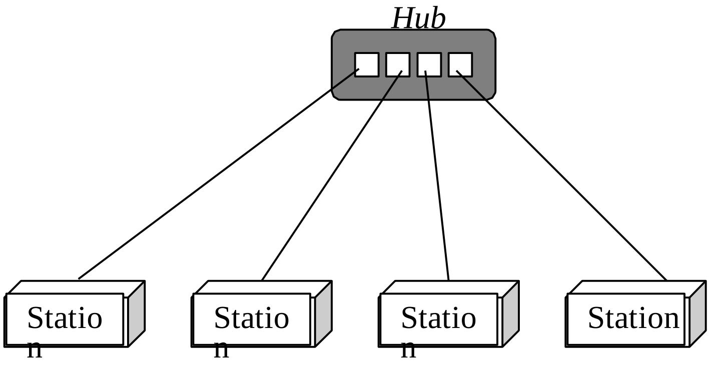


Figure 3 A star topology connecting four station

A Star topology is less expensive than a mesh topology. In a star, each device needs only one link and one I/O port to connect to any number of others.

This factor also makes it easy to install and reconfigure. Far less calling needs to be housed, and additions, moves and deletions involve only one connection between that device and the hub.

Other advantages include robustness. If one link fails, only that link is affected. All other remain active. This factor also lends itself to easy fault identification and fault isolation. As long as the hub is working, it can be used to monitor link problems and bypass detective links.

One big disadvantage of a star topology is the dependency of the whole topology on one single point, the hub. If the hub goes down, the whole system is dead.

Although a star requires far less cable than a mesh, each node must be linked to a central hub. For this reason, often more cabling is registered in a star then in some other topologies (such as ring or bus)

3.1.3 Bus Topology

A bus topology is multipoint. One long cable acts as a backbone to link all the devices in the network (see figure 4)

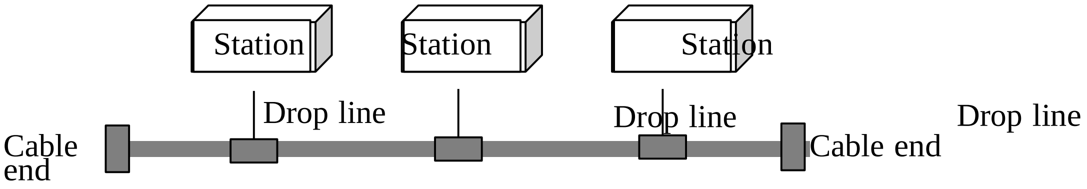


Figure 4 A bus topology connecting three stations

Nodes are connected to the bus cable by drop lines and taps. A drop line is a connection running between the device and the main cable. A tap is a connector that either splices into the main cable or punctures the sheathing of a cable to create a contact with the metallic core. As a signal travels along the backbone, some of its energy is transformed into heat. Therefore, it becomes weaker and weaker as it travels farther and farther. For this reason, there is a limit on the number of taps a bus can support and on the distance between those taps.

Advantages of a bus topology include ease of installation. Backbone cable can be laid along the most efficient path and then connected to the nodes by drop lines of various lengths. In this way, a bus uses less cabling than mesh or star topologies.

Disadvantages include difficult reconnection and fault isolation. A bus is usually designed to be optimally efficient at installation. It can therefore be difficult to add new devices. Signal reflection at the taps can cause degradation of quality.

In addition, a fault or break in the bus cable stops all transmission, even between devices on the same side of the problem. The damaged area reflects signals back in the direction of origin, creating noise in both directions.

Bus topology was the one of the first topologies used in the design of early local area networks

3.4 Ring Topology

In a ring topology, each device has a dedicated point to point connection with only the two devices on either side of it. A signal is passed along the ring in one direction from device to device, until it reaches its destination. Each device in the ring incorporates a repeater. When a device receives a signal intended for another device, its repeater regenerates the bits and passes them along (see figure 5).

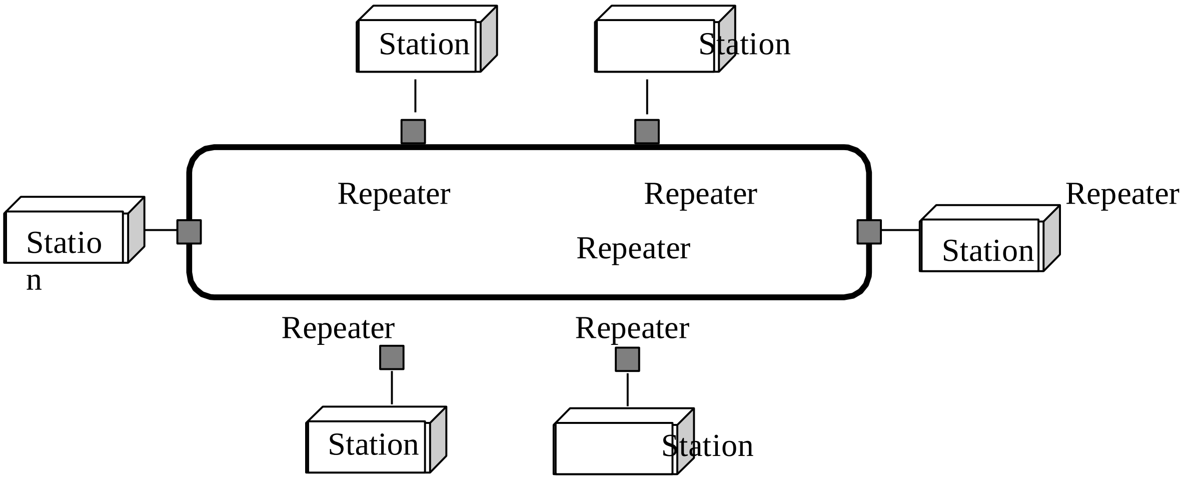


Figure 5 A ring topology connecting six stations

A ring is relatively easy to install and reconfigure. Each device is linked to only its immediate neighbors (either physically or logically). To add or delete a device requires changing only two connections. The only constraints are media and traffic considerations (maximum ring length and number of devices). In addition, fault isolation is simplified. Generally in a ring, a signal is circulating at all times. If one device does not receive a signal within a specified period, it can issue an alarm. The alarm alerts the network operator to the problem and its location.

However, unidirectional traffic can be a disadvantage. In a simple ring, a break in the ring (such as a disabled station) can disable the entire network. This weakness can be solved by using a dual ring or a switch capable of closing off the break.

Today, the need for higher speed LANS has made this topology less popular

3.5 Hybrid Topology

A network can be hybrid. For example, we can have a main star topology with each branch connecting several stations in a bus topology as shown in Figure 6.

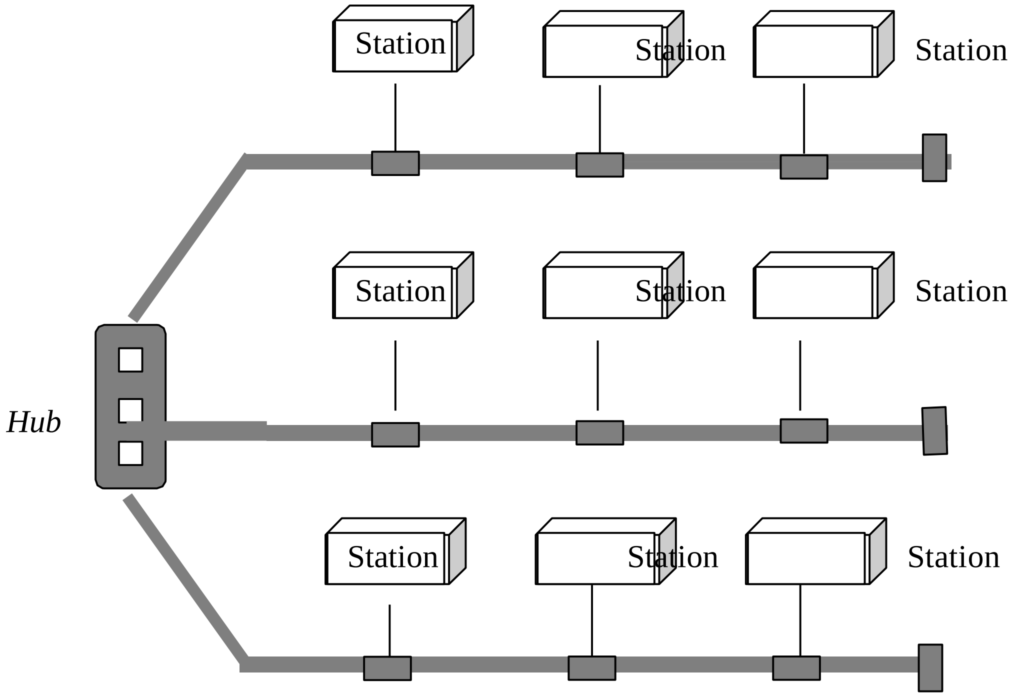


Figure 6 A hybrid topology: a star backbone with three bus networks

SELF-ASSESSMENT EXERCISES

a) What are the three criteria necessary for an effective and efficient network?

b) What is network topology?

4.0 CONCLUSION

In the context of a communication network, the term topology refers to the very in which the end points, or stations attached to the network are interconnected. Topologies are the important part of the network design theory. A better network can be built if you have the knowledge of these topologies and if you know the difference between each topology.

5.0 SUMMARY

Topology refers to the physical or logical arrangement of a network. Devices may be arranged in a mesh, star, bus or ring topology.

A mesh offers several advantages over other network topologies. The main disadvantages of a mesh are the number of I/O ports required.

Star topology is less expensive than a mesh topology. One big disadvantage of a star topology is the dependency of the whole topology on one single point, the hub.

A bus topology is multipoint unlike mesh and star topologies that are point to point connections. An advantage of a bus topology is ease of installation. Disadvantages include difficult reconnection and fault isolation.

Ring topology is relatively easy to install and reconfigure. However, unidirectional traffic can be a disadvantage.

Hybrid topology is complex which can be built of two or more above networked topologies.

6.0 TUTOR-MARKED ASSIGNMENT

1. For each of the following four networks, discuss the consequences if a connection fails.

a) Five devices arranged in a bus topology

b) Five devices arranged in a ring topology

2. For n devices in a network, what is the number of cable links required for a mesh, ring, bus and star topology?

3. Name the four basic networking topologies and cite on advantage of each type.

7.0 REFERENCES/FURTHER READING

1. Burgess, M. (2004). Principles of Network and System Administration. (2nd Ed.). Chichester, West Sussex , England: Wiley.

2. Forouzan, B.A, & Fegan, S.C. (2007). Data communications and Networking (4th Ed). Mc Graw Hill.

3. Limoncelli, T. A.,Hogan, C. J. & Chalup, S. R (2007}. The Practice of System and Network Administration. (2nd Ed.). Upper Saddle River, NJ: Addison-Wesley

4. Stallings, W. (2009). Data and computer communications (8th ed.). Upper saddle River, NJ.: Pearson Education Inc.

5. Subramanian, M. (2000). Network Management: Principles and Practice, Addison-Wesley

Module 2: Network Devices & Technology

UNIT 1: NETWORK TECHNOLOGY

1.0 INTRODUCTION

This unit looks at what constitutes a local area network (LAN), then a wide area network (WAN) and then discusses the differences between the two. We then discuss the technologies for implementing WAN.

2.0 OBJECTIVES

- explain the categories of networks
- state the distinctions between LAN and WAN
- explain the technologies used in implementing WAN.

3.0 MAIN CONTENT

3.1 Categories of Networks/Network Technologies

Today when we speak of networks, we are generally referring to two primary categories: local area networks (LANs) and wide area networks (WANs).

The category into which a network falls is determined by its size. A LAN normally covers an area less than 2 miles; a WAN can be worldwide. Networks of a size in between are normally referred to as metropolitan area networks (MANs) and span tens of miles.

3.1.1 LOCAL AREA NETWORK

One type of network that becomes ubiquitous is the local area network. Indeed, LAN is to be found in virtually all medium and large size office buildings. Depending on the needs of an organization and the type of technology used, a LAN can be as simple as two PCs and a printer in someone's home office; or it can extend throughout a company and include audio and video peripherals. Currently LAN size is limited to a few kilometers. LANs are designed to allow resources to be shared between personal computers or workstations. The resources to be shared can include hardware (e.g. a printer), software (e.g. an application program) or data. In addition to size, LANs are distinguished from other types of networks by transmission media

and topology. In general, a given LAN will use only one type of transmission medium. The most common LAN topologies are bus, ringed star.

Early LANS had data rates in the 4 to 16 megabits per seconds (mbps) ranges. LANs come in a parallel of different configurations. The most common is switched LANs and wireless LANs. The most switched LAN is a switched Ethernet LAN, which may consist of a single switch with a parallel of attached devices, or parallel of interconnected switches. Today, however, speeds are normally 100 or 1000 mbps. Wireless LANs are the newest evolution in LAN technology.

3.1.2 WIDE AREA NETWORK (WAN)

A wide area network (WAN) provides long distance transmission of data, image, audio, video information over large geographic area that may comprise a country, a continent or even the whole world. WAN can be as complex as the backbones that connect the Internet or as simple as a dial-up line that a home computer to the Internet. We normally refer to the first as a switched WAN and to the second as a point to point WAN. The switched WAN connects the end systems which usually comprise a router (internet – working connecting devices) that connects together LAN or WAN. The point to point WAN is normally a line leased from a telephone or cable T.V provider that connects a home computer or a small LAN to an Internet service provider (ISP). This type of WAN is often used to provide Internet access. Wireless WANs are become more and more popular. Traditionally, WANs have been implemented using one of two technologies: Circuit switching and packet switching. More recently, frame relay and asynchronous transfer mode (ATM) networks have assumed major roles.

CIRCUIT SWITCHING

In a circuit- switching network, a dedicated communications path is established between two stations through the nodes of the network. That path is a connected sequence of physical links between nodes. On each link, a logical channel is dedicated to the connection. Data generated by the source station are transmitted along the dedicated path as rapidly as possible. At each mode, incoming data are routed or switched to the appropriate outgoing channel without delay. The most common example of circuit switching is the telephone network.

PACKET SWITCHING

A quite different approach is used in a packet switching network. In this case, it is not necessary to dedicate transmission capacity along a path through the network. Rather, data are sent out in a sequence of small chunks, called packets. Each packet is passed through the network from node to node along some path leading from source to destination. At each node, the entire packet is received, stored briefly, and then transmitted to the next node. Packet switching networks are commonly used for terminal to computer communications.

FRAME RELAY

Packet switching was developed at a time when digital long distance transmission facilities exhibited a relatively high error rate compared to today's facilities. As a result, there is a considerable amount of overhead built into packet switching schemes to compensate for errors. The overhead includes additional bits added to each packet to introduce redundancy and additional processing at the end stations and the intermediate switching nodes to detect and recover from errors.

With modern high-speed communication systems, this overhead is unnecessary and counterproductive. It is unnecessary because the rate of errors has been dramatically lowered and any remaining errors can easily be caught in the end systems by logic that operates above the level of the packet-switching logic. It is counterproductive because the overhead involved soaks up a significant fraction of the high capacity provided by the network.

Frame relay was developed to take advantage of these high data rates and low error rates whereas the original packet-switching networks were designed with a data rate to the end user of about 64 kbps. Frame relay networks are designed to operate efficiently at user data rate of up to 2mbps. The key to achieving these high data rates is to strip out most of the overhead involved with errors control.

ASYNCHRONOUS TRANSFER MODE (ATM)

Sometimes referred to as cell relay is a culmination of developments in circuit switching and packet switching. ATM can be viewed as an evolution from frame relay. The most obvious difference between frame relay and ATM is that frame relay uses variable length packets called frames and ATM uses fixed length packets, called cells. As with frame relay, ATM provides little overhead for error control depending on the inherent reliability of the transmission system and on higher layers of logic in the end of systems to catch and correct errors. By wiring a fixed packet length, the processing overhead is reduced even further for ATM compared to frame relay. The result is that ATM is designed to work in the range of 10s and 100s of mbps and in the Gbps range. ATM can also be viewed as an evolution from circuit switching; only fixed-data- rate circuits are available to the end system. ATM allows the definition of multiple virtual channels with date rate that are dynamically defined at the time the virtual channel is created. By using small, fixed-size cells, ATM is so efficient that it can offer a contant-data rate channel even though it is using a packet-switching technique. Thus ATM extends circuit switching to allow multiple channels with the data rate on each channel dynamically set on demand.

3.1.3 Distinctions between LANs and WANs

There are several key distinctions between LANs and WANs. Among which are:

1. The scope of the LAN is small, typically a single building or a cluster of buildings. This difference in geographic scope leads to different technical solution.

2. It is usually the case that the LAN is owned by the same organization that owns the attached devices. For WANs, this is less often the case, or at least a significant fraction of the network assets is not owned. This has two implications. First, care must be taken in the choice of LAN, because there may be a substantial capital investment (compared to dial-up or leased charges of WANs) for both purchase and maintenance. Second, the network management responsibility for a LAN falls solely on the user.

3. The internal data rates of LANs are typically much greater than those of WANs.

3.1.4 Metropolitan Area Network (MAN)

A MAN is a network with a size between a LAN and WAN. It normally covers the area inside a town or a city. It is designed for customers who need a high-speed connectivity, normally to the Internet and have end points spread over a city or part of a city.

3.1.5 Interconnection of Networks: Internetwork

Today, it is very rare to see a LAN, a MAN or a WAN in isolation; they are connected to one another. When two or more networks are connected, they become an internetwork or internet.

SELF ASSESSMENT EXERCISES

What is an internet?

What is the Internet?

4.0 CONCLUSION

Whereas wide area networks may be public or private, LANs usually are owned by the organization that is using the network to interconnect equipment. LANs have much greater capacity than WANS to carry what is generally a greater internal communication load.

5.0 SUMMARY

In general terms, communications networks can be categorized as local area networks (LANs) and wide area networks (WANs).

A LAN consists of a shared transmission medium and a set of hardware and software for interfacing devices to the medium and regulating the orderly access of the medium. LAN size is limited. In addition to size, LANs are distinguished from other types of networks by their transmission media and topology. A WAN provides long-distance transmission over large geographic areas. WAN is often used to provide Internet access. Traditionally, WANs have been implemented using one of two technologies: circuit switching and packet switching. Wireless WANs are becoming more and more popular.

6.0 TUTOR-MARKED ASSIGNMENTS

1. What are the advantages of packed switching compared to circuit switching?
2. What are some of the factors that determine whether a communication system is LAN or WAN?
3. Outline the distinctions between LAN and WAN.
4. Discuss circuit-switching network.

7.0 REFERENCES/FURTHER READING

1. Burgess, M. (2004). Principles of Network and System Administration. (2nd Ed.). Chichester, West Sussex, England: Wiley.

2. Forouzan, B.A, & Fegan, S.C. (2007). Data communications and Networking (4th Ed). McGraw Hill.

3. Limoncelli, T. A.,Hogan, C. J. & Chalup, S. R (2007}. The Practice of System and Network Administration. (2nd Ed.). Upper Saddle River, NJ: Addison-Wesley

4. Stallings, W. (2009). Data and computer communications (8th ed.). Upper saddle River, NJ.: Pearson Education Inc.

UNIT 2: NETWORK DEVICES–I

1.0 INTRODUCTION

As corporations grow, network designers need to extend the area of a network, the number of users on a particular network, and the bandwidth available to the network users. To solve these problems, network designers break a network into smaller portions and connect them with networking devices such as bridges, switches and gateways etc. Depending on the complexities of each of the networks being connected, a choice is made between these different network devices.

In this unit, and the next unit, we will examine features of several network devices.

2.0 OBJECTIVES

By the end of this unit, you will be able to explain:

- repeaters
- bridges
- witches
- hubs.

3.0 MAIN CONTENT

3.1 Network Devices

Most common features of network devices are to interconnect networks, boost signals etc. The basic difference between them is that they operate at different layers. Now let us examine each device separately.

3.1.1 Repeaters

When a signal is sent over a long network cable, signal gets weakened due to attenuation. This results in some data getting lost in the way. In order to boost the data signal, Repeaters are needed to amplify the weakened signal. They are known as signal boosters or amplifiers. They are physical layer devices. They are like small boxes that connect two segments of networks, refine and regenerate the digital signals on the cable and send them on their way.

Repeaters help in increasing the geographical coverage of networks i.e. LAN for example, IEEE802.3 Standard allows for up to four repeaters connecting five cable segments to a maximum of 3000 metres distance.

Repeaters use different physical media as:

This Ethernet cable and fibre optic cable: Token ring networks translate between electrical signals on shielded or unshielded twisted pair wiring and light pulse on fibre–optic cabling.

In modern installations, repeaters are housed in the central wiring hubs of 10 Base–T and fibre optic cable systems.

Repeaters send every bit of data appearing on either cable segment through to the other side, even if the data consist of malformed packets from a malfunctioning Ethernet adapter or packets not destined for use of the local LAN segment.

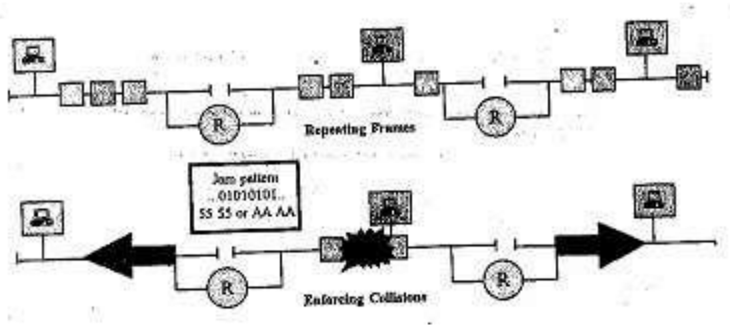


Figure 1 : Repeater Action

Bridges

Segmenting a large network with a device has numerous benefits. Among these are reduced collisions (in an Ethernet network), contained bandwidth utilization, and the ability to filter out unwanted packets. However, if the addition of the interconnect device required extensive reconfiguration of stations, the benefits of the device would be outweighed by the administrative overhead required to keep the network running. Bridges were created to allow network administrators to segment their networks transparently. This means that individual stations need not know whether there is a bridge separating them or not. It is up to the bridge to make sure that packets get properly forwarded to their destinations. This is the fundamental principle underlying all of the bridging behaviours.

Bridges work at the Data Link Layer of the OSI model. Since bridges work in the Data Link Layer they do not examine the network layer addresses. They just look at the MAC addresses for Ethernet and Token Ring, Token Bus and determine whether or not to forward or ignore a packet.

Purpose of a Bridge

The purposes of a Bridge are as followings:

1. Isolates networks by MAC addresses
2. Manages network traffic by filtering packets
3. Translates from one protocol to another.

Now let us examine each functionality of a bridge in detail.

1. Isolates networks by MAC addresses

A bridge divides a network into separate collision domains (Fig. 2). This reduces congestion as only frames that need to be forwarded are sent across interfaces. All transmissions between nodes connected to same segment are not forwarded and therefore, do not load the rest of the network.

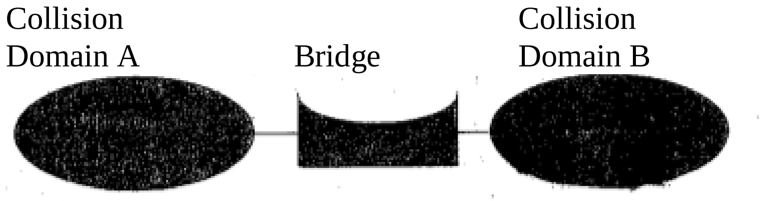


Figure 2: Bridge

Thus, bridges effectively improve the bandwidth of the network by reducing the unnecessary traffic in the network.

For example, if you have one segment called 100: it has 50 users (in several departments) using this network segment. The Engineering Department is CAD (Computer–Aided Design)–oriented, while the Accounting Department is into heavy number crunching (year end reports, month end statements, etc.). On this network, any traffic between clients of Accounting Department and the Accounting File Server (in the Accounting Department) will be heard across the Segment 100. Likewise, any traffic between the Engineering Dept clients (to the CAD File Server) will be heard throughout the Network Segment. The result is that “Other” Departments accesses to the Generic File Server are incredibly slow: this is because of the unnecessary traffic that’s being generated from other departments (Engineering and Accounting).

The solution is to use one bridge to isolate the Accounting Department and another bridge to isolate the Engineering Department. The Bridges will only allow packets to pass through that are not on the local segment. The bridge will first check its “routing” table to see if the packet is on the local segment. If it is, it will ignore the packet, and not forward it to the remote segment. If a client of Accounting Department sends a packet to the Accounting File Server, then Bridge #1 will check its routing table (to see if the Accounting File Server is on the local port). If it is on the local port, then Bridge #1 will not forward the packet to the other segments. If a client of Accounting Department sends a packet to the Generic File Server, Bridge #1 will again check its routing table to see if the Generic File Server is on the local port. If it is not, then Bridge #1 will forward the packet to the remote port.

2. Manages network traffic by filtering packets

Bridges listen to the network traffic, and build an image of the network on each side of the bridge. This image of the network indicates the location of each node (and the bridge’s port that accesses it). With this information, a bridge can make a decision whether to forward the packet across the bridge – if the destination address is not on the same port – or, it can decide not to forward the packet (if the destination is on the same port).

3. Translates from one protocol to another

The MAC layer also contains the Bus Arbitration method used by the network. This can be CSMA/CD, as used in Ethernet, or Token Passing, as used in Token Ring. Bridges are aware of bus arbitration and special translation bridges can be used to translate between Ethernet and Token Ring.

Bridges physically separate a network segment by managing the traffic (that’s based on the MAC address). Bridges are store and forward devices. They receive a packet on the local segment, store it, and wait for the remote segments to be clear before forwarding the packet. The two physical types of bridges are Local and Remote bridges.

4. Local Bridges

Local Bridges are used (as in the previous examples) where the network is being locally (talking about physical location now) segmented. The 2 segments are physically close together: same building, same floor, etc. Only one bridge is required.

5. Remote Bridges

Remote Bridges are used in pairs, and also used where the network is remotely segmented (again, talking of physical locations). The two segments are physically far apart: different buildings, different floors, etc. 2 x half–bridges are required; one at each segment. The remote bridges are half of a normal bridge, and may use several different communications media in between.

6. Bridging Methodologies

Transparent Bridges examine the MAC address of the frames to determine whether the packet is on the local segment or on the distant segment. Early bridges required the system administrator to manually build the routing table to tell a bridge which addresses were on which side of the bridge. Manually building a routing table is called *fixed* or *static* routing. Modern bridges are self–learning: they listen to the network in **promiscuous mode,** meaning that they accept all packets, regardless of the packets' addressing. The bridge then looks up each packet's destination DLC Address in its internal tables to find out which port the destination NIC is attracted to. Finally, it forwards the packet onto only the necessary port. In the case of a broadcasting message, the bridge forwards the packet onto every port except the port that the packet came from. **Promiscuous listening** is the key to the bridge's transparent operation. Since the bridge effectively "hears" all packets that are transmitted, it can decide whether forwarding is necessary without any special behaviour from the individual stations**.**

Consider a situation where there are two bridges, Bridge A and B. As frames flow on Bridge A's local port, Bridge A examines the source address of each frame. Any frames with a destination address (other than the nodes on the local port) are forwarded to the remote port. As far as Bridge A is concerned, nodes on Bridge B's local port appear as if they were on Bridge A's remote port and therefore are mapped in the table accordingly. Similarly, Bridge B also develops its routing table for various nodes.

The algorithm used by transparent bridges is *backward learning*. As mentioned above, the bridges operate in **promiscuous mode** and track the source addresses of different frames. Because it knows what ports different addresses come from, it also knows onto what port to send packets going to those addresses. The backward learning algorithm can be written in Pseudo Code as follows:

```
if the address is in the tables then
forward the packet onto the necessary port.
if the address is not in the tables, then
forward the packet onto every port except for the port that
the packet was received on, just to make sure the destination
gets the message. add an entry in your internal tables
linking the source address of the packet to whatever port the
packet was received from.
```

Take, for example, a simple network consisting of a four–port transparent bridge with five stations attached to it. The ports on the bridge shall be numbered one through four, with Station A and Station B on port 1, no station on port 2, Station C on port 3, and Station D and Station E on port 4. The bridge has just been brought on–line, and its tables are empty.

Station B transmits a packet destined for station C. Since the bridge doesn't know what port station B is on yet, it puts the packet out onto every port except Port 1 (the packet came from Port 1, so the bridge knows that the packet has already been seen by stations on port 1). This behaviour is known as flooding. The bridge also examines the source address in the packet and determines that Station B is attached to Port 1. It updates its stables to reflect this.

Now that the bridge knows where Station B is, it will forward packets destined for Station B only onto Port 1. As stations transmit packets, the bridge will learn the location of more and more stations until, finally, it knows the location of every station that is attached to its ports. The beauty of the system is that even if the bridge doesn't know the location of a station, packets still get sent to their destination, just with a tiny bit of wasted bandwidth.

Finally, the bridge ages each entry in its internal tables and deletes the entry if, after a period of time known as the aging time, the bridge has not received any traffic from that station. This is just an extra safeguard to keep the bridge's tables up–to–date.

7. Advantages of Transparent Bridges

- Self learning: Requires no manual configuration, considered plug and work.
- Independent of higher level protocols (TCP/IP, IPX/SPX, Netbeui, etc.).
- No hardware changes required, no software changes required.

8. Disadvantages of Transparent Bridges

Can only work with one path between segments: loops are not allowed: A loop would confuse the bridges as to which side of the bridge a node was really on (i.e., local or remote)? Transparent Bridges are not suitable for use on MANs on WANs, because many paths can be taken to reach a destination. In a LAN, it is simple to determine that a loop occurs, but in a large corporate network (with several hundred bridges), it may be next to impossible to determine. As such, bridges are most commonly used in LAN–to–LAN connectivity (and not in MANs or WANs).

9. Spanning Tree Bridges

The Spanning Tree Protocol was developed to address the problem of loops in Transparent Bridging. The IEEE 802.ID (Institute of Electrical and Electronic Engineers) committee formed the Spanning Tree Protocol.

The Spanning Tree Protocol (STP) converts a loop into a tree topology by disabling a bridge link. This action ensures that there is a unique path from any node to every other node (in a MAN or WAN). Disabled bridges are kept in a stand–by–mode of operation until a network failure occurs. At a time, the Spanning Tree Protocol will attempt to construct a new tree, using any of the previously disabled links.

The Spanning Tree Protocol is a Bridge–to–Bridge communication where all bridges cooperate to form the overall bridge topology. The Spanning Tree algorithm is dynamic, and periodically checks every one to four seconds to see if the bridge topology has changed.

Each bridge is assigned an arbitrary number that assigns priority to the bridge in the Internetwork. The number is concatenated with the bridge MAC address. If 2 bridges have the same priority, the MAC address is used as a tie breaker mechanism. The lower the assigned number, the higher the bridge priority.

During initial power–up, a Bridge Protocol Data Unit (BPDU) is flooded out each network port of the bridge. The BPDU contains the following: the current spanning tree root, the distance to the root (measured in hops through other bridges), the bridge address information, and the age of the information in the BPDU. Bridges priorities are usually controlled manually so as to configure the traffic flow – over the Internetwork – on a preferred path.

Problems can arise where, for example, the Spanning Tree Algorithm may select a path from Los Angeles to New York City – and back to San Francisco rather than the preferred route of Los Angeles to San Francisco.

10. Source Routing Bridges

Source–Routing is mostly used to interconnect token ring LANs. In Source–Routing, the source station must determine, in advance, the route to the LAN of the destination station, and include this route in the header of each frame. To determine the routing information, the source station first issues a search frame which is generally an LLC Test command, on its ring. If a response is received from the desired destination station, it indicates that both source and destination stations are on the same ring and that no routing information is required.

However, if no response is received, the source station issues a route discovery frame, which fans–out on every ring in the LAN segment. As the frame is forwarded from one ring to another, each bridge updates the routing information in the search frame. When the search frame reaches the destination, it contains the route between the source and destination stations. The destination station then sends a response frame to the source station, with the routing information. Both stations then use the routing information in each subsequent frame sent to each other.

Source–Routing uses two key parameters to identify a route between a source station and a destination station. These parameters are ring numbers and bridge numbers. Each ring is assigned a unique number.

These numbers generally range between 1 and FFF (hex). Each bridge is assigned a bridge number, ranging between 0 and F (hex). The only restriction when assigning bridge numbers is that parallel bridges connecting identical rings, must have different bridge numbers. The route between the source and the destination stations consists of LAN numbers and bridge numbers. The route is obtained by thus: each bridge which receives the route discovery frame adds to the existing route, its number and the ring number that it forwards this frame to.

The Pseudo Code for Source Routing Bridges can be written as:

- The host uses its known path to the destination if it has one that is not old.
- Else, the host sends a probe message.
- The probe will be forwarded by every bridge that sees it, on every LAN to which the bridge is attached (except the one the probe came in on).
- If the bridge sees its own ID already in the path the probe is accumulating, it will drop the probe without forwarding it (preventing a loop).
- The probe will eventually get to the destination by every possible path, including the shortest.
- The destination will return the probe to the sender, using the discovered route as its source routing path.
- The source will then send its “real” message using the newly discovered route.

3.1.3 Switches

A switch is a device that incorporates bridge functions as well as point– to–point ‘dedicated connections’. They connect devices or networks, filter, forward and flood frames based on the MAC destination address of each frame. Switch operates at Data link layer of the OSI model.

They are technically called bridges. They move data without contention. Ethernet switches provide a combinations of shaed/dedicated 10/100/1000 Mbps connection. Some E–net switches support cut–through switching: frame forwarded immediately to destination without waiting for assembling of the entire frame in the switch buffer. They significantly increase throughput. It provides express lane for traffic.

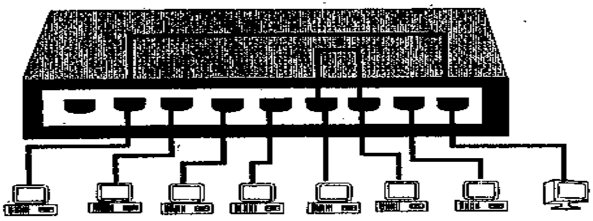


Figure 3: Switch

3.1.4 Hubs

If multiple incoming connections need to be connected with multiple out–going connections, then a hub (Figure 4) is required. In data communications, a hub is a place of convergence where data arrive from one or more directions and are forwarded out in one or more other directions. Hubs are multi–port repeaters and as such, they obey the same rule as repeaters. They operate at the OSI Model Physical Layer.

Hubs are used to provide a Physical Star Topology. At the centre of the star is the Hub, with the network nodes located on the tips of the star.

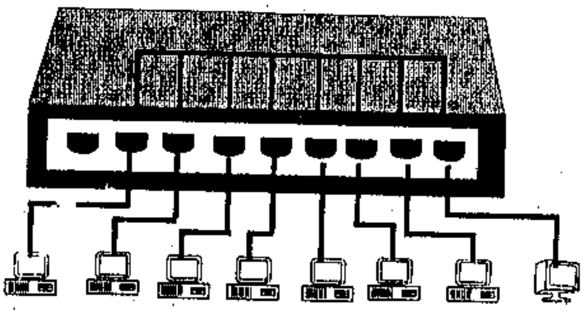


Figure 4: Hub

Star Topology

The hub is installed in a central wiring closet, with all the cables extending out to the network nodes. The advantage of having a central wiring location is that it’s easier to maintain and troubleshoot large networks. All of the network cables come to the central hub. This way, it is especially easy to detect and fix cable problems. You can easily move a workstation in a star topology by changing the connection to the hub at the central wiring closet.

The disadvantages of a star topology are shown below:

- Failure of the Hub can disable a major section of the network.
- The Star Topology requires more cabling than does the Ring or the Bus topology because all stations must be connected to the hub, not to the next station.

Hub's Segment–to–Segment Characteristics

To understand the Ethernet segment–to–segment characteristics of a hub, let us first determine how the Ethernet Hubs operate. Logically, they appear as Topology, and physically, as a Star Topology. Looking inside an Ethernet, we can see that it consists of an electronic printed circuit board. Understating that inside the Hub is only more repeaters, we can draw the conclusion that all connections attached to a Hub are on the same segment (and have the same segment number). A single repeater is said to exist from any port to any port, even though it is indicated as a path of 2 repeaters.

Cascaded Hub Network

Connecting hubs together through ports creates Cascading Hubs. One Master Hub (Level 1) is connected to many Level 2 (slave) Hubs, which are masters to Level 3 (slave) Hubs in a hierarchical tree (or clustered star). The maximum number of stations in a Cascaded Hub Network is limited to 128.

Backbone Networks

In a Backbone Network, there is no master Hub. The level 1 Hubs are connected through their AUI port to a Coax backbone. For thin coax, up to 30 hubs can be connected together. For thick coax, up to 100 hubs can be connected to the backbone. The backbone is considered to be a populated segment.

Level 2 Hubs are allowed to be connected to Level 1 Hubs' 10 Base T ports. This connection between the two hubs is considered an unpopulated segment, or link segment. Up to 1,024 stations (or nodes) can be attached to the Level 2 Hubs' 10 BaseT ports.

All stations and segments would appear as 1 Logical segment, with 1 Network Number. In the real world, 1024 stations are never attached to 1 segment; as the resulting traffic would slow the network to a crawl.

Hub’s Addressing

Because a Hub is just many repeaters in the same box, any network traffic between nodes is heard over the complete network. As far as the stations are concerned, they are connected on 1 long logical bus (wire).

Half–Duplex and Full–Duplex Ethernet Hubs

Normal Ethernet operation is Half–Duplex: only 1 station or node is talking at a time. The stations take turns talking on the bus (CSMA/CD–bus arbitration). Full–Duplex Ethernet Hubs are hubs which allow two–way communication, thus doubling the available bandwidth from 10 Mbps to 20 Mbps. Full–duplex hubs are proprietary products, and normally only work within their own manufacturer’s line.

For example, if A wanted to talk to C, a direct 10 Mbps line would be connected through the two switching hubs. Simultaneously, if D wanted to talk to B, another direct 10 Mbps line (in the opposite direction) would be connected through the two switching hubs (doubling the available bandwidth to 20 Mbps).

There are no official standards for Full–Duplex Ethernet although proprietary standards do exist.

Switching Hubs

Switching hubs are hubs that will directly switch ports to each other. They are similar to full duplex hubs, except that they allow dedicated 10 Mbps channels between ports.

If A wanted to communicate with B, a dedicated 10 Mbps connection would be established between the two. If C wanted to communicate with D, another dedicated 10 Mbps connection would be established.

3.1.5 Comparison Of Switches And Hubs

	HUBS	SWITCHES
1.	Collision Domain	Broadcast Domain
2.	All of the parts on a hub are part of the same Ethernet	Each part on a switch may be regarded as a separate Ethernet (but all are part of the same local area network).
3.	All parts on a hub share the same 10Mb (100 Mb) bandwidth)	Each part on a switch has its own 10Mb (100 Mb) bandwidth
4.	Any frame appearing on one port of a hub is repeated to all other ports on the hub	A directed frame appearing on one part of a switch is forwarded only to the destination port.
5.	A sniffer on any hub port can see all of the traffic on the network	
6.	A hub will repeat defective frames	Switched networks are difficult to sniff.

4.0 CONCLUSION

In this unit, we have examined the features of several network devices such as repeaters, bridges, switches, hubs, etc. and their various purposes in networks.

This unit has exposed you to when and how to use any of these devices. But you should note that all the network devices discussed in this unit are used at physical layer and Data link layer.

5.0 SUMMARY

In this unit we have studied about features of different network devices namely: repeaters, bridges, hubs and switches.

Repeaters are used in long distance network cable to enhance the signals that get weakened due to attenuation.

Bridges are used to interconnect multiple LANs two devices at the data link layers of the OSI model.

Switches are used for performing the functions of bridges as well as point–to–point dedicated connections.

Hubs are used to interconnect various incoming connections with different outgoing connections at the Physical layer of the OSI Model.

6.0 TUTOR–MARKED ASSIGNMENT

1. Which of the following network devices is used at the physical layer? (a) Routers (b) Bridges (c) Repeaters (d) Switches
2. List the major functionality of a bridge
3. Compare Switches and Hubs.
4. What are Switching Hubs?

In the next unit, we will examine another set of network devices.

7.0 REFERENCES/FURTHER READING

UNIT 3 NETWORK DEVICES–II

1.0 INTRODUCTION

In the previous unit, we studied some of the network devices which are used at physical layer and data link layer. In this unit, we continue our discussion about devices/operating at lower layers, and also look at higher layer devices. Routers and Gateways work at network layers and above, whereas modem work at a lower layer. We will also examine the differences between bridges and routers.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- routers
- gateways
- modems.

3.0 MAIN CONTENT

3.1 Network Devices

3.1.1 Routers

In an environment consisting of several network segments with different protocols and architecture, a bridge may not be adequate for ensuring fast communication among all of the segments. A complex network needs a device which not only knows the address of each segment, but also can determine the best path for sending data and filtering broadcast traffic to the local segment. Such a device is called a Router.

Routers are both hardware and software devices. They can be cards that plug into a collapsed backbone, stand–alone devices or software that would run on a file server.

Purpose of Routers

The purpose of a router is to connect nodes across an Internetwork, regardless of the Physical Layer and Data Link Layer protocol that is used. Routers are hardware and topology–independent. Routers are not a ware of the type of medium or frame that is being used (Ethernet, Token Ring, FDDI, X.25, etc.). Routers are a ware of the Network Layer protocol that is used (e.g., Novell's IPX, UNIX's IP, XNS, Apple's DDP, and so on).

Router OSI Operating Layer

Routers operate on the OSI Model's Network Layer. The Internetwork must use the same Network Layer protocol. Routers allow the transportation of the Network Layer PDU through the Internetwork, even though the Physical and Data Link Frame size and addressing scheme may change.

Routers that only know Novell IPX (Internetwork Packet Exchange) will not forward Unix's IP (Internetwork Packet) PDUs, and vice versa. Routers only see the Network Layer protocols that they have been configured for. This means that a network can have multiple protocols running on it (e.g., SPX/IP, TCPIIP AppleTalk, XNS, etc.).

For example, a Novell SPX/IPX router; only sees the Network Layer protocol, IPX. This means that any TCP/IP PDUs will not pass through: the router does not recognise the PDUs, and doesn't know what to do with them. Therefore, routers allow network traffic to be isolated – or segmented – based on the Network Layer Protocol. This provides a functional segmentation of the network.

Routers that can only see one protocol are called Protocol–Dependent Routers. Routers that can see many different protocols (two or more) are called Multi–protocol Routers.

Routing Protocols

Routing Protocols are a "sub–protocol" of the Network Layer Protocol. They deal specifically with the routing of packets from the source, to the destination (across an Internetwork). Examples of Routing Protocols are: RIP, IGRP and OSPF. Let us look at each of these protocols in some more detail.

RIP–Routing Information Protocol

RIP was one of the first routing protocols to gain widespread acceptance. It is described in RFC1058, which is an internet standard. Commercial NOS, such as Novell, Apple, Banyan Vines, and 3Com, use RIP as the base routing algorithm for their respective protocol suites.

RIP is a distance vector algorithm. Routers maintain a detailed view of locally–attached network segments, and a partial view of the remainder of the routing table. The routers contain information on the umber of Hop counts of each segment. A hop is considered to be one transverse through a router. Pass through a router and the hop count increases by 1.

The routers are updated every 30 seconds, when each router sends out a RIP broadcast. This advertisement process is what enables RIP routing to be dynamic. Dynamic routers can change routing tables on the fly (as the network configuration changes). By using the Hop Count information from their routing tables, routers can select the shortest path (the least number of hops) to the destination.

Apple uses RTMP (Routing Table Maintenance Protocol):

This adds a good, bad or suspect route status indicator, depending on the age of the route information.

Novell adds Ticks to the RIP Algorithm:

Ticks are dynamically assigned values that represent the delay associated with a given route. Each tick is considered 1/18 of a second. LAN segments are typically assigned a value of 1 tick. A T1 link may have a value of 5 to 6 ticks and a 56 Kbps line may have a value of 20 ticks. A larger number of ticks indicate a slower routing path.

Three commonest problems that can occur with RIP are shown below:

1. Routing loops

The router indicates that the shortest path is going back the way the packet came from

2. Slow Route Convergence

Routers have delay timers that start counting after the RIP advertising packet is broadcast. This gives the routers time to receive and formulate a proper routing table from the other routers. If the delay timer is too short, the routing table can be implemented with incomplete data causing routing loops.

3. Hop Count Exceeded

The maximum number of hop counts is 15 for RIP. A hop count of 15 is classified as unreachable which makes RIP unsuitable for large networks where hop counts of 15 and above are normal.

EGRP–Exterior Gateway Routing Protocol

EGRP was created to solve many of the problems with RIP, and has become the default routing protocol across the internet. EGRP is an enhanced distance vectoring protocol; it uses up to5 metrics (conditions) to determine the best route as shown below:

1. Bandwidth
2. Hop Count (Delay)–maximum of 255
3. Maximum Packet size
4. Reliability
5. Traffic (Load).

These routing metrics are much more realistic indicators (of the best routes) than simple hop counts.

OSPF–Open Shortest Path First

OSPF is a link state premises:

It has several states of routers that are linked together in a hierarchical routing model. This means that each router maintains link status information and this is exchanged between routers wishing to build routing tables. Unlike RIP, OSPF uses IP directly, OSPF packets being identified by a special value in the IP datagram protocol field.

The top of the root is the Autonomous Router that connects to the autonomous systems (the Internet). The next is the Backbone Routers, the highest area in the OSPF system. Border routers are attached to multiple areas and they run multiple copies of the routing algorithm. Last are internal routers that run a single routing database for one area.

Basically, by dividing the network into a routing hierarchy, both substantial reduction of routing update traffic and faster route convergence – result on a local basis. Each level has a smaller routing table and less to update.

3.1.2 Comparison of Bridges and Routers

- Both are stored–and forward devices, but Routers are Network Layer devices (examine network layer headers) and Bridges are Link Layer devices.

- Routers maintain routing tables (hierarchical, aggregatable addresses) and implement routing algorithms, bridges maintain filtering tables (flat addresses) and implement filtering, learning and spanning tree algorithms.

3.1.3 Gateways

This device (Figure 1) is used to connect totally dissimilar networks. They function at a high end of OSI model. They perform protocol conversion for all seven layers of the OSI model. They are commonly used to connect a LAN and a main frame computer. Gateways handle conversions of messages, addresses and protocol, to deliver a message from one network to another. They offer greatest flexibility in internetworking communications. Gateway's decision – making is more complex than Routers. They are very costly and their implementation, maintenance and operations, are also very complex. They are slower than other devices. They can recover e–mail messages in one format and convert them into another format.

Gateways provide an interface between IPX–based LANs and the IP protocols of the internet. This provides a centralised and secure way to connect IPX–based LANs to IP networks. Because of this, a single IP address can be used for an entire network. Therefore, this eliminates configuration and maintenance problems.

Dual–homed Gateway is also present in the network. It is a system that has two or more network interfaces. It acts to block or filter some or all of the traffic trying to pass between the networks in firewall configuration.

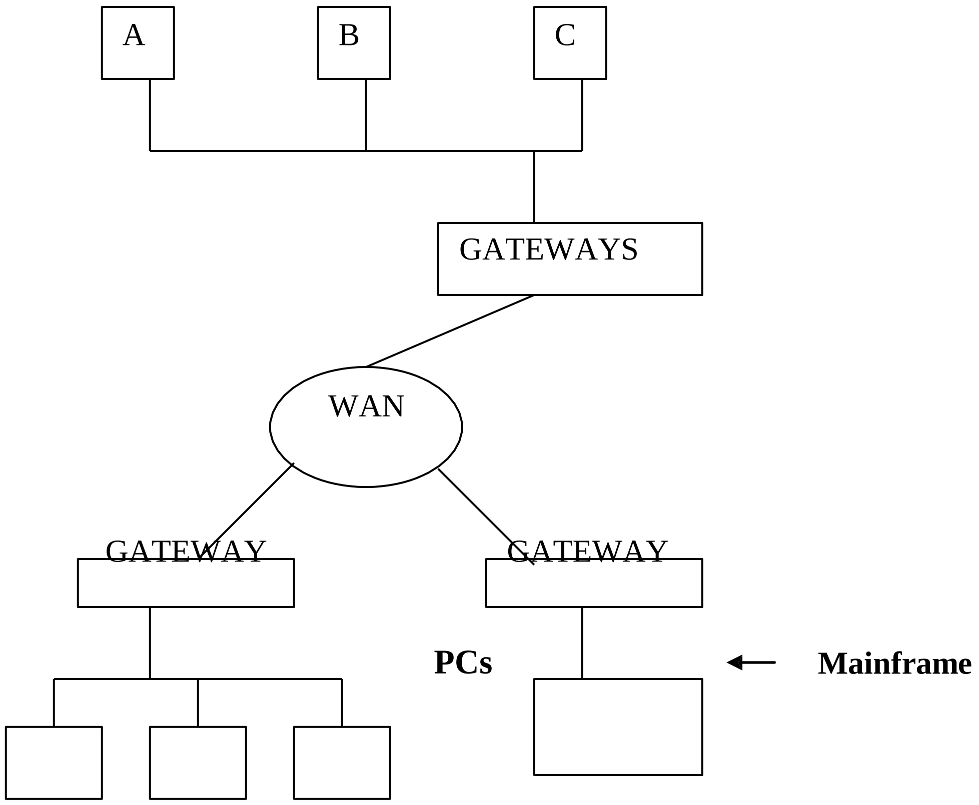


Figure 1: Gateway

Gateway has its main memory and processor to perform protocol conversion.

Typical corporate gateways connect the PC world of token Ring, Ethernet and AppleTalk LANs to IBM's main frame SNA environment with x.25 packet switched networks or DECnet networks.

At the lowest level, gateway provides terminal emulation so all LAN workstations can emulate varies considerably depending on the gateway.

Second level of gateway functionality includes file sharing between LAN & host. Novell has developed a platform – independent version of netware that will run on several different platforms, including several traditional mini–computer platforms.

At the higher level of functionality, a gateway would provide peer–to–peer communications between micro computer programs running on the LAN, and mainframe programs running on the host. These types of client/server relationships will become more and more important in the near future as programs are written to distribute databases among LAN's mini–computers and mainframes, with the machine users communicating with the programs, using the same type of user interface.

How Do Gateways Link Hosts and LANs?

Using gateway's micro–mainframe connection is much more cost effective than other types of connections like using coaxial cable via PC 3270 emulation card etc. The gateway board emulates a cluster controller so each network workstation is seen by the mainframe as a terminal linked to the cluster controller. The gateway's multiple mainframe sessions are split among the network's workstations, so the channel rarely sits idle. Only the gateway needs to have a circuit card and the software necessary for protocol conversion and terminal emulation.

Remote LAN Gateways

These gateways (Figure 2) are becoming very common because of the evolution of enterprise networks and WAN. A PC on the remote site's LAN functions as a gateway and runs gateway software. This gateway PC functions as a cluster controller and communications with a front- end processor using IBM's Synchronous Data Link Control (SDLC) protocol via synchronous modems located at both sites.

The limitation of these gateways has speed. A synchronous modem can dial up a front–end processor at speeds up to 64Kbps. Companies with heavy micro–mainframe traffic might require multiple remote gateways to solve this congestion problem.

X.25 Gateways

Remote LAN can also communicate with IBM mainframe viz., x.25 gateway. A gateway PC with an adapter card functions as a cluster controller and runs special gateway software that Contains the QLLC protocol, an IBM defined protocol that runs over the X.25 suite. The other LAN workstations emulate IBM 3270 terminals. The IBM host simply assumes it's communicating with the remote cluster controller.

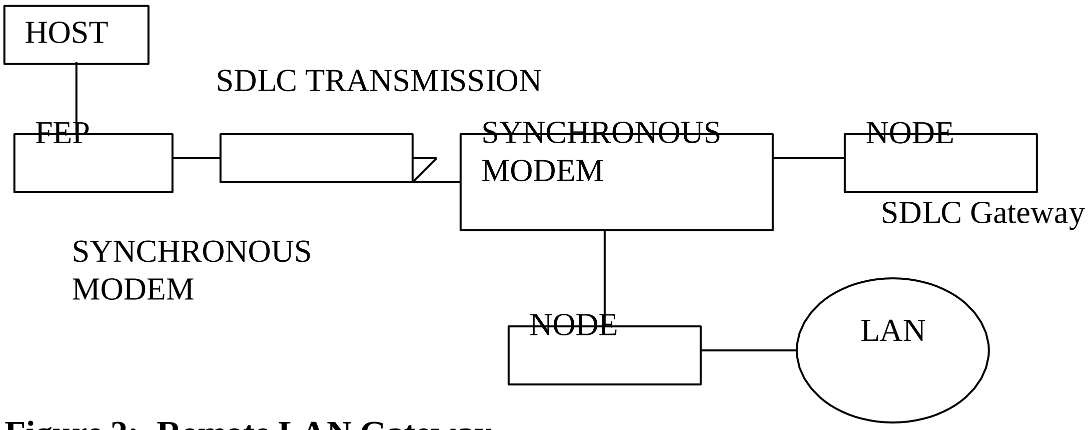


Figure 2: Remote LAN Gateway

Netware Workstation running 3270
Terminal Emulation Software

3.14 Modem

This is a device which is used to convert digital signals generated by the computer into an analog signal to be carried by a public access telephone line. It is also the device that converts the analog signal received over a phone line into digital signal usable by the computer. A modem derives it meaning from a modulation, and demodulation is a composite word that refers to two functional units that make up a device. A signal modulator and a signal demodulator. A modulator converts digital signal into an analog signal. A demodulator converts analog signal into digital signal.

Modem can be classified into many categories to include the mode of transmission and their techniques, as well as by the application features they contain and the type of lines they are built to service.

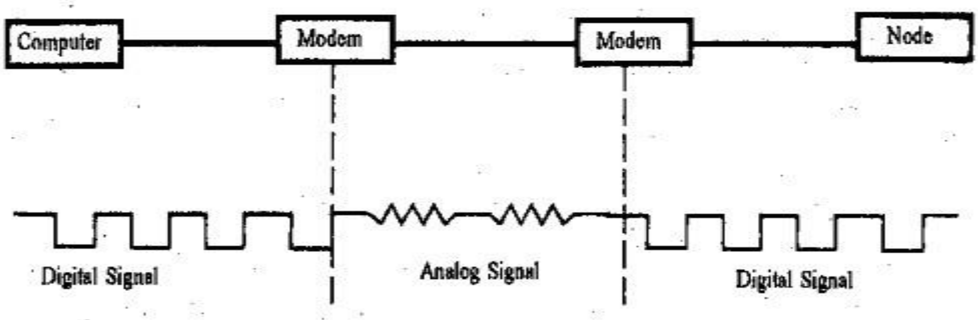


Figure 3: Signal conversion by modems i.e Modulation and Demodulation

Speed

Modem speed ranges from 300 bps to 56kbps. It normally transmits about 10 bits/character (each character has 8 bits); maximum rate of characters for a high speed modem is 2,880 characters/sec. For example, a compressed image of 20KB (equivalent to 20,000 characters) will take nearly 6 seconds to load on the fastest modem. The tasks which a modem can perform are:

1. Automatically dials another modem using either touch–tone or pulse dialing.
2. Auto answer i.e., automatically answers another modem for making connection.
3. Disconnects a telephone connection when data transfer has been completed or if an error occurs.
4. Automatic speed negotiation between two modems
5. Converts bits into the form suitable for the line (Modulator)
6. Transfer data reliably with the correct type of handshaking
7. Convert received signals back into bits (demodulator)

Modem standards

The CCIT (now known as ITU) has defined standards for modem communication. Each uses v number to define their type.

v.22 bis – It operates at 1200 or 2400bps v.32
– Operates at 9600 bps
v.32 bis – Operates at 19,200 bps v. 33
– Operates at 14,400 bps v. 34 –
Operates at 28,800 bps

Modem Commands

They are provided by Hayes Company that pioneered Modems and defined the standard method of programming the mode of modem, which is the AT command language. A computer gets the attention of the modem by sending "AT" command. For example, 'ATDT' is the touch–tone dial command. Initially, a modem is in the command mode and accepts commands from the computer. These commands are sent at either 300 bps or 1200bps.

Most commands are sent with AT prefix. Each command is followed by carriage return character; a command without this is ignored. More than one command can be placed in a single line and spaces can be entered to improve readability, either character case can be used.

Modem can enter two states; the normal state and command state. In the normal state, the modem transmits or receives characters from the computer an in the command state, characters sent to the modem are interpreted as commands. Once a command is interpreted, the modem goes into the normal state. Any character sent to the modem is then sent along with line. To interpret the modem or to end a connection so that it goes back into command mode, three consecutive '+' characters are sent i.e. '+++'.

Example:

When a computer wants to make a connection using telephone no. 17325, it sends the command. 'ATCH 17325' using tone dialing. The modem then replies with an OK response i.e., 'O' value and it tries to make connection with remote modem. If it is not able to make connection, it sends a message in form of a code as (3) for no carrier, (7) for busy (6) for no dial tone etc. If it gets connected then it returns a connect code as it sends '+++' and then waits for a command from host computer. In this case, command is "hang–up the connection" (ATH). The modem will then return an OK response when it has successfully cleared the connection.

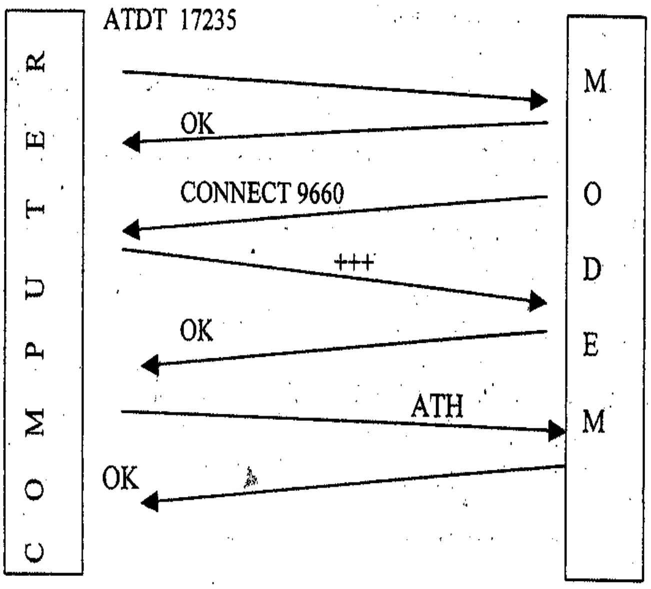
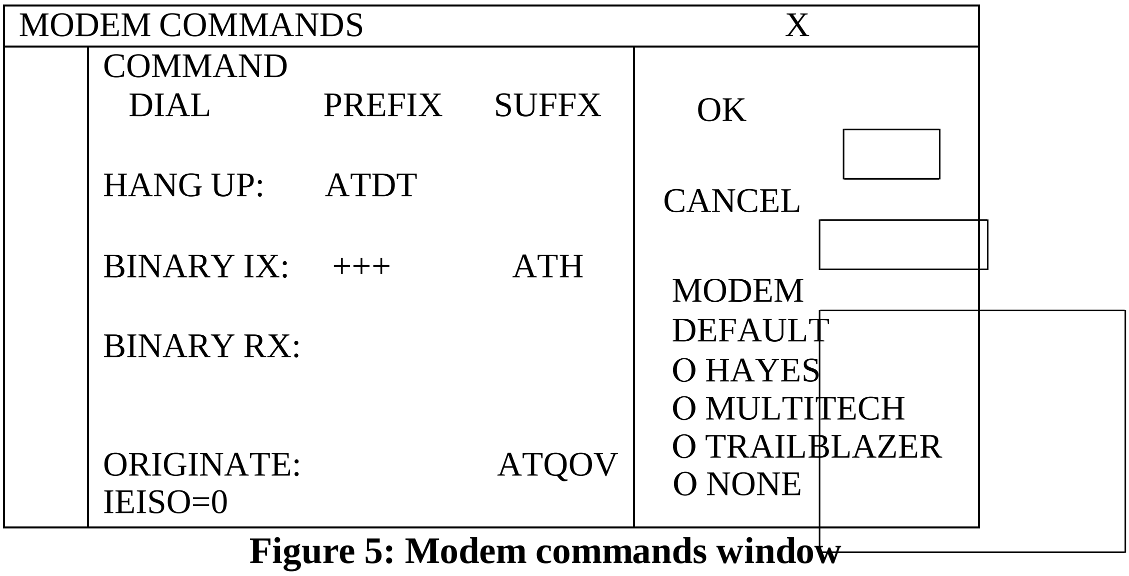


Figure 4: Connection establish & release

The modem contains various status registers called s–register which store modem settings.

Modem Set Up

The following figure shows a sample window from the MS Windows terminal program (in both MS Windows 3.x and Windows 95/98). It shows the modem command window. It can be seen that when the modem dials a number, the prefix to the number dialed is ‘ATDT’. The hang–up command sequence is ‘+++’ ATH.



MODEM COMMANDS X

COMMAND

DIAL	PREFIX	SUFFX
HANG UP:	ATDT	
BINARY IX:	+++	ATH
BINARY RX:		
ORIGINATE: IEISO=0		ATQOV

OK

CANCEL

MODEM DEFAULT

○ HAYES

○ MULTITECH

○ TRAILBLAZER

○ NONE

Figure 5: Modem commands window

Modem Indicator

These are used to inform the user about current status of a connection. Typically the indicator lights are:

- AA – ON when receiving call. OFF when not receiving calls, flash when call is incoming.
- CD – ON when modem detects the remote modem’s carrier, else it is off.
- OH – ON when modem is on the hook else off.
- RD – Flashes when modem is getting data or a command from the computer.
- SD – Flashes when Modem is sending data.
- TR – Shows that DTR line is active i.e., computer is ready to send or receive data.
- MR – Shows that modem is powered up.

The following table illustrates widely used modems with bit rates & modulation techniques

Typical Modems:

ITU Recommendations	Bit rate (bps)	Modulation
V.21	300	FSK
V.22	1200	PSK
V.22 bis	2400	ASK/PSK
V.27 ter	4800	PSK
V.29	9600	ASK/PSK
V.32	9600	ASK/PSK
V.32 bis	14400	ASK/PSK
V.34	28800	ASK/PSK

Most modems operate with V .22 bis (2400bps), V.32 (9600bps), V.32 bis (14400bps) The V.32 and V.32 bis modems can be enhanced with echo cancellation. They also typically have built–in compression using either the V.42 bis standards or MNPC (Microcom Networking Protocol) level 5.

4.0 CONCLUSION

In this unit you have been taken through network devices such as routers and gateways that work at network layers and above, and modems that work at a lower layer.

This unit has also exposed you to the differences between bridges and routers.

5.0 SUMMARY

In this unit, we have studied about some networking devices which are used at higher layers of OSI model. The devices which were covered are the following:

1. Router

Used to connect two devices at the network layer of the OSI Model

2. Gateway

Used to connect totally dissimilar networks because they can perform protocol conversion for all seven layers of the OSI Model.

3. Modem

Used to connect the computer with the telephone lines. A Modem can convert digital signal of a computer to analog signals, so that it can be transferred through the telephone lines.

6.0 TUTOR–MARKED ASSIGNMENT

1. Which layer does the Router operates?
   (a) Physical Layer (b) MAC Layer (d) Network Layer
   (c) Session Layer
2. List few standards of modems.
3. List the names of routing protocols.

7.0 REFERENCES/FURTHER READING

Unit 5 Asynchronous Transfer Mode (ATM)

1.0 INTRODUCTION

Asynchronous Transfer Mode (ATM) is a form of data transmission that allows voice, video and data to be sent along the same network. In contrast to ATM, in the past, voice, video and data were transferred using separate networks. For example, voice was transmitted over the phone, video over cable networks, and data over an internetwork. ATM is the ultimate culmination of all the developments in switching and transmission in the last twenty years and has the best of circuit switching and packet switching (discussed, in the previous block).

Asynchronous Transfer Mode (ATM) is a technology that has its history in the development of broadband ISDN in the 1970s and 1980s. In this unit, first we will have a re–look at different types of switching techniques (technologies) and then we will examine how ATM is compatible with the existing technologies and then compare the architectural difference between ATM and the OSI model and finally, spend some time on how ATM protocol works.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- explain the term ATM
- discuss the compatibility of ATM as technology
- compare ATM-layered architecture with OSI Model
- describe how ATM protocol works
- describe the structure of ATM cell
- identify the various ATM classes of services
- define the various ATM classes of service
- discuss the approach and tools used for ATM traffic control
- discuss the benefits of ATM technology
- explain the various applications of ATM technology

3.0 MAIN CONTENT

3.1 Switching Techniques

In this section, we will discuss different types of switching techniques.

Circuit Switching

This was the first type of data transfer mechanism used. Circuit switching is used in the telephone networks to transmit voice and data signals. In a synchronous transmission, which involves transmission of voice, a synchronized connection must be made between the sender and receiver because there must be a constant time interval between each successive bit, character, or event. To enable synchronized transmission, circuit switching establishes a dedicated connection between the sender and the receiver involved in the data transfer over the network. As a result, the connection consumes network capacity whether or not there is an active transmission taking place; for example, the network capacity is used even when a caller is put on hold. For different applications, utilisation of the line can vary enormously. However, there is little delay and effective transparency for the user. It is very efficient for Constant Bit Rate (CBR).

Packet Switching

In contrast to circuit switching, packet switching ensures that the network is utilised at all times. It does this by sending signals even in the small unused segments of the transmission – for example, between the words of a conversation or when a caller is put on hold. However, in packet switching, there can be variations in the timing when the digital bits are received. For normal voice and data communications, this is not a problem, but for broadband signals, such as television, it is a huge problem that causes the picture to jerk and the audio to be out of synchronization with the picture. Data to be sent is broken down into chunks or packets. Each packet contains data and header information for control e.g., routing. At each node the packet is received, stored briefly and passed on. At each node, the packets may be put on a queue for further movement into the network.

There are two approaches to transport–

1. **Datagram,** where each packet can take any path through the network as long as they all reach the destination.

2. **Virtual Circuit,** where all the packets are routed through the same path without having the path dedicated.

Datagram allows for dynamic handling of congestion and no call setup is necessary. Virtual channels allow for sequencing, error and flow control.

Though, Packet switching is much more efficient than circuit switching, Packet–switched networks have been slow. The public data networks that use the x.25 standard for public switching allow users to operate typically at speeds of 9.6 kbps. The standard leased lines that large companies use for their high-speed data communications, operate at 56 kbps. ATM can transmit bits through the network at speeds up to 622 Mbps.

Multirate Circuit Switching

This is an enhancement of the synchronous **Time–Division Multiplexing (TDM)** approach used initially in circuit switching. In circuit switching, a station must operate at which must be used regardless of application. In multirate switching, multiplexing is introduced. A station attaches to the network by means of a single physical link which carries multiple fixed data–rate channels (B–channel @ 64kbps). Traffic on each channel can be switched independently through the network to various destinations. This is used for simple ISDN. So the user has a number of data rate choice but they are fixed so Variable Bit Rate (VBR) is difficult to accommodate efficiently.

Frame Relay

Frame relay is essentially identical to packet switching. Frame relay saw its development as a result of high data rates and low error rates in modern high–speed communications systems. In old packet switching, there was considerable overhead involved in error recovery, redundancy enhancement and routing information. With Frame relay, the packets are now of variable length and not fixed length, meaning that they were designed to operate at up to 2Mbps. This was very good for VBR.

Cell Relay

This is an evolution from Frame relay and multirate circuit switching. Cell relay uses fixed sized packets called calls. Multirate circuit switching also has fixed channels. Cells relay allows for the definition of virtual channels with data rates dynamically defined. Using a small cell size allows almost constant data rate even though it uses packets. From frame relay, cell relay takes improved error control into account, and allows more errors to be handled at a higher logical level. The fixed–size cells reduce overhead even more and thus allow rates of tens to hundreds of Mbps.

So, in the evolution of switching technology there has been a change from two areas – circuit switching for CBR, and packet switching for VBR

3.2 How Compatible is ATM as Technology?

ATM is emerging as a viable technology. Some of its application are as follows:

- ATM is used in many networks today including both private and public environments. ATM is used extensively by most public service providers today to integrate different types of traffic into one network.
- ATM can be used in existing twisted pair, fibre–optic, coaxial, and hybrid fibre/coax (HFC) networks for local area network (LAN) and wide area network (WAN) communications. Because ATM was developed to have such a wide range of compatibility with existing networks, its implementation does not require replacement or over–building of telephone, data, or cable networks.
- ATM is also compatible with wireless and satellite communications.

3.3 ATM Layered Architecture in Comparison with OSI Model

ATM has a layered structure that is similar to the 7–layered OSI model. However, ATM only addresses the functionality of the two lowest layers of the OSI, i.e;

- The physical layer, and
- The data link layer.

Apart from these two layers, all other layers of the OSI model are irrelevant in ATM, as these layers are only part of the encapsulated information portion of the cell which is not used by the ATM network.

In ATM, the functionality of the two lower OSI layers is handled by three layers.

Application Layer
User Layers
ATM Adaptation Layer (AAL): ***Convergence sublayer***
Segmentation and Reassembly sublayer
ATM Layer
Physical Layer
Transmission Convergence Sub layer

ATM Protocol Model

i) Physical Layer

The Physical layer defines the specification of a transmission medium (copper, fibre optic, coaxial, HFC, wireless) and a signal–encoding scheme and electrical to optical transformation. It provides Convergence with physical transport protocols such as SONET, as well as the mechanism for transforming the flow of cells into a flow of bits.

The ATM form has left most of the specification for this level to the implementer.

ii) The ATM layer deals with cells and cell transport. It defines the layout of a cell and tells what the header fields mean. The size of a cell is 53 bytes (5 bytes of header and 48 bytes of payload). Because each cell is the same size and all are relatively small, delay and other problems with multiplexing different sized packets are avoided.

It also deals with establishment and release of virtual circuits. Congestive control is also located here. It resembles the network layer of the OSI model as it has got the characteristics of the network layer protocol of OSI model like;

- Routing
- Switching
- End-to-end virtual circuit set up
- Traffic management

Switches in ATM provides both switching and multiplexing cell format of ATM Layer are distinguished as

- UNI (User Network Interface)
- UNI (Network–Network Interface

In both cases, the cell consists of a 5–byte header followed by a 48–byte pay–load but the two headers are slightly different.

iii) ATM Adaptation Layer

The ATM Adaptation Layer (AAL) maps the higher-level data into ATM cells to be transported over the ATM network, i.e., this layer segments the data and adds appropriate error control information as necessary. It is dependent on the type of services (voice, data, etc.) being transported by the higher layer.

This is the adaptation layer that divides all types of user data into 48–byte cells. The ATM layer that adds the five–byte header information to direct the user data to its destination.

Depending on the type of data, several AAL protocols have been defined. However, no AAL is restricted to a specific data class or type; all types of data could conceivably be handled by any of the AALs. The various AAL protocols define are:

1. AAL 1
2. AAL 2
3. AAL ¾
4. AAL 5

It is divided into two sublayers

- SAR (Segmentation & Reassembly)
- CS (Convergence Sublayer)

Segmentation & Reassemble

This is the lower part of the AAL. The SAR sublayer breaks packets up into cells on the transmission side and puts them back together again at the destination. It can add headers and trailers to the data units given to it by the CS to form payloads. It is basically concerned with cells.

Convergence Sublayer

The CS sublayer makes it possible to have ATM system offer different kinds of services to different applications. The CS is responsible for accepting bit streams or arbitrary length messages from the application and breaking them into units of 44 or 48 bytes for transmission.

How ATM Protocol Works

When a user sends data over the ATM network, the higher–level data unit is passed down to the Convergence Sublayer of the AAL Layer, which prepares data for the ATM Layer according to the designated AAL protocol. The data is then passed down to the Segmentation and Reassembly sublayer of the AAL Layer, which divides the data unit into appropriately sized segments.

These segments are then passed down to the ATM Layer, which defines an appropriate cell header for each segment and encapsulates the header and payload segment into a 53–byte ATM cell. The cells are then passed down to the Physical Layer; which streams the cells at an appropriate pace for the transmission medium being used, adding empty cells as needed.

ATM circuits are of two types:

1. Virtual Paths and,
2. Virtual Channels.

A virtual channel is a unidirectional pipe made up from the concatenation of a sequence of connection elements.

A virtual path consists of a set of these channels.

Each virtual channel and virtual path has an identifier associated with it. Virtual path is identified by Virtual Path Identifiers (VPI) and a virtual channel is identified by a Virtual Channel Identifier (VCI). All channels within a single path must have distinct channel identifiers but may have the same channel identifier as channels in different virtual paths.

An individual channel can, therefore, be uniquely identified by its virtual channel and virtual path number. Cell sequence is maintained through a virtual channel connection.

ATM connections can be categorised into two types:

i) **Point–to–point connections: –** These are the connections which connect two ATM end–systems. Such connections can be unidirectional or bidirectional.

ii) **Point–to–multipoint connection:** These are the connections which connect a single source end–system known as the root node, to multiple destination end–systems (known as leaves).

The basic operation of an ATM switch is very simple to understand.

1. The ATM switch receives a cell across a link on a known VCI or VPI value.

2. The ATM switch looks up to the connection value in a local translation table to determine the outgoing port (or ports) of the connection and the new VPI/VCI value of the connection on that link.

3. The ATM switch then retransmits the cell on that outgoing link with the appropriate connection identifiers.

The manner in which the local translation tables are set up determines the two fundamental types of ATM connections:

- **Permanent Virtual Connections (PVC):** A PVC is a connection set up by some external mechanism, typically network management, in which a set of switches between an ATM source and destination ATM system are programmed with the appropriate VPI/VCI values.

- **Switched Virtual Connections (SVC):** An SVC is a connection that is set up automatically through a signal protocol. SVCs do not require the manual to set up PVCs and, as such, are likely to be much more widely used.

The ATM Network

An ATM networks consist of a set of ATM switches interconnected by point–to–point ATM links or interfaces. ATM switches support three kinds of interfaces:

- User–Network Interfaces (UNI)
- Network–Node Interfaces (NNI)
- Inter–Carrier Interface (ICI)

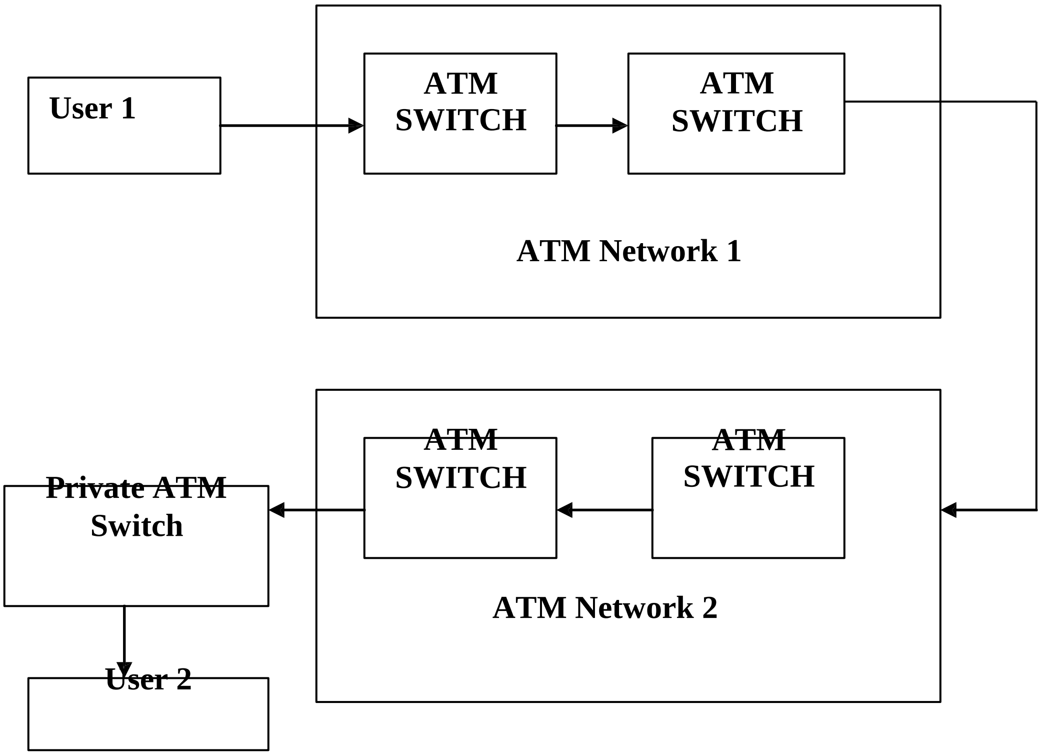


Figure 1: ATM Network

- The UNI exists between a single end-user and a public ATM network, between a single end-user and a private ATM switch, or between a private ATM switch and the public ATM network.
- The NNI exists between switches in a single public ATM network. NNIs may also exist between two private ATM switches.
- The ICI is located between two public ATM networks.

The major differences between these two types of interfaces are administrative and signaling related. The only type of signaling exchanged across the UNI is that required to set up a **Virtual Channel** for the transmission.

Communication across the NNI and the ICI will require signaling for virtual–path and virtual–channel establishment, together with various exchange mechanisms for the exchange of information such as routing tables, etc.

Let us take an example to understand how the ATM network works

- Let there be a user 1 in Delhi who wishes to transfer a data file-to user 2 in Bangalore. A virtual channel is created and a virtual path is established from switch to switch within the public ATM network in Delhi (ATM Network 1) which, in turn, establishes contact with the public ATM network in Bangalore (ATM Network 2).
- ATM Network 2 also establishes a virtual path from switch to switch within the network and with the private ATM Switch at the destination. The private ATM network completes the virtual path by establishing a virtual channel with User 2 in Bangalore.
- At each interface in this network, a unique virtual path identifier (VPI) and the virtual channel identifier (VCI) is established for this transmission. These identifiers are significant only for a specific switch and two nodes adjacent to it in the virtual path. Each node within the virtual path (including both the end-users and the switches) maintain a pool of inactive identifiers to be used as needed.
- User 1 or User 2 terminates the cell and the virtual path is discontinued. The VCI and VPI values are returned to the pool of available values for each switch.

Notice that only the user at either end of the transmission deal with the 48–byte information load within the cell. At each stage of the transmission, the switch is only concerned with accepting the cell from one port, changing the VPI/VCI according to its tables, and routing the cell out the appropriate switch port.

The ATM Cell

ATM transmits all the information in small, fixed–size packets called cells. Each individual ATM cell consists of a 5–byte cell header and 48 bytes of data. The ATM network uses the header to support the virtual path and the virtual channel routing, and to perform a quick error check for corrupted cells.

Bytes

5	48
Header	User data

Figure 2: An ATM Cell

The Header Format

The structure of the header is different in UNI and NNI. In the network–network interface, the virtual path identifier field is expanded from 8 to 12 bits.

8	7	6	5	4	3	2	1
Generic Flow Control*				Virtual Path Identifier			
Virtual Path Identifier				Virtual Channel Identifier			
Virtual Channel Identifier							
Virtual Channel Identifier				Payload Type ID			CLP
Header Error Control							
INFORMATION PAYLOAD (48 Bytes)							

Figure 3: User–Network Interface

8	7	6	5	4	3	2	1
Virtual Path Identifier							
Virtual Path Identifier				Virtual Channel Identifier			
Virtual Channel Identifier							
Virtual Channel Identifier				Payload Type ID CLP			
Header Error Control							
INFORMATION PAYLOAD (48 Bytes)							

Figure 4: Network–Network Interface

Let’s now look at the characteristics of each of the fields of the header format of an ATM cell.

Generic Flow Control (GFC)

The GFC field of the header is only defined across the UNI and does not appear in the NNI.

Function

- It controls the traffic flow across the UNI.

Virtual Path Identifier (VPI)

The VPI is an 8–bit field for the UNI and a 12–bit field for the NNI

Function

- It constitutes a routing field for the network and is used to identify virtual paths. In an idle cell, the VPI is set to all 0’s.

- Together with the Virtual Channel Identifier, the VPI provides a unique local identification for the transmission.

Virtual Channel Identifier (VCI)

It is a 16–bit field used to identify a virtual channel. For idle cells, the VCI is set to all 0’s.

Function

- It functions as a service access point it and is used for routing to and from the end-user.
- Together with the Virtual Path Identifier, the VCI provides a unique local identification for the transmission.

Payload Type Identifier (PTI)

The PTI field indicates the type of information in the information field. The value in each of the three bits of PTI indicates different conditions.

Bit 1 is set to 1 to identify operation, administration, or maintenance cells (i.e., anything other than data cells).

Bits 2 is set to 1 to indicate that congestion was experienced by a data cell in transmission and is only valid when bit 4 is set to 0.

Bit 3 is used to convey information between end-users.

Cell Loss Priority (CLP)

The 1–bit CLP field is used for indication of the priority of the cell. It is used to provide guidance to the network in the event of congestion. When set to value 1, it indicates that the cell is subject to discard within the network. When the CLP value is set to 0, it indicates that the cell is of relatively high priority and should be discarded only in situations when no alternative is available.

Header Error Control (HEC)

Each ATM cell includes an 8–bit HEC that is calculated based on the remaining 32 bits of the header.

Function:

- It detects all single–bit errors and some multiple–bit errors. As an ATM cell is received at a switch, the HEC of the cell is compared and all cells with HEC discrepancies (errors) are discarded. Cells with single–bit errors may be subject to error correction if supported or discarded. When a cell is passed through the switch and the VPI/VCI values are altered, the HEC is recalculated for the cell prior to being passed out the port.

Advantages of small, fixed-sized cells

Here is a list of some advantages of a cell.

1. Reduced queuing delay for a high priority cell;
2. Easy to implement the switching mechanism in hardware;
3. The fixed cell size ensures that time–critical information such as voice or video, is not adversely affected by long data frames or packets;
4. The header is organised for efficient switching in high–speed hardware implementations and carries pay–load–type information, virtual – circuit identifiers, and header error check.

ATM Classes Of Service

ATM is connection oriented and allows the user to specify the resources required on a per–connection basis (per SVC) dynamically. There are five classes of service defined for ATM (as per ATM Forum UNI 4.0 specification).

Service class	**Quality of Service Parameter**
Constant bit rate (CBR)	CBR class is used for emulating circuit switching. The cell rate is constant with time. CBR applications are sensitive to cell–delay variation. Examples of applications that can use CBR are telephone traffic (i.e., nx64 kbps), video conferencing, and television.
Variable bit rate –real time (VBR – RT)	VBR–NRT class allows users to send traffic at a rate that varies with time depending on the availability of user information. Statistical multiplexing is provided to make optimum use of network resources. Multimedia e–mail is an example of VBR–NRT.
Variable bit rate–non real time (VBR–NRT)	This class is similar to VBR–NRT but is designed for applications that are sensitive to cell–delay variation. Examples of real–time VBR are voice with speech activity detection (SAD) and interactive compressed video.
Available bit rate (ABR)	ABR class provides rate–based flow control and is aimed at data traffic such as file transfer and e–mail. Although the standard does not require the cell transfer delay and cell–loss ratio to be guaranteed or minimised, it is desirable for switches to minimise delay and loss as much as possible. Depending upon the state of congestion in the network, the source is required to control its rate. The users are allowed to declare a minimum cell rate, which is guaranteed to the connection by the network.
Unspecified bit Rate (UBR)	UBR class is the catch–all other class and is widely used today for TCP/IP.

The ATM Forum has identified certain technical parameters to be associated with a connection.

ATM Technical Parameters

Technical Parameters	Definition
Cell loss ratio (CLR)	CLR is the percentage of cells not delivered at their destination because they were lost in the network due to congestion and buffer overflow.
Cell transfer delay (CTD)	The delay experienced by a cell between network entry and exit points is called the CTD. It includes propagation delays, queuing delays at various intermediate switches, and service times at queuing points.
Cell delay variation (CVD)	CVD is a measure of the variance of the cell transfer delay. High variation implies larger buffering for delay–sensitive traffic such as voice and video.
Peak cell rate (PCR)	The maximum cell rate at which the user will transmit. PCR is the inverse of the minimum cell inter–arrival time.
Sustained cell rate (SCR)	This is the average rate, as measured over a long interval, in the order of the connection lifetime.
Burst tolerance (BT)	This parameter determines the maximum burst that can be sent at the peak rate. This is the bucket–size parameter for the enforcement algorithm that is used to control the traffic entering the network.

ATM Technical Parameters

Finally, there are a number of ATM classes of service. These classes are:

ATM Classes of Service

Classes of Service	CBR	VBR –NRT	VBR – RT	ABR	UBR
CLR	Yes	Yes	Yes	yes	No
CTD	Yes	No	Yes	no	No
CDV	Yes	Yes	Yes	No	No
PRC	Yes	Yes	Yes	No	Yes
SCR	No	Yes	Yes	No	No
BT @ PCR	No	Yes	Yes	No	No
Flow control	No	No	No	Yes	No

Its extensive class–of–service capabilities make ATM the technology of choice for multimedia communications.

3.8 ATM Traffic Control

An ATM network needs efficient traffic control mechanisms to allocate network resources in such a way as to separate traffic flows according to the various service classes and to cope with potential errors within the network at anytime. The network should have the following traffic control mechanisms:

- Network Resource Management
- Connection Admission Control
- Usage Parameter Control and Network Parameter Control
- Priority Control
- Congestion Control.

Network Resource Management

Network Resource management deals with allocation of network resources in such a way that traffic is separated on the basis of the service characteristics. A tool of network resource management which can be used for traffic control is the **virtual path technique.** A Virtual Path Connection (VPC) groups several Virtual Channel Connections (VCCs) together such that only the collective traffic of an entire virtual path has to be handled. In this type of set up, priority can be supported by re–aggregating traffic types requiring different qualities of service through virtual paths. Messages for the operation of traffic control can be more easily distributed, a single message referring to all the virtual channels within a virtual path will do.

Connection Admission Control

Connection Admission Control is the set of actions taken by the network in protecting itself from excessive loads. When a user requests a new virtual path connection or virtual channel connection, the user needs to specify the traffic characteristics in both directions for that connection. The network establishes such a connection only if sufficient network resources are available to establish the end–to–end connection with the required quality of service. The agreed quality of service for any of the existing channels must not be affected by the new connection.

Usage Parameter Control and Network Parameter Control

After a connection is accepted by the Connection Admission Control function, the UPC function of network monitors the connection to check whether the traffic conforms to the traffic contract.

The main purpose of UPC/NPC is to protect the network resources from an overload on one connection that would affect the quality of service of other already established connections.

Usage Parameter Control (UPC) and Network Parameter Control (NPC) do the same job at different interfaces. The UPC function is performed at the user network interface, while the NPC function is performed at the network node interface.

Functions performed by the Usage Parameter Control include:

- Checking the validity of VPI/VCI values
- Monitoring the traffic volume entering the network from all active VP and VC connections to ensure that the agreed parameters are not violated.
- Monitoring the total volume of the accepted traffic on the access link.
- Detecting violations of assigned parameters and taking appropriate actions.

Priority Control

Priority control is an important function as its main objective is to discard lower priority cells in order to protect the performance of higher–priority cells.

Congestion Control

Congestion is a state of network wherein the network resources are overloaded. This situation indicates that the network is not able to guarantee the negotiated quality of service to established connections and to the new connection requests. ATM Congestion Control refers to the measures taken by the network to minimise the intensity, spread and duration of network congestion.

3.9 Benefits of ATM

1. As a high–bandwidth medium with low delay and the capability to be switched or routed to a specific destination, ATM provides a uniformity that meets the needs of the telephone, cable television, video, and data industries. This universal compatibility makes it possible to interconnect the networks – something that is not currently possible because of the various transmission standards used by each industry.
2. One of the key advantages of ATM is its ability to transmit video

without creating a jittery picture of losing the synchronization of the sound and picture.

3. ATM is also extremely fast and provides dynamic bandwidth for bursty traffic.

4. AT&T has developed ATM switches capable of transmitting 20 gigabits of data per second (Gbps) and a shared switch that can transmit up to 662 Gbps.

5. Telephone networks connect every telephone to every other telephone using a dedicated path, but carry narrow bandwidth signals. Cable networks carry broadband signals, but only connect subscribers to centralised locations. To build a network that would provide a dedicated connection between sender and receiver for broadband communications would be prohibitively expensive. For this reason, ATM seems to be the best hope since it can use existing networks to deliver simple voice and data as well as complex and time–sensitive television signals. ATM can also handle bi–directional communications easily.

6. Unlike packet switching, ATM is designed for high–performance multimedia networking.

3.10 ATM Applications

ATM technologies, standards, and services are being applied in a wide range of network environments.

ATM Services

Service providers globally are introducing or already offering ATM services to their business users.

ATM Work Group and Campus Networks

Enterprise users are deploying ATM campus networks based on the ATM LANE standards. Workgroup ATM is more of a niche market with the wide acceptance of switched–Ethernet desktop technologies.

ATM Enterprise Network Consolidation

A new class of products has evolved as an ATM multimedia network–consolidation vehicle. It is called an ATM Enterprise Network switch. A full–featured ATM ENS offers a broad range of in–building (e.g., voice, video. LAN, and ATM) and wide-area interfaces (e.g leased line, circuit switched, frame relay and ATM at narrowband and broadband speeds) and supports ATM switching, voice networking, frame–relay SVCs, and integrated multi-protocol routing.

Multimedia Virtual Private Networks and Managed Services

Service providers are building on their ATM networks to offer a broad range of services. Examples include managed ATM, LAN, voice and video services.

Frame–Relay Backbones

Frame–relay service providers are deploying ATM backbones to meet the rapid growth of their frame–relay services to use as a networking infrastructure for a range of data services, and to enable frame relay to ATM service interworking services.

Internet Backbones

Internet service providers are likewise deploying ATM backbones to meet the rapid growth of their frame–relay services, to use as a networking infrastructure for a range of data services, and to enable Internet class–of–service offerings and virtual private intranet services.

Residential Broadband Networks

ATM is the networking infrastructure of choice for carriers establishing residential broadband services, driven by the need for highly scalable solutions.

Carrier Infrastructure for the Telephone and Private–Line Networks

Some carriers have identified opportunities to make more effective use of their SONET/SDH fibre infrastructure by building an ATM infrastructure to carry their telephony and private–line traffic.

4.0 CONCLUSION

This unit has discussed the different types of switching techniques and how the Asynchronous Transfer Mode (ATM) – layered architecture compares with the OSI model.

Also, you have been taken through how the ATM protocol works, together with a detailed discussion on the ATM network, the ATM cell, the ATM traffic control and classes of services, benefits of ATM and its applications.

5.0 SUMMARY

- Asynchronous Transfer Mode (ATM) is a high–performance, cell–oriented switch and multiplexing technology that utilises fixed–length packets to carry different types of traffic.

- ATM is a technology defined by protocol standards created by the ITU–T, ANSI, ETSI and the ATM Forum

- ATM is asynchronous because cells are not transferred periodically. Cells are given time slots on demand.

- ATM is a technology that will enable carriers to capitalise on a number of revenue opportunities through multiple ATM classes of service; high speed Local–Area Network (LAN) interconnection; voice, video, and future multimedia applications in business markets in the short term; and in community and residential markets in the longer term.

- ATM reduces infrastructure costs through efficient bandwidth management, operational simplicity, and the consolidation of overlay networks.

6.0 TUTOR–MARKED ASSIGNMENT

i. Fill in the blanks;

1. ATM cells are ………………….with ……………header formats.
2. Asynchronous Transfer Mode (ATM) is also known as …………..
3. ATM is fundamentally a ………….. switching technology
4. Multirate circuit switching is an enhancement of …..............
5. The two main approaches to packet switching are………….and ………….

ii. What is ATM?
iii. Differentiate between Datagram and Virtual circuit.
iv. List the types of ATM connections
v. ATM switches support three kinds of interfaces. List and explain each of them.
vi. Draw and explain the structure of an ATM cell.

7.0 REFERENCES/FURTHER READING

UNIT 6 DATA TRANSMISSION AND MULTIPLEXING

1.0 INTRODUCTION

In the previous unit, the basics of computer network were discussed. This unit covers topics related to the physical layer, which will comprise the difference between data rate and bandwidth, analog and digital and finally, characteristics of different transmission media.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- types of transmission
- domain concepts – time and frequency
- difference between analog & digital signal
- type of transmission impairments
- transmission media.

3.0 MAIN CONTENT

3.1 Transmission Terminology

Data transmission occurs between transmitters and receivers over some transmission medium.

Transmission media may be classified as:

- Guided
- Unguided
- In both cases, communication is in the form of electromagnetic waves.

With guided media, the waves are guided along a physical path. Examples of guided media are twisted pair, coaxial cable, and optical fibre. Unguided media provide a means for transmitting electromagnetic waves but do not guide them; examples are propagation through air, vacuum and seawater. In this unit, we will discuss about guided media only.

A transmission may be

- Simplex
- Half–duplex
- Full duplex

In simplex transmissions, signals are transmitted in only one direction; one station is a transmitter and the other is the receiver. In the half–diplex operation, both stations may transmit but only one at a time. In full–duplex, operation, both stations may transmit simultaneously. In the latter case, the medium is carrying signals in both directions at the same time.

3.1.1 Time–Domain Concept

As a function of time, an electromagnetic signal can be either continuous or discrete. A continuous signal is one in which the signal intensity varies in smooth fashion over time. There are no breaks or discontinuities in the signal. A discrete signal is one in which the signal intensity maintains a constant level for some period of time and then changes to another constant level.

3.1.2 Frequency Domain Concepts

In practice, an electromagnetic signal will be made up of many frequencies.
It can be shown, using a discipline known as Fourier analysis, that any signal is made up of components at various frequencies, in which each
component is sinusoidal.

So, we can say that for each signal, there is a time–domain function (t) that specifies the amplitude of the signal at each instance of time. Similarly, there is a frequency–domain function S(t) that specifies the constituent frequency of the signal. The spectrum of the signal is the range of frequencies that it contains.

3.1.3 Relationship between Data Rate and Bandwidth

The concept of effective bandwidth is somewhat a fuzzy one. It is the band within which most of the energy is confined. The term “most” in this context is somewhat arbitrary. The important issue here is that, although a given waveform may contain frequencies over a very broad range, as a practical matter, any transmission medium that is used will be able to accommodate only a limited band of frequencies. This, in turn, limits the data rate that can be carried on the transmission.

3.2 Analog and Digital Data Transmission

The terms ‘analog’ and ‘digital’ correspond, roughly, to continuous and discrete, respectively. These two terms are used frequently in data communications at least in three contexts:

- Data
- Signaling
- Transmission

3.2.1 Data

Analog signal takes on continuous values on some interval. For example, voice and video are continuously varying patterns of intensity. Most data collected by sensors, such as temperature and pressure, take on continuous values. Digital data take on discrete values; examples are text and integers.

3.2.2 Signals

In a communication system, data are propagated from one point to another by means of electrical signals. An analog signal is a continuously varying electromagnetic wave that may be propagated over a variety of media, depending on spectrum.

A digital signal is a sequence of voltage pulses that may be transmitted over a wire medium; for example, a constant positive voltage level may represent binary 1, and a constant negative voltage level may represent binary 0.

3.2.3 Transmissions

Both analog and digital signals may be transmitted on suitable transmission media. Analog transmission is a means of transmitting analog signal without regard to their context.

	Analog signal	**Digital signal**
Analog data	• Signal occupies the same spectrum as the analog data. • Analog data are encoded to occupy different portions of spectrum.	• Analog data are encoded using a codec to produce a digital bit stream
Digital Data	• Digital data are encoded using a modem to produce Analog signal	• Signal consists of two voltage levels to represent the two binary values • Digital data are encoded to produce a digital signal with desired properties.

3.3 Transmission Media

The purpose of the physical layer is to transport a raw bit stream from one machine to another. Various physical media can be used for the actual transmission. Each one has its own niche in terms of bandwidth, delay, cost, and ease of installation and maintenance. Media are roughly grouped into guided media, such as copper wire and fibre optics, and unguided media, such as radio and lasers through the air. We will look at these in this section and next one.

3.3.1 Twisted Pair

Although the bandwidth characteristic of magnetic tape is excellent, the delay characteristics are poor. Transmission time is measured in minutes or hours, not milliseconds. For many applications, an on–line connection is needed. The oldest and still most common transmission medium is twisted pair. A twisted pair consists of two insulated copper wires, typically about 1mm thick. The wires are twisted together in a helical form, just like a DNA molecule. The purpose of twisting the wires is to reduce electrical interference from similar pairs close by. The common application of the twisted pair is the telephone systems.

Twisted pairs can be used for either analog or digital transmission. The bandwidth depends on the thickness of the wire and the distance travelled, but several megabytes/sec can be achieved for a few kilometres in many cases. Due to their adequate performance and low cost, twisted pairs are widely used and are likely to remain so for years to come.

Twisted pair cabling comes in several varieties, two of which are important for computer networks Category 3 twisted pairs consist of two insulated wires gently twisted together. Four such pairs are typically grouped together in a plastic sheath for protection and to keep the eight wires together.

Starting around 1988, the more advanced category 5 twisted pairs were introduced. They are similar to Category 3 pairs, but with more twists per centimetres and insulation, which result in less cross talk and a better quality signal over longer distances, making them more suitable for high–speed computer communication. Both of these wiring types are often referred to as UTP (Unshielded Twisted Pair, to contrast them with the bulky, expensive, shielded twisted pair cables IBM introduced in the early 1980s, but which have not proven popular outside of IBM installations.

3.3.2 Baseband Coaxial Cable

Another communication transmission medium is the coaxial cable. It has better shielding than twisted pairs, so it can span longer distances at higher speeds. Two kinds of coaxial cable are widely used. One kind,
50–ohm cable is commonly used for digital transmission and is the subject of this section. The other kind, 75–ohm cable, is commonly used for analog transmission and will be described in the next section. This distinction is based on historical, rather than technical factor, (e.g., early dipole antennas had an impedance of 300 ohms, and it was easy to build
4:1 impedance matching transformers).

A coaxial cable consists of a stiff copper wire as the core, surrounded by an insulating material. The construction and shielding of the coaxial cable give it a good combination of high bandwidth and excellent noise immunity. The bandwidth possible depends on the cable length**.** For 1 km cables, a data rate of 1 or 2 Gbps is feasible. Longer cables can also be used, to be widely used within the telephone systems but have now largely been replaced by fibre optics on long–haul routes. In the United States alone, 1000 km of fibre is installed every day (counting a 100 km bundle with 10 strands of fibre as 1000 km). Sprint is already 100 per cent fibre, and the other major carriers are rapidly approaching that. Coax is still widely used for cable television and some local area networks.

3.3.3 Broadband Coaxial Cable

The other kind of coaxial cable system uses analog transmission on standard cable television cabling. It is cabled broadband. Although the term “broadband” comes from the telephone world, where it refers to anything wider than 4kHz, in the computer networking world, “broadband cable” means any cable network using analog transmission.

Since broadband networks use standard cable television technology, the cables can be used up to 300 MHz (and up to 450 MHz) and can run for nearly 100 km due to the analog signaling, which is much less critical than digital signaling. To transmit digital signals on an analog network, each interface must contain electronics to convert the outgoing bit stream to an analog signal, and the incoming analog signal to a bit stream. Depending on the type of these electronics, 1 bps may occupy roughly 1 Hz of bandwidth. At higher frequencies, many bits per Hz are possible using advanced modulation techniques.

Broadband systems are divided up into multiple channels frequently, the 6MHz channels used for television broadcasting. Each channel can be used for analog television, CD–quality audio or a digital bit stream at, say, 3 Mbps, independent of the others. Television and data can be mixed on one cable.

3.4 Multiplexing

In communication, multiplexing is a technique that transmits signals from several sources over a single communication channel. So in order to minimize the cost of communication bearer, various means of sharing a communication channel between several users, have been devised; these are known as multiplexing. In this section, we will discuss about two multiplexing techniques: FDM & TDM.

Frequency Division Multiplexing (FDM)

In FDM, the frequency spectrum is divided among the logical channels with each user having exclusive possession of some frequency band.

Time Division Multiplexing (TDM)

In TDM, the users take turns (in a round robin), each one is periodically getting the entire bandwidth for a little burst of time. Television broadcasting provides an example of both kinds of multiplexing. Each TV channel operates in a different frequency range, which is a portion of the allocated spectrum, with the inter–channel separation great enough to prevent interference. This system is an example of FDM. During the transmission of any program (Serial/film), there is an advertisement as well. These two alternate in time on the same frequency. This is an example of TDM.

4.0 CONCLUSION

This unit covered topics relating to the physical layer, which comprises the difference between data rate and bandwidth, analog and digital transmission and characteristics of transmission media except wireless transmission.

Also, the unit has taken you through the types of transmission impairments and how to minimize them.

5.0 SUMMARY

Transmissions can take place through the media of guided and unguided type and it can be simplex, half–duplex and full–duplex. In simplex, the data/signals are transmitted in one direction by a station i.e., by the sender; in half–duplex, the transmission can be done in one direction at a time whereas in full–duplex, the transmission can take place in directions. The concept of time domain and frequency domain deals with the electromagnetic signals and components at various frequencies spectrum. The concept of analog and digital transmission deals with data signaling and transmission which can be analog data i.e., signal occupies same spectrum and digital data are encoded using a modem to produce analog signal. The other type of signal is digital, which uses a bit stream.

Media used in transmission are of the magnetic type and it is one of the most common ways to store data physically on tapes, floppy disks and hard disks. Twisted pairs are used both for analog as well as digital transmission. Twisted pair can be Cat 3, or Cat 5. Both of them are UTP cables. Baseband cable is used for longer distances at high–speed 50 ohm and 75 ohm are normally used. Broadband Coaxial cable refers to anything wider than 4KHz. Broadband is divided into multiple channels and each channel can be used for analog signal also. It is used for CD– quality audio or a bit stream.

6.0 TUTOR MARKED ASSIGNMENT

i. What is the difference between data rate and bandwidth?
…………………………………………………………………………
…………………………………………………………………………
………………………………………………………………………….
………………………………………………………………………….

ii. List the characteristics of broadband coaxial cable
……………………………………………………………………………
……………………………………………………………………………..
……………………………………………………………………………..
………………………………………………………………………………

7.0 REFERENCES/FURTHER READING

UNIT 7 MEDIUM ACCESS CONTROL AND DATA LINK LAYER

1.0 INTRODUCTION

This unit introduces the design of Data Link Layer and its Medium Access Control Sublayer. This includes various protocols for achieving reliable, efficient communication. It also covers the study of nature of errors, causes and how they can be detected and corrected.

The MAC sublayer contains protocols which determine who goes next on a multi access channel.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- understand various error–handling methods.
- explain the various flow control methods.
- Identify MAC sublayer protocols like CSMA/CD, Token Passing

3.0 MAIN CONTENT

3.1 Data Link Layer

To exchange digital information between devices A and B, we require an interconnecting transmission medium to carry the electrical signals; a

standard interface and the physical layer to convert bits into electrical signals and vice–versa.
This has certain limitations:

- If the electrical signal gets impaired due to the encountered interference with other signals or electromagnetic waves from external sources, error may be introduced in the data bits.

- Errors can also be introduced if the receiving device is not ready for the incoming signal, hence resulting in the loss of some information.

The Data Link Layer constitutes the second layer of the hierarchical OSI Model. The Data Link Layer together with Physical Layer and the interconnectivity medium provide a data link connection for reliable transfer of data bits over an imperfect physical connection.

It accomplishes the task by having the sender break the input data up into data frames, transmit the frames sequentially and process the acknowledgement frames sent back by the receiver. It is up to the Data Link Layer to create and recognise frame boundaries.

Another issue that arises in Data Link Layer is how to keep a fast transmitter from overflowing a slow receiver in data.

The Data Link Layer incorporates certain processes, which carry out error control, flow control and the associated link management functions. The data block along with the control bits is called a frame.

Data Link Layer is divided into two sublayers:

Logical Link Control (LLC) concerned with providing a reliable communication path between two devices. It is also involved with flow control and sequencing of class. The LLC is non–architecture–specific which is the same for all IEEE–defined LANs.

Medium Access Control focuses on methods of sharing a single transmission medium.

3.1.1 Services Provided By Data Link Layer (Logical Link Control)

- **Framing:** Some control bits are added to the data packets from network layer to mark the start and end of a frame. This is done using character count, character or bit stuffing.

- **Flow Control:** Flow Control deals with how to keep the fast sender from overflowing a slow receiver by buffer at the receiver sides and acknowledgement.

- **Error Detection and Correction Codes:** Various methods used for error detection and corrections are: parity bit, cyclic redundancy check, checksum, hamming code, etc.

3.1.2 Retransmission Strategies.

In this section, we will discuss several retransmission strategies, which are also considered as a flow control and error control mechanism.

Stop and Wait

The sender allows one message to be transmitted, checked for errors and an appropriate ACK (Positive Acknowledgement) or NAK (Negative Acknowledgement) returned to the sending station. No other data messages can be transmitted until the receiving station sends back a reply, thus the same STOP & WAIT is derived from the originating station sending a message, stopping further transmission and waiting for a reply.

Its major drawback is the idle line time that results when the stations are in the waiting period. If the ACK is lost, then the master station retransmits the same message to the receiver side. The redundant transmission could possibly create a duplicate frame. A typical approach to solving this problem is the provision for a sequence number in the header of the message. The receiver can then check for the sequence number to determine if the message is a duplicate. The Stop and Wait mechanism requires a very small sequence number, since only one message is outstanding at any time. The sending and the receiving station only use a one bit alternating sequence of 0 and 1 to maintain the relationship of the transmitted message and its ACK/NAK status.

Sliding Window

The data control signals flow from sender to receiver in a more continuous manner and several messages can be outstanding at any one time.

The transmitting station maintains a sending window that maintains the number of frames it is permitted to send to the receiving station and the receiving station also maintains a receiving window that performs complementary functions. The two sides use the window to coordinate the flow of frames between each other. The window wrap around is used to reuse the same set of numbers for different frames. There are sliding window techniques:

(1) Go Back N
(2) Selective Repeat

The following two diagrams (Figure 1 and Figure 2) explain the function of Go Back N and Selective Repeat respectively.

GO Back N

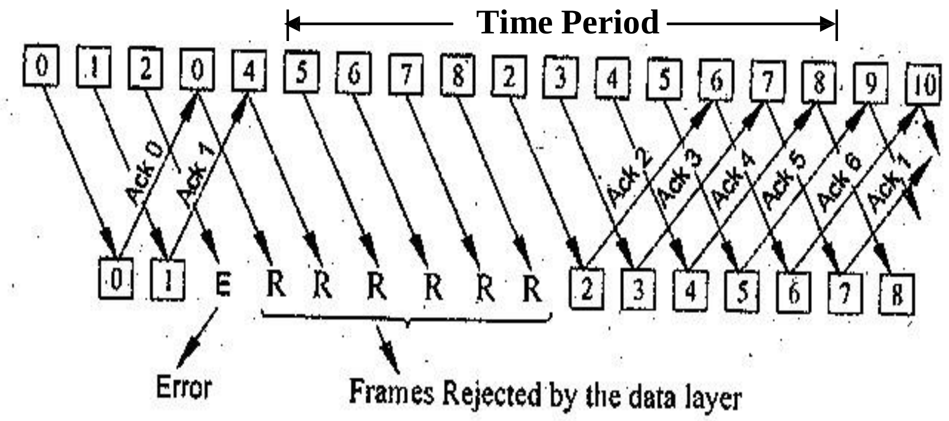


Figure 1: Go Back N

This is a sliding window technique. It allows data and control messages to be transmitted continuously without waiting for its acknowledgement from the receiver. In the event, if an error is detected at the receiving side, the erroneous message is retransmitted, as well as all other frames that were transmitted after the erroneous message.

Selective Repeat

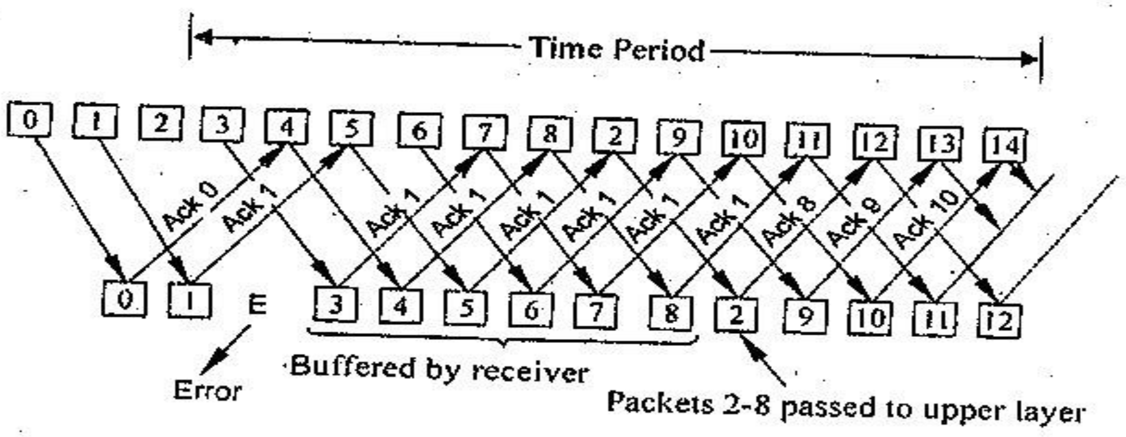


Figure 2: Selective Repeat

This method provides for a more refined approach. In contrast to the Go Back N, the only messages retransmitted are those for which negative acknowledgement is received.

Studies reveal that the selective repeat mechanism obtains greater throughout than the Go Back N. Selective Repeat mechanism requires additional logic to maintain the sequence of the recent message and merge it into the proper place as they queue at the proper site.

3.2 Medium Access Control Sublayer

In any broadcast network, the key issue is how to determine who gets to use the channel when there is competition for it. The protocols used to determine who goes next on a multi–access channel belong to a sub–layer of a Data Link Layer called MAC sublayer.

3.2.1 Contention Based Media Access Protocols

Contention is what happens at a staff meeting when several people start to talk at the same time. In contention protocol, no policeman controls usage of the communication channel.

All workstations on a contention Network share a common transmission channel. Messages are broadcast on that channel and may be overheard by all attached workstations. A workstation responds only to a message with its address: Message intended for different modes are ignored.

Messages to be transmitted are converted to packets and are sent when ready, without verifying the availability of the channel. When transmission of a channel overlap with that of another, collision occurs. Colliding packets with their messages are destroyed.

3.2.2 Polling–Based MAC Protocols

Polling involves the channel control of all workstations in a network. The primary workstation acts like a teacher going down the rows of the classroom asking each student for homework. When one student has answered, the next is given a chance to respond.

A polling network contains two classes of workstations, the primary workstation and the multiple secondary workstations connected to it. A buffer that can temporarily store messages is associated with each secondary workstation. When a workstation has information to transmit, the data is passed to the buffer. The frames are held until the central controller polls the workstation.

These are the two possibilities for the path of a message from source to destination workstation:

- All messages may require passing to the central workstation, which route them to their destination.

- Messages may be sent directly.

Polling technique can be said to maintain a tight control over the network than do contention–based protocols.

Token Passing

The network continuously circulates a special bit pattern known as a token, among all the modes in the network.

Each token contains network information, comprising a header, a data field and a trailer. Any mode willing to send a frame has to grass a token first. Now, let us talk about some standards.

3.2.3 IEEE Standard 802.3 and Ethernet

It is for CSMA/CD LAN. When a station wants to transmit, it listens to the cable. If the cable is busy, the station waits until it goes idle, otherwise, it transmits immediately. If two or more stations simultaneously begin transmitting on an idle cable, they will collide. All colliding stations then terminate their transmissions, wait a random time and repeat the whole process all over again.

3.2.4 IEEE Standard 802.4 Token Bus

Token Bus combines features of Ethernet and token ring (discussed in the next section). It combines the physical configuration of Ethernet (bus topology) and collision–free (predictable delay) feature of token ring. Token bus is a physical bus that operates as logical ring using tokens.

It is a linear cable onto which the stations are attached. When the logical ring is initialized, the highest numbered station may send the first frame after it is done, it passes permission to its immediate neighbour by sending the neighbour a special control frame called a token.

3.2.5 IEEE Standard 802.5 Token Ring

In a token ring, the token circulates around the ring whenever all stations are idle. When a station wants to transmit a frame, it is required to seize the token and remove it from the ring before transmitting. This action is done by inverting a single bit in the 3–byte token which instantly changes it into the first 3 bytes of a normal data frame. Because there is only one token, only one station can transmit at a given instant, thus solving the channel access problem.

4.0 CONCLUSION

In this unit, you have been introduced to a number of protocols such as IEEE Standard 802.3 and Ethernet, IEEE Standard 802.4 Token Bus, IEEE Standard 802.5 Token Ring; etc. for MAC sub–layer, which is a part of data link layer.

Also, the unit has discussed issues relating to error handling methods.

5.0 SUMMARY

In this unit, an introduction to issues of DLL and various methods for allocation of a common channel to the competing users were discussed. LANS are dominated by four types of architecture: Ethernet, Token Bus, Token Ring and Fibre Distributed Data Interface (FDDI). Each of them has its own advantages and disadvantages. Depending upon the requirements, the choice is made.

6.0 TUTOR MARKED ASSIGNMENT

i. How is Selective Repeat better than Go Back N?
…………………………………………………………………
…………………………………………………………………
…………………………………………………………………
…………………………………………………………………

ii. What are the four LAN architectures?
…………………………………………………………………
…………………………………………………………………
…………………………………………………………………
…………………………………………………………………
…………………………………………………………………
……………………………………………………………….

7.0 REFERENCES/FURTHER READING

Module 3: Network Administration

Unit1: Network Administration: Scope, Goals, Philosophy and Standards

1.0 INTRODUCTION

Network and distribution processing systems are of critical and growing importance in business, government and other organizations. Therefore, networks must be managed for effectiveness and efficiency. This unit discusses fundamental aspects of network administration.

2.0 OBJECTIVES

At the end of this unit, you should be

also to: Define network

administration

Know the scope of network administration

State the goals of system administration

Understand the challenges of system administration

State the Meta principles of system administration

3.0 MAIN CONTENT

3.1 What is network and system administration

Network and system administration is a branch of *engineering* that concerns the operational management of human–computer systems. It is about putting together a network of computers (workstations, PCs and supercomputers), getting them running and then *keeping* them running in spite of the activities of *users* who tend to cause the systems to fail.

A system administrator works for users, so that they can use the system to produce work. However, a system administrator should not just cater for one or two selfish needs, but also work for the benefit of a whole community. Today, that community is a global community of machines and organizations, which spans every niche of human society and culture, thanks to the Internet. It is often a difficult balancing act to determine the best policy, which accounts for the different needs of everyone with a stake in a system. Once a computer is attached to the Internet, we have to consider the consequences of being directly connected to all the other computers in the world.

In the future, improvements in technology might render system administration a somewhat easier task – one of pure resource administration – but, today, system administration is not just an administrative job, it is an extremely demanding engineer's job. It's about hardware, software, user support, diagnosis, repair and prevention. System administrators need to know a bit of everything: the skills are technical, administrative and socio-psychological.

The terms *network administration* and *system administration* exist separately and are used both variously and inconsistently by industry and by academics.

System administration is the term used traditionally by mainframe and Unix engineers to describe the management of computers whether they are coupled by a network or not. To this community, network administration means the management of network infrastructure devices (routers and switches). The world of personal computers (PCs) has no tradition of managing individual computers and their subsystems, and thus does not speak of system administration, *per se*. To this community, network administration is the management of PCs in a network. In this material, we shall take the first view, since this is more precise.

Network and system administration are increasingly challenging. The complexity of computer systems is increasing all the time. Even a single PC today, running Windows NT, and attached to a network, approaches the level of complexity that mainframe computers had ten years ago. We are now forced to think *systems* not just computers.

3.2 Scope of Network administration

The management of a network, usually called network administration, consists of procedures and services that keep the network running properly. An important part of network management entails making sure that the network is available (or up and running as IT professionals say) when employees and managers need it. Other admin activities are:

- Monitoring the network capacity to ensure that all transmission requirements can be met.
- Adding capacity to the network by increasing band width interconnecting additional modes, or creating and interconnecting additional networks.
- Training people to use the network effectively
- Assisting IT professionals in organizational applications that will make good use of the network's capabilities.

Backing up the network software and data regularly to protect against the failure of network or any of its components

- Putting security procedures in place to make certain that only authorized users have access to the network and ensuring that all security procedures are followed
- Making sure the network personnel can respond quickly and effectively in the event of a network operational or security failure.
- Diagnosing and troubleshooting problems on the network and determining the best course of action to take to solve them.

3.3 The goal of Network administration

The goal is to keep the network running properly and configuring and managing services that are provided over the network.

There are many services that we use regularly. There are some which work in the background enabling other services to run smoothly.

3.4 The challenges of system administration

System administration is not just about installing operating systems. It is about planning and designing an efficient *community* of computers so that real *users* will be able to get their jobs done. That means:

- Designing a network which is logical and efficient.
- Deploying large numbers of machines which can be easily upgraded later.
- Deciding what services are needed.
- Planning and implementing adequate security.
- Providing a comfortable environment for users.
- Developing ways of fixing errors and problems which occur.
- Keeping track of and understanding how to use the enormous amount of knowledge which increases every year.

Some system administrators are responsible for both the hardware of the network and the computers which it connects, i.e. the cables as well as the computers. Some are only responsible for the computers. Either way, an understanding of how data flow from machine to machine is essential as well as an understanding of how each machine affects every other.

3.5 The Meta principles of system administration

Many of the principles in this course material derive from a single overriding issue: they address the *predictability* of a system. The term system clearly implies an operation that is *systematic*, or predictable – but, unlike simple mechanical systems, like say a clock, computers interact with humans in a complex cycle of feedback, where uncertainty can enter at many levels. That makes human–computer systems difficult to predict, unless we somehow fix the boundaries of what is allowed, as a matter of policy.

Principle (Policy is the foundation). *System administration begins with a policy – a decision about what we want and what should be, in relation to what we can afford.*

Policy speaks of what we wish to accomplish with the system, and what we are willing to tolerate of behavior within it. It must refer to both the component parts and to the environment with which the system interacts.

If we cannot secure predictability, then we cannot expect long-term conformance with a policy.

Principle (Predictability). *The highest level aim in system administration is to work towards a predictable system. Predictability has limits. It is the basis of reliability, hence trust and therefore security.*

Policy and predictability are intertwined. What makes system administration difficult is that it involves a kind of 'search' problem. It is the hunt for a stable region in the landscape of all policies, i.e. those policies that can lead to stable and predictable behavior. In choosing policy, one might easily promote a regime of cascading failure, of increasing unpredictability that degenerates into chaos. Avoiding these regimes is what makes system administration difficult.

As networks of computers and people grow, their interactions become increasingly complex and they become *non-deterministic*, i.e. not predictable in terms of any manageable number of variables. We therefore face another challenge that is posed by inevitable growth:

Principle (Scalability). *Scalable systems are those that grow in accordance with policy; i.e. they continue to function predictably, even as they increase in size.*

These meta-themes will recur throughout this material. The important point to understand about predictability is that it has limits. Human–computer systems are too complex and have too many interactions and dependencies to be deterministic. When we speak of predictability, it must always be within a margin of error. If this were not the case, system administration would not be difficult.

3.6 Advice to the students

To study this subject, we need to cultivate a way of thinking which embodies a basic scientific humility and some core principles:

- Independence or self-sufficiency in learning. We cannot always ask someone for the right answer to every question.

- Systematic and tidy work practices.
- An altruistic view of the system. Users come first: collectively and only then individually.
- Balancing a fatalistic view (the inevitability of errors) with a determination to gain firmer control of the system.

Some counter-productive practices could be avoided:

- The belief that there exists a right answer to every problem.
- Getting fraught and upset when things do not work the way we expect.
- Expecting that every problem has a beginning, middle and an end (some problems are chronic and cannot be solved without impractical restructuring).

We can begin with a checklist:

- Look for answers in manuals and newsgroups.
- Use controlled trial and error to locate problems.
- Consider all the information; listen to people who tell you that there is a problem. It might be true, even if you can’t see it yourself.
- Write down experiences in an A–Z so that you learn how to solve the same problem again in the future.
- Take responsibility for your actions. Be prepared for accidents. They are going to happen and they will be your fault. You will have to fix them.
- Remember tedious jobs like vacuum cleaning the hardware once a year.
- After learning about something new, always pose the question: *how does this apply to me?*

SELF-ASSESSMENT EXERCISES

1. Is system administration management or engineering?

2. Why does the physical environment play a role in system administration?

4.0 CONCLUSION

Network administration is concerned with establishing and administering overall goals, policies and procedures of network management. This requires a well rounded skills set and not just technical skills.

5.0 SUMMARY

Network and system administration is a branch of *engineering* that concerns the operational management of human–computer systems. System administration is not just about installing operating systems. It is about planning and designing an efficient *community* of computers so that real *users* will be able to get their jobs done. *System administration begins with a policy – a decision about what we want and what should be, in relation to what we can afford.* Policy speaks of what we wish to accomplish with the system, and what we are willing to tolerate of behavior within it. To study this subject, we need to cultivate a way of thinking which embodies a basic scientific humility and some core principles:

6.0 TUTOR-MARKED ASSIGNMENTS

1. State the top-most principles that guide network and system administrators

2. What kinds of issues does system administration cover?

3. State the meta principles of system administration

4. What are the challenges of system administration?

7.0 REFERENCES/FURTHER READING

1. Burgess, M. (2004). Principles of Network and System Administration. (2nd Ed.). Chichester, West Sussex , England: Wiley.

2. Forouzan, B.A, & Fegan, S.C. (2007). Data communications and Networking (4th Ed). Mc Graw Hill.

3. Limoncelli, T. A.,Hogan, C. J. & Chalup, S. R (2007}. The Practice of System and Network Administration. (2nd Ed.). Upper Saddle River, NJ: Addison-Wesley.

4. Stallings, W. (2009). Data and computer communications (8th ed.). Upper saddle River, NJ.: Pearson Education Inc.

UNIT 2: NETWORK PROTOCOL

1.0 INTRODUCTION

This unit discusses packets and protocols which are the fundamental building blocks of data transmission over the network.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- state why are protocols and standards needed?
- state the function of a packet
- state the principal function of protocols in a network
- explain the layered approach to computer communications
- discuss some of the different protocols and their uses
- state how protocols work
- state the advantages and disadvantages of standards.

3.0 MAIN CONTENT

3.1 PROTOCOLS, PACKETS AND STANDARDS

All data that is transmitted across the network is put into packets containing information about the source and destination of the data. These packets are created using standards or protocols. Since there are many different network configurations, there are many different protocols. By having a variety of protocols, you can choose the one that best fulfills the needs of your network.

3.1.1 Function of packets

The function of a packet is to carry data from one point to another. Protocols require that packet contain some basic information about their source and their destination and in many cases, protocols require that the packet include a *checksum*. A checksum is a number that can be used to verify that the packet has been transferred across the network without being corrupted.

3.1.2 Packet Structure

The structure of the packet is extremely important. Useless a packet is structured exactly as it is supposed to be; it is ignored by the receiving party and assumed to be corrupted. Basic packet structure requires that the packet include a *header* section, a *data* section, and in most cases, a *cyclic redundancy check* (CRC) section (also called a *trailer*). Not every protocol requires that a CRC be attached.

3.1.2.1 Header

The header section of a packet contains the routing information. This information includes the source and destination of the packet. The header also contains the number of the packet, which is generated when the packet is created. In addition, the header can contain a protocol version number, the length of the header, the type of service, the length of the entire packet, the flags, the time to live, and other information.

3.1.2.2 Data

The data is the actual information that is being transmitted over the network from one application to another. Each protocol has a predefined maximum data size. If the data is larger than this maximum data size, the data is broken into smaller pieces and transmitted in multiple packets.

3.1.2.3 CRC

A CRC (Cyclic redundancy check) is calculated prior to the data being sent and attached to the bottom of a packet. At the destination, a new CRC is computed and compared to the original to verify that the packet was not corrupted. A CRC is usually attached to the bottom of a packet, but some protocols include CRC within the header.

3.1.3 Creating packets

Before data is transmitted across the network, it is broken into smaller, more manageable pieces called *packets*. All packets are numbered so they can be put back together when they reach their destination. The header, which contains the source address, destination address, and packet number, along with other information, is attached to the beginning of the packet. A CRC is then calculated and added to the end of the packet.

3.1.4 Encapsulation

Encapsulation is the process of encoding data for transmitting it across the network. Once a packet is created as described previously, in order for the packet to be transmitted to its final destination, it may need to use a protocol in addition to the one that it is currently using. A header and CRC are then added to the newly created packet. This packet is an encapsulated packet. Figure 1 illustrates an encapsulated packet.

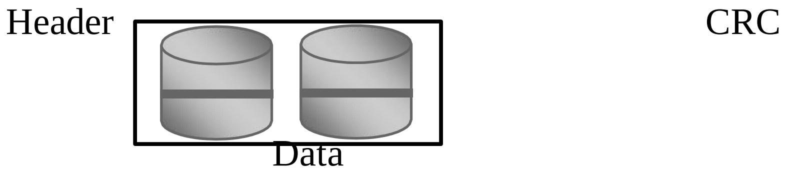


Figure 1 Encapsulated packet

3.2 PROTOCOLS

In computer networking, communication occurs between entities in different systems. An entity is anything capable of sending or receiving information. However, two entities cannot simply send bit streams to each other and expect to be understood. For communication to occur, the entities must agree on a protocol. A protocol is a set of nodes that govern data communications. A protocol defines what is communicated, how it is communicated, and when it is communicated. The key elements of protocol are syntax, semantics and timing.

Syntax: The term syntax refers to the structure or format of the data, meaning the order in which they are presented. For example, a simple protocol might expect the first 8 bits of data to be the address of the sender, the second 8 bits to be the address of the receiver, and the rest of the stream to be message itself.

SEMANTICS:

The word semantics refers to the meaning of each section of bits. How is a particular pattern to be interpreted, and what action is to be taken based on that interpretation? For example, does the address identify the route to be taken or the final destination of the message?

TIMING:

The term timing refers to two characteristics: When data should be sent and how fast they can be sent. For example, if a sender produces data of 100mbps and the receiver can process data
at only 1 mbps, the transmission will overload the receiver and some data will be lost.

3.2.1 How Protocols Work

When an application first attempts to transfer data across the network, the data is usually t 24 large to be sent in a single transmission. To meet the need of networking, the protocol th governs the transmission of the data first breaks the data into packets. The protocol numbers
each of the packets so can later be put back together when they arrive at their destination and transmits each of the packets across the network. In addition to this

numbering, information on the source, destination, and the protocol used is added to the header of the packet.

Protocols are the first software layer to receive data that has been transmitted across the network. After all packets have been received, they are put back together using the numbers that were placed in the header at the origin of the packet. Once the data has all been put back together, it can be used by the application that the data was sent across the network to.

3.2.2 Functions of Protocols

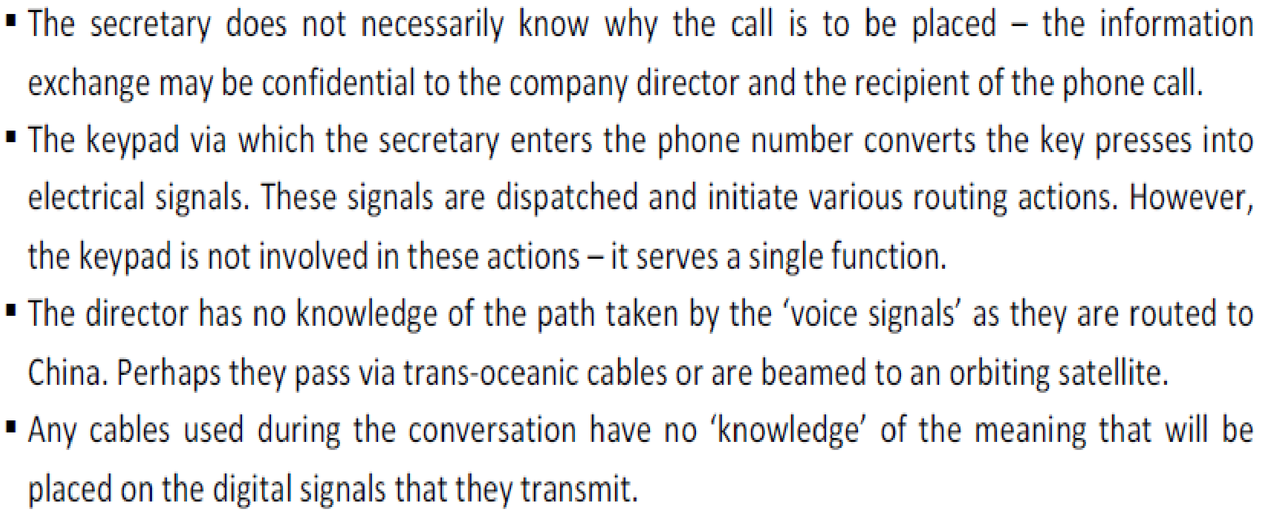
The principal functions of protocols in a network are line access and collision avoidance. Line access concerns how the sending device gains access the network to send a message. Collision avoidance refers to managing message transmission so that the messages do not collide with each other on the network. Other functions of protocols are to identify each device in the communication path, to secure the attention of the other device, to verify correct receipt of the transmitted message, to verify that a message requires transmission because it cannot be correctly interpreted and to perform recovery when errors occur.

3.3 The layered approach to computer communications

In order to enable two or more computers to communicate in a meaningful manner, we must define with great care all aspects of the communication process (i.e. we must define a ‘communications protocol’). By way of a useful analogy, let us consider the situation in which the director of a company in the UK wishes to communicate with a person in another company located in China. The director may ask a secretary to put a call through and will provide sufficient information for the secretary to identify the person who is to be contacted. Here, the director will not give the actual phone number- it may be left to the secretary to obtain this information. From this point, the director has no further involvement until the phone connection is in place. The secretary will locate and dial the number and this will initiate various electronic/software activities. Neither the director nor the secretary has any interest in knowing how the electronic and software systems will route the call. It may be carried by electronic cables, fiber optic cables, or be routed via a satellite. Additionally, it may use communications systems that route the call across the Atlantic through the US and then across the Pacific Ocean, or it may be routed in an easterly direction. These low-level issues are of little interest to the secretary – a number is dialed and processes occur that result in a phone ringing in an office somewhere in China. Hopefully, the intended recipient is available and the secretary notifies the director. Both parties must now adopt/agree on a common language and must exercise a degree of hand-shaking (in this sense we mean that only one person should talk

at any one time). Finally, at the end of the conversation, an acceptable convention is used to bring the call to a conclusion. All these issues form part of the 'communications protocol' that is needed to enable a useful dialogue and it is important to note that the elements that underpin

the communication do not need to have any knowledge of the overall purpose that they will serve. For example:



- The secretary does not necessarily know why the call is to be placed – the information exchange may be confidential to the company director and the recipient of the phone call.
- The keypad via which the secretary enters the phone number converts the key presses into electrical signals. These signals are dispatched and initiate various routing actions. However, the keypad is not involved in these actions – it serves a single function.
- The director has no knowledge of the path taken by the 'voice signals' as they are routed to China. Perhaps they pass via trans-oceanic cables or are beamed to an orbiting satellite.
- Any cables used during the conversation have no 'knowledge' of the meaning that will be placed on the digital signals that they transmit.

The establishment of a communications protocol that enables computers (and other digital systems) to communicate is, in many ways, similar to the protocols used to support the sort of phone conversation referred to in the above analogy (although computer communications are perhaps more complex). To handle the design implementation and maintenance of such systems, a 'layered' approach is adopted. In figure 2, we indicate two computers that need to communicate. Perhaps, for example, an applications program running on Node A wishes to send a data file to a similar program running on Node B (just as in the same way the company director mentioned above wishes to talk to a person in a remote location). In order to transmit the data a number of tasks must be performed, and these are carried out by layers of software located on both nodes.

Each layer carries out a number of specific tasks and directly communicates with the immediately adjacent software layers. However, from a logical point of view each layer communicates with a corresponding layer on the remote computer – i.e. corresponding software layers located on the two nodes have similar/equivalent functionality. The lowest layer on either node is responsible for interfacing with the physical interconnect.

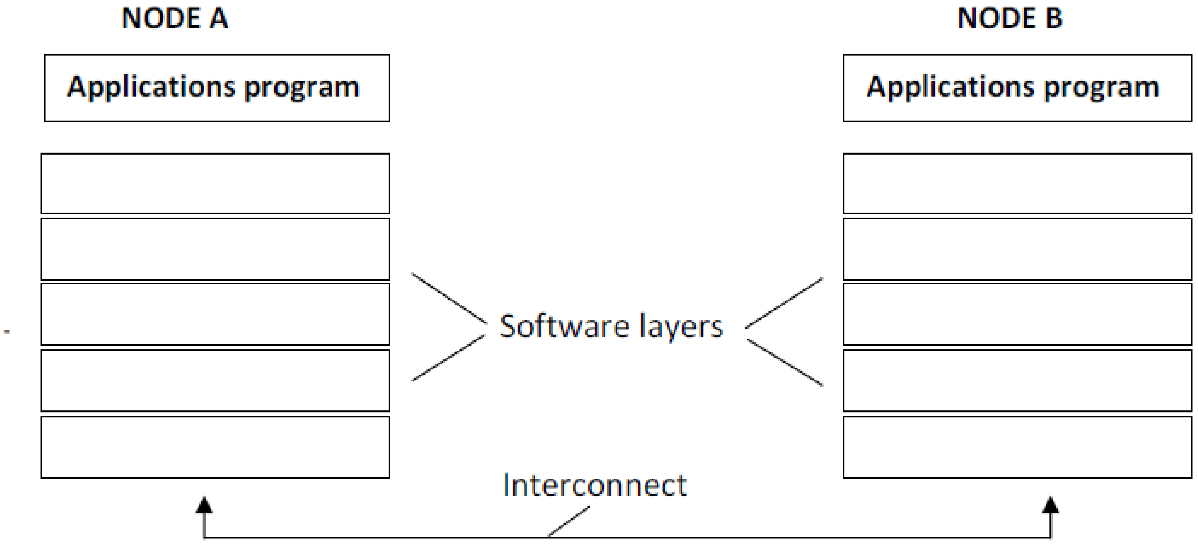


Figure 2: An applications program on Node A wishes to send data to an equivalent program on Node B

In order for Node A to transmit a data file to Node B, various events

must take place. For example:

Node A must inform the communications network of the identity of the destination system (Node B)

Node A must ensure that Node B is prepared to receive the data.

The file transfer applications program on Node A must ensure that the file management program on the destination system is ready to accept and store the file.

If the file formats used on the two systems are incompatible, one or other system must perform a format translation function.

File transfer must be carried out in an orderly manner and in such a way that the two communicating machines do not block other network traffic. This will involve splitting the data file into packets (chunks) and appending various information to each packet.

Node B provides acknowledgement of receipt

Node B reassembles the packet in other to reconstruct the original data file

Node B must attempt to detect any errors in data it has received. In some cases Node B may be able to correct errors.

In the case that secure transmission is required, the data may be encrypted by Node A prior to transmission. Node B must then perform the reverse process.

To achieve this high degree of cooperation between computers, the tasks are broken into subtasks that are implemented individually using a layered approach. These layers form the data communication protocol architecture. Example of such layer architectures are: the Open System Interconnection (OSI) model, and the Transmission Control Protocol/ Internet Protocol (TCP/IP). Key advantages of a layered structure include:

The complex communication protocol is divided into subtasks and these are implemented within a layered structure. Each layer has limited functionality and this ‘divide and conquer’ approach facilitates the design and the implementation of the system.

Higher-level layers need have no knowledge of tasks performed by the lower layers. Thus, for example, a higher-level layer needs no knowledge of the type of interconnect that is in use. Again, this facilitates the design process.

When changes are made to the communications protocol, only certain relevant layers need to be modified/replaced. This makes it easier to upgrade software and undertake software testing.

Structuring software using a layered approach tends to result in *larger programs* which run *more slowly* than if a non-layered approach were to be adopted. However, these two weaknesses are outweighed by the benefits that are associated with the layered approach - especially in terms of providing a structured framework within which the complex issues associated with computer communications may be resolved.

3.4 Standards

Are essential in creating and maintaining an open and competitive market for equipment manufacturers and in guaranteeing national and international interoperability of data and telecommunications technology and processes. Standards provide guidelines to manufacturers, vendors, government agencies and other service providers to ensure the kind of interconnectivity necessary in today's market place and in international communication.

Standards play an important role in our everyday lives and facilitate the operation of products produced by different manufacturers. For example:

Countries adopt a standard type of mains plug and socket. Without such a standard, we would find that we had to continually rewire mains plugs or employ some form of adaptor. This provides an example of national standard.

Car manufacturers adopt a standard for the relative placement of the clutch, brake and accelerator pedals. This provides an example of global standard.

Computers are equipped with standard interface sockets (e.g. serial, parallel and USB) via which they are able to connect to peripheral devices. This provides an example of global standard.

Standards may come into being in various ways.
For example:

A standard may be established (imposed) by the company that plays the most dominant role in any particular area. For example, the serial and parallel ports employed by today's PC were implemented on the earliest PCs introduced by IBM. They soon became standard for desktop computing

A standard may gradually evolve

A standard may be developed/defined by a committee of experts.

Although standardization can facilitate our use of technologies and products, standards seldom reflect an optimal solution. For example, the VHS videotape format became a standard, while 28 other superior and equally cost-effective formats fell by the wayside. Furthermore, in the case of standards developed by committees, these often reflect many technological compromises and take long periods to develop. Such standards are often out of date even before they are released.

From a computer user's perspective, standards are extremely important because they allow a combination of products from different manufacturers to be used together. Standards ensure greater compatibility and interoperability between various types of equipment and technologies.

In data communications, standards provide guidelines to manufacturers and service providers to ensure compatibility, connectivity, and interoperability of technologies – an essential requirement in today's global market. Key advantages of standards are:

- To ensure a large market for hardware or software products – thus encouraging mass production
- To allow products from different vendors to communicate, thus giving customers more flexibility in the selection and use of equipment.

On the other hand, standards do have limitations:

- They tend to slow down technological change. This is due to the fact that, in some cases, by the time a standard is developed, subjected to scrutiny, reviewed, compromised and endorsed by all concerned parties – and then disseminated, more efficient technologies could have developed.
- Many standards may exist for the same thing. It is often difficult to decide which standard will provide better compatibility and remain in place for the greatest amount of time.

Many official computer-related standards are defined by the following organizations:

- **ANSI** (America National Standards Institute)
- **ITU** (International Telecommunication Union)
- **IEEE** (Institute of Electrical and Electronic Engineers)
- **ISO** (International Organization for Standardization)
- **VESA** (Video Electronics Standards Association).

Car drivers generally use agreed signals when turning left or right. Aero plane pilots follow specific standardized rules for communicating throughout the world. Similarly, for any computer-based systems to communicate successfully, they need to use ‘the same language’.

This means that what is communicated, how it is communicated, and when it is communicated must conform to some mutually acceptable conventions agreed between the parties involved. These conventions are known as a ‘protocol’, which can be defined as a set of rules governing the exchange of data between two or more devices.

Typical tasks performed by protocols are as follows:

To make sure that the source device activates the data communication line

To inform the transmission system about the destination system.

To make sure that the source device communicates with the destination device before sending data

To make sure the destination device is ready to accept the data

To make sure that the destination file management system is ready to store incoming files

To ensure compatibility between source and destination, and to perform format translation.

In the 1980s, many companies entered the desktop computing market and this led to a rich diversity of products. Unfortunately, these systems would often not operate together, nor could software developed for use on one particular type of machine necessarily be used on another. In short, although the lack of standards enabled product diversity, it hampered computer usage. Quickly, standards were developed (and/or evolved) and these impacted on many areas of computing. For example:

Compatibility improved. By conformance to standards, hardware and software systems developed by different manufacturers could be used together (although there were often unforeseen problems)

The diversity of available products decreased

Backwards compatibility became an important issue. For example, a new model of computer, or a new release of an operating system should support the function of older products. This has greatly increased hardware and software complexity and retarded the development of radically new computer products.

3.5 The OSI model

The Open System Interconnection (OSI) reference model was developed by the International Standards Organization (ISO) and provides a framework for protocol development. By implementing a communication protocol that adheres to the OSI model, systems developed by different manufacturers are able to communicate. The tasks that must be performed to enable machines to communicate in an effective and efficient manner are incorporated within a seven-layer hierarchy, as indicated in figure 3. Although the protocols detailed within this reference 0 model are seldom used, the model provides us with an excellent conceptual framework for understanding the tasks performed by the various software layers. Below we briefly summarize aspects of the functionality of the various layers.

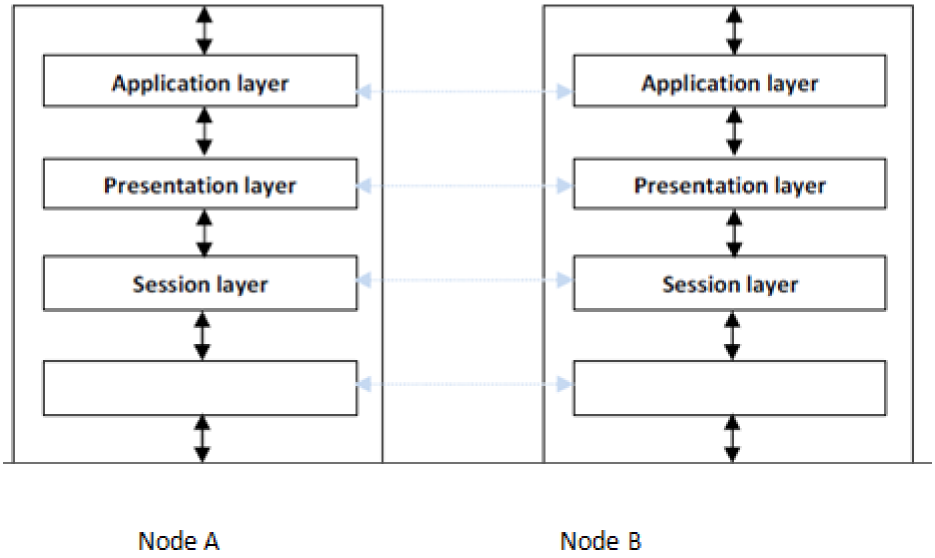


Figure 3: The layers within the OSI reference model

3.5.1 Application layer

This should not be confused with the applications programs that may be running on a computer. The application layer provides network access to the user and to applications programs. This layer passes data to (and receives data from) the presentation layer, and logically communicates directly to the application layer on the remote computer. This is indicated in figure 3 where the horizontal lines indicate the logical communication of each layer with its remote counterpart. The application layer needs know nothing of the tasks carried out by the lower layers – it needs only interface with the user (and applications programs) and with the presentation layer.

3.5.2 Presentation layers

Different computers may employ different character set formats. A user is not interested in 31 such differences and one of the tasks undertaken by the presentation layer is to translate between different formats that may be used to represent numbers, characters and other symbols. Additionally, the presentation layer is also involved in ensuring secure data transmission (consequently, when data is being transmitted the presentation layer undertakes encryption, and when data is being received it performs decryption).

3.5.3 Session layer

A user applications program may need to open a 'session' with a remote machine. For example, a user may wish to log on to a remote computer and carry out various tasks and this will involve the transmission and reception of data over a period of time. This necessitates synchronization whereby each node knows when it can transmit and when it is to receive data (i.e. when it must 'listen'). The session layer deals with this synchronization and additionally is involved in error recovery. Consider the case that a file is being transmitted between two nodes, and during this process the network fails. Without the support of the session layer it would be necessary to start the transmission process again from the beginning. However, the session layer inserts checkpoints into the transmitted data stream and these are used to efficiently recover from such failures. Following a failure, transmission can be recommenced from the point at which the last checkpoint was successfully delivered to the destination node. The session layer carries out various other activities, such as bracketing a set of related and non-independent activities. For example, there may be a need to carry out a task on a remote machine, which involves the execution of series of commands. Perhaps if only some of these commands are executed (i.e. they are not carried out in their entirety) problems will ensue. If the individual commands are executed as each arrives at the remote machine then, in the case that the network connection fails, there is the likelihood of incomplete execution. One task performed by the session layer relates to the buffering of such commands – as each arrives it is temporarily stored and not passed to higher layers until all commands (and any associated data) have been received. The series of commands may then execute in full.

3.5.4 Transport layer

This acts as the intermediary between the lower layers (whose implementation is dependent on the underlying network architecture) and the three upper layers which provide user services and whose architecture is (at least in principle) independent of the detailed network characteristics .

The type of transport service that is provided to the session layer is determined by the transport layer. Suppose a node wished to send an extremely large file to a remote machine via a shared network (or set of interconnected networks). Without judicious design (in relation to the type of transport service used), there is the possibility that such a transmission could block

the network(s) in such a way that whilst the transmission is in progress no other machines could communicate. The approach commonly used to prevent such a situation is to split the data into chunks (‘packets’) which are individually encapsulated within a frame containing all the necessary data needed to enable a packet delivery to the intended destination. The splitting of the data into smaller units is carried out by the transport layer. These packets may traverse a set of networks by different routes and so arrive at their destination out of order. The transport layer reorders packets and so enables them to be correctly reassembled.

3.5.5 Network Layer

This layer decides on routing issues, determining the path that should be followed by packets when they traverse networks. In fact, in such a situation the path taken is not defined solely by the source node but by all the nodes (network devices) through which packets pass on their way to the destination. Consider the situation illustrated in figure 4

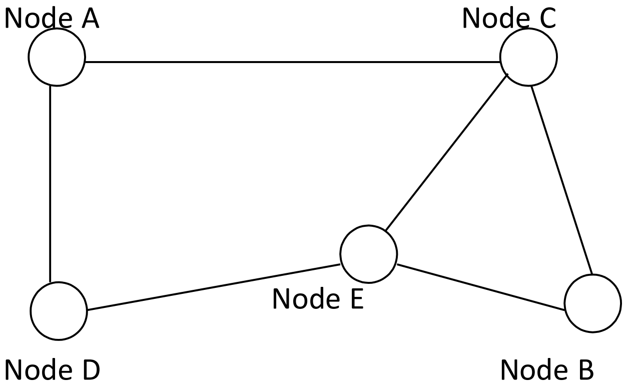


Figure 4: A simple network in which a packet may be sent from Node A to B via different routes. The circles represent nodes, and the lines network interconnects.

Suppose that a packet is to be sent from Node A to Node B. The packet will have to pass through at least one intermediate node (network device). These nodes may simply forward the packet, or may decide on the direction of the next step in its voyage. Thus, for example, Node D simply performs a forwarding function, whereas Nodes C and E are able to make routing decisions. The transport layer plays a critical role in determining the time it will take for packets to reach their destination and in this sense the actions of the transport layer impact on transmission latency.

3.5.6 The data link layer

This layer is responsible for various low-level network specific tasks and plays a crucial part in the detection and correction of errors that may occur during the transmission process.

Correction may be achieved by means of additional information inserted into messages prior to their transmission that can be used to modify bits corrupted during the transmission process.

Alternatively, correction may involve requesting re-transmission. Additionally, the data link layer plays a pivotal role in managing network access and ensuring that network ‘collisions’ (which occur when two or more nodes attempt to transmit onto the same LAN at the same time) are handled correctly. Devices connected together via networks do not necessarily demonstrate the same transmission/reception characteristics.

Thus a device able to transmit at high speed (i.e. that has high bit-rate) could readily swamp a slower recipient. Buffering techniques are used to circumvent this problem and this necessitates a protocol that ensures that the capacity of the buffer is not exceeded. This is referred to as flow control.

3.5.7 The physical layer

This layer deals with the transmission of the bit stream through the transmission medium, and the connection strategy used to enable the transfer of bits from one node to another. Thus the physical layer defines the signal levels, the type of transmission medium employed (e.g. twisted pair cable, coaxial cable, fiber optic cable), and also the techniques that will be used to permit the passage of data, such as circuit switching (in which a dedicated path is set up between two communicating nodes), packet switching, etc.

3.6 The TCP/IP protocol

In the late 1960s, the US Department of Defence's Advance Research Project Agency (ARPA) initiated a project that centered upon the interconnection of geographically dispersed computing systems. Gradually a large-scale network of university and government computing facilities evolved (this network was named ARPANET), which used packet switching techniques and initially employed leased phone lines. Early networking protocols were slow and unreliable and in 1974 a new set of protocols were proposed. These formed the basis for TCP/IP (Transmission Control Protocol/Internet Protocol (TCP/IP) which today underpins the operation of the Internet.

A protocol such as TCP/IP must support a number of essential requirements such as:

Reliability: in terms of both data integrity and timely delivery

Fault tolerance: the failure of a network segment should not seriously disrupt overall network operation; it must be possible to route packets along different paths so that th 34 can still reach their destination

Transparent communications: different computer systems and LANs should be able to communicate transparently.

It is convenient to employ a layered model in order to most readily conceptualize TCP/IP. We can therefore consider TCP/IP within a four-layer framework (a five–layer model is sometimes preferred). In figure 5 these layers are depicted, and are placed alongside the layers that comprise the OSI model. Below we briefly summarize aspects of their role

3.6.1 Application layer

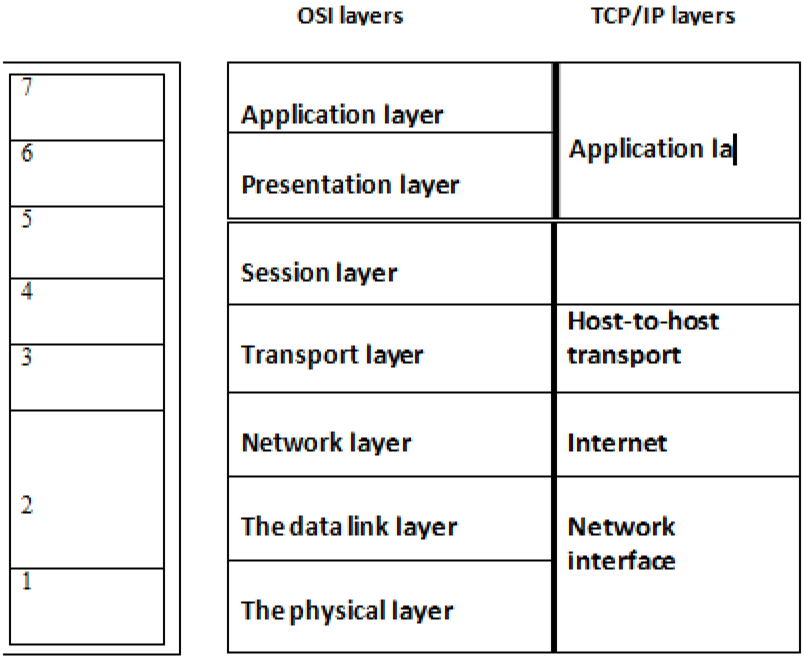
This layer provides communication services to the user and to applications programs. It can be viewed as corresponding to the application, presentation and session layers found in the OSI model. The application layer contains all the high-level protocols (such as those that we commonly encounter when accessing the Internet – such as DNS (Domain Name System) and HTTP).

3.6.2 Transport layer

Two different protocols are defined in this layer (TCP and UDP (User Datagram Protocol)). These differ in a number of important respects. For example:

Reliability: in the case of UDP, error correction is not implemented – the onus for this activity is placed on the applications program. This contrast with TCP in which error detection and correction form an integral part. Free from error correction overheads, UDP can (under some circumstances) demonstrate high performance

Flow control: in the case of TCP, flow control is implemented and this prevents a faster machine from swamping a recipient that operates more slowly.

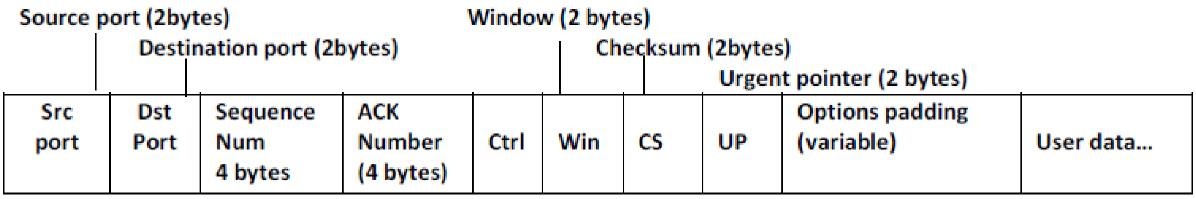


	OSI layers	TCP/IP layers
7	Application layer	Application la
6	Presentation layer	
5	Session layer	
4	Transport layer	Host-to-host transport
3	Network layer	Internet
2	The data link layer	Network interface
1	The physical layer	

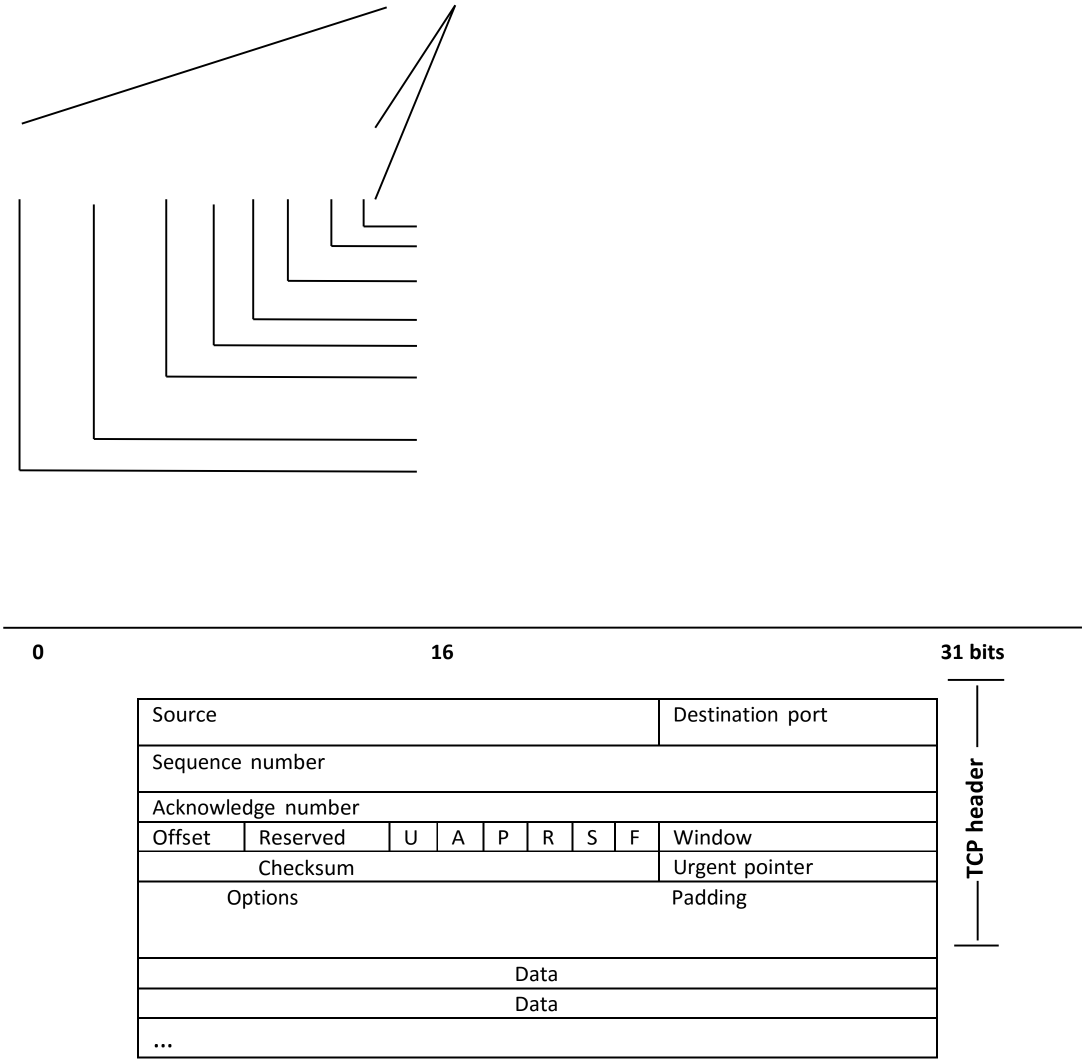
Figure 5: A conceptual model of TCP/IP set alongside the layers that comprise the OSI model

A stream of data that is to be transmitted is fragmented into chunks and the transport layer appends various information, before passing these to the internet layer. At the receiving node, the transport layer reassembles these data chunks. In the case of TCP, the transport layer encapsulates the data chunks into a TCP segment (in the case of UDP, the encapsulated data is usually referred to as a packet. There are differences between the information contained in the UDP and TCP headers.) Here the data is provided with a 'header' containing various important information; see Figure 6. It is instructive to consider the purpose of several pieces of information contained in the header:

Source and Destination ports: many well known (widely used) application protocols are designated by unique identification numbers provided by the 'Internet Assigned Numbers Authority'. For example, the File Transfer Protocol (FTP) is identified as "port21', and the Simple Mail Transfer Protocol (SMTP) as 'port 25'. TCP inserts this information into the header and thereby provides information on the source and destination applications protocol associated with the data to be transferred. The source port and destination port fields are each two bytes long, and values below 256 are used to reference 'well-known' ports.



Src port	Dst Port	Sequence Num 4 bytes	ACK Number (4 bytes)	Ctrl	Win	CS	UP	Options padding (variable)	User data...



Source								Destination port
Sequence number								
Acknowledge number								
Offset	Reserved	U	A	P	R	S	F	Window
Checksum								Urgent pointer
Options								Padding
Data								
Data								
...								

Figure 6: Information contained within a TCP segment

Sequence number: TCP numbers each byte of data that is transmitted between two nodes during the transfer process. The sequence number references the first byte of data encapsulated within frame. This is most readily understood by means of an example. Suppose that a set of frames are transmitted between node A and node B, and that each contains 256 bytes of data. Then the sequence numbers contained in the first four frames transmitted by Node A could be 1,257,513, 769 (the process is slightly more complex since the sequence number of the first frame need not be 1). Node B these sequence numbers to reconstruct the data chunks and correct for frames being received out of their transmitted order

Header length: this enables the receiving node to determine the point at which the header ends and the data starts. It is necessary to specify this length as not of fixed size.

Checksum: this enables the transport layer to perform error detection

Options: various options can be included. For example, one option enables the recipient to inform the source node about the maximum segment size that it is able to accept. This is indicated during the establishment of a communication and ensures that the recipient's buffer will not be swamped by a high-speed transfer.

Internet layer

At the sending node, the Internet layer takes packets or segments generated by the transport layer, and further encapsulate these to produce datagrams. The additional information appended by the Internet layer (the 'IP header') is intended to enable the datagrams to be injected onto any network and travel (via intermediate networks) to the intended destination. During their transit, intermediate network devices will use this information to determine the direction they should take. Since the routing of packets is fundamental to the Internet layer, it may be considered to be equivalent to the network layer used in the OSI model.

Network interface layer

In terms of its functionality, this layer is equivalent to the lowest two layers used in the OSI model. It further encapsulates a datagram received from the Internet layer producing a 'frame'. This layer makes the connection to the transmission medium and employs the appropriate protocol for launching and receiving frames.

The process of encapsulation referred to above is summarized in figure 7 and in Table 1 an overview of the functionality of the layers that have been conceptualized in connection with TCP/IP is presented.

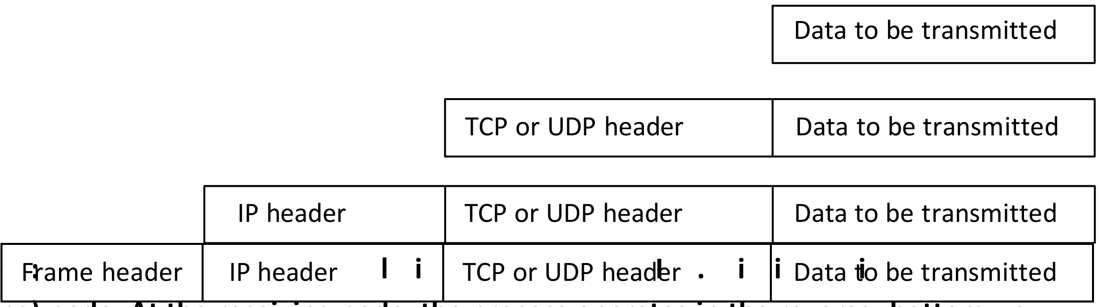


Figure rce (sending) node. At the receiving node, the process operates in the reverse: bottom up.

Application (4)

Similar to OSI application layer
Serves as communication interface by providing specific application services
Examples include email, virtual terminal, file transfer, WWW

Transport (3)

Defined by two protocols:

User Datagram protocol (UDP)

- a connectionless protocol
- provides unreliable datagram service (no end -to-end error detection or correction)
- does not retransmit any unreceived data
- requires little overhead
- application protocols include Trivial File Transfer Protocol (TFTP), Network File System (NFS), Simple Network Management Protocol (SNMP), Bootstrap Protocol (BOOTP), and Domain Name Service (DNS)

Transmission Control Protocol (TCP)

- (the TCP of TCP/IP)
- connection–oriented protocol
- provides reliable data transmission via end-to-end detection and correction
- guarantees data is transferred across a network accurately and in correct order
- retransmits any data not received by destination node
- guarantees against data duplication between sending and receiving nodes
- application protocols include Telnet, FTP, SMTP and POP

Internet (2)

(The IP of TCP/IP)
Transfers user messages from source host to destination host
Connectionless datagram service
Route selection is based on a metric
Uses Internet or IUP addresses to locate a host within the Internet
Relies on routers or switches
Integral part is Internet Control Message Protocol (ICMP); this uses an IP datagram to carry messages about state of communications environment

Network Interface (1)

Connects host to the local network hardware
Makes a connection to the physical medium
Uses a specific protocol for accessing the medium
Places data into frames
Effectively performs all functions of the first two layers of the OSI model

Table 1: A summary of some aspects of the functionally of the conceptualized four-layer TCP/IP model

4.0 CONCLUSION

You have been introduced to the basic concepts of packets, protocols and standards. You are now in a position to relate these concepts to how communication occurs between entities in different systems in your environment.

5.0 SUMMARY

Packets and protocols are the fundamental building blocks of data transmission over the network. A packet is a segment of data that has a header with destination and addressing information attached to it.

A protocol is a set of rules that govern data communication; the key elements of a protocol are syntax, semantics and timing.

Standards are necessary to ensure that products from different manufacturers can work together as expected. The ISO, ITU-T, ANSI, IEEE and EIA are some of the organizations involved in standards creation.

To enable two or more computers to communicate in a meaningful manner, a communication protocol must be defined.

We briefly summarize aspects of the functionality of the various layers of the OSI model and layers of the TCP/IP protocol.

TCP/IP today underpins the operation of the Internet.

6.0 TUTOR-MARKED ASSIGNMENTS

1 List the major disadvantages with the layered approach to protocols.

2a. Why are protocols needed?

b Why are standards needed?

3. How does the protocol travel through the OSI model?

4. What does OSI stand for and what do we use it for?

7.0 REFERENCES/FURTHER READING

1. Burgess, M. (2004). Principles of Network and System Administration. (2nd Ed.). Chichester, West Sussex , England: Wiley.

2. Forouzan, B.A, & Fegan, S.C. (2007). Data communications and Networking (4th Ed). Mc
Graw Hill.

3. Limoncelli, T. A.,Hogan, C. J. & Chalup, S. R (2007}. The Practice of System and
Network Administration. (2nd Ed.). Upper Saddle River, NJ: Addison-Wesley

4. Stallings, W. (2009). Data and computer communications (8th ed.). Upper saddle River, NJ.: Pearson Education Inc.

Unit 3 Network, Transport and Application Layers

1.0 INTRODUCTION

This unit covers three important layers in networking, namely: the network, transport and the application layers. Important algorithms and mechanisms related to each layer are introduced. We will also discuss certain types of network devices.

2.0 OBJECTIVES

By the end of this unit, you will be able to:

- to explain the difference between the network, transport and application layers
- to make use of the routing and congestion algorithms and explain transport control mechanisms.

3.0 MAIN CONTENT

3.1 Network Layer

The Network Layer provides services to the Transport Layer. It can be based on either virtual circuits or datagrams. In both cases, its main job is routing packets from the source to the destination. In virtual circuit subnets, a routing decision is made when the virtual circuit is set up. In datagram subnets, it is made on every packet.

The network layer services have been designed with the following goals:

- The services should be independent of the subnet technology.

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

- The network addresses made available to the transport layer should use a uniform numbering plan, even across LANs and WANs.

The discussion centres on the question of whether the network layer should provide connection – oriented service or connectionless service.

3.1.1 Routing Algorithms

Many routing algorithms are used in computer networks. Static algorithms include shortest path routing, flooding, and flow–based routing. Dynamic algorithms include distance vector routing and link state routing. Most actual networks use one of these. Other important routing topics are hierarchical routing, routing for mobile hosts, broadcast routing, and multicast routing.

The function of the network layer is routing packets from the source machine to the destination machine. In most subnets, packets will require multiple hops to make the journey. The only notable exception is for broadcast networks, but even here, routing is an issue if the source and destination are not on the same network. The algorithms that choose the routes and the data structures that they use are a major area of network layer design.

The routing algorithm is that part of the network layer software responsible for deciding which output line an incoming packet should be transmitted on.

Regardless of whether routes are chosen independently for each packet or only when new connection is established, there are certain properties that are desirable in a routing algorithm: correctness, simplicity, robustness, stability, fairness and optimality.

Stability is also an important goal for the routing algorithm. Routing algorithms can be grouped into two major classes: *non–adaptive* and *adaptive*. Non–adaptive algorithms do not base their routing decisions on measurements or estimates of the current traffic and topology. Instead, the choice of the route to use to get from I to J is computed in advance, off–line, and downloaded to the routers when the network is booted. This procedure is something called static routing.

Adaptive algorithms, in contrast, change their routing decisions to reflect changes in the topology, and usually the traffic as well. Adaptive algorithms differ in where they get their information (e.g., every sec, when the load changes, or when the topology changes), and what metric is used for optimization (e.g., distance, number of hops, or estimated transit time). In the following sections, we will introduce a variety of routing algorithms, both static and dynamic.

3.1.1.1 Shortest Path Routing

Let us begin our study of routing algorithms with a technique that is widely used in many forms because it is simple and easy to understand. The idea is to build a graph of the subnet, with each node of the graph representing a router and each arc of the graph representing a communication line. To choose a route between a given pair of routers, the algorithm just finds the shortest path between them on the graph. One way of measuring path length is the number of hops.

3.1.1.2 Flooding

Another static algorithm is flooding, in which every incoming packet is sent out on every outgoing line except the one it arrived on. Flooding obviously generates vast numbers of duplicate packets, in fact an infinite number unless some measures are taken, but it is one of the simplest routing algorithms.

Routers need to communicate with other routers so they can exchange routing information. Most network operating systems have associated routing protocols which support the transfer of routing information. Typical routing protocols and their associated network operating systems are:

- BGP (Border Gateway Protocol)–TCP/IP.
- EGP (Exterior Gateway Protocol)–TCP/IP.
- IS–IS (Immediate System to Intermediate System) – DECent, OSI.
- NLSP (NetWare Link State Protocol) – Net Ware 4.1
- OSPF (Open Shortest Path First) – TCP/IP.
- RIP (Routing Information Protocols) – XNS, Net Ware, TCP/IP.
- RTMP (Routing Table Maintenance Protocol) – Apple Talk.

3.1.2 Congestion Control Algorithms

The situation in which when too many packets are present in the subnet, performance degrades. This situation is called congestion.

Subnet can become congested, increasing the delay and lowering the throughput for packets. Network designers attempt to avoid congestion by proper design. Techniques include traffic shaping, flow specifications, and bandwidth reservation. If congestion does occur, it must be dealt with. Choke packets can be sent back, load can be shed, and other methods applied.

Congestion can be brought about by several factors.

- If all of a sudden, streams of packets begin arriving on three or four input lines and all need the same output line, a queue will build up. If there is insufficient memory to hold all of them, packets will be lost. If the router has an infinite amount of memory, congestion gets worse, not better, because by the time packets get to the front of the queue, they have already timed out and duplicates have been sent. All these packets will be dutifully forwarded to the next router, increasing the load all the way to the destination.

- Slow processors can also cause congestion. If the routers; CPUs are slow at performing the book–keeping tastes required of them, queues can build up, even though there is excess line capacity. Similarly, low–bandwidth lines can also cause congestion.

- Congestion control has to do with making sure the subnet is able to carry the offered traffic. It is a global issue, involving the behavior of all the hosts, all the routers, the store and forwarding process within the routers, and all the other factors that tend to diminish the carrying capacity of the subnet.

3.1.3 Comparison of Virtual Circuit and Datagram Subnets.

Inside the subnet, several trade–offs exist between virtual circuits and datagrams.

- One trade–off is between router memory space and bandwidth.
  Virtual circuits allow packets to contain circuit numbers instead of full destination addresses. If the packets tend to be fairly short,
  a full destination address in every packet may represent a
  significant amount of overhead, and hence, wasted bandwidth. The price paid for using virtual circuits internally is the table space within the routers. Depending upon the relative cost of communication circuits versus router memory, one or the other may be cheaper.

- Another trade–off is setup time versus address passing time.
  Using virtual circuit requires a setup phase, which takes time and consumes resources. However, figuring out what to do with a data packet in a virtual circuit subnet is easy: the router just uses the circuit number to index into a table to find out where the packet goes. In datagram subnet, a more complicated procedure is required to determine where the packet goes.

- Virtual circuits have some advantages in avoiding congestion within the subnet because resources can be reserved in advance, when the connection is established. Once the packets start arriving, the necessary bandwidth and router capacity will be there. With a datagram subnet, congestion avoidance is more difficult.

- For transaction processing systems, the overhead is required to set up and clear a virtual circuit. If the majority of the traffic is expected to be of this kind, the use of switched virtual circuits inside the subnet makes little sense. On the other hand, permanent virtual circuits, which are set up manually and last for months or years, may be useful here.

- Virtual circuits also have vulnerability problem. If a router crashes and loses its memory, even if it comes back up a second later, all the virtual circuits passing through it will have to be aborted. In contrast, if a datagram router goes down, only those users whose packets were queued up in the router at the time, will suffer, and may be not even all those, depending upon whether they have already been acknowledged or not. The loss of communication line is fatal to virtual circuits using it but can be easily compensated for if datagrams are used.

- Datagrams also allow the routers to balance the traffic throughout the subnet, since routes can be changed halfway through a connection.

- It is worth explicitly pointing out that the service offered is a separate issue from the subnet structure. In theory, all four combinations are possible. Obviously, a virtual circuit implementation of a connection–oriented service and datagram implementation of a connectionless service are reasonable. Implementing connections using datagrams also makes sense when the subnet is trying to provide a highly robust service.

3.1.4 Internetworking

In internetworking modes are connected together through different network equipment. Networks connect to other networks through repeaters, bridges or routers. A repeater corresponds to the physical layer and always routes signals from one network segment to another. Bridges route using Data Link Layer and routers route using the network layer.

Networks differ in various ways, so when multiple networks are connected together, problems can occur. Sometimes the problems can be finessed by tunneling a packet through a hostile network, but if the source and destination networks are different, this approach fails. When different networks have different maximum packet sizes, fragmentation may be called for.

3.1.4.1 Repeaters

All types of network connections suffer attenuation and pulse distortion. For a given cable specification and bit rate, each has a maximum length of cable. Repeaters can be used to increase the maximum interconnection length and will do the following:

- Clean signal pulses.
- Passes all signals between attached segments
- Boost signal power
- Possibly translate between two different media types (e.g., fibre – optic to twisted–pair cable).

3.1.4.2 Bridges

Bridges filter input and output traffic so that only packets intended for a network are actually routed into the network and only packets intended for the outside are allowed out of the network.

3.1.4.3 Routers

Routers examine the network address field and determine the best route for the packet. They have the great advantage that they normally support several different types of network layer protocols.

Routers, which only read one type of protocol, will normally have high filtering and forwarding rates. If they support multiple protocols, then there is normally an overhead in that the router must detect the protocol and look into the correct place for the destination address.

3.2 Transport Layer

The Transport Layer provides reliable cost effective data transport from the source machine to destination machine.

3.2.1 Transport Service And Mechanism

The Transport Layer provides various services, the most important being an end–to–end, reliable, connection–oriented byte stream from sender to receiver. It is accessed through service primitives that permit the establishment, use and release of connections.

Transport protocols must be able to do connection management over unreliable networks. Connection establishment is complicated by the existence of delayed duplicate packets that can reappear at opportune moments. To deal with them, three–way handshakes are needed to establish connections. Releasing a connection is easier than establishing one but is still far from trivial due to the two–army problem.

3.2.2 Types of Service/Quality of Service (QoS)

The need to define quality of service arises from the realisation that users require different quality presentations at different times. The different quality presentations map onto different parameter values. When a multimedia presentation is transmitted via a network, it translates into different requirements of network performance. To be able to specify QoS aspects concisely and to request them of a network, QoS must be specified as a set of parameters that can be assigned numerical values. In a multimedia presentation, the ultimate user of the system is a human being. Thus, the quality of the presentation is a matter of the user's perception, which is limited by the response of the human vision and auditory senses. This perceptual nature of QoS makes it subjective and difficult to quantify precisely. Thus, it is easier to specify a range of values rather than a single value.

3.2.3 Transport Control Mechanism

The transport control service is implemented by a transport protocol used between the two transport entities. It is similar to the Data Link Protocol, but with some differences:

- Environments in which they operate (at the data link layer, two routers communicate directly via a physical channel, whereas at the transport layer, this physical channel is replaced by the entire subnet.

- In Data Link Protocol, it is not necessary for a router to specify which router to talk to–each outgoing line uniquely specifies a particular route.

3.2.3.1 Addressing

When an application process wishes to set up a connection to a remote application process, it must specify which one to connect to. In Internet, these end–points are (IP address + Local Port) pairs. The end–point in this context:

- **TSAP** (Transport Service Access Point)

- **NSAP** (Network Service Access Point)

A transport entity supports multiple TSAPs.

3.2.3.2 Flow Control And Buffering

Flow control of transport layer is similar to that of Data Link Layer, but in Transport layers, the number of connections open is numerous as compared to Data Link Layer.

If the subnet provides datagram service, the sending transport entity must also be buffered, for re–transmitting in the case of loss. If the receiver knows that the sender buffer all TPDUs (Transport Protocol Data Units) until they are acknowledged, the receiver may or may not dedicate specific buffers to specific connections.

In summary, if the network service is unreliable, the sender must buffer all TPDUs. However, with reliable network services, other trade–off becomes possible.

The optimum trade–off between source buffering and destination buffering depends on the type of traffic carried by the connection. For low–bandwidth bursty 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.

3.2.3.3 Multiplexing

Multiplexing is multiple things on to one i.e., multiplexing several conversations onto connections Virtual circuits and physical links play a role in several layers of the network architecture.

Need for Multiplexing

- A number of virtual circuits are open by the users or one user opening more than one, which requires a lot of buffer in the router; this gives a solid reason for packet switched network.

- To bill the users based on the amount of data sent, not the connection time.

Upward Multiplexing: Multiplexing of different transport connections onto the same network connection is attractive.

Downward Multiplexing: The transport layer opens multiple network connections and distributes the traffic among them on a round–robin basis.

3.2.3.4 Connection Establishment and Management

Connection Establishment is not as easy as it sounds, but it is in fact, a complicated task, we have to take care of the losses that occur during transmission. At first glance, it would seem sufficient for one transport entity to just send a CONNECTION REQUEST TPDU to the destination and wait for a CONNECTON ACCEPTANCE reply. The problem occurs when the network is not reliable.

The solution could be to give each connection a connection identifier, chosen by the initiating party, and put it in each TPDU, including the one requesting the connection. After each connection is released, each transport entity could update a table listing absolute connection as (peer transport entity, connection identifier) pair.

Unfortunately, this scheme has a basic flaw: it requires each transport entity to maintain a certain amount of history information indefinitely. If a machine crashes and loses its memory, it will no longer know which connection identifiers have already been used.

3.2.3.5 Crash Recovery

If host and routers are subject to crashes, recoveries from these crashes become an issue. If the transport entity is entirely within the hosts, recovery from network and router crashes is straightforward. If the network later provides datagram services, the transport entity expects lost TPDUs all the time and knows how to cope with them. If the network layer provides connection–oriented service, then loss of virtual circuits is handled by establishing a new one and then probing the remote transport entity to ask it which TPDUs it has received and which one it has not received.

3.2.4 TCP/UDP

In this section we will discuss two important transport layer protocols: TCP and UDP.

3.2.4.1 Transmission Control Protocol (TCP)

TCP provides a highly reliable, connection–oriented, end–to–end transport service between processes in end systems connected to the subnet. TCP only assumes that the layer below offers an unreliable datagram service. TCP provides the types of facility associated with the ISO Class 4 transport service, including error recovery, sequencing of packets, flow control by the windowing method, and the support of multiplexed connections from the layer above.

3.2.4.2 Format of TCP Header

The sender's TCP layer communicates with the receiver's TCP layer using the TCP protocol data unit. It defines parameters such as the source port, destination port, sequence number and so on. Its descriptions are given below:

- Source and destination port number – which are 16 bit values to identify the local port number.

- Sequence number – which identifies the current sequence number of the data segment. This allows the receiver to keep track of the data segments received. Any segment that is missing can be easily identified.

- Data offset – which is a 32–bit value and identifies the start of the data.

- Flags – the flag field is defined as UAPRSE, where U is the urgent flag, a the acknowledgement flag, P the push function, R the reset flag, S the sequence synchronize flag and E the end of transmission flag.

- Windows – is a 16 bit values and gives the number of data blocks that the receiving host can accept at a time.

- Checksum – is a 16 bit checksum for the data and header.

- UrgPtr – is the urgent pointer and is used to identify an important area of data.

3.2.4.3 User Data Protocol (UDP)

The internet protocol suite also supports a connectionless transport protocol, UDP (User Data Protocol). UDP provides a way for applications to send encapsulated raw IP datagrams and send them without having to establish a connection. Many client–server applications that have one request and one response use UDP rather than go through the trouble of establishing and later releasing a connection.

A UDP segment consists of an 8–byte header followed by the data. The two ports serve the same function as they do in TCP: to identify the end–points within the source and destination machines. The UDP length field includes the 8–byte header and the data. The UDP checksum includes the same format pseudo–header, the UDP header, and the UDP data, padded out to an even number of bytes if need be. It is optional and stored as 0 if not computed.

3.3 Application Layer

The Application Layer contains a variety of protocols that are commonly needed. For example, there are hundreds of incompatible terminal types in the world. To solve this problem there is need to define an abstract network virtual terminal that editors and other programs can be written to deal with.

Another application layer function is file transfer. Different file systems have different file naming conventions.

3.3.1 The Domain Name System (DNS)

DNS is a scheme for assigning meaningful high–level name to a large set of machines, and discusses a mechanism that maps between high–level machine names and IP addresses. It considers both the translation from–high level name to IP addresses and the translation from IP addresses to high–level machines names. It has been used to assign machine names throughout the global Internet. It uses a geographically distributed set of servers to map names to addresses, the implementation of the name mapping mechanism provides a large scale example of the client sever paradigm.

In a TCP/IP internet, hierarchical machine names are assigned according to the structure of organisations that obtain authority for parts of the namespace, not necessarily according to the structure of the physical network interconnections.

3.3.2 TCP/IP Internet Domain Name

The mechanism that implements a machine name hierarchy for TCP/IP Internets is called the Domain Name Systems (DNS). DNS has two conceptually independent aspects. The first is abstract; it specifies the name syntax and rules for delegating authority over names. The second is concrete; it specifies the implementation of a distributed computing system that efficiently maps names to addresses. This section considers the name syntax, and later sections examine the implementation.

The domain name system uses a hierarchical naming scheme as domain names. As in our earlier examples, a domain name consists of a sequence of sub names separated by a delimiter character, the period. In our examples we said that individual sections of the name might represent sites or groups, but the domain system simply calls each section, a label. Thus, the domain name cs.ignou.org contains three labels: CS, IGNOU, and ORG. Any suffix of a label in a domain name is called a domain. In the above example, the lowest level domain is CS. ignou.org (the domain name for ignou) and the top level domain is org. As the example shows, domain names are written with the local label first and the top domain last. As we will see, writing them in this order makes it possible to compress messages that contain multiple domain names.

Domain Name	**Meaning**
Com	Commercial Organisations
Edu	Educational Institutes Gov
Government Institutions Mil	Military
Groups	
Net	Major network support centres
Org	Organisations other than those above
Arpa	Temporary ARPANET domain (obsolete) INT
International Organisations	
Country code	Each country (geographic scheme)

3.3.3 Electronic Mail

This is the most widely used service facilitating users to send and receive messages electronically in a store and forward manner. Different E–mail standards, viz., SMTP, UUCP and X400 Message Handling system, are supported on ERNET.

Electronic Mail is a system whereby a computer user can exchange messages with other computer users or group of users via a communications network.

The backbone of an electronic mail system is a communication network that connects remote terminals to a central system or a local area network that interconnects personal computers. Users can send mails to a single recipient or they can broadcast it to any number of selected users on the systems. When multi–tasking personal computer and workstation are used, mail can be delivered to users while they are working on something else. Otherwise, users have to interrogate their mail boxes in a central system, or file server.

Many users first encounter computer networks when they send or receive electronic mail to or from a remote site. E–mail is the most widely used application service. Indeed, many computer users access networks only through electronic mail.

E–mail is popular because it offers a fast, convenient method of transferring information. E–mail can accommodate small notes or large voluminous memos with a single mechanism. It should not surprise you to learn that more users send files with electronic mail than with file transfer protocols.

Characteristics

- Store and forward
- Delivery time ranging from few seconds to hours
- Largely textual
- Binary files may be appended or “uuencoded”
- Multimedia (“mime” standard)
- Distribution lists with “cc:”, “bcc:”, “fcc:”
- Mail forwarding
- Auto–processing
- Statistics collection
- Secure email
- Several mailers: smtp,uucp (smtp requires IP connectivity; uucp works with dial–up).

3.3.4 WWW (World Wide Web)

The World Wide Web is a system for linking up hypertext documents. Each document is a page written in HTML, possible with hyperlinks to other documents. A browser can display a document by establishing a TCP connection to its server, asking for the document, and then closing the connection. When a hyperlink is selected by the user, that document can also be fetched in the same way. In this manner, documents all over the world are linked together in a giant web. Some facts about WWW:

- Fastest growing discovery and retrieval system
- Presently 10,000 servers, growing at an astounding rate
- Retrieve “hypermedia” documents, with text, graphics, audio, video, and links to other hypermedia documents.
- A navigational system based on “hyperlink”
- State–less interaction between client and server, conforming to “http” protocol.

3.3.5 Mail–Based Applications

Plain–old mail
Notices
Auto–save and processing
News dissemination through LISTSRV

Archival search and retrieval
Access to network–wide news (bulletin boards)

3.4 Remote Procedure Call (RPC)

The designers chose to build three independent pieces; the NFS protocol itself, a general purpose Remote Procedure Call (RPC) mechanism, and a general purpose External Data Representation (XDR). Their intent was to separate the three to make it possible to use RPC and XDR in other software, including application programs as well as other protocols. For example, a programmer can divide a program into a client side and a server side that use RPC as the chief communication mechanism can one of the client sides, the programmer designates some procedures as remote, forcing the compiler to incorporate RPC code into those procedures. On the server side, the programmer implements the desired procedures and uses other RPC facilities to declare them to be part of a server. When the executing client program calls one of the remote procedures, RPC automatically collects values for argument, from a message, sends the message to the remote server, awaits a response, and stores returned values in the designated arguments. In essence, communication with the remote server occurs automatically as a side– effect of a remote call. The RPC mechanism hides all the details of protocols, making it possible for programmers who know little about the underlying communication protocols to write distributed programs.

3.5 File Transfer Protocol (FTP)

FTP (File Transfer Protocol) is the primary method of transferring files over the Internet. “FTP” transfers files to and from a remote network site. Some sites maintain Anonymous accounts on the system for retrieval of public domain softwares stored on the system.

The ftp protocol is used to access files by FTP, the Internet’s file transfer protocol. FTP has been around more than two decades and is well entrenched. Numerous FTP servers all over the world allow people anywhere on the internet to log in and download whatever files have been placed on the FTP server. The web does not change this; it just makes obtaining files by FTP easier, as FTP has a somewhat arcane interface.

3.6 Telnet

Telnet is a program that allows you to establish a virtual terminal connection between two machines using TCP/IP. For this, you must have its internet address or host name of computer.

4.0 CONCLUSION

In this unit, you have been taken through Network Layer, Transport Layer and Application layer. Also the various algorithms, mechanisms, and protocols relating to each of these layers have been discussed.

It has also discussed other concepts like multiplexing, crash recovery, electronic mail, WWW, etc.

5.0 SUMMARY

In this unit, you have been introduced to the network, transport and application layers, their features, services offered by them and the algorithms used by them. Other concepts covered include internetworking, repeaters, routes, bridges, multiplexing, addressing and transport control mechanisms. Standards and definitions of commonly used terms in Application Layer are covered briefly to familiarise you with current trends.

6.0 TUTOR-MARKED ASSIGNMENT

i. What is the difference between virtual circuit and Datagram subnets?……………………………………………………………...
………………………………………………………………………
………………………………………………………………………
………………………………………………………………………

ii. Distinguish between non–adaptive and adaptive algorithms.
…………………………………………………………………......
………………………………………………………………………
………………………………………………………………………
……………………………………………………………………….

7.0 REFERENCES/FURTHER READING

Course Marking Scheme