JAWAHARLAL COLLEGE OF ENGINEERING AND TECHNOLOGY

(Approved by AICTE, Affiliated to APJ Abdul Kalam Technological University, Kerala)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

(NBA Accredited)

COURSE MATERIAL

CST 303 COMPUTER NETWORKS

VISION OF THE INSTITUTION

Emerge as a centre of excellence for professional education to produce high quality engineers and entrepreneurs for the development of the region and the Nation

MISSION OF THE INSTITUTION

- To become an ultimate destination for acquiring latest and advanced knowledge in the multidisciplinary domains.
- To provide high quality education in engineering and technology through innovative teaching-learning practices, research and consultancy, embedded with professional ethics.

- To promote intellectual curiosity and thirst for acquiring knowledge through outcome based education.
- To have partnership with industry and reputed institutions to enhance the employability skills of the students and pedagogical pursuits.
- To leverage technologies to solve the real life societal problems through community services.

ABOUT THE DEPARTMENT

- Established in: 2008
- Courses offered: B.Tech in Computer Science and Engineering
- Affiliated to the A P J Abdul Kalam Technological University.

DEPARTMENT VISION

To produce competent professionals with research and innovative skills, by providing them with the most conducive environment for quality academic and research oriented undergraduate education along with moral values committed to build a vibrant nation.

DEPARTMENT MISSION

- Provide a learning environment to develop creativity and problem solving skills in a professional manner.
- Expose to latest technologies and tools used in the field of computer science.
- Provide a platform to explore the industries to understand the work culture and expectation of an organization.
- Enhance Industry Institute Interaction program to develop the entrepreneurship skills.
- Develop research interest among students which will impart a better life for the society and the nation.

PROGRAMME EDUCATIONAL OBJECTIVES

Graduates will be able to

- Provide high-quality knowledge in computer science and engineering required for a computer professional to identify and solve problems in various application domains.

- Persist with the ability in innovative ideas in computer support systems and transmit the knowledge and skills for research and advanced learning.
- Manifest the motivational capabilities, and turn on a social and economic commitment to community services.

PROGRAM OUTCOMES (POS)

Engineering Graduates will be able to:

1. **Engineering knowledge**: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

2. **Problem analysis**: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

3. **Design/development of solutions**: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

4. **Conduct investigations of complex problems**: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

5. **Modern tool usage**: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

6. **The engineer and society**: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

7. **Environment and sustainability**: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

8. **Ethics**: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

9. **Individual and team work**: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

10. **Communication**: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

11. **Project management and finance**: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

12. **Life-long learning**: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

OURSE OUTCOMES

SUBJECT CODE: C303		
COURSE OUTCOMES		
C302.1	To identify the basic structure and functional units of a digital computer. And analyze the effect of addressing modes on the execution time of a program	**L3**
C302.2	To explain the design issues of data link layer, link layer protocols, bridges & switches and to Illustrate wired and wireless LAN protocols.	**L2**
C302.3	To select appropriate routing algorithms, congestion control techniques, and Quality of Service requirements for a network.	**L3**
C302.4	To Understand the importance of network layer in internet and various network layer protocols.	**L2**
C302.5	To Illustrate the functions and protocols of transport layer and application layer in inter-networking	**L2**

Outcome Number **Outcome** **Knowledge Level**

PROGRAM SPECIFIC OUTCOMES (PSO)

The students will be able to

- Use fundamental knowledge of mathematics to solve problems using suitable analysis methods, data structure and algorithms.
- Interpret the basic concepts and methods of computer systems and technical specifications to provide accurate solutions.
- Apply theoretical and practical proficiency with a wide area of programming knowledge, design new ideas and innovations towards research.

CO PO MAPPING

Note: H-Highly correlated=3, M-Medium correlated=2,L-Less correlated=1

CO'S	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
C302.1	3	3	3									3
C302.2	3	3	3									3
C302.3	3	3	3	3								3
C302.4	3	3	3	3								3
C302.5	3	3	3	3		3						3
C302	3	3	3	3		3						3

CO PSO MAPPING

CO'S	PSO1	PSO2	PSO3
C302.1	3		
C302.2	3		
C302.3	3		
C302.4	3		
C302.5	3		
C302.6	3	3	
C302	2.83	3	

Reference Materials

Module 1

Introduction – Uses – Network Hardware – LAN –MAN – WAN, Internetworks – Network Software – Protocol hierarchies – Design issues for the layers – Interface & Service – Service Primitives.

Reference models – OSI – TCP/IP.

Computer network/data network: A network can be defined as a group of computers and other devices connected in some way so as to be able to exchange data. A network allows nodes to share resources. In computer networks, computing devices exchange data with each other using connections between nodes (data links). These data links are established over cable media such as wires or optic cables, or wireless media such as Wi-Fi. Each of the devices on the network can be thought of as a node, each node has a unique address.

Advantages of computer network:

□ It enhances communication and availability of information: Networking, especially with full access to the web, allows ways of communication that would simply be impossible before it was developed. Instant messaging can now allow users to talk in real time and send files to other people wherever they are in the world, which is a huge boon for businesses. Also, it allows access to a vast amount of useful information, including traditional reference materials and timely facts, such as news and current events.

□ It allows for more convenient resource sharing: This benefit is very important, particularly for larger companies that really need to produce huge numbers of resources to be shared to all the people. Since the technology involves computer-based work, it is assured that the resources they wanted to get across would be completely shared by connecting to a computer network which their audience is also using.

□ It makes file sharing easier : Computer networking allows easier accessibility for people to share their files, which greatly helps them with saving more time and effort, since they could do file sharing more accordingly and effectively.

□ It is highly flexible. This technology is known to be very flexible, as it gives users the opportunity to explore everything about essential things, such as software without affecting their functionality. Plus, people will have the accessibility to all information they need to get and share.

□ It is an inexpensive system. Installing networking software on your device would not cost too much, as you are assured that it lasts and can effectively share information to your peers. Also, there is no need to change the software regularly, as mostly it is not required to do so.

□ It increases cost efficiency. With computer networking, you can use a lot of software products available on the market which can just be stored or installed in your system or server and can then be used by various workstations.

□ It boosts storage capacity. Since you are going to share information, files and resources to other people, you have to ensure all data and content are properly stored in the system. With this networking technology, you can do all of this without any hassle, while having all the space you need for storage.

Disadvantages of computer network

□ It lacks independence. Computer networking involves a process that is operated using computers, so people will be relying more of computer work, instead of exerting an effort for their tasks at hand. Aside from this, they will be dependent on the main file server, which means that, if it breaks down, the system would become useless, making users idle.

□ It poses security difficulties. Because there would be a huge number of people who would be using a computer network to get and share some of their files and resources, a certain user's security would be always at risk. There might even be illegal activities that would occur, which you need to be careful about and aware of.

□ It allows for more presence of computer viruses and malware. There would be instances that stored files are corrupt due to computer viruses. Thus, network administrators should conduct regular check-ups on the system, and the stored files at the same time.

□ Its light policing usage promotes negative acts. It has been observed that providing users with internet connectivity has fostered undesirable behaviour among them. Considering that the web is a minefield of distractions—online games, humor sites. The huge network of machines could also encourage them to engage in illicit practices, such as instant messaging and file sharing, instead of working on work-related matters. While many organizations draw up certain policies on this, they have proven difficult to enforce and even engendered resentment from employees.

□ It requires an efficient handler. A computer network to work efficiently and optimally, it requires high technical skills and know-how of its operations and administration. A person just having basic skills cannot do this job. Take note that the responsibility to handle such a system is high, as allotting permissions and passwords can be daunting. Similarly, network configuration and connection is very tedious and cannot be done by an average technician who does not have advanced knowledge.

□ It requires an expensive set-up. Though computer networks are said to be an inexpensive system when it is already running, its initial set up cost can still be high depending on the number of computers to be connected. Expensive devices, such as routers, switches, hubs, etc., can add up to the cost. Aside from these, it would also need network interface cards (NICs) for workstations in case they are not built in.

- It lacks robustness. If a computer network’s main server breaks down, the entire system would become useless. Also, if it has a bridging device or a central linking server that fails, the entire network would also come to a standstill.

USES OF COMPUTER NETWORK

Business Applications: Many companies have a substantial number of computers. For example, a company may have separate computers to monitor production, keep track of inventories, and do the payroll. Initially, each of these computers may have worked in isolation from the others, but at some point, management may have decided to connect them to be able to extract and correlate information about the entire company. The issue here is resource sharing, and the goal is to make all programs, equipment, and especially data available to anyone on the network without regard to the physical location of the resource and the user.

A computer network can provide a powerful communication medium among employees.

A third goal for increasingly many companies is doing business electronically with other companies, especially suppliers and customers.

A fourth goal that is starting to become more important is doing business with consumers over the Internet. Airlines, bookstores, and music vendors have discovered that many customers like the convenience of shopping from home.

Home Applications

Some of the more popular uses of the Internet for home users are as follows:

1. Access to remote information.

2. Person-to-person communication. 3.Interactive entertainment.

4. Electronic commerce.

Mobile Users

Mobile computers, such as notebook computers and personal digital assistants (PDAs), are one of the fastest growing segments of the computer industry.

There are two types of transmission technology that are in widespread use. They are as follows:

1. Broadcast links.

2. Point-to-point .

Broadcast networks have a single communication channel that is shared by all the machines on the network.

Point-to-point networks consist of many connections between individual pairs of machines. To go from the source to the destination, a packet on this type of network may have to first visit one or more intermediate machines. Machine are received by all the others.

Types of computer network

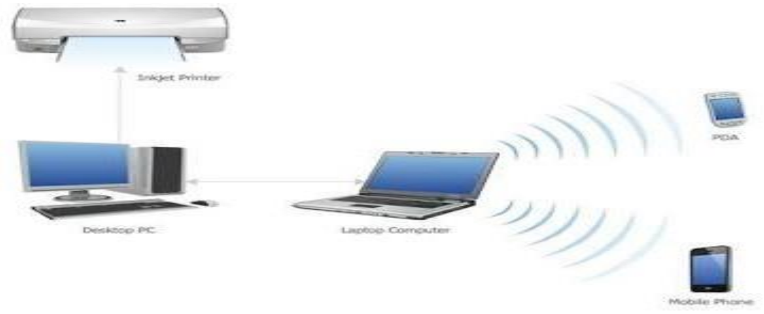
1. Personal Area Network (PAN)

The smallest and most basic type of network, a PAN is made up of a wireless modem, a computer or two, phones, printers, tablets, etc., and revolves around one person in one building. These types of networks are typically found in small offices or residences and are managed by one person or organization from a single device.

Eg : wireless computers ,keyboard &Mouse Bluetooth embedded headphones

2. Local Area Network (LAN)

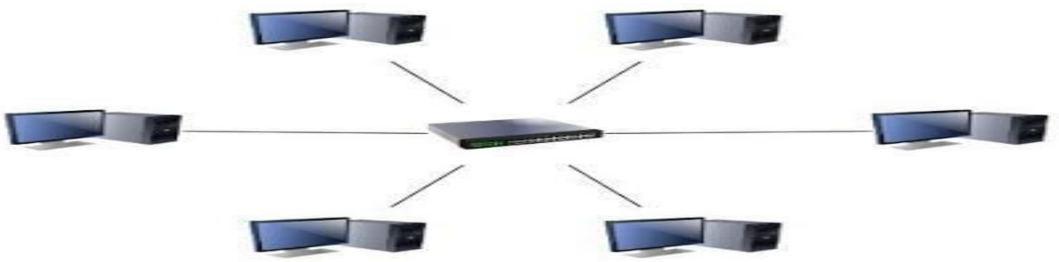
LANs connect groups of computers and low-voltage devices together across short distances (within a building or between a group of two or three buildings in close proximity to each other) to share information and resources. Enterprises typically manage and maintain LANs



LAN (Local Area Network) links the devices in local areas such as in your campus, building etc. It provides useful way of sharing resources such as printer &scanner, sharing of file server.

3. Wireless Local Area Network (WLAN)

Functioning like a LAN, WLANs make use of wireless network technology, such as Wi-Fi. Typically seen in the same types of applications as LANs, these types of networks don’t require that devices rely on physical cables to connect to the network.



4. Campus Area Network (CAN)

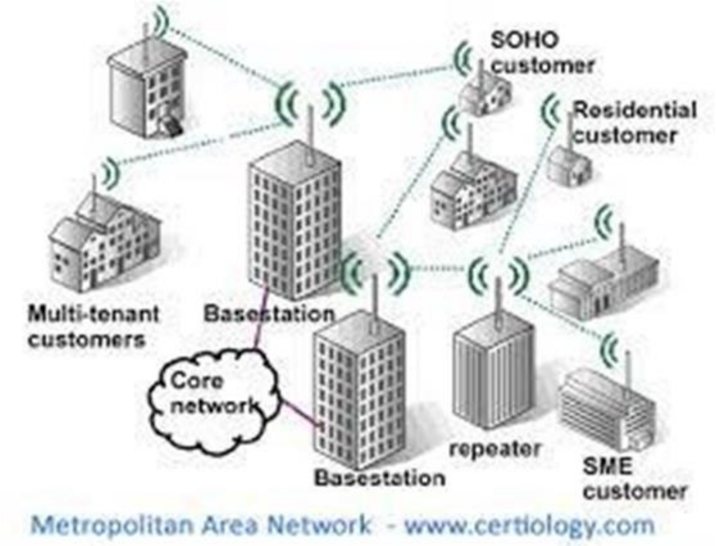
Larger than LANs, but smaller than metropolitan area networks (MANs, explained below), these types of networks are typically seen in universities, large K-12 school districts or small businesses. They can be spread across several buildings that are fairly close to each other so users can share resources.

5. Metropolitan Area Network (MAN)

These types of networks are larger than LANs but smaller than WANs

– and incorporate elements from both types of networks. MANs span an entire geographic area (typically a town or city, but sometimes a campus). Ownership and maintenance is handled by either a single person or company (a local council, a large company, etc.).

MAN covers a particular tower or large city. It exceeds to 32 to 40 km or 20 to 25 miles. Multiple LAN connected to form LAN. MAN provides uplink for LAN to WAN. They provide faster communication using optic cable. The backbone of MAN is high capacity &high- speed fibre optics.

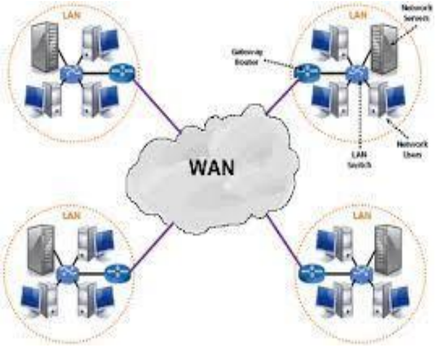


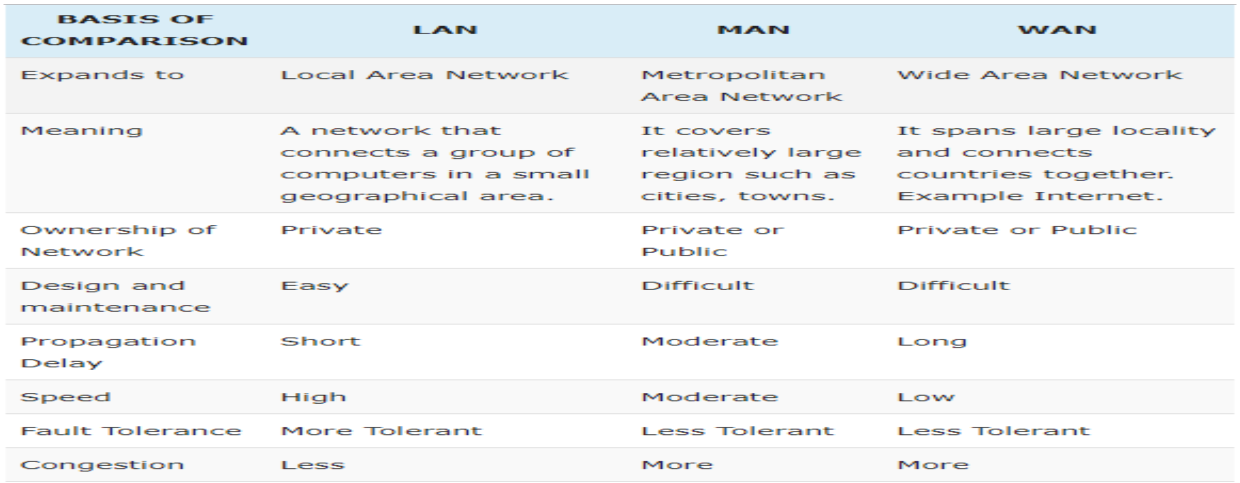
Metropolitan Area Network - www.certiology.com

.

6. Wide Area Network (WAN)

Slightly more complex than a LAN, a WAN connects computers together across longer physical distances. This allows computers and low-voltage devices to be remotely connected to each other over one large network to communicate even when they're miles apart. WAN covers a particular country. A WAN connects small network LAN & MAN.A computer user in one location can communicate with computer user in other location. It connects more than one LAN'S. It is used for large geographical area. It is active larger than 30miles.





BASIS OF COMPARISON	LAN	MAN	WAN
Expands to	Local Area Network	Metropolitan Area Network	Wide Area Network
Meaning	A network that connects a group of computers in a small geographical area.	It covers relatively large region such as cities, towns.	It spans large locality and connects countries together. Example Internet.
Ownership of Network	Private	Private or Public	Private or Public
Design and maintenance	Easy	Difficult	Difficult
Propagation Delay	Short	Moderate	Long
Speed	High	Moderate	Low
Fault Tolerance	More Tolerant	Less Tolerant	Less Tolerant
Congestion	Less	More	More

INTERNETWORKS

Many networks exist in the world, often with different hardware and software. People connected to one network often want to communicate with people attached to a different one. The fulfilment of this desire requires that different, and frequently incompatible networks, be connected, sometimes by means of machines called gateways to make the connection and provide the necessary translation, both in terms of hardware and software. A collection of interconnected networks is called an internetwork or internet.

An internetwork is formed when distinct networks are interconnected. Internetworking is the practice of connecting a Computer network with the networks through the use of gateways that provide a common method of routing information packets between the networks. The resulting system of interconnected networks is called an internetwork. The most notable example of internetworking is the internet, a network of networks based on many underlying hardware technologies, but unified by an internetworking protocol standard, the internet protocol suite, often also referred to as TCP/IP.

Network Software

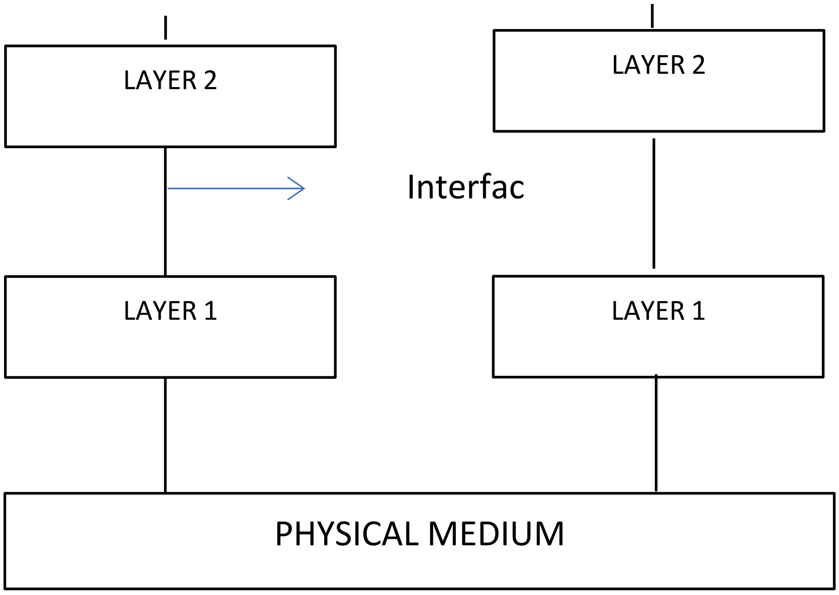
The first computer networks were designed with the hardware as the main concern and the software as an afterthought. This strategy no longer works. Network software is now highly structured.

PROTOCOL

In computer networks, 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 rules that govern data communications. A protocol defines what is communicated, how it is communicated, and when it is communicated. The key elements of a protocol are syntax, semantics, and timing.

PROTOCOL HIERARCHIES

To reduce their design complexity, most networks are organized as a stack of layers or levels, each one built upon the one below it. The number of layers, the name of each layer, the contents of each layer, and the function of each layer differ from network to network. The purpose of each layer is to offer certain services to the higher layers, shielding those layers from the details of how the offered services are actually implemented. Layer n on one machine carries on a conversation with layer n on another machine. The rules and conventions used in this conversation are collectively known as the layer n protocol.



In order to understand how the actual communication is achieved between two remote hosts connected to the same network, a general network diagram is shown above divided into a series of layers. As it seen later on the on the course the actual number as well as their function of each layer differs from network to network.

Each layer passes data and control information to the layer below It. As soon as the data are collected form the next layer, some functions are performed there and the data are upgraded and passed to the next layer. This continues until the lowest layer is reached. Actual communication occurs when the information passes layer 1 and reaches the Physical medium. This is shown with the solid lines on the diagram.

Theoretically layer n on one machine maintains a conversation with the same layer in the other machine. The way this conversation is achieved is by the protocol of each layer. Protocol is collection of rules and conventions as agreement between the communication parties on how communication is to proceed. The latter is known as virtual communication and is indicated with the dotted lines on the diagram above.

Layer n of one machine carries a conversion with the layer n of another machine. The rules and conversion are collectively known as protocol.

The data and information is passed by each layer to the lower layer. When the lower layer is reached it is passed to the physical medium which actual communication occurs. Between the pair of adjacent layer their lies the interface. The interface defines which type of services the lower layer offers to the upper layer. Protocols are together called protocol stack or set of protocols.

As far as the above diagram is concerned another important issue to be discussed is the interface between each layer. It defines the services and operation the lower layer offers to the one above It. When a network is built decisions are made to decide how many layers to be included and what each layer should do. So each layer performs a different function and as a result the amount of information past from layer to layer is minimized.

Design issues for a layer

- Every layer has a mechanism of connection establishment.

Since a network has many computers, some having multiple process. A machine has to specify with whom it has to establish connection. Because of having consequence of multiple destination, the addressing is needed in order to specify the specific destination.

- Another set of design decisions concerns the rules for data transfer.

In some data transfer take place in one direction and in some other it travel in both direction but not simultaneously. And there are situations were data transfer takes place simultaneously. Protocol determines how many logical channels are needed per connection.

- Error control: The physical communication circuits are not perfect. There are error detecting and error correcting codes. Both ends of the connection should agree which one is being used. The receiver must someway tell the sender which message is have been correctly received and which is not.

- Speed of sender is greater than the receiver.

There will be some kind of access from the receiver to the sender directly & indirectly about the receiver current situation.

- Inability of process to accept long message.

- Very expensive to set up connection for each communication process.

- Reliability: It is a design issue of making a network that operates correctly even when it is made up of unreliable components.

- Addressing: There are multiple processes running on one machine. Every layer needs a mechanism to identify senders and receivers.

- Error Control: It is an important issue because physical communication circuits are not perfect. Many error detecting and error correcting codes are available. Both sending and receiving ends must agree to use any one code.

- Flow Control : If there is a fast sender at one end sending data to a slow receiver, then there must be flow control mechanism to control the loss of data by slow receivers. There are several mechanisms used for flow control such as increasing buffer size at receivers, slow down the fast sender, and so on. Some process will not be in position to accept arbitrarily long messages. This property leads to mechanisms for disassembling, transmitting and the reassembling messages.

- Multiplexing and De-multiplexing : If the data has to be transmitted on transmission media separately, it is inconvenient or expensive to setup separate connection for each pair of communicating processes. So, multiplexing is needed in the physical layer at sender end and de-multiplexing is need at the receiver end.

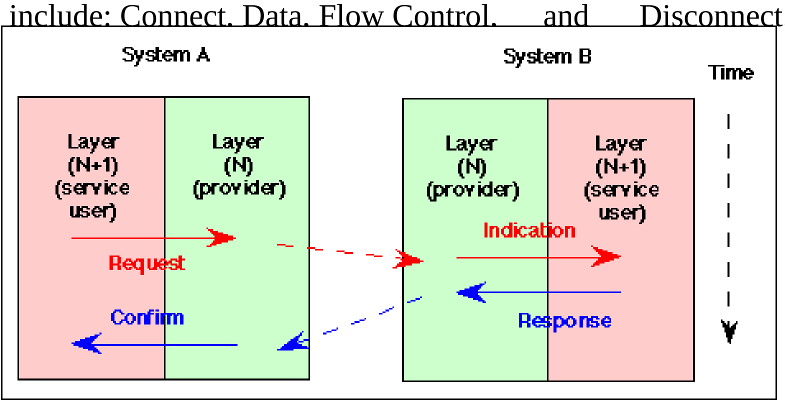
- Scalability : When network gets large, new problem arises. Thus scalability is important so that network can continue to work well when it gets large.

- Routing : When there are multiple paths between source and destination, only one route must be chosen. This decision is made on the basis of several routing algorithms, which chooses optimized route to the destination.

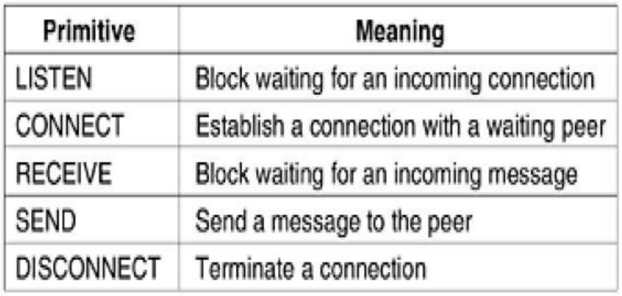
- Confidentiality and Integrity: Network security is the most important factor.

Service Primitives : Each protocol which communicates in a layered architecture communicates in a peer to peer manner with its remote protocol entity. Communication between adjacent protocol layers (i.e. within the same communications node) are managed by calling functions, called Primitives, between the layers. A service is formally specified by a set of primitives (operations) available to a user process to access the service. These primitives tell the service to perform some action or report on an action taken by a peer .

- There are various types of actions that may be performed by primitives. Examples of primitives include: Connect, Data, Flow Control, and Disconnect.



Some of the services are:



Primitive	Meaning
LISTEN	Block waiting for an incoming connection
CONNECT	Establish a connection with a waiting peer
RECEIVE	Block waiting for an incoming message
SEND	Send a message to the peer
DISCONNECT	Terminate a connection

:

Each primitive specifies the action to be performed or advises the result of a previously requested action. A primitive may also carry the parameters needed to perform its functions. One parameter is the packet to be sent/received to the layer above/below (or, more accurately, includes a pointer to data structures containing a packet, often called a "buffer").

There are four types of primitive used for communicating data. The four basic types of primitive are :

Request : A primitive sent by layer (N + 1) to layer N to request a service. It invokes the service and passes any required parameters. Indication : A primitive returned to layer (N + l) from layer N to advise of activation of a requested service or of an action initiated by the layer N service.

Response : A primitive provided by layer (N + 1) in reply to an indication primitive. It may acknowledge or complete an action previously invoked by an indication primitive.

Confirm : A primitive returned to the requesting (N + l)st layer by the Nth layer to acknowledge or complete an action previously invoked by a request primitive.

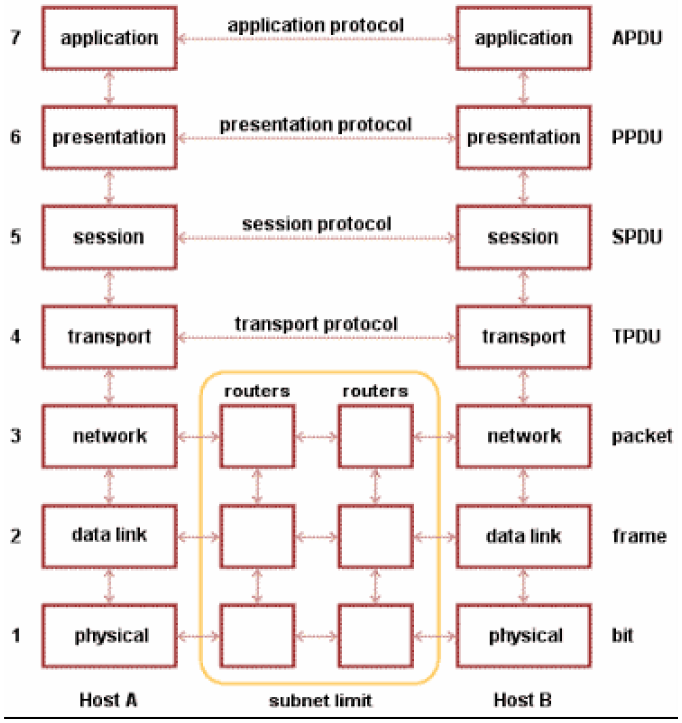
To send Data, the sender invokes a Data. Request specifying the packet to be sent, and the Service Access Point (SAP) of the layer below. At the receiver, a Data. Indication primitive is passed up to the corresponding layer, presenting the received packet to the peer protocol entity.

REFERENCE MODELS

1. **OSI MODEL**

2. **TCP/IP MODEL**

OSI REFERENCE MODEL



OSI stands for Open Systems Interconnection. It has 7-layers and attempts to abstract common features common to all approaches to data communications, and organize them into layers so that each layer only worries about the one above it and the one directly below it.

Although the actual data transmission is vertical, starting from the Application layer of the clients' computer all the way to the Application layer of the destination computer, each layer is programmed as though the data transmission were horizontal. This can be observed by above figure. In this figure peers are entities comprising the corresponding layers on each machine meaning that the peers that communicate using the protocol. In reality, as I stated above, no data are directly transferred from layer n on one machine to the corresponding layer on another machine.

Physical Layer

The physical layer has as a main function to transmit bits over a communication channel as well as to establish and terminate a connection to a communications medium. It is also responsible to make sure that when one side sends a '1' bit the other side will receive '1' bit and not '0' bit. The physical layer is combination of 1s and 0s.it is concerned with transmitting raw bits over a communication channel. Voltage needs for transfer.

Data Link Layer

Data link layer provides means to transfer data between network entities. Network Layer breaks into frames and passes them to the physical layer. At the receiving end data link layer detects and possibly corrects the errors that may occur during the transmission and passes the correct stream to the network layer. It's also concerned with flow control techniques. It controls the flow of transmission, and error detection. That is the main task of this layer is to take raw transmission facility and transform it into a line that appears free of transmission errors to the network layer. It accomplishes the task by having the sender break the input data up into data frames, transmit the data sequentially and process the acknowledgement frames sent back to the receiver. The data link layer creates and recognize the frame boundaries. This can be accomplished by attaching special bit patterns at the beginning and end of the frame.

Network Layer

This layer performs network routing, flow control and error control functions. Network routing simply means the way packets are routed from source to destination and flow control. prevents the possibility of congestion between packets which are present in the subnet simultaneously and form bottlenecks. The main task of this layer is to decide the path from multiple paths. That is the key design issue is determining how packets are routed from source to destination. They are highly dynamic.

If too many packets are present in a subnet at the same time they will get in each other's way forming bottlenecks. The control of such congestion also belongs to this layer. When a packet has to travel from one network layer to its destination, many problems may arise. The addressing used by the second network may be different from the first one. The second packet may not accept the packet because it is too large. The protocols may differ. It is up to the network layer to overcome all these problems to allow heterogenous networks interconnected.

Transport Layer

The Transport Layer has as a main task to accept data from the Session layer, split them up into smaller units an passes them to the Network layer making sure that all the pieces arrive correctly to the destination. It is the first end-to-end layer all the way from source machine to destination machine unlike the first three layers which are chained having their protocols between each machine. This is shown clearly in the diagram above. This layer will ensure that the data reached to receiver without any error. And the basic function of this layer is to accept data from the session layer, split it up to smaller units if need be, pass to the network layer and ensure that all pieces arrive correctly at the other end. Also the transport layer creates a distinct network connections for each transport connection required by the session layer. If the transport connection requires high throughput the transport layer might create multiple network connections dividing the data among the network connections to improve throughput. Transport layer might multiplex several transport connections on to the same network connection in order to reduce the cost.

Session Layer

Session layer is responsible for controlling exchange information and for synchronization. This layer is for creating a session dialog control and which allows the user on different machines to establish sessions between them. session layer has the total management f synchronization .also allow a user to log into a time sharing system.one of the services of the session layer is to manage the dialogue box.

Presentation Layer

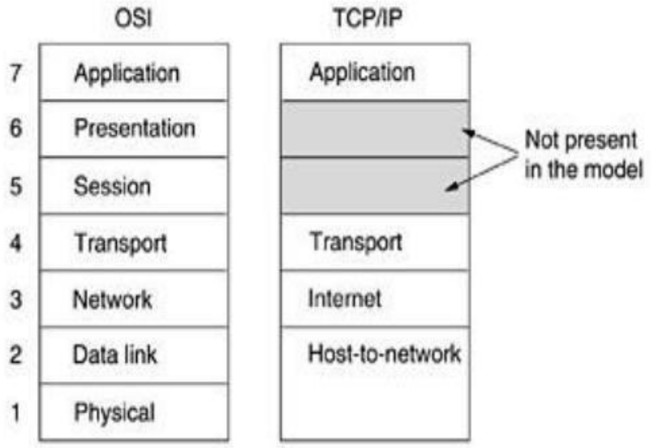
It is responsible to translate different data formats from the representation used inside the computer (ASCII) to the network standard representation and back. Computers use different codes for representing character strings so a standard encoding must be used and is handled by the presentation layer. Generally, in a few words this layer is concerned with the syntax and semantics of the information transmitted. The presentation layer performs encapsulation, description, compression and decompression. Also certain functions that are requested sufficiently often to warrant finding a general solution for them.

Application layer

The upper layer of this model performs common application service for the application processes meaning that software programs are written in the application layer to handle the many different terminal types that exist and map the virtual terminal software onto the real terminal. It contains a variety of protocols and is concerned with file transfer as well as electronic mail, remote job entry and various other services of general interest. This layer has a particular application. It contains a variety of protocols that are commonly needed. Another application of this layer is file transfer. Different file system has different file naming conventions, different ways of representing text lines and on.

TCP/IP REFERENCE MODEL

TCP/IP reference model was named after its two main protocols: TCP (Transmission Control Protocol) and IP (Internet Protocol). This model has the ability to connect multiple networks together in a way so that data transferred from a program in one computer are delivered safely to a similar program on another computer.



Internet Layer

This layer is a connectionless internetwork layer and defines a connectionless protocol called IP. Its concerned with delivering packets from source to destination. These packets travel independently each taking a different route so may arrive in a different order than they were send. Internet layer does not care about the order the packets arrive at the destination as this job belongs to higher layers. This layer, called the internet layer, is the linchpin that holds the whole architecture together. Its job is to permit hosts to inject packets into any network and have they travel independently to the destination (potentially on a different network). They may even arrive in a different order than they were sent, in which case it is the job of higher layers to rearrange them, if in-order delivery is desired. Note that "internet" is used here in a generic sense, even though this layer is present in the Internet.

The internet layer defines an official packet format and protocol called IP (Internet Protocol). The job of the internet layer is to deliver IP packets where they are supposed to go. Packet routing is clearly the major issue here, as is avoiding congestion. For these reasons, it is reasonable to say that the TCP/IP internet layer is similar in functionality to the OSI network layer. Fig. shows this correspondence.

Transport Layer

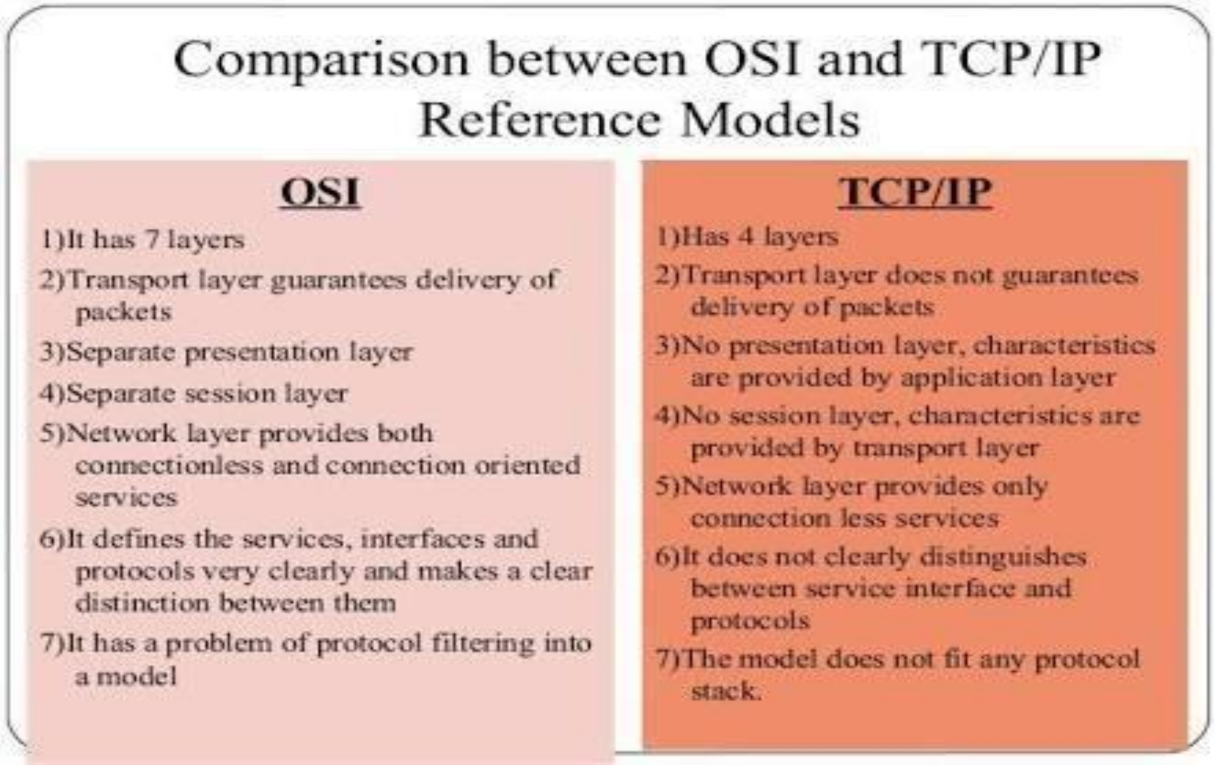
It contains two end-to-end protocols. TCP is a connection-oriented protocol and is responsible for keeping track of the order in which packets are sent and reassemble arriving packets in the correct order. It also ensures that a byte stream originating on one machine to be delivered without error on any other machine on the internet. The incoming byte stream is fragmented into discrete messages and is

passed to the internet layer. With an inverse process, at the destination, an output stream is produced by reassembling the received massage.

UDP is the second protocol in this layer and it stands for User Datagram Protocol. In contrast to TCP, UDP is a connectionless protocol used for applications operating on its own flow control independently from TCP. It is also an unreliable protocol and is widely used for applications where prompt delivery is more important than accurate delivery. such as transmitting speech or video. The layer above the internet layer in the TCP/IP model is now usually called the transport layer. It is designed to allow peer entities on the source and destination hosts to carry on a conversation, just as in the OSI transport layer. Two end-to- end transport protocols have been defined here. The first one, TCP (Transmission Control Protocol), is a reliable connection-oriented protocol that allows a byte stream originating on one machine to be delivered without error on any other machine in the internet. It fragments the incoming byte stream into discrete messages and passes each one on to the internet layer. At the destination, the receiving TCP process reassembles the received messages into the output stream. TCP also handles flow control 26 to make sure a fast sender cannot swamp a slow receiver with more messages than it can handle.

Application Layer

It is the upper layer of the model and contains different kinds of protocols used for many applications It includes virtual terminal TELNET for remote accessing on a distance machine, File Transfer Protocol FTP and e-mail (SMTP). It also contains protocols like HTTP for fetching pages on the www and others. The TCP/IP model does not have session or presentation layers. On top of the transport layer is the application layer. It contains all the higher-level protocols. The early ones included virtual terminal (TELNET), file transfer (FTP), and electronic mail (SMTP), as shown in Fig.6.2. The virtual terminal protocol allows a user on one machine to log onto a distant machine and work there. The file transfer protocol provides a way to move data efficiently from one machine to another. Electronic mail was originally just a kind of file transfer, but later a specialized protocol (SMTP) was developed for it. Many other protocols have been added to these over the years: the Domain Name System (DNS) for mapping host names onto their networDk aodwdrnelsosaesd, eNdNTfrPo,mtheKptruontooctoel sfo.rinmoving USENET news articles around, and HTTP, the protocol for fetching pages on the World Wide Web, and many others.



Comparison between OSI and TCP/IP Reference Models

OSI	TCP/IP
1)It has 7 layers	1)Has 4 layers
2)Transport layer guarantees delivery of packets	2)Transport layer does not guarantees delivery of packets
3)Separate presentation layer	3)No presentation layer, characteristics are provided by application layer
4)Separate session layer	4)No session layer, characteristics are provided by transport layer
5)Network layer provides both connectionless and connection oriented services	5)Network layer provides only connection less services
6)It defines the services, interfaces and protocols very clearly and makes a clear distinction between them	6)It does not clearly distinguishes between service interface and protocols
7)It has a problem of protocol filtering into a model	7)The model does not fit any protocol stack.

MODULE II

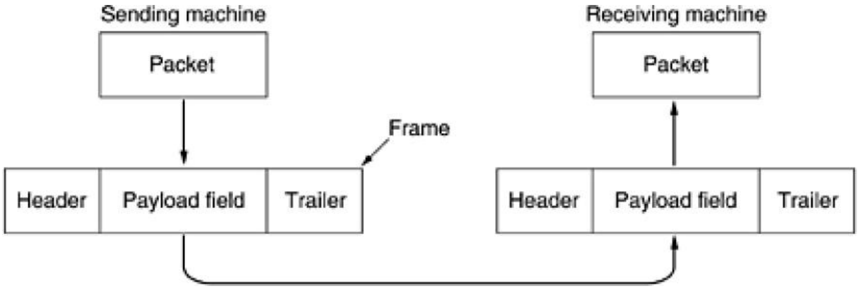
Syllabus: Data Link layer Design Issues – Flow Control and ARQ techniques. Data link Protocols – HDLC. DLL in Internet. MAC Sub layer – IEEE 802 FOR LANs & MANs, IEEE 802.3, 802.4, 802.5. Bridges - Switches – High Speed LANs - Gigabit Ethernet. Wireless LANs - 802.11 a/b/g/n, 802.15.PPP

Data Link Layer Design Issues

The data link layer has a number of specific functions it can carry out. These functions include

1. Providing a well-defined service interface to the network layer.
2. Dealing with transmission errors.
3. Regulating the flow of data so that slow receivers are not swamped by fast senders.

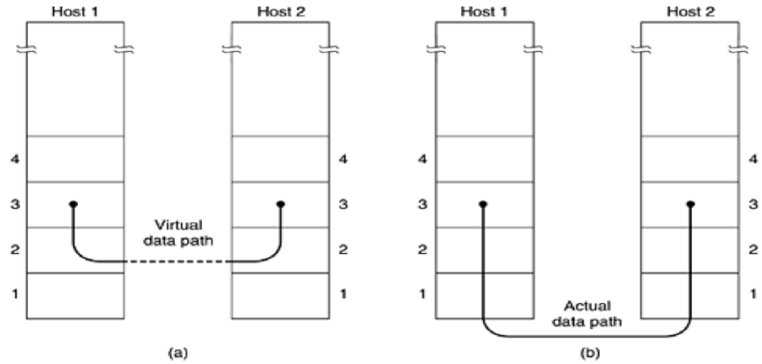
To accomplish these goals, the data link layer takes the packets it gets from the network layerand encapsulates them into frames for transmission. Each frame contains a frame header, a payload field for holding the packet, and a frame trailer, as illustrated in Fig.



Relationship between packets and frames.

2.1 Services Provided to the Network Layer

The function of the data link layer is to provide services to the network layer. The principal service is transferring data from the network layer on the source machine to the network layer onthe destination machine. On the source machine is an entity, call it a process, in the network layerthat hands some bits to the data link layer for transmission to the destination. The job of the data link layer is to transmit the bits to the destination machine so they can be handed over to the network layer there *(a) Virtual communication. (b) Actual communication.*



The data link layer can be designed to offer various services. The actual services offered can varyfrom system to system. Three reasonable possibilities that are commonly provided are

1. Unacknowledged connectionless service.
2. Acknowledged connectionless service.
3. Acknowledged connection-oriented service.

Unacknowledged connectionless service

- Consists of having the source machine send independent frames to the destination machine without having the destination machine acknowledge them.
- No logical connection is established beforehand or released afterward. If a frame is lost due to noise on the line, no attempt is made to detect the loss or recover from it in thedata link layer.
- This class of service is appropriate when the error rate is very low so that recovery is leftto higher layers.
- It is also appropriate for real-time traffic, such as voice, in which late data are worse than bad data. Most LANs use unacknowledged connectionless service in the data link layer.

Acknowledged connectionless service.

- When this service is offered, there are still no logical connections used, but each frame sent is individually acknowledged.
- In this way, the sender knows whether a frame has arrived correctly. If it has not arrived within a specified time interval, it can be sent again.
- This service is useful over unreliable channels, such as wireless systems.

Connection-oriented service.

- With this service, the source and destination machines establish a connection before any data are transferred.
- Each frame sent over the connection is numbered, and the data link layer guarantees that each frame sent is indeed received.
- Furthermore, it guarantees that each frame is received exactly once and that all frames are received in the right order.

2.2 Framing

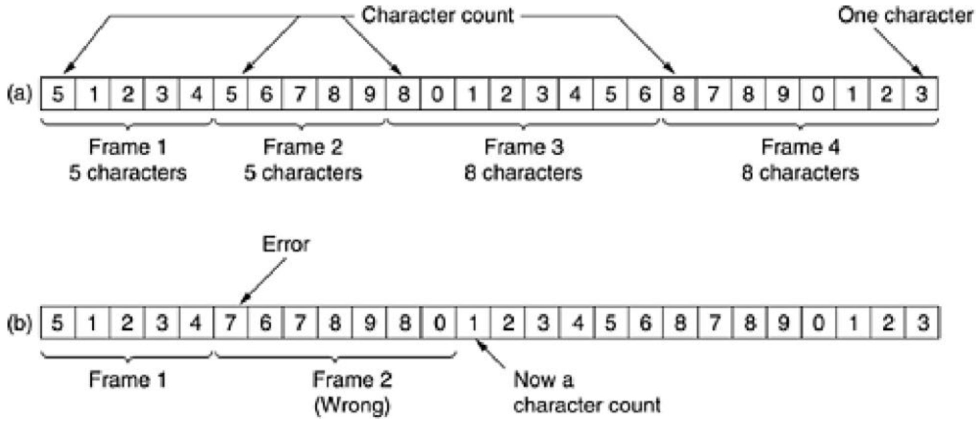
To provide service to the network layer, the data link layer must use the service provided to it by the physical layer. What the physical layer does is accept a raw bit stream and attempt to deliver it to the destination. This bit stream is not guaranteed to be error free. The number of bits received may be less than, equal to, or more than the number of bits transmitted, and they may have different values. It is up to the data link layer to detect and, if necessary, correct errors.

The usual approach is for the data link layer to break the bit stream up into discrete frames and compute the checksum for each frame. When a frame arrives at the destination, the checksum is recomputed. If the newly-computed checksum is different from the one contained in the frame, the data link layer knows that an error has occurred and takes steps to deal with it. for breaking the bit stream up into frames, various methods have been devised.

1. Character count.
2. Flag bytes with byte stuffing.
3. Starting and ending flags, with bit stuffing.
4. Physical layer coding violations.

1. Character count.

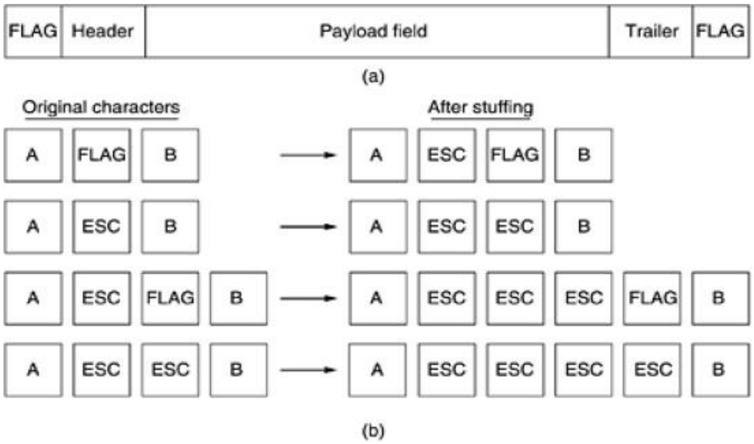
The first framing method uses a field in the header to specify the number of characters in the frame. When the data link layer at the destination sees the character count, it knows how many characters follow and hence where the end of the frame is.



A character stream. (a) Without errors. (b) With one error.

2. Flag bytes with byte stuffing.

The second framing method start and end a frame with special bytes, called a flag byte, as both the starting and ending delimiter.



(a) frame delimited by flag bytes. (b) Four examples of byte sequences before and after byte stuffing.

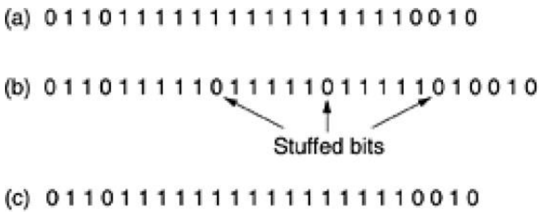
A serious problem occurs with this method when binary data, such as object programs or floating-point numbers, are being transmitted. It may easily happen that the flag byte's bit pattern occurs in the data. This situation will usually interfere with the framing. One way to solve this problem is to have the sender's data link layer insert a special escape byte (ESC) just before each "accidental" flag byte in the data. The data link layer on the receiving end removes the escape byte before the data are given to the network layer. This technique is called byte stuffing or character stuffing. Thus, a framing flag byte can be distinguished from one in the data by the absence or presence of an escape byte before it.

- A major disadvantage of using this framing method is that it is closely tied to the use of 8-bit characters. Not all character codes use 8-bit characters. For example. UNICODE uses 16-bit characters

3. Starting and ending flags, with bit stuffing.

- This allows data frames to contain an arbitrary number of bits and allows character codes with an arbitrary number of bits per character.
- It works like this. Each frame begins and ends with a special bit pattern, 01111110 (flag byte).
- Whenever the sender's data link layer encounters five consecutive 1s in the data, it automatically stuffs a 0 bit into the outgoing bit stream.
- When the receiver sees five consecutive incoming 1 bits, followed by a 0 bit, it automatically destuffs (i.e.,deletes) the 0 bit. Just as byte stuffing is completely transparent to the network layer in both computers, so is bit stuffing. If the user data contain the flag pattern, 01111110, this flag is transmitted as 011111010 but stored in

the receiver's memory as 01111110.Following Figure gives an example of bit stuffing.


(a) 011011111111111111110010
(b) 011011111011111011111010010
Stuffed bits
(c) 011011111111111111110010

Bit stuffing. (a) The original data. (b) The data as they appear on the line. (c) The data as theyare stored in the receiver's memory after destuffing.

4. Physical layer coding violations.

The last method of framing is only applicable to networks in which the encoding on the physical medium contains some redundancy. For example, some LANs encode 1 bit of data by using 2 physical bits. Normally, a 1 bit is a high-low pair and a 0 bit is a low-high pair. The scheme means that every data bit has a transition in the middle, making it easy for the receiver to locate the bit boundaries. The combinations high-high and low-low are not used for data but are used for delimiting frames in some protocols.

Many data link protocols use a combination of a character count with one of the other methods for extra safety. When a frame arrives, the count field is used to locate the end of the frame. Onlyif the appropriate delimiter is present at that position and the checksum is correct is the frame accepted as valid. Otherwise, the input stream is scanned for the next delimiter.

Error control

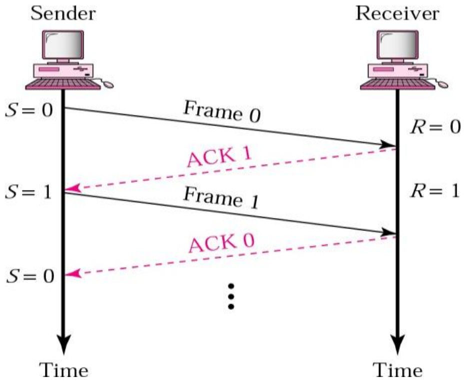
The bit stream transmitted by the physical layer is not guaranteed to be error free. The data link layer is responsible for error detection and correction. The most common error control method is to compute and append some form of a checksum to each outgoing frame at the sender's data linklayer and to recompute the checksum and verify it with the received checksum at the receiver's side. If both of them match, then the frame is correctly received; else it is erroneous. The checksum may be of two types.
# Error detecting : Receiver can only detect the error in the frame and inform the sender about it.# Error detecting and correcting : The receiver can not only detect the error but also correct it.
Error control in the data link layer is based on automatic repeat request (ARQ). Whenever anerror is detected, specified frames are retransmitted.

Flow control

- Flow control coordinates the amount of data that can be sent before receiving acknowledgement
- It is one of the most important functions of data link layer.
- Flow control is a set of procedures that tells the sender how much data it can transmit before it must wait for an acknowledgement from the receiver.
- Receiver has a limited speed at which it can process incoming data and a limited amount of memory in which to store incoming data.
- Receiver must inform the sender before the limits are reached and request that the transmitter to send fewer frames or stop temporarily.
- Since the rate of processing is often slower than the rate of transmission, receiver has a block of memory (buffer) for storing incoming data until they are processed.

Mechanisms for Flow Control:

1. **Stop and Wait Protocol:** This is the simplest file control protocol in which the sender transmits a frame and then waits for an acknowledgement, either positive or negative, from the receiver before proceeding. If a positive acknowledgement is received, the sender transmits the next packet; else it retransmits the same frame. However, this protocol has one major flaw in it. If a packet or an acknowledgement is completely destroyed in transit due to a noise burst, a deadlock will occur because the sender cannot proceed until it receives an acknowledgement. This problem may be solved using timers on the sender's side. When the frame is transmitted, the timer is set. If there is no response from the receiver within a certain time interval, the timer goes off and the frame may be retransmitted.



- However, generally large block of data split into small frames Called "Fragmentation"
- Advantages are
  - Limited buffer size at receiver
  - Errors detected sooner (when whole frame received)

– On error, retransmission of smaller frames is needed

  - Prevents one station occupying medium for long periods

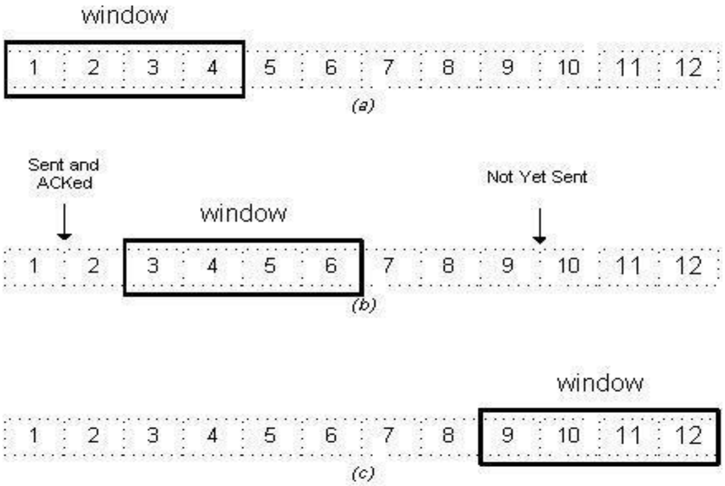
- Channel Utilization is higher when
- the transmission time is longer than the propagation time
- frame length is larger than the bit length of the link

2. **Sliding Window Protocols:**

- Inspite of the use of timers, the stop and wait protocol still suffers from a few drawbacks.
- Firstly, if the receiver had the capacity to accept more than one frame, its resources are being underutilized.
- Secondly, if the receiver was busy and did not wish to receive any more packets, it may delay the acknowledgement. However, the timer on the sender's side may go off and cause an unnecessary retransmission.

These drawbacks are overcome by the **sliding window protocols.**

- In sliding window protocols the sender's data link layer maintains a 'sending window' which consists of a set of sequence numbers corresponding to the frames it is permitted to send.
- Similarly, the receiver maintains a 'receiving window' corresponding to the set of frames it is permitted to accept.
- The window size is dependent on the retransmission policy and it may differ in values for the receiver's and the sender's window.
- The sequence numbers within the sender's window represent the frames sent but as yet not acknowledged.
- Whenever a new packet arrives from the network layer, the upper edge of the window is advanced by one. When an acknowledgement arrives from the receiver the lower edge is advanced by one.
- The receiver's window corresponds to the frames that the receiver's data link layer may accept.
- When a frame with sequence number equal to the lower edge of the window is received, it is passed to the network layer, an acknowledgement is generated and the window is rotated by one.
- If however, a frame falling outside the window is received, the receiver's data link layer has two options.
  - It may either discard this frame and all subsequent frames until the desired frame is received **or**
  - it may accept these frames and buffer them until the appropriate frame is received and then pass the frames to the network layer in sequence.

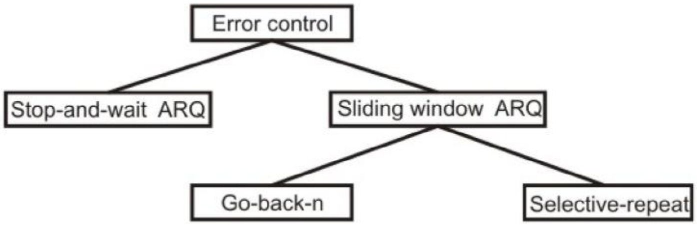


Hence, Sliding Window Flow Control
oAllows transmission of multiple frames

o Assigns each frame a k-bit sequence number
o Range of sequence number is [0…2k-1], i.e., frames are counted modulo 2k.

Error Control Techniques

When an error is detected in a message, the receiver sends a request to the transmitter to retransmit the ill-fated message or packet. The most popular retransmission scheme is known as Automatic-Repeat-Request (ARQ). Such schemes, where receiver asks transmitter to re-transmitif it detects an error, are known as reverse error correction techniques. There exist three popular ARQ techniques, as shown in Fig.

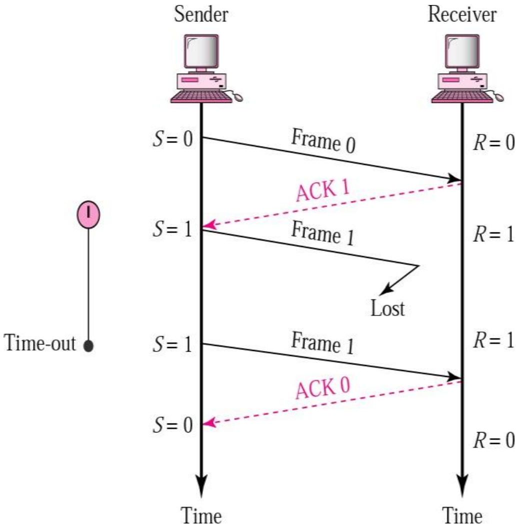


Error control techniques

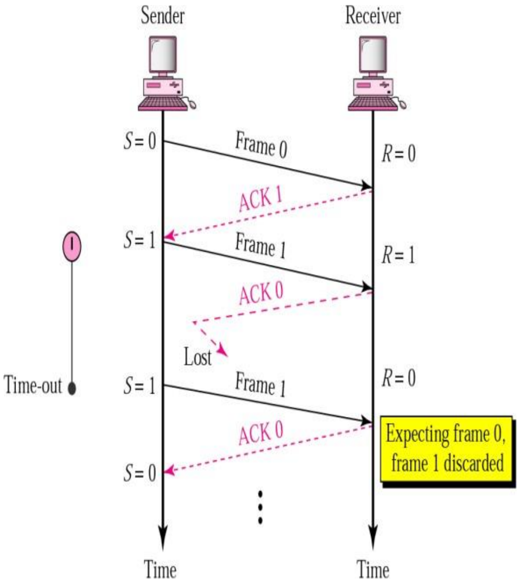
Stop-and-Wait ARQ

In Stop-and-Wait ARQ, which is simplest among all protocols, the sender (say station A) transmits a frame and then waits till it receives positive acknowledgement (ACK) or negative acknowledgement (NACK) from the receiver (say station B). Station B sends an ACK if the frame is received correctly, otherwise it sends NACK. Station A sends a new frame after receiving ACK; otherwise it retransmitsthe old frame, if it receives a NACK. This is illustrated

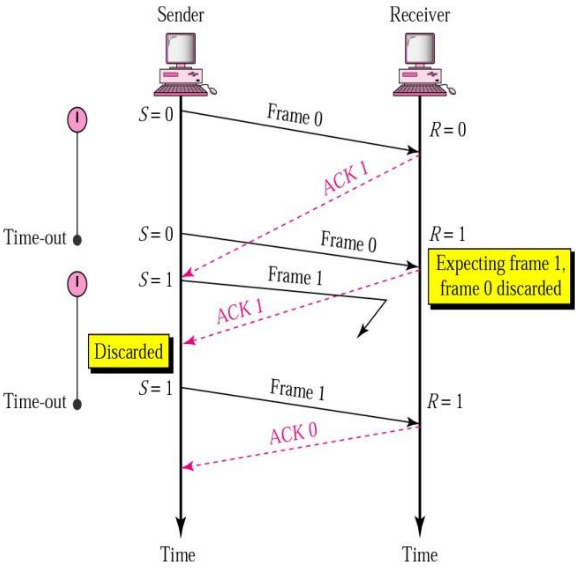
Stop-and-Wait ARQ, lost ACK frame



Stop-and-Wait ARQ, lost ACK frame



Stop-and-Wait, delayed ACK frame



Go-back-N ARQ

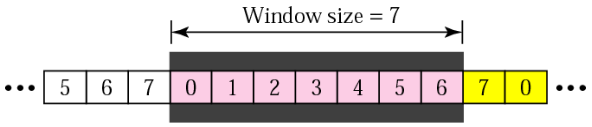
The most popular ARQ protocol is the go-back-N ARQ, where the sender sends the frames continuously without waiting for acknowledgement. That is why it is also called as *continuous ARQ*. As the receiver receives the frames, it keeps on sending ACKs or a NACK, in case a frameis incorrectly received. When the sender receives a NACK, it retransmits the frame in error plus all the succeeding frames. Hence, the name of the protocol is go-back-N ARQ. If a frame is lost, the receiver sends NAK after receiving the next frame. In case there is long delay before sendingthe NAK, the sender will resend the lost frame after its timer times out. If the ACK frame sent bythe receiver is lost, the sender resends the frames after its timer times out.

- We can send up to W frames before worrying about ACKs.
- We keep a copy of these frames until the ACKs arrive.
- This procedure requires additional features to be added to Stop-and-Wait ARQ.
- Frames from a sender are numbered sequentially.
- We need to set a limit since we need to include the sequence number of each frame in the header.
- If the header of the frame allows m bits for sequence number, the sequence numbers range from 0 to $2^{m} – 1$. for m = 3, sequence numbers are: 1, 2, 3, 4, 5, 6, 7.
- We can repeat the sequence number.

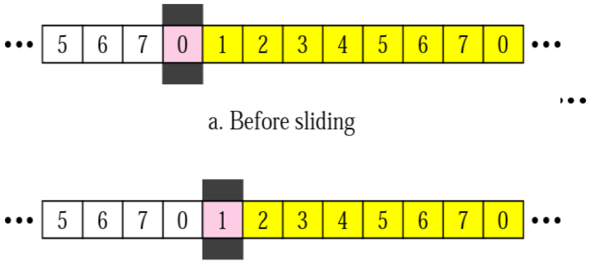
- Sequence numbers are: 0, 1, 2, 3, 4, 5, 6, 7, 0, 1, 2, 3, 4, 5, 6, 7, 0, 1, …

Sender Sliding Window

Receiver Sliding Window



a. Before sliding

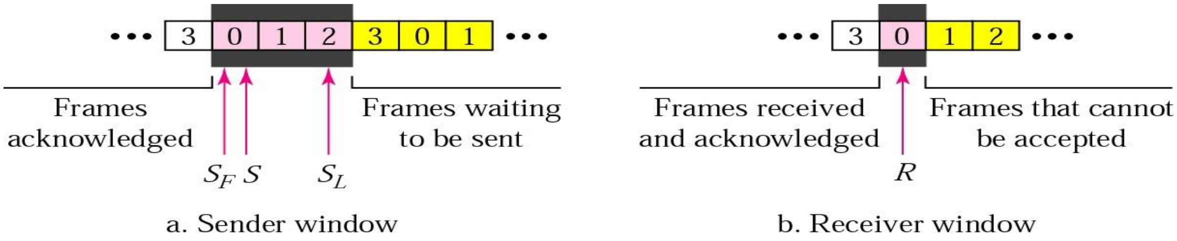


a. Before sliding

b. After sliding

Control Variables

- Sender has 3 variables: S, S_F, and S_L
- S holds the sequence number of recently sent frame
- S_F holds the sequence number of the first frame
- S_L holds the sequence number of the last frame
- Window size is W, where W= S_L- S_F+1
- Receiver only has the one variable, R that holds the sequence number of the frame it expects to receive. If the seq. no. is the same as the value of R, the frame is accepted, otherwise rejected.


3
0
1
2
3
0
1
Frames acknowledged
Frames waiting to be sent
S_F S
S_L
3
0
1
2
Frames received and acknowledged
Frames that cannot be accepted
R

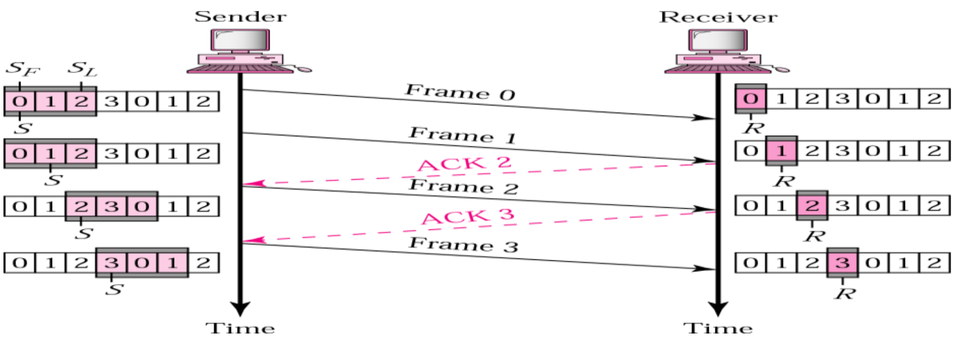
a. Sender window

b. Receiver window

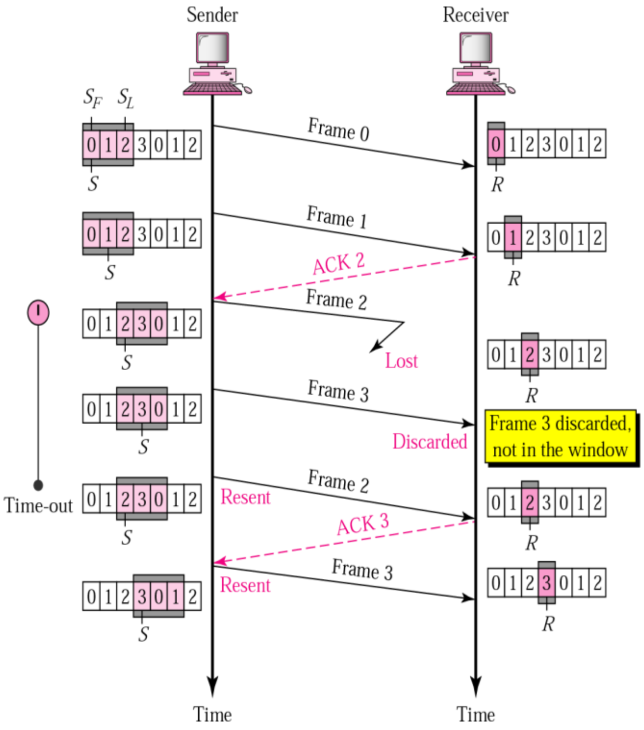
Acknowledgement

- Receiver sends positive ACK if a frame arrived safe and in order.
- If the frames are damaged/out of order, receiver is silent and discards all subsequent frames until it receives the one it is expecting.
- The silence of the receiver causes the timer of the unacknowledged frame to expire.
- Then the sender resends all frames, beginning with the one with the expired timer.
- For example, suppose the sender has sent frame 6, but the timer for frame 3 expires (i.e. frame 3 has not been acknowledged), then the sender goes back and sends frames 3, 4, 5, 6 again. Thus it is called Go-Back-N-ARQ
- The receiver does not have to acknowledge each frame received, it can send one cumulative ACK for several frames.

Go-Back-N ARQ, normal operation- The sender keeps track of the outstanding frames andupdates the variables and windows as the ACKs arrive.



Go-Back-N ARQ, lost frame



Go-Back-N ARQ, damaged/lost/delayed ACK

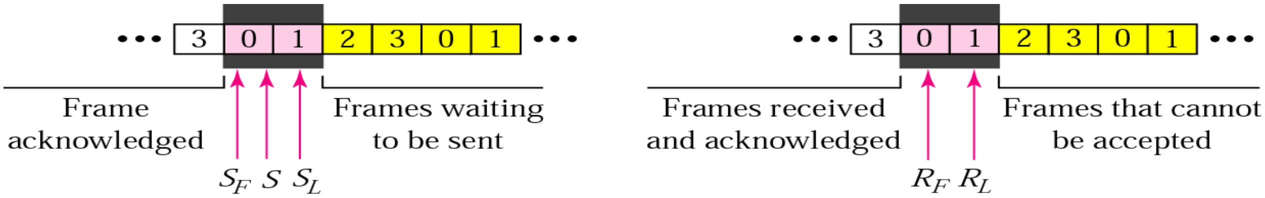
- If an ACK is damaged/lost, we can have two situations:
- If the next ACK arrives before the expiration of any timer, there is no need for retransmission of frames because ACKs are cumulative in this protocol.
- If ACK1, ACK2, and ACk3 are lost, ACK4 covers them if it arrives before the timer expires.
- If ACK4 arrives after time-out, the last frame and all the frames after that are resent.
- Receiver never resends an ACK.
- A delayed ACK also triggers the resending of frames

Size of the sender window must be less than 2^m. Size of the receiver is always 1. If m = 2,window size = $2^m - 1 = 3$.

Selective Repeat ARQ

- Go-Back-N ARQ simplifies the process at the receiver site. Receiver only keeps track of only one variable, and there is no need to buffer out-of-order frames, they are simply discarded.
- However, Go-Back-N ARQ protocol is inefficient for noisy link. It bandwidth inefficient and slows down the transmission.
- In Selective Repeat ARQ, only the damaged frame is resent. More bandwidth efficient but more complex processing at receiver.
- It defines a negative ACK (NAK) to report the sequence number of a damaged frame before the timer expires.

Sender and receiver windows

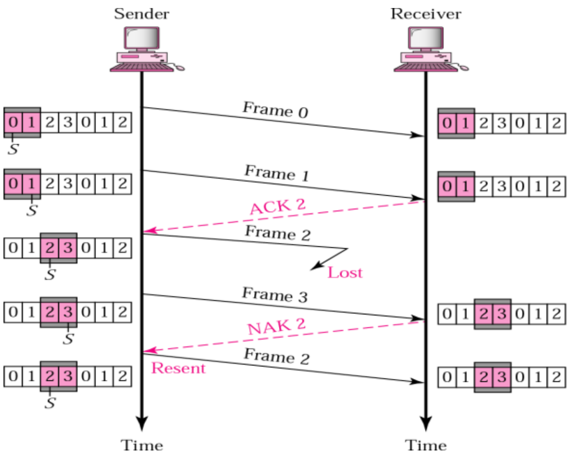


3
0
1
2
3
0
1
Frame acknowledged
Frames waiting to be sent
S_F S S_L
3
0
1
2
3
0
1
Frames received and acknowledged
Frames that cannot be accepted
R_F R_L

a. Sender window

b. Receiver window

Selective Repeat ARQ, lost frame



- Size of the sender and receiver windows must be at most one-half of 2^m. If m = 2, window size should be $2^m/2 = 2$. Fig compares a window size of 2 with a window size of

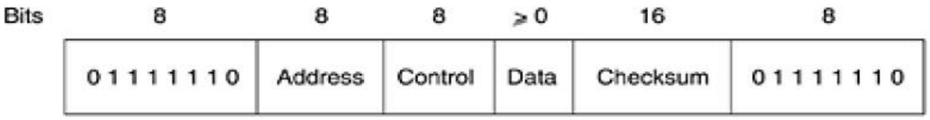
3.   Window size is 3 and all ACKs are lost, sender sends duplicate of frame 0, window of the receiver expect to receive frame 0 (part of the window), so accepts frame 0, as the 1st frame of the next cycle – an **error**.

Example datalink protocols:

High-level Data Link Control Procedures: HDLC

- HDLC is a bit-oriented protocol.
- It was developed by the International Organization for Standardization (ISO).
- It has been so widely implemented because it supports both half-duplex and full-duplex communication lines, point-to-point (peer to peer) and multi-point networks, and switched or non-switched channels.
- HDLC supports several modes of operation, including a simple sliding-window mode forreliable delivery.
- Since Internet provides retransmission at higher levels (i.e., TCP), most Internet applications use HDLC's unreliable delivery mode, Unnumbered Information.
- Other benefits of HDLC are that the control information is always in the same position, and specific bit patterns used for control differ dramatically from those in representing data, which reduces the chance of errors.
- It has also led to many subsets. Two subsets widely in use are Synchronous Data Link Control (SDLC) and Link Access Procedure-Balanced (LAP-B).

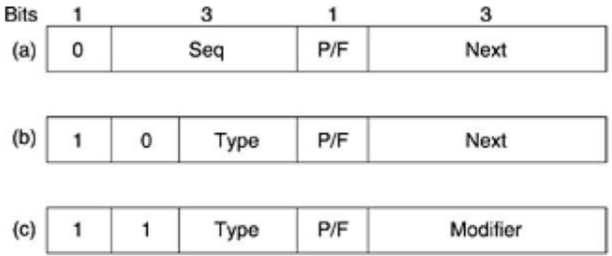
HDLC Frame Structure



Bits	8	8	8	≥ 0	16	8
	01111110	Address	Control	Data	Checksum	01111110

Fig. frame format for bit-oriented protocols

- The Control field is used for sequence numbers, acknowledgements, and other purposes, as discussed below.
- The Data field may contain any information. It may be arbitrarily long
- The Checksum field is a cyclic redundancy
- The frame is delimited with another flag sequence (01111110).
- On idle point-to-point lines, flag sequences are transmitted continuously. The minimum frame contains three fields and totals 32 bits, excluding the flags on either end.
- There are three kinds of frames: Information, Supervisory, and Unnumbered. The contents of the Control field for these three kinds are shown in the following Fig
- The protocol uses a sliding window, with a 3-bit sequence number. Up to seven unacknowledged frames may be outstanding at any instant.
- The Seq field in Fig. (a) is the frame sequence number.
- The Next field is a piggybacked acknowledgement. However, all the protocols adhere to the convention that instead of piggybacking the number of the last frame received correctly, they use the number ofthe first frame not yet received (i.e., the next frame expected). The choice of using the last frame received or the next frame expected is



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

arbitrary; it does not matter which convention is used, provided that it is used consistently.

Fig . Control field of (a) an information frame, (b) a supervisory frame, (c) an unnumbered frame.

- The P/F bit stands for Poll/Final. It is used when a computer (or concentrator) is polling a group of terminals. When used as P, the computer is inviting the terminal to send data.
- All the frames sent by the terminal, except the final one, have the P/F bit set to P. The final one is set to F.
- In some of the protocols, the P/F bit is used to force the other machine to send a Supervisory frame immediately rather than waiting for reverse traffic onto which to piggyback the window information.
- The bit also has some minor uses in connection with the Unnumbered frames.
- The various kinds of Supervisory frames are distinguished by the Type field.
- Type 0 is an acknowledgement frame (officially called RECEIVE READY) used to indicate the next frame expected. This frame is used when there is no reverse traffic to use for piggybacking.
- Type 1 is a negative acknowledgement frame (officially called REJECT). It is used to indicate that a transmission error has been detected. The Next field indicates the first frame in sequence not received correctly (i.e., the frame to be retransmitted). The sender is required to retransmit all outstanding frames starting at Next. This strategy is similar to our protocol 5 rather than our protocol 6.
- Type 2 is RECEIVE NOT READY. It acknowledges all frames up to but not including Next, just as RECEIVE READY does, but it tells the sender to stop sending. RECEIVE NOT READY is intended to signal certain temporary problems with the receiver, such as a shortage of buffers, and not as an alternative to the sliding window flow control. When the condition has been repaired, the receiver sends a RECEIVE READY, REJECT, or certain control frames.
- Type 3 is the SELECTIVE REJECT. It calls for retransmission of only the frame specified. In this sense it is like our protocol 6 rather than 5 and is therefore most useful when the sender's window size is half the sequence space size, or less. Thus, if a receiver wishes to buffer out-of-sequence frames for potential future use, it can force the retransmission of any specific frame using Selective Reject.
- HDLC and ADCCP allow this frame type, but SDLC and LAPB do not allow it (i.e., there is no Selective Reject), and type 3 frames are undefined.
- The third class of frame is the Unnumbered frame. It is sometimes used for control purposes but can also carry data when unreliable connectionless service is called for. The various bit-oriented protocols provide a command, DISC (DISConnect), that allows a machine to announce that it is going down (e.g., for preventive maintenance).
- They also have a command that allows a machine that has just come back on-line to announce its presence and force all the sequence numbers back to zero. This command is called SNRM (Set Normal Response Mode). Unfortunately, "Normal Response Mode" is anything but normal. It is an unbalanced (i.e., asymmetric) mode in which one end of the line is the master and the other the slave.

The DLL in the Internet

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

In practice, point-to-point communication is primarily used in two situations. First, thousands of organizations have one or more LANs, each with some number of hosts (personal computers, user workstations, servers, and so on) along with a router (or a bridge, which is functionally similar).

The second situation in which point-to-point lines play a major role in the Internet is the millions of individuals who have home connections to the Internet using modems and dial-up telephone lines.

For both the router-router leased line connection and the dial-up host-router connection, some point-to-point data link protocol is required on the line for framing, error control, and the other data link layer functions

The Point- to- point protocol –PPP

The Internet needs a point-to-point protocol for a variety of purposes, including router-to-router traffic and home user-to-ISP traffic. This protocol is PPP (Point-to-Point Protocol), which is defined in RFC 1661 and further elaborated on in several other RFCs (e.g., RFCs 1662 and 1663). PPP handles error detection, supports multiple protocols, allows IP addresses to be negotiated at connection time, permits authentication, and has many other features. PPP providesthree features:

1. A framing method that unambiguously delineates the end of one frame and the start of the next one. The frame format also handles error detection.
2. A link control protocol for bringing lines up, testing them, negotiating options, and bringing them down again gracefully when they are no longer needed. This protocol is called LCP (Link Control Protocol). It supports synchronous and asynchronous circuits and byte-oriented and bit-oriented encodings.
3. A way to negotiate network-layer options in a way that is independent of the network layer protocol to be used. The method chosen is to have a different NCP (Network Control Protocol) for each network layer supported.

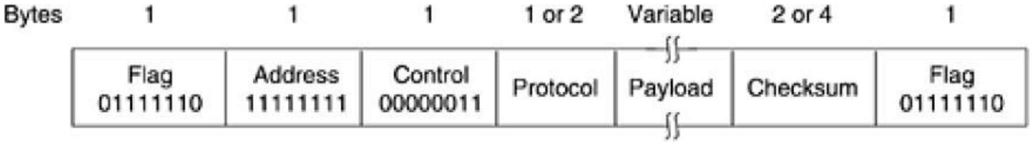
To see how these pieces fit together, let us consider the typical scenario of a home user calling up an Internet service provider to make a home PC a temporary Internet host.

- The PC first calls the provider's router via a modem.
- After the router's modem has answered the phone and established a physical connection, the PC sends the router a series of LCP packets in the payload field of one or more PPP frames. These packets and their responses select the PPP parameters to be used.
- Once the parameters have been agreed upon, a series of NCP packets are sent to configure the network layer. Typically, the PC wants to run a TCP/IP protocol stack, so it needs an IP address.
- There are not enough IP addresses to go around, so normally each Internet provider gets a block of them and then dynamically assigns one to each newly attached PC for the duration of its login session. If a provider owns n IP addresses, it can have up to n machines logged in simultaneously, but its total customer base may be many times that. The NCP for IP assigns the IP address. At this point, the PC is now an Internet host and can send and receive IP packets, just as hardwired hosts can. When the user is finished,

NCP tears down the network layer connection and frees up the IP address. Then LCP shuts down the data link layer connection. Finally, the computer tells the modem to hang up thephone, releasing the physical layer connection.

PPP frame format

The PPP frame format was chosen to closely resemble the HDLC frame format. The major difference between PPP and HDLC is that PPP is character oriented rather than bit oriented.



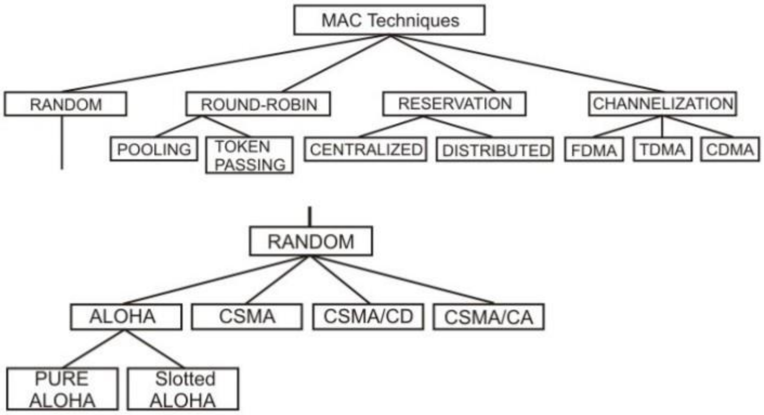
- All PPP frames begin with the standard HDLC flag byte (01111110), which is byte stuffed if it occurs within the payload field.
- Next comes the Address field, which is always set to the binary value 11111111 to indicate that all stations are to accept the frame. Using this value avoids the issue of having to assign data link addresses.
- The Address field is followed by the Control field, the default value of which is 00000011. This value indicates an unnumbered frame.
- In other words, PPP does not provide reliable transmission using sequence numbers and acknowledgements as the default. In noisy environments, such as wireless networks, reliable transmission using numbered mode can be used. The exact details are defined in RFC 1663, but in practice it is rarely used.

- Since the Address and Control fields are always constant in the default configuration, LCP provides the necessary mechanism for the two parties to negotiate an option to just omit them altogether and save 2 bytes per frame.
- The fourth PPP field is the Protocol field. Its job is to tell what kind of packet is in the Payload field. Codes are defined for LCP, NCP, IP, IPX, AppleTalk, and other protocols. Protocols starting with a 0 bit are network layer protocols such as IP, IPX, OSI CLNP, XNS. Those starting with a 1 bit are used to negotiate other protocols.
- These include LCP and a different NCP for each network layer protocol supported. The default size of the Protocol field is 2 bytes, but it can be negotiated down to 1 byte using LCP.
- The Payload field is variable length, up to some negotiated maximum. If the length is not negotiated using LCP during line setup, a default length of 1500 bytes is used. Padding may follow the payload if need be. After the Payload field comes the Checksum field, which is normally 2 bytes, but a 4-byte checksum can be negotiated.
- In summary, PPP is a multiprotocol framing mechanism suitable for use over modems, HDLC bit-serial lines, SONET, and other physical layers. It supports error detection, option negotiation, header compression, and, optionally, reliable transmission using an HDLC-type frame format.

THE MEDIUM ACCESS CONTROL SUBLAYER

- A network of computers based on multi-access medium requires a protocol for effective sharing of the media.
- As only one node can send or transmit signal at a time using the broadcast mode, the main problem here is how different nodes get control of the medium to send data, that is “who goes next?”. The protocols used for this purpose are known as Medium Access Control (MAC) techniques.
- The key issues involved here are - Where and how the control is exercised.
- ‘Where’ refers to whether the control is exercised in a centralised or distributed manner.
- In a centralized system a master node grants access of the medium to other nodes. A centralized scheme has a number of advantages as mentioned below:
  - Greater control to provide features like priority, overrides, and guaranteed bandwidth.
  - Simpler logic at each node.
  - Easy coordination. Although this approach is easier to implement, it is vulnerableto the failure of the master node and reduces efficiency.
- On the other hand, in a distributed approach all the nodes collectively perform a medium access control function and dynamically decide which node to be granted access. This approach is more reliable than the former one.
- ‘How’ refers to in what manner the control is exercised. It is constrained by the topology and trade off between cost-performance and complexity.

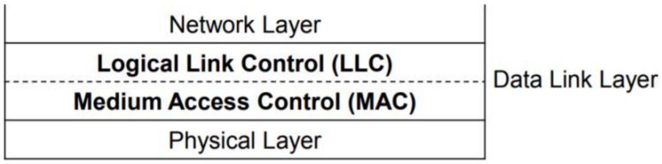
Various approaches for medium access control are shown in Fig.



In broadcast networks, several stations share a single communication channel. The major issuein these networks is, which station, should transmit data at a given time. This process of deciding the turn of different stations is known as Channel Allocation. To coordinate the accessto the channel, multiple access protocols are required. All these protocols belong to the MAC sublayer.

Data Link layer is divided into two sublayers:

- Logical Link Control (LLC)- LLC is responsible for error control
- Medium Access Control (MAC) • & flow control. MAC is responsible for multiple access resolutions.



Channel Allocation Problem

In broadcast networks, single channel is shared by several stations. This channel can be allocated toonly one transmitting user at a time. There are two different methods of channel allocations:

- Static Channel Allocation
- Dynamic Channel Allocation

Static Channel Allocations

In this method, a single channel is divided among various users either on the basis of frequencyor on the basis of time. It either uses

- FDM (Frequency Division Multiplexing) - In FDM, fixed frequency is assigned to each user or
- TDM (Time Division Multiplexing). ,whereas, in TDM, fixed time slot is assigned to each user.

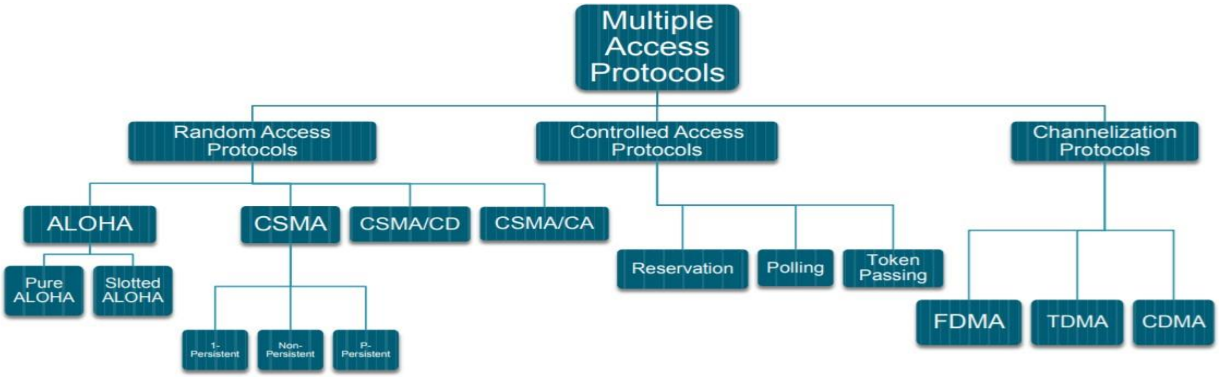
Dynamic Channel Allocation

In this method, no user is assigned fixed frequency or fixed time slot. All users are dynamicallyassigned frequency or time slot, depending upon the requirements of the user.

Multiple Access Protocols

Many protocols have been defined to handle the access to shared link. These protocols areorganized in three different groups.

- Random Access Protocols
- Controlled Access Protocols
- Channelization Protocols



Random Access Protocols

It is also called Contention Method. In this method, there is no control station. Any station can send the data. The station can make a decision on whether or not to send data. This decision depends on the state of the channel, i.e. channel is busy or idle. There is no scheduled time for a stations to transmit. They can transmit in random order.

There is no rule that decides which station should send next. If two stations transmit at the same time, there is collision and the frames are lost. The various random access methods are:

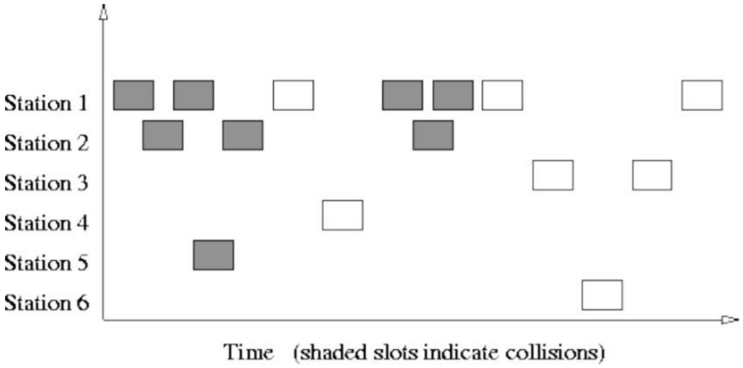
- ALOHA
- CSMA (Carrier Sense Multiple Access)
- CSMA/CD (Carrier Sense Multiple Access with Collision Detection)
- CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance)

ALOHA

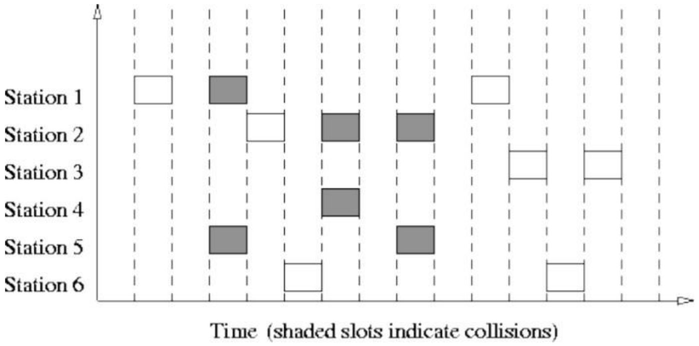
ALOHA was developed at University of Hawaii in early 1970s by Norman Abramson. It was used for ground based radio broadcasting. In this method, stations share a common channel. When two stations transmit simultaneously, collision occurs and frames are lost. There are two different versions of ALOHA:

- Pure ALOHA
- Slotted ALOHA

Pure ALOHA: In pure ALOHA, stations transmit frames whenever they have data to send. When two stations transmit simultaneously, there is collision and frames are lost. In pure ALOHA, whenever any station transmits a frame, it expects an acknowledgement from the receiver. If acknowledgement is not received within specified time, the station assumes that the frame has been lost. If the frame is lost, station waits for a random amount of time and sends it again. This waiting time must be random, otherwise, same frames will collide again and again. Whenever two frames try to occupy the channel at the same time, there will be collision and boththe frames will be lost. If first bit of a new frame overlaps with the last bit of a frame almost finished, both frames will be lost and both will have to be transmitted.

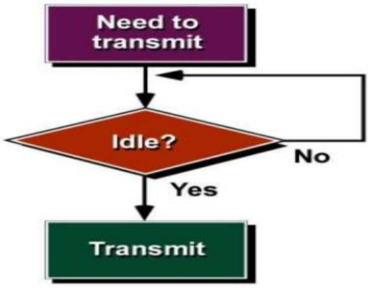


Slotted ALOHA : Slotted ALOHA was invented to improve the efficiency of pure ALOHA. In slotted ALOHA, time of the channel is divided into intervals called slots. The station can send a frame only at the beginning of the slot and only one frame is sent in each slot. If any station is not able to place the frame onto the channel at the beginning of the slot, it has to wait until the next time slot. There is still a possibility of collision if two stations try to send at the beginning of the same time slot.



Carrier Sense Multiple Access (CSMA)

CSMA was developed to overcome the problems of• ALOHA i.e. to minimize the chances of collision. CSMA is based on the principle of “carrier sense”. The station sense the carrier or channel before transmitting a frame. It means the station checks whether the channel is idle or busy. The chances of collision reduces to a great extent if a station checks the channel before trying to use it.



The chances of collision still exists because of propagation delay. The frame transmitted by onestation takes some time to reach the other station. In the meantime, other station may sense the channel to be idle and transmit its frames. This results in the collision. There are three differenttypes of CSMA protocols:

- 1-Persistent CSMA
- Non-Persistent CSMA
- P-Persistent CSMA

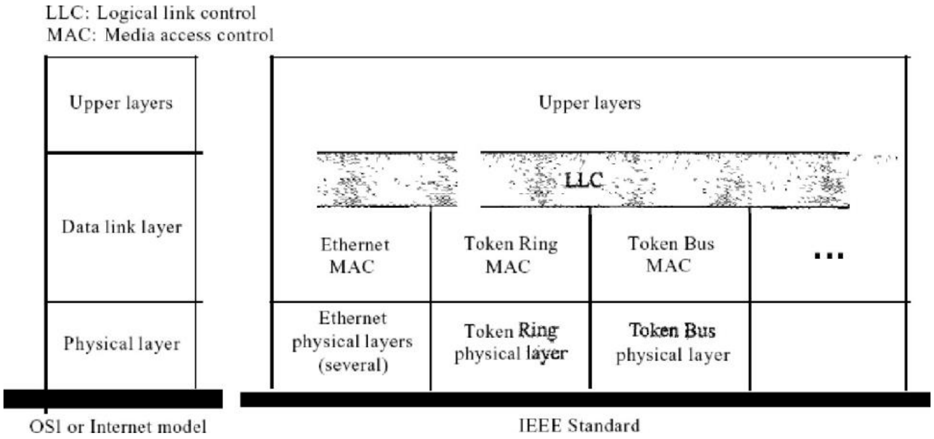
1- **Persistent CSMA**: In this method, station that wants to transmit data, continuously senses the channel to check whether he channel is idle or busy. If the channel is busy, station waits until it becomes idle. When the station detects an idle channel, it immediately transmits the frame. This method has the highest chance of collision because two or more stations may find channel to be idle at the same time and transmit their frames.

Non-Persistent CSMA: A station that has a frame to send senses the channel. If the channel is idle, it sends immediately. If the channel is busy, it waits a random amount of time and then senses the channel again. It reduces the chance of collision because the stations wait for a random amount of time. It is unlikely that two or more stations will wait for the same amount of time and will retransmit at the same time.

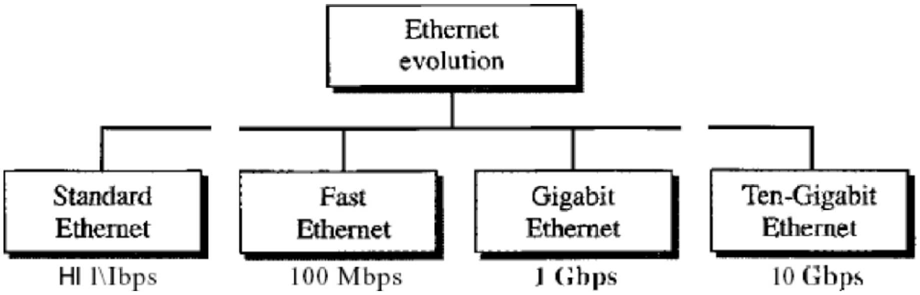
P-Persistent CSMA: In this method, the channel has time slots such that the time slot duration is equal to or greater than the maximum propagation delay time. When a station is ready to send, it senses the channel. If the channel is busy, station waits until next slot. If the channel is idle, it transmits the frame. It reduces the chance of collision and improves the efficiency of the network.

IEEE 802

IEEE Project 802 has created a sublayer called media access control that defines the specific access method for each LAN. For example, it defines *CSMA/CD* as the media access method forEthernet LANs and the token passing method for Token Ring and Token Bus LANs. Also, part of the framing function is also handled by the MAC layer. In contrast to the LLC sublayer, the MAC sublayer contains a number of distinct modules; each defines the access method and the framing format specific to the corresponding LAN protocol. IEEE has also created several physical layer standards for different LAN protocols.



IEEE 802.3 Ethernet



Standard Ethernet is also known as CSMA with Collision Detection (CSMA/CD)
Ethernet access protocol:

– 1-Persistent CSMA/CD with Binary Exponential Backoff Algorithm

Layers specified by 802.3

– Ethernet Physical Layer

– Ethernet Medium Access (MAC) Sublayer

Ethernet architecture can be divided into two layers:

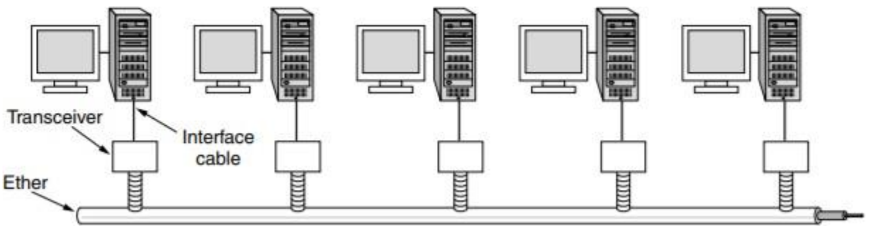
- Physical layer: this layer takes care of following functions.

  - Encoding and decoding
  - Collision detection
  - Carrier sensing
  - Transmission and receipt

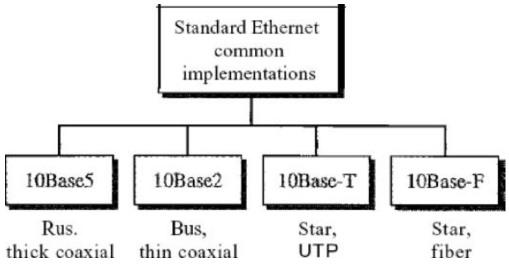
- Data link layer: Following are the major functions of this layer.

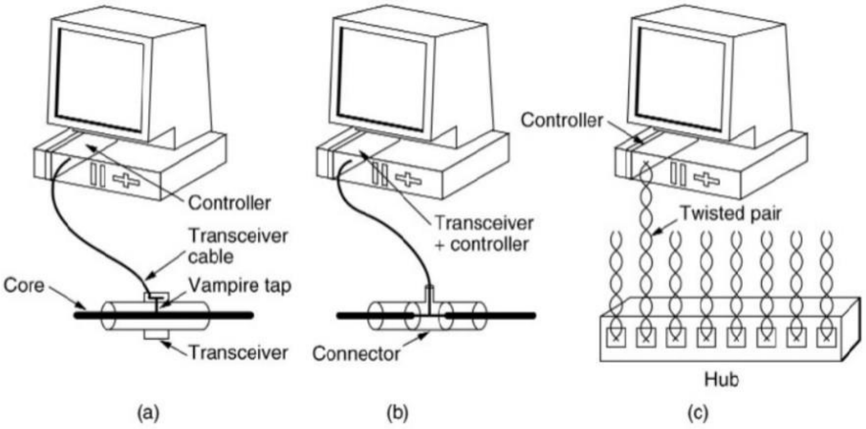
  - Station interface
  - Data Encapsulation /Decapsulation
  - Link management
  - Collision Management

Architecture of classic Ethernet: Classic Ethernet snaked around the building as a single long cable to which all the computers were attached. This architecture is shown in Fig.. The first variety, popularly called **thick Ethernet**, resembled a yellow garden hose, with markings every
2.5 meters to show where to attach computers. It was succeeded by **thin Ethernet**, which bent more easily and made connections using industry-standard BNC connectors. Thin Ethernet was much cheaper and easier to install, but it could run for only 185 meters per segment (instead of 500 m with thick Ethernet), each of which could handle only 30 machines (instead of 100). Each version of Ethernet has a maximum cable length per segment (i.e., unamplified length) over which the signal will propagate. To allow larger networks, multiple cables can be connected by repeaters. A repeater is a physical layer device that receives, amplifies (i.e., regenerates), and retransmits signals in both directions.



Architecture of classic Ethernet



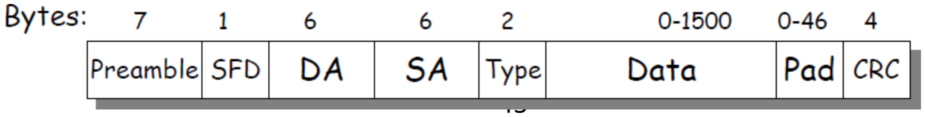


(a) 10Base5 (b)10Base2 (c)10Base-T

Three kinds of Ethernet cablingEthernet uses Manchester encoding.

The Ethernet MAC Sublayer :

Ethernet frame format:



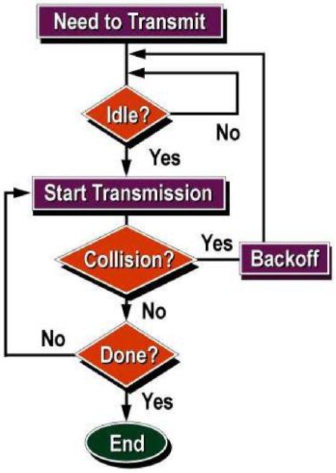
Bytes:	7	1	6	6	2	0-1500	0-46	4
	Preamble	SFD	DA	SA	Type	Data	Pad	CRC

1. Preamble: trains clock-recovery circuits
2. Start of Frame Delimiter: indicates start of frame
3. Destination Address: 48-bit globally unique address assigned by manufacturer.

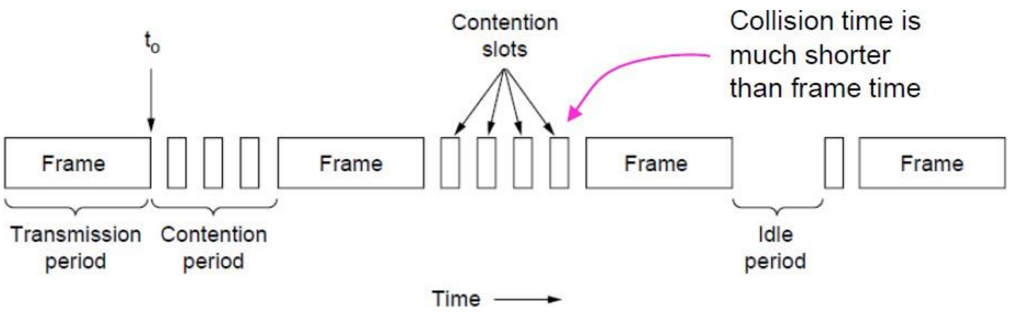
1b: unicast/multicast 1b:
local/global address

4. Type: Indicates protocol of encapsulated data (e.g. IP = 0x0800)
5. Pad: Zeroes used to ensure minimum frame length
6. Cyclic Redundancy Check: check sequence to detect bit errors.

In CSMA/CD protocol, the station senses the channel before transmitting the frame. If the channel is busy, the station waits. Additional feature in CSMA/CD is that the stations can detect collisions. The stations abort their transmission as soon as they detect collision. This feature is not present in CSMA. The stations continue to transmit even though they find that collision has occurred.



- Whenever the channel is found idle, the station does not transmit immediately.
- It waits for a period of time called Interframe Space (IFS).
- When channel is sensed idle, it may be possible that some distant station may have already started transmitting.
- Therefore, the purpose of IFS time is to allow this transmitted signal to reach its destination.
- If after this IFS time, channel is still idle, the station can send the frames.
- Contention window is the amount of time divided into slots.
- Station that is ready to send chooses a random number of slots as its waiting time.
- The number of slots in the window changes with time.
- It means that it is set of one slot for the first time, and then doubles each time the station cannot detect an idle channel after the IFS time.
- In contention window, the station needs to sense the channel after each time slot.
- Despite all the precautions, collisions may occur and destroy the data.
- Positive acknowledgement and the time-out timer helps guarantee that the receiver has received the frame.

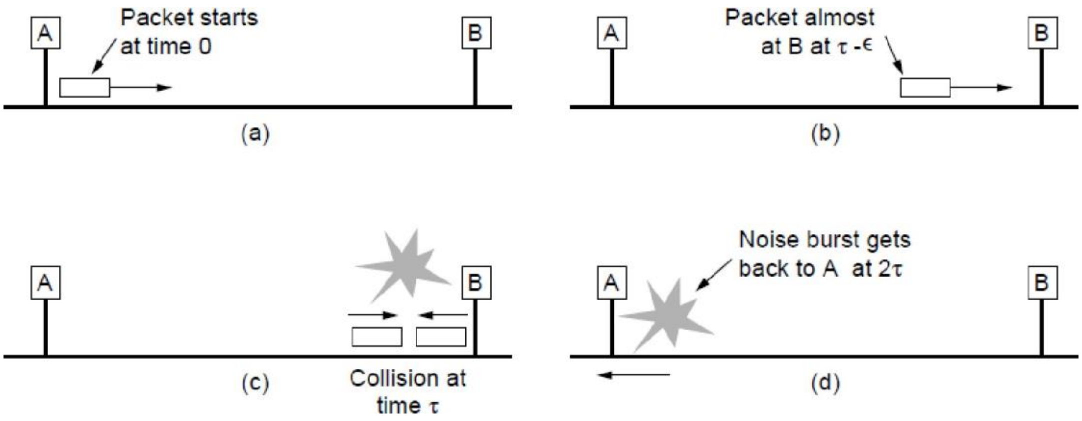


Other important issues -There are some more important issues, which are briefly discussed below. **Inter-frame Gap:** There is mandatory requirement of 9.6 ms interval between two frames to enable other stations wishing to transmit to take over after a frame transmission is over. In other words, a 96 bit-time delay is provided between frame transmissions.

How are collisions detected? A station sends frame and continues to sense the medium. If the signalstrength sensed by a station exceeds the normal signal strength, it is treated as collision detection.

What the station does? The transmitting station sends a jamming signal after collision is detected. 32-bit jam signal: 10101010 --- 10101010
48-bit jam signal: 10101010 --- 10101010

The jam signal serves as a mechanism to cause non-transmitting stations to wait until the jam signal ends. **Minimum Frame Size:** A frame must take more than 2τ time to send, where τ is the propagation time for preventing the situation that the sender incorrectly concludes that the frame was successfully sent. This slot time is 51.2µsec corresponding to 512 bit = 64 bytes. Therefore the minimum frame length is 64 bytes (excluding preamble), which requires that the data field must have a minimum sizeof 46 bytes.



Binary Exponential Backoff:

It refers to a collision resolution mechanism used in random access MAC protocols. Thisalgorithm is used in Ethernet (IEEE 802.3) wired LANs.

- In Ethernet networks, this algorithm is commonly used to schedule retransmissions after collisions.

- After a collision, time is divided into discrete slots whose length is equal to 2τ, where τ is the maximum propagation delay in the network.

- The reason for this choice is that 2 τ is the minimum amount of time a source needs to listen to the channel to always detect a collision.

- The stations involved in the collision randomly pick an integer from the set {0,1}. This set is called the contention window. If the sources collide again because they picked the same integer, the contention window size is doubled and it becomes {0,1,2,3}. Now the sources involved in the second collision randomly pick an integer from the set {0,1,2,3} and wait that number of slot times before trying again. Before they try to transmit, they listen to the channel and transmit only if the channel is idle. This causes the source which picked the smallest integer in the contention window to succeed in transmitting its frame.

- In general, after collisions, a random number between 0 and is chosen.

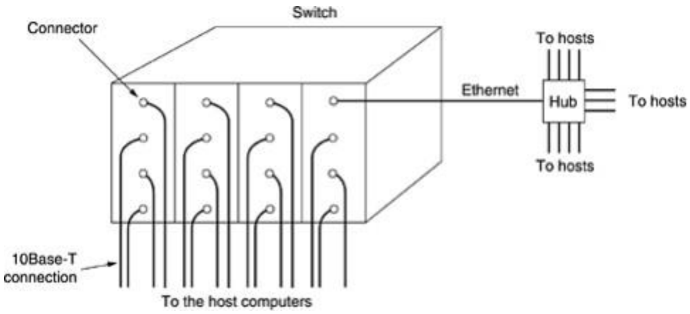
- After a station detects collision, it aborts its transmission in the slot duration itself in which it started transmitting.

- In Ethernet, the doubling of the contention window stops after 10 collisions and the contention window remains {0,1,...,1023}.

- After 16 collisions , the process is aborted and the source stops trying.

Switched Ethernet

As more and more stations are added to an Ethernet, the traffic will go up. Eventually, the LAN will saturate. One way out is to go to a higher speed, say, from 10 Mbps to 100 Mbps. But with the growth of multimedia, even a 100-Mbps or 1-Gbps Ethernet can become saturated. There is an additional way to deal with increased load: switched Ethernet, as shown in Fig. The heart of this system is a switch containing a high-speed backplane and room for typically 4 to 32 plug-in line cards, each containing oneto eight connectors. Most often, each connector has a 10Base-T twisted pair connection to a single hostcomputer.



When a station wants to transmit an Ethernet frame, it outputs a standard frame to the switch. The plug-in card getting the frame may check to see if it is destined for one of the other stations connected to the same card. If so, the frame is copied there. If not, the frame is sent over the high-speed backplane to the destination station's card. The backplane typically runs at many Gbps, using a proprietary protocol.

What happens if two machines attached to the same plug-in card transmit frames at the same time? It depends on how the card has been constructed. One possibility is for all the ports on the card to be wired together to form a local on-card LAN. Collisions on this on-card LAN will be detected and handled the same as any other collisions on a CSMA/CD network—with retransmissions using the binary exponential backoff algorithm. With this kind of plug-in card, only one transmission per card is possible at any instant, but all the cards can be transmitting in parallel. With this design, each card forms its own collision domain, independent of the others. With only one station per collision domain, collisions are impossible and performance is improved. With the other kind of plug-in card, each input port is buffered, so incoming framesare stored in the card's on-board RAM as they arrive. This design allows all input ports to receive (and transmit) frames at the same time, for parallel, full-duplex operation, something not possiblewith CSMA/CD on a single channel. Once a frame has been completely received, the card can then check to see if the frame is destined for another port on the same card or for a distant port. In the former case, it can be transmitted directly to the destination. In the latter case, it must be transmitted over the backplane to the proper card. With this design, each port is a separate collision domain, so collisions do not occur. The total system throughput can often be increased by an order of magnitude over 10Base5, which has a single collision domain for the entire system. Since the switch just expects standard Ethernet frames on each input port, it is possible to use some of the ports as concentrators. In the above Fig., the port in the upper-right corner is connected not to a single station, but to a 12-port hub. As frames arrive at the hub, they contend for the ether in the usual way, including collisions and binary backoff. Successful frames make itto the switch and are treated there like any other incoming frames: they are switched to the correct output line over the high-speed backplane. Hubs are cheaper than switches, but due to falling switch prices, they are rapidly becoming obsolete. Nevertheless, legacy hubs still exist.

Fast Ethernet (802.3u)

- 100 Mbps bandwidth
- Uses same CSMA/CD media access protocol and packet format as in Ethernet.
- 100BaseTX (UTP) and 100BaseFX (Fiber) standards
- Physical media :-
    - 100 BaseTX - UTP Cat 5e
    - 100 BaseFX - Multimode / Singlemode Fiber
- Full Duplex/Half Duplex operations.
- Provision for Auto-Negotiation of media speed:

10 Mbps or 100Mbps (popularly available for copper media only).

- Maximum Segment Length
- 100 Base TX - 100 m

- 100 Base FX - 2 Km (Multimode Fiber)
- 100 Base FX - 20 km (Singlemode Fiber)

IEEE 802.5: Token Ring Network

Developed by IBM, adopted by IEEE as 802.5 standard

- Token rings latter extended to FDDI (Fiber Distributed Data Interface) and 802.17 (Resilient Packet Ring) standards
- Nodes connected in a ring – Data always flows in one direction around ring – Like Ethernet, all nodes see all frames, and protocol is necessary

Tokens

- Token ring named because token (a special sequence of bits) is passed around the ring
- Each node receives and retransmits token
- A node with something to transmit can take token off ring and insert frame
- Destination node copies frame, but sends on
- When sender receives frame, node drops it and reinserts the token
- All nodes get chance to transmit (round-robin)

Physical properties

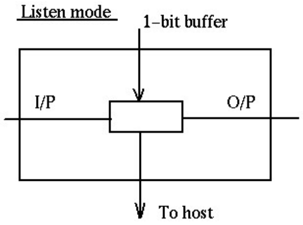
- Nodes connected to ring using electromechanical relay
- Prevents node failure from crashing ring
- Several relays often packed into one multistation access unit (MSAU)
- Provide easy addition and removal of nodes
- Data rate either 4 or 16 Mbps
- Uses Manchester encoding
- IBM Token Rings can have up to 260 stations per ring, 802.5 up to 250
- Physical medium for IBM is twisted pair, not specified for 802.5

Media Access Control

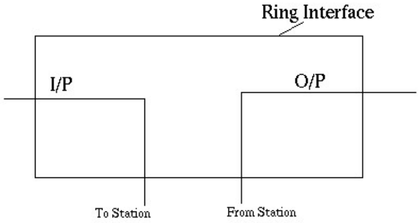
- Each node includes receiver, transmitter, one or more bits of memory between
- Ring must contain enough memory to hold entire ring
    - 802.5 token is 24 bits long, so if each station can hold only 1 bit, must have at least 24 stations, or more than one bit-time distance between stations
    - Alternative is monitor station that adds delay

Modes of Operation

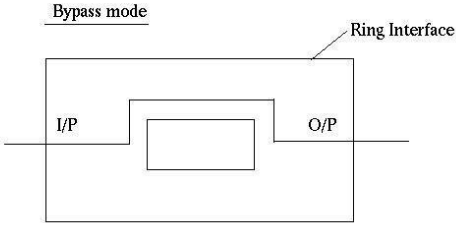
1. **Listen Mode:** In this mode the node listens to the data and transmits the data to the next node. In this mode there is a one-bit delay associated with the transmission.



2.**Transmit Mode:** In this mode the node just discards the any data and puts the data onto the network.



3.**By-pass Mode:** In this mode reached when the node is down. Any data is just bypassed. There is no one-bit delay in this mode



Seizing Token -When node wants to send, modifies one bit in second byte of token – thischanges first two bytes into transmission preamble

- Station then inserts one or more packets onto ring
- Each packet contains destination address, can also contain multicast or broadcast address

Receiving- If node recognizes address, copies data from transceiver into buffer, but still forwardsit

- Sending station is responsible for removing packet from the ring

- Station might be draining first part of packet from ring while transmitting end of packet, if ring is small enough

Transmission Limits: Control how long sender can transmit (token hold time, THT)

- The more bits a node can send, the better the ring utilization, but the poorer the response time for other nodes
- In 802.5, THT defaults to 10 ms – Sender responsible for knowing how long it has already held token, how long next packet will take to transmits
- Nodes also compute token rotation time (TRT)
    - – TRT ≤ ActiveNodes * THT + RingLatency
    - – ActiveNodes are # that have data to send
    - – RingLatency is time to send token around ring if no node has anything to send

Reliable Delivery: Receiver sets the A bit in packet trailer if it recognizes itself as addressednode

• Sets the C bit in the trailer when it finishes copying packet into its buffer

• Sender can check for missing A or C bits when it gets packet back to verify that sender was there and was able to buffer entire packet

Priority: Token includes 3-bit priority field

– Each device assigns priority to each packet it needs to send

– Only captures token if packet priority >= token's

• Frame header includes 3 reservation bits

– Node X can set reservation bits to priority of its packet if bits don't already have >= value

– Token holder escalates priority to that value when it releases token

– Node X must reset priority to old value when done

Ring Maintenance: Each token ring has a monitor that oversees the ring. Among the monitor's responsibilities are seeing that the token is not lost, taking action when the ring breaks, cleaningthe ring when garbled frames appear and watching out for orphan frames. An orphan frame occurs when a station transmits a short frame in it's entirety onto a long ring and then crashes oris powered down before the frame can be removed. If nothing is done, the frame circulates indefinitely.

- **Detection of orphan frames:** The monitor detects orphan frames by setting the monitor bit in the Access Control byte whenever it passes through. If an incoming frame has this bit set, something is wrong since the same frame has passed the monitor twice. Evidentlyit was not removed by the source, so the monitor drains it.

- **Lost Tokens:** The monitor has a timer that is set to the longest possible tokenless interval : when each node transmits for the full token holding time. If this timer goes off, themonitor drains the ring and issues a fresh token.
- **Garbled frames:** The monitor can detect such frames by their invalid format or checksum, drain the ring and issue a fresh token.

Any node can become monitor

– Procedure defined to elect monitor when ring first connected or monitor fails

• Monitor periodically announces its presence

• If this is missed, another station will send claim token to attempt to become monitor

– Tie broken by rule like "highest address wins"

– If sender gets claim back, it becomes monitor

Ring monitor: Monitor can insert delay into ring if needed

• Makes sure that there is always a token

– Timeout after NumStations*THT+RingLatency

– Generate new token

• Removes packets if sender dies

– Monitor bit in header set first time packet passes monitor

– Packet with bit set is removed by monitor

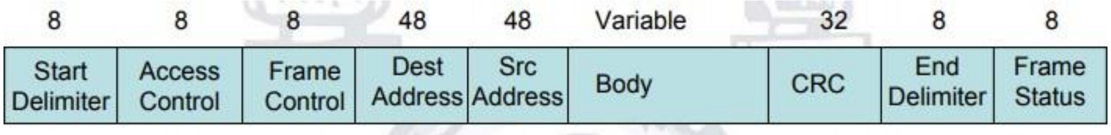
• Also detects dead stations

– Catch more subtle errors than MSAU switch

– Send beacon packet to suspected dead node

– Instruct MSAU to bypass malfunctioning node

Frame format:

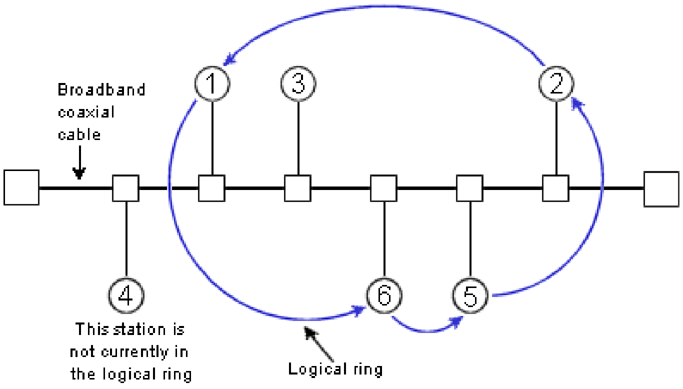


8	8	8	48	48	Variable	32	8	8
Start Delimiter	Access Control	Frame Control	Dest Address	Src Address	Body	CRC	End Delimiter	Frame Status

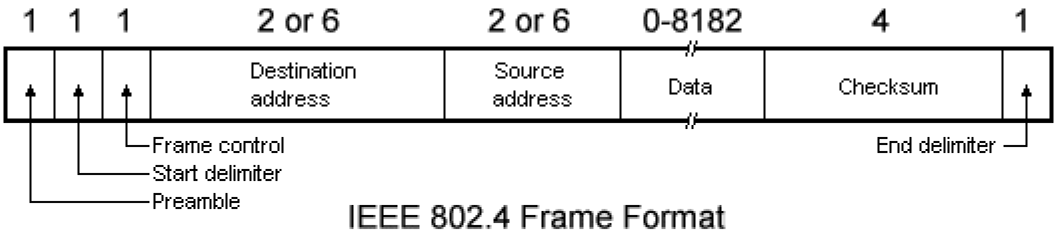
- Start and end delimiters use invalid Manchester codes
- Access control includes priority and reservation bits
- Frame control indicates higher level protocol
- Addresses identical to Ethernet addresses – Standard allows for 16-bit addresses, but they are not typically used
- Frame status byte includes A and C bits

IEEE 802.4: Token Bus Network

In this system, the nodes are physically connected as a bus, but logically form a ring with tokens passed around to determine the turns for sending. It has the robustness of the 802.3 broadcast cable and the known worst case behavior of a ring. The structure of a token busnetwork is as follows:



Frame Structure



IEEE 802.4 Frame Format

802.4 frame has the following fields:

- Preamble: The Preamble is used to synchronize the receiver's clock.

- Starting Delimiter (SD) and End Delimiter (ED): The Starting Delimiter and Ending Delimiter fields are used to mark frame boundaries. Both of them contain analog encoding of symbols other than 1 or 0 so that they cannot occur accidentally in the user data. Hence no length field is needed.
- Frame Control (FC): This field is used to distinguish data frames from control frames. For data frames, it carries the frame's priority as well as a bit which the destination can set as an acknowledgement. For control frames, the Frame Control field is used to specify the frame type. The allowed types include token passing and various ring maintenance frames.
- Destination and Source Address: The Destination and Source address fields may be 2 bytes (for a local address) or 6 bytes (for a global address).
- Data: The Data field carries the actual data and it may be 8182 bytes when 2 byte addresses are used and 8174 bytes for 6 byte addresses.
- Checksum: A 4-byte checksum calculated for the data. Used in error detection.

Ring Maintenance:

Mechanism:

When the first node on the token bus comes up, it sends a **Claim_token** packet to initialize the ring. If more than one station send this packet at the same time, there is a collision. Collision is resolved by a contention mechanism, in which the contending nodes send random data for 1, 2, 3and 4 units of time depending on the first two bits of their address. The node sending data for thelongest time wins. If two nodes have the same first two bits in their addresses, then contention is done again based on the next two bits of their address and so on.

After the ring is set up, new nodes which are powered up may wish to join the ring. For this a node sends **Solicit_successor_1** packets from time to time, inviting bids from new nodes to join the ring. This packet contains the address of the current node and its current successor, and asks for nodes in between these two addresses to reply. If more than one nodes respond, there will be collision. The node then sends a **Resolve_contention** packet, and the contention is resolved using a similar mechanism as described previously. Thus at a time only one node gets to enter thering. The last node in the ring will send a **Solicit_successor_2** packet containing the addresses ofit and its successor. This packet asks nodes not having addresses in between these two addresses to respond.

A question arises that how frequently should a node send a Solicit_successor packet? If it is sent too frequently, then overhead will be too high. Again if it is sent too rarely, nodes will have to wait for a long time before joining the ring. If the channel is not busy, a node will send a Solicit_successor packet after a fixed number of token rotations. This number can be configured by the network administrator. However if there is heavy traffic in the network, then a node woulddefer the sending of bids for successors to join in.

There may be problems in the logical ring due to sudden failure of a node. What happens when a node goes down along with the token? After passing the token, a node, say node A, listens to the channel to see if its successor either transmits the token or passes a frame. If neither happens, it resends a token. Still if nothing happens, A sends a **Who_follows** packet, containing the address of the down node. The successor of the down node, say node C, will now respond with a **Set_successor** packet, containing its own address. This causes A to set its successor node to C, and the logical ring is restored. However, if two successive nodes go down suddenly, the ringwill be dead and will have to be built afresh, starting from a **Claim_token** packet.

When a node wants to shutdown normally, it sends a **Set_successor** packet to its predecessor,naming its own successor. The ring then continues unbroken, and the node goes out of the ring.

The various control frames used for ring maintenance are shown below:

Frame Control Field	*Name*	*Meaning*
0	Claim_token	Claim token during ring maintenance
1	Solicit_successor_1	Allow stations to enter the ring
10	Solicit_successor_2	Allow stations to enter the ring
11	Who_follows	Recover from lost token
100	Resolve_contention	Used when multiple stations want to enter
1000	Token	Pass the token
1100	Set_successor	Allow the stations leave the ring

Priority Scheme:

Token bus supports four distinct priority levels: 0, 2, 4 and 6.

0 is the lowest priority level and 6 the highest. The following times are defined by the token bus:

- THT: Token Holding Time. A node holding the token can send priority 6 data for a maximum of this amount of time.
- TRT_4: Token Rotation Time for class 4 data. This is the maximum time a token can take to circulate and still allow transmission of class 4 data.
- TRT_2 and TRT_0: Similar to TRT_4.

When a station receives data, it proceeds in the following manner:

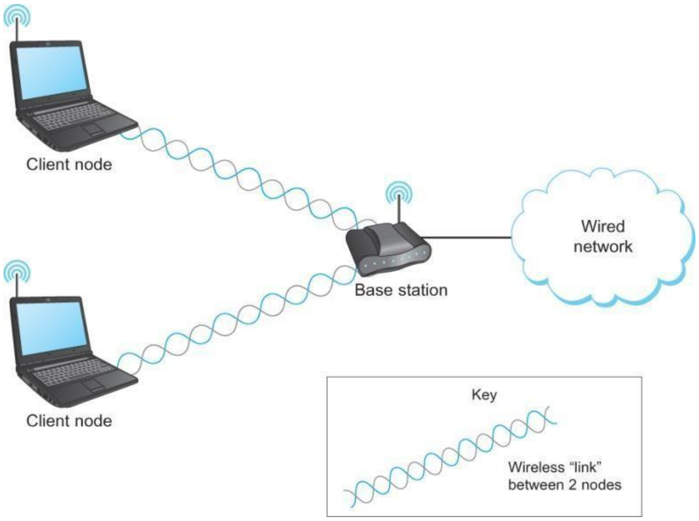
- It transmits priority 6 data for at most THT time, or as long as it has data.
- Now if the time for the token to come back to it is less than TRT_4, it will transmit priority 4 data, and for the amount of time allowed by TRT_4. Therefore the maximum time for which it can send priority 4 data is= Actual TRT - THT - TRT_4
- Similarly for priority 2 and priority 0 data.

This mechanism ensures that priority 6 data is always sent, making the system suitable for realtime data transmission. In fact this was one of the primary aims in the design of token bus.

Wireless LAN

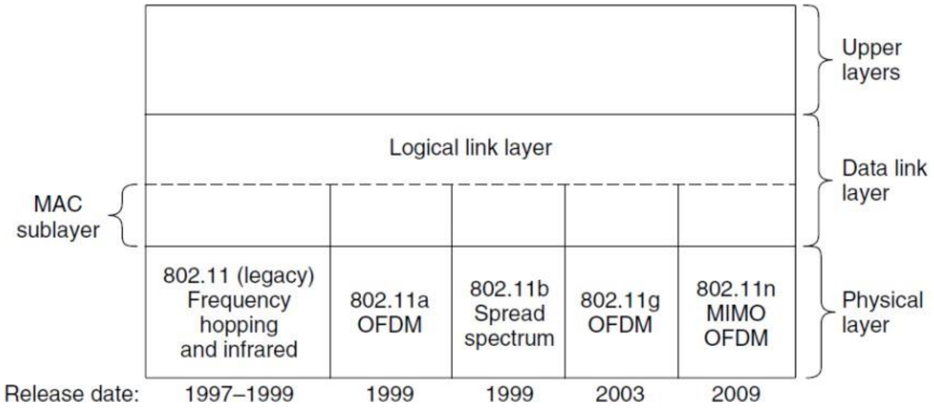
- A system of notebook computers that communicate by radio can be regarded as a wireless LAN
- A common configuration for a wireless LAN is an office building with base stations (also called access points) strategically placed around the building.
- All the base stations are wired together using copper or fiber.

- A simplifying assumption that all radio transmitters have some fixed range will be used follow.



Wireless clients associate to a wired AP (Access Point) Called infrastructure mode; there is alsoad-hoc mode with no AP, but that is rare.

802.11 Architecture/Protocol Stack



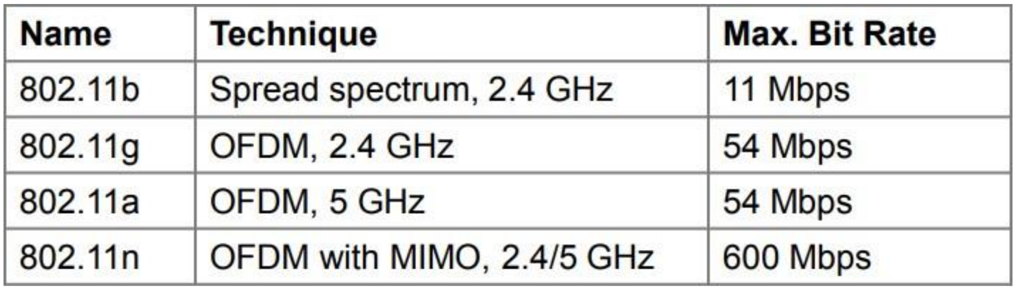
In 802.11, the MAC (Medium Access Control) sublayer determines how the channel is allocated, that is, who gets to transmit next. Above it is the LLC (Logical Link Control) sublayer, whose job it is to hide the differences between the different 802 variants and make them indistinguishable as far as the network layer is concerned.

- The 1997 802.11 standard specifies three transmission techniques allowed in the physical layer.
- The infrared method uses much the same technology as television remote controls do.
- The other two use short-range radio, using techniques called FHSS and DSSS.
- Both of these use a part of the spectrum that does not require licensing (the 2.4-GHz ISM band).
- All of these techniques operate at 1 or 2 Mbps and at low enough power that they do not conflict too much.

- In 1999, two new techniques were introduced to achieve higher bandwidth. These are called OFDM and HR-DSSS.
- They operate at up to 54 Mbps and 11 Mbps, respectively. In 2001, a second OFDM modulation was introduced, but in a different frequency band from the first one.

The 802.11 Physical Layer

- Each of the five permitted transmission techniques makes it possible to send a MAC frame from one station to another. They differ, however, in the technology used and speeds achievable.



Name	Technique	Max. Bit Rate
802.11b	Spread spectrum, 2.4 GHz	11 Mbps
802.11g	OFDM, 2.4 GHz	54 Mbps
802.11a	OFDM, 5 GHz	54 Mbps
802.11n	OFDM with MIMO, 2.4/5 GHz	600 Mbps

The 802.11 MAC Sublayer Protocol

- When a receiver is within range of two active transmitters, the resulting signal will generally be garbled and useless.
- A naive approach to using a wireless LAN might be to try CSMA: just listen for other transmissions and only transmit if no one else is doing so.
- The trouble is, this protocol is not really appropriate because what matters is interference at the receiver, not at the sender.

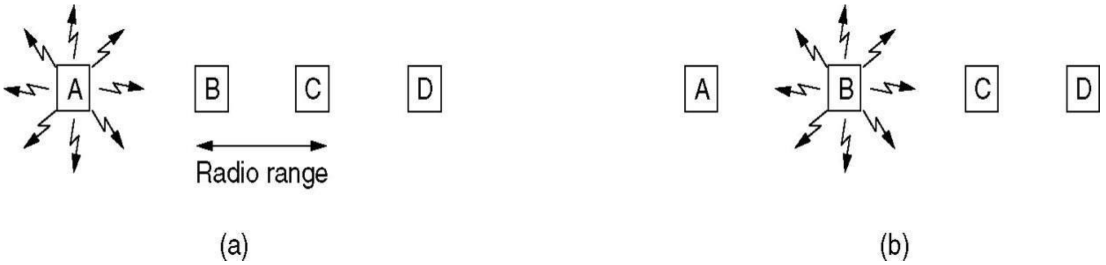
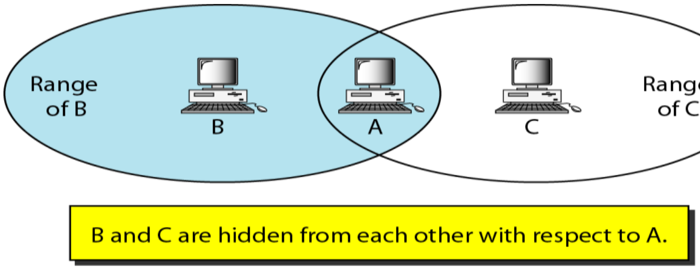


Fig. A wireless LAN. (a) A transmitting. (b) B transmitting.

When A is transmitting to B (previous figure part a)
If C senses the medium, it will not hear A because A is out of range, and thus falselyconclude that it can transmit to B.
If C does start transmitting, it will interfere at B, wiping out the frame from A.
The problem of a station not being able to detect a potential competitor for the mediumbecause the competitor is too far away is called the **hidden station problem**.



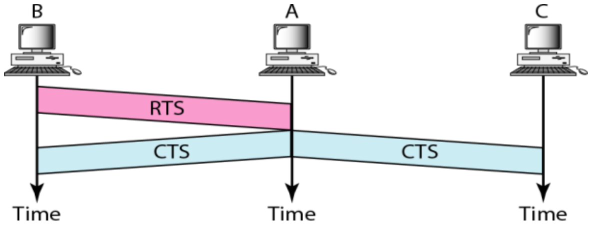
B and C are hidden from each other with respect to A.

Fig. Hidden station problemWhen B transmitting to A (previous figure part b)

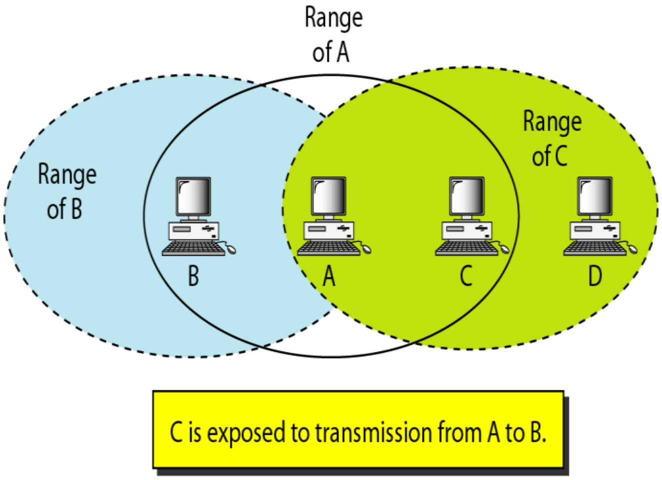
If C senses the medium, it will hear an ongoing transmission and falsely conclude that itmay not send to D, when in fact such a transmission would cause bad reception only in the zone between B and C, where neither of the intended receivers is located.

This is called the **exposed station problem**.

The CTS frame in CSMA/CA handshake can prevent collision from a hidden station.

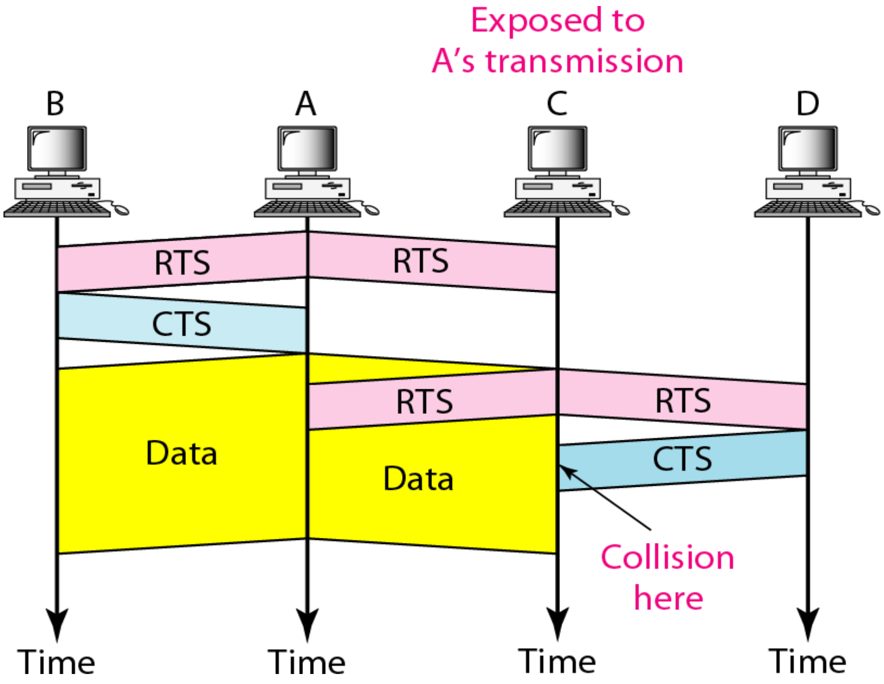


Use of handshaking to prevent hidden station problem



C is exposed to transmission from A to B.

Exposed station problem



Use of handshaking in exposed station problem

The problem is that before starting a transmission, a station really wants to know whether there isactivity around the receiver.

An early protocol designed for wireless LANs is MACA (Multiple Access with Collision Avoidance) (Karn, 1990).

The basic idea behind it is for the sender to stimulate the receiver into outputting a short frame, so stations nearby can detect this transmission and avoid transmitting for the duration of the upcoming (large) data frame.

Let us now consider how A sends a frame to B.

- A starts by sending an RTS (Request To Send) frame to B. This short frame (30 bytes) contains the length of the data frame that will eventually follow.
- Then B replies with a CTS (Clear to Send) frame. The CTS frame contains the data length (copied from the RTS frame). Upon receipt of the CTS frame, A begins transmission.

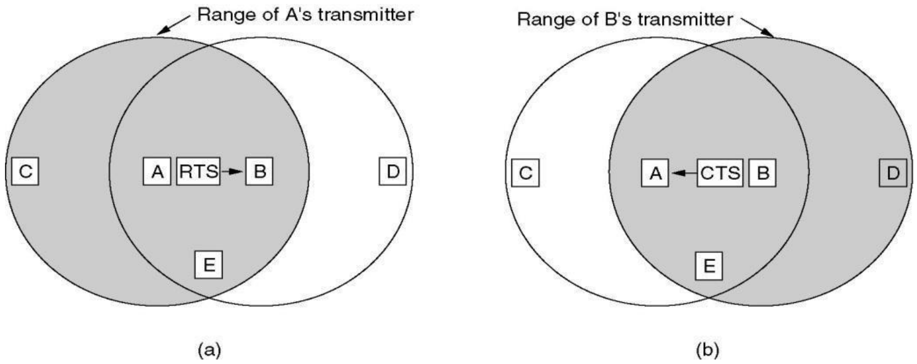


Fig. The MACA protocol. (a) A sending an RTS to B. (b) B responding with a CTS to A.

- Any station hearing the RTS is clearly close to A and *must remain silent* ***long enough*** for the CTS to be transmitted back to A without conflict.
- Any station hearing the CTS is clearly close to B and *must remain silent* ***during*** the upcoming data transmission, whose length it can tell by examining the CTS frame.
- C is within range of A but not within range of B. Therefore, it hears the RTS from A but not the CTS from B. As long as it does not interfere with the CTS, it is *free to transmit* while the data frame is being sent.
- D is within range of B but not A. It does not hear the RTS but does hear the CTS. Hearing the CTS tips it off that it is close to a station that is about to receive a frame, so it *defers sending* anything until that frame is expected to be finished.
- E hears both control messages and, like D, must be silent until the data frame is complete.
- Collisions can still occur:

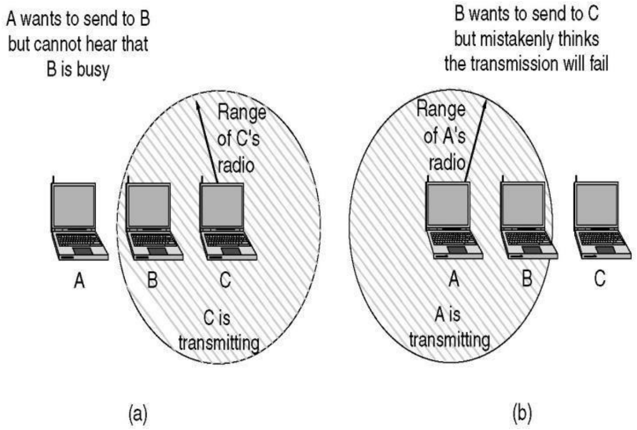
For example, B and C could both send RTS frames to A at the same time. These will collide and be lost. In the event of a collision, an unsuccessful transmitter (i.e., one that does not hear a CTS within the expected time interval) waits a random amount of time and tries again later.

- New improvements have been made to MACA to improve its performance and the new protocol named MACAW (MACA for Wireless).

The 802.11 MAC Sublayer Protocol

- Two problems have been explained before: the hidden station problem and the exposed station problem.
- 802.11 supports two modes of operation:
- The first, called DCF (Distributed Coordination Function), does not use any kind of central control (in that respect, similar to Ethernet).
- The other, called PCF (Point Coordination Function), uses the base station to control all activity in its cell.

- All implementations must support DCF but PCF is optional.



(a) The hidden station problem. (b) The exposed station problem.

802.11 Services

- There are five distribution services are provided by the base stations and deal with station mobility as they enter and leave cells, attaching themselves to and detaching themselves from base stations.

1. Association.

This service is used by mobile stations to connect themselves to base stations. Typically, it is used
just after a station moves within the radio range of the base station.
Upon arrival, it announces its identity and capabilities.
The capabilities include the data rates supported, need for PCF services, and powermanagement requirements.
The base station may accept or reject the mobile station.
If the mobile station is accepted, it must then authenticate itself.

2. Disassociation.

Either the station or the base station may disassociate
A station should use this service before shutting down or leaving, but the base station mayalso use it before going down for maintenance.

3. Reassociation.

A station may change its preferred base station using this service.
This facility is useful for mobile stations moving from one cell to another.

4. Distribution.

This service determines how to route frames sent to the base station.
If the destination is local to the base station, the frames can be sent out directly over the air.Otherwise, they will have to be forwarded over the wired network.

5. Integration.

If a frame needs to be sent through a non-802.11 network with a different addressing scheme or frame format, this service handles the translation from the 802.11 format to the format required by the destination network.

The remaining four services are intracell (i.e., relate to actions within a single cell). They are used after association has taken place and are authentication, deauthentication, privacy and data delivery.

BLUE TOOTH

It is a wireless standard for interconnecting computing and communication devices and accessories using short-range, low-power, inexpensive wireless radios.

Bluetooth Architecture:

The basic unit of a Bluetooth system is a **piconet**, which consists of a master node and up to seven active slave nodes within a distance of 10 meters. Multiple piconets can exist in the same (large) room and can even be connected via a bridge node that takes part in multiple piconets. Aninterconnected collection of piconets is called a **scatternet**.

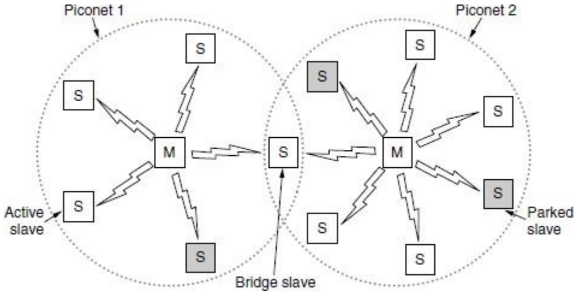
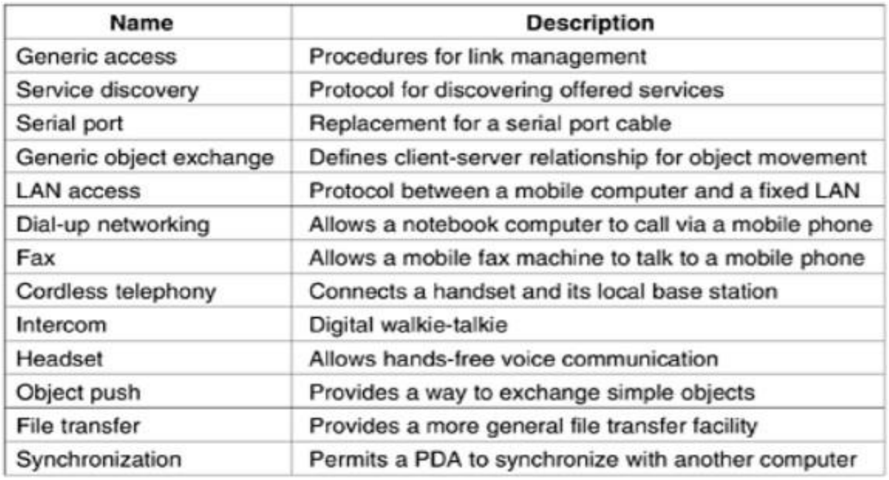


FIG.Two piconets can be connected to form a scatternet.

In addition to the seven active slave nodes in a piconet, there can be up to 255 parked nodes in the net. These are devices that the master has switched to a lowpower state to reduce the drain ontheir batteries. In parked state, a device cannot do anything except respond to an activation or beacon signal from the master. Two intermediate power states, hold and sniff, also exist, but these will not concern us here. At its heart, a piconet is a centralized TDM system, with the master controlling the clock and determining which device gets to communicate in which time slot. All communication is between the master and a slave; direct slave-slave communication is not possible.

Bluetooth Profile



Name	Description
Generic access	Procedures for link management
Service discovery	Protocol for discovering offered services
Serial port	Replacement for a serial port cable
Generic object exchange	Defines client-server relationship for object movement
LAN access	Protocol between a mobile computer and a fixed LAN
Dial-up networking	Allows a notebook computer to call via a mobile phone
Fax	Allows a mobile fax machine to talk to a mobile phone
Cordless telephony	Connects a handset and its local base station
Intercom	Digital walkie-talkie
Headset	Allows hands-free voice communication
Object push	Provides a way to exchange simple objects
File transfer	Provides a more general file transfer facility
Synchronization	Permits a PDA to synchronize with another computer

The Bluetooth Protocol Stack

The Bluetooth standard has many protocols grouped loosely into the layers shown in Fig.

- The bottom layer is the physical radio layer, which corresponds fairly well to the physical layer in the OSI and 802 models. It deals with radio transmission and modulation. It specifics details of the air interface, including frequency, frequency hopping, modulation scheme, and transmission power.
- The baseband layer is somewhat analogous to the MAC sublayer but also includes elements of the physical layer. It deals with how the master controls time slots and how these slots are grouped into frames. It is concerned with connection establishment within a piconet, addressing, packet format, timing and power control.
- Next comes a layer with a group of somewhat related protocols.
- The link manager handles the establishment of logical channels between devices, including power management, authentication, and quality of service.
- The logical link control adaptation protocol (often called L2CAP) shields the upper layers from the details of transmission. It is analogous to the standard 802 LLC sublayer, but technically different from it. It adapts upper layer protocols to the baseband layer. Provides both connectionless and connection-oriented services.
- As the names suggest, the audio and control protocols deal with audio and control, respectively. The applications can get at them directly, without having to go through the L2CAP protocol.
- The next layer up is the middleware layer, which contains a mix of different protocols. The 802 LLC was inserted here by IEEE for compatibility with its other 802 networks. The RFcomm, telephony, and service discovery protocols are native.
- RFcomm (Radio Frequency communication) is the protocol that emulates the standard serial port found on PCs for connecting the keyboard, mouse, and modem, among other devices. It has been designed to allow legacy devices to use it easily.
- Finally, the service discovery protocol is used to locate services within the network.
- The top layer is where the applications and profiles are located. They make use of the protocols in lower layers to get their work done. Each application has its own dedicated subset of the protocols. Specific devices, such as a headset, usually contain only those protocols needed by that application and no others.

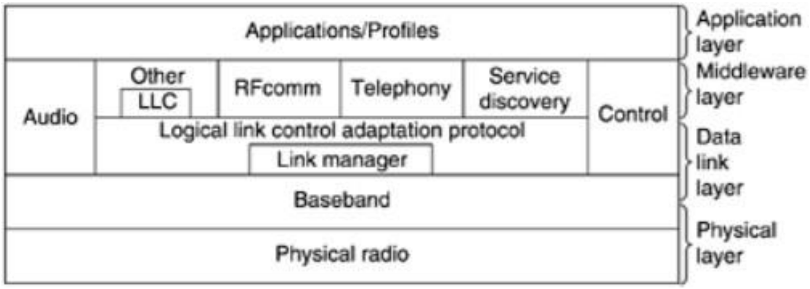


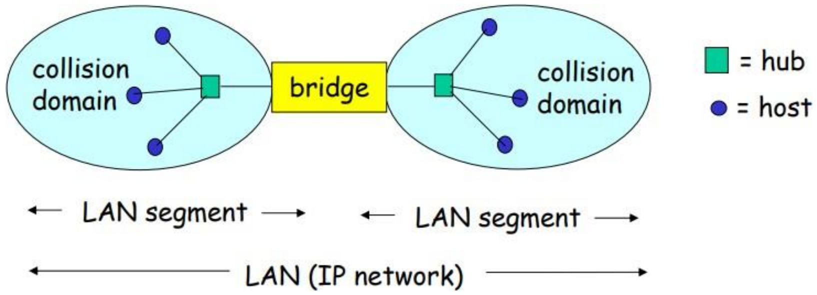
Fig. The Bluetooth protocol architecture.

BRIDGES

- MAC layer (Data Link Layer) device
- stores and forwards Ethernet frames
- examines frame header and selectively forwards frame based on MAC dest address
- when frame is to be forwarded on a segment, uses CSMA/CD to access segment
- transparent – hosts are unaware of presence of bridges
- plug-and-play, self-learning – bridges do not need to be configured

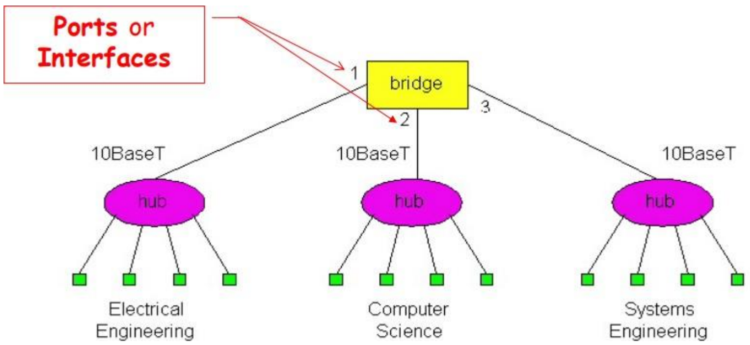
Bridges: traffic isolation

- Bridge installation breaks LAN into LAN segments
- Bridges filter packets: – same-LAN-segment frames are not (usually) forwarded onto other LAN segments
- Segments become separate collision domains



Forwarding

How does the bridge determine to which LAN segment to forward frame?



Self learning

- A bridge has a Bridge Table
- Entry in Bridge Table: – (Node LAN Address, Bridge Interface, Time Stamp) – stale entries in table dropped (TTL can be 60 min)
- Bridges learn which hosts can be reached through which interfaces – when frame received, bridge "learns" location of sender: incoming LAN segment – records sender/location pair in Bridge Table

Filtering/Forwarding

When bridge receives a frame:

The routing procedure for an incoming frame depends on the LAN it arrives on (the source

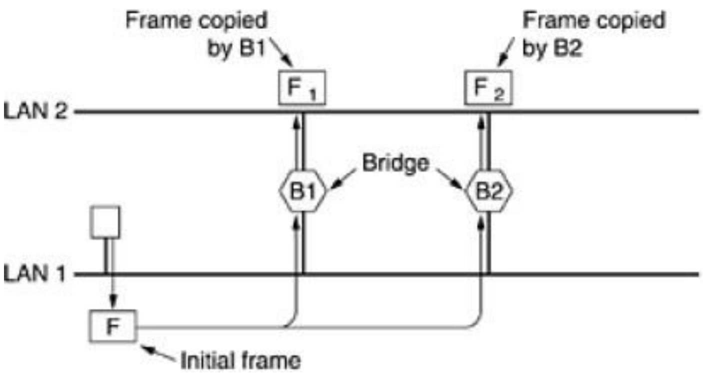
LAN) and the LAN its destination is on (the destination LAN), as follows:

1. If destination and source LANs are the same, discard the frame.
2. If the destination and source LANs are different, forward the frame.
3. If the destination LAN is unknown, use flooding.

Spanning Tree Bridges

To increase reliability, some sites use two or more bridges in parallel between pairs of LANs, asshown in Fig. This arrangement, however, also introduces some additional problems because it creates loops in the topology.

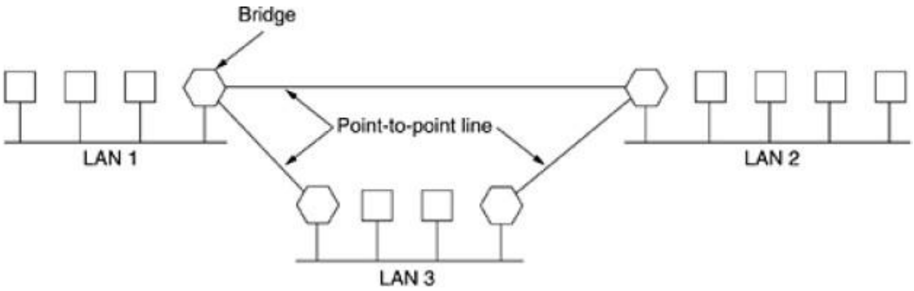
Two parallel transparent bridges.



The solution to this difficulty is for the bridges to communicate with each other and overlay the actual topology with a spanning tree that reaches every LAN. To build the spanning tree, first thebridges have to choose one bridge to be the root of the tree. They make this choice by having each one broadcast its serial number, installed by the manufacturer and guaranteed to be unique worldwide. The bridge with the lowest serial number becomes the root. Next, a tree of shortest paths from the root to every bridge and LAN is constructed. This tree is the spanning tree. If a bridge or LAN fails, a new one is computed. The result of this algorithm is that a unique path is established from every LAN to the root and thus to every other LAN. Although the tree spans all the LANs, not all the bridges are necessarily present in the tree (to prevent loops). Even after the spanning tree has been established, the algorithm continues to run during normal operation in order to automatically detect topology changes and update the tree.

Remote Bridges

A common use of bridges is to connect two (or more) distant LANs. For example, a company might have plants in several cities, each with its own LAN. Ideally, all the LANs should be interconnected, so the complete system acts like one large LAN. This goal can be achieved by putting a bridge on each LAN and connecting the bridges pairwise with point-to-point lines (e.g.,lines leased from a telephone company).



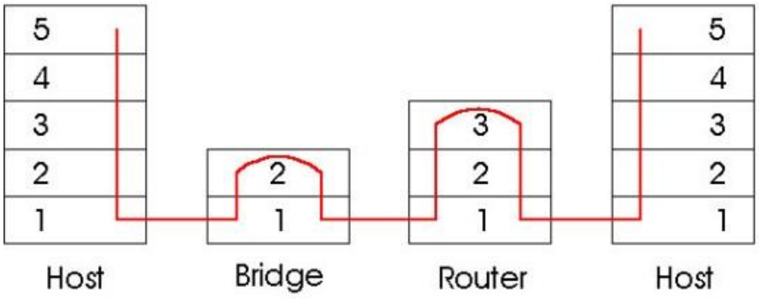
Remote bridges can be used to interconnect distant LANs.

Some Bridge Features

- Isolates collision domains resulting in higher total max throughput
- Limitless number of nodes and geographical coverage
- Can connect different Ethernet types
- Transparent (“plug-and-play”): no configuration necessary

Bridges vs. Routers

- Both are Store-and-Forward devices – routers: network l ayer devices (examine network layer headers) – bridges are link layer devices
- Routers maintain Routing Tables – implement routing algorithms
- B ridges maintain Bridge Tables – implement filtering, learning and spanning tree algorithms



Bridges advantages and disadvantages

Advantages:

-Bridge operation is simpler requiring less packet processing

- Bridge tables are self learning

Disadvantages:

- All traffic confined to spanning tree, even when alternative bandwidth is available
- Bridges do not offer protection from broadcast storms

Routers advantages and disadvantages

Advantages:

+ arbitrary topologies can be supported, cycling is limited by TTL counters (and goodrouting protocols)

+ p r o vid e protection against broadcast stormsDisadvantages:

- require I P configuration (not plug and play)

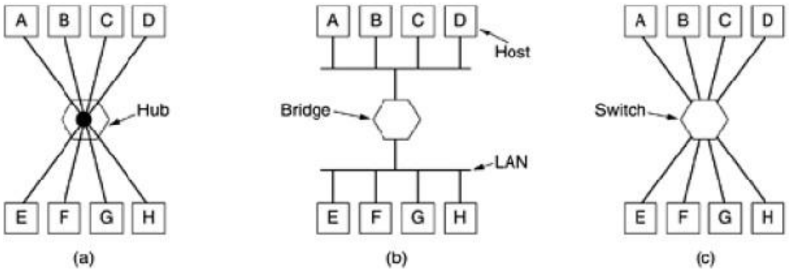
- require higher packet processing

- bridges do well in small networks (few hundred hosts)
- routers are used in large networks (thousands of hosts)

SWITCHES

Switches are similar to bridges in that both route on frame addresses. In fact, many peopleuses the terms interchangeably. The main difference is that a switch is most often used to connect individual computers, as shown in Fig. (c).

(a) A hub. (b) A bridge. (c) A switch.



- As a consequence, when host *A* in Fig. (b) wants to send a frame to host *B*, the bridge gets the frame but just discards it.
- In contrast, in Fig.(c), the switch must actively forward the frame from *A* to *B* because there is no other way for the frame to get there.
- Since each switch port usually goes to a single computer, switches must have space for many more line cards than do bridges intended to connect only LANs.
- Each line card provides buffer space for frames arriving on its ports.
- Since each port is its own collision domain, switches never lose frames to collisions.
- However, if frames come in faster than they can be retransmitted, the switch may run outof buffer space and have to start discarding frames.
- To alleviate this problem slightly, modern switches start forwarding frames as soon as the destination header field has come in, but before the rest of the frame has arrived. These switches do not use store-and-forward switching. Sometimes they are referred to as **cut-through switches**.

Repeaters, Hubs, Bridges, Switches, Routers, and Gateways

1. **Repeater** – A repeater operates at the physical layer. Its job is to regenerate the signal over the same network before the signal becomes too weak or corrupted so as to extend the length to which the signal can be transmitted over the same network. An important point to be noted about repeaters is that they do not amplify the signal. When the signal becomes weak, they copy the signal bit by bit and regenerate it at the original strength. It is a 2 port device.

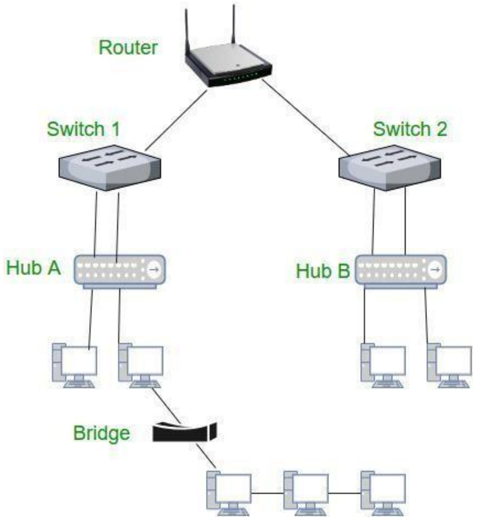
2. **Hub** – A hub is basically a multiport repeater. A hub connects multiple wires coming from different branches, for example, the connector in star topology which connects different stations. Hubs cannot filter data, so data packets are sent to all connected devices. In other words, collision domain of all hosts connected through Hub remains one. Also, they do not have intelligence to find out best path for data packets which leads to inefficiencies and wastage.

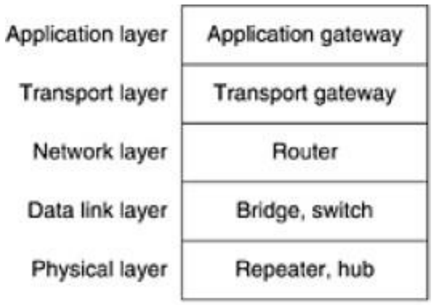
3. Bridge – A bridge operates at data link layer. A bridge is a repeater, with add on functionality of filtering content by reading the MAC addresses of source and destination. It is also used for interconnecting two LANs working on the same protocol. It has a single input and single output port, thus making it a 2 port device.

4. Switch – A switch is a multi port bridge with a buffer and a design that can boost its efficiency(large number of ports imply less traffic) and performance. Switch is data link layer device. Switch can perform error checking before forwarding data, that makes it very efficient asit does not forward packets that have errors and forward good packets selectively to correct port only. In other words, switch divides collision domain of hosts, but broadcast domain remains same.

5. Routers – A router is a device like a switch that routes data packets based on their IP addresses. Router is mainly a Network Layer device. Routers normally connect LANs and WANs together and have a dynamically updating routing table based on which they make decisions on routing the data packets. Router divide broadcast domains of hosts connected through it.

6. Gateway – A gateway, as the name suggests, is a passage to connect two networks together that may work upon different networking models. They basically works as the messenger agents that take data from one system, interpret it, and transfer it to another system. Gateways are also called protocol converters and can operate at any network layer. Gateways are generally more complex than switch or router.





Application layer	Application gateway
Transport layer	Transport gateway
Network layer	Router
Data link layer	Bridge, switch
Physical layer	Repeater, hub

(a)

MODULE III

Network layer – Routing – Shortest path routing, Flooding, Distance Vector Routing,Link State Routing, RIP, OSPF, Routing for mobile hosts.

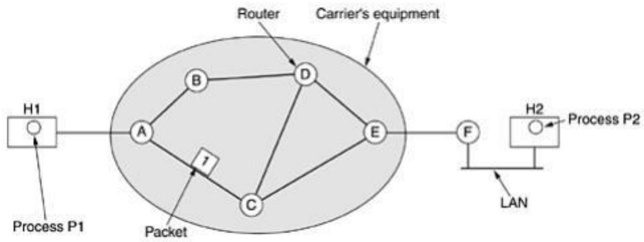
THE NETWORK LAYER

The network layer is concerned with getting packets from the source all the way to the destination. Getting to the destination may require making many hops at intermediate routers along the way. This function clearly contrasts with that of the data link layer, which has the more modest goal of just moving frames from one end of a wire to the other. Thus, the network layer is the lowest layer that deals with end-to-end transmission. To achieve its goals, the network layer must know about the topology of the communication subnet (i.e., the set of all routers) and choose appropriate paths through it. It must also take care to choose routes to avoid overloading some of the communication lines and routers while leaving others idle. Finally, when the source and destination are in different networks, new problems occur. It is up to the network layer to deal with them.

Network Layer Design Issues

1. **Store-and-Forward Packet Switching:**

Figure 5-1. The environment of the network layer protocols.



The major components of the system are:

- the carrier's equipment (routers connected by transmission lines), shown inside the shaded oval, and
- the customers' equipment, shown outside the oval.
- Host H1 is directly connected to one of the carrier's routers, A, by a leased line.
- H2 is on a LAN with a router, F, owned and operated by the customer. This router also has a leased line to the carrier's equipment.

A host with a packet to send transmits it to the nearest router, either on its own LAN or over a point-to-point link to the carrier. The packet is stored there until it has fully arrived so the checksum can be verified. Then it is forwarded to the nextrouter along the path until it reaches the destination host, where it is delivered.

2. **Services Provided to the Transport Layer:**

The network layer provides services to the transport layer at the network layer/transport layer interface. The network layer services have been designed with the following goals in mind:

1. The services should be independent of the router technology.

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

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

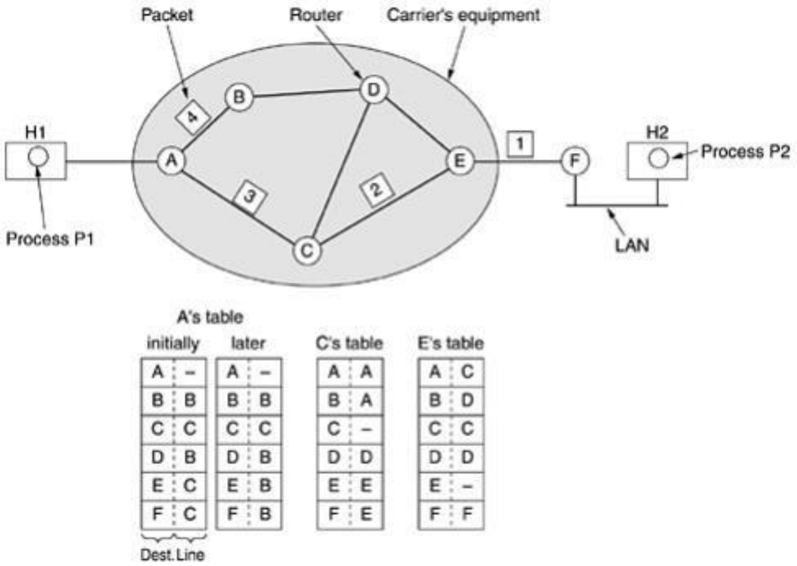
The two classes of service the network layer can provide to its users:

- The network service should be connectionless, with primitives SEND PACKET and RECEIVE PACKET. No packet ordering and flow control should be done, because the hosts are going to do that. Each packet must carry the full destination address, because each packet sent is carried independently of its predecessors.
- The subnet should provide a reliable, connection-oriented service. Quality of service is the dominant factor, and without connections in the subnet, quality of service is very difficult to achieve, especially for real-time traffic such as voice and video.

3. Implementation of Connectionless Service.

If connectionless service is offered, packets are injected into the subnet individually and routed independently of each other. No advance setup is needed. The packets are frequently called **datagrams** (in analogy with telegrams) and the subnet is called a **datagram subnet**. If connection-oriented service is used, a path from the source router to the destination router must be established before any data packets can be sent. This connection is called a **VC (virtual circuit).**

Figure 5-2. Routing within a datagram subnet.



	A's table			
	initially	later	C's table	E's table
A	–	–	A	C
B	B	B	A	D
C	C	C	–	C
D	B	B	D	D
E	C	B	E	–
F	C	B	E	F

Suppose that the process P1 in Fig. 5-2 has a long message for P2. It hands the message to the transport layer with instructions to deliver it to process P2 on host H2. The transport layer code runs on H1, typically within the operating system. It prepends a transport header to the front of the message and hands the result to the network layer.

the message is four times longer than the maximum packet size, so the network layer has to break it into four packets, 1, 2, 3, and 4 and sends each of them in turn to router A using some point-to-point protocol, for example, PPP. At this point the carrier takes over. Every router has an internal table telling it where to send packets for each possible destination. Each table entry is a pair consisting of a destination and the outgoing line to use for that destination. Only directly-connected lines can be used. For example, in Fig. 5-2, A has only two outgoing lines—to B and C—so every incoming packet must be sent to one of these routers, even if the ultimate destination is some other router. A's initial routing table is shown in the figure under the label "initially."

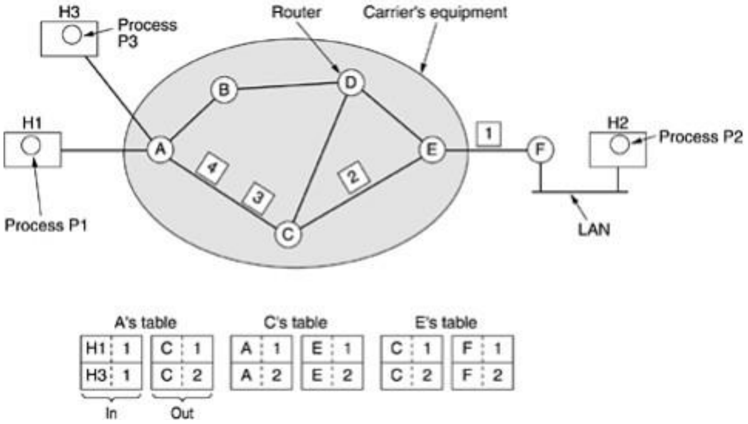
As they arrived at A, packets 1, 2, and 3 were stored briefly (to verify their checksums). Then each was forwarded to C according to A's table. Packet 1 was then forwarded to E and then to F. When it got to F, it was encapsulated in a data link layer frame and sent to H2 over the LAN. Packets 2 and 3 follow the same route.

Packet 4, When it got to A it was sent to router B, even though it is also destined for F. For some reason, A decided to send packet 4 via a different route than that of the first three. Perhaps it learned of a traffic jam somewhere along the ACE path and updated its routing table, as shown under the label "later." The algorithm that manages the tables and makes the routing decisions is called **the routing algorithm**.

4. Implementation of Connection-Oriented Service.

For connection-oriented service, we need a virtual-circuit subnet.When a connection is established, a route from the source machine to the destination machine is chosen as part of the connection setup and stored in tables inside the routers. That route is used for all traffic flowing over the connection. When the connection is released, the virtual circuit is also terminated.

Figure 5-3. Routing within a virtual-circuit subnet.



A's table			
In		Out	
H1	1	C	1
H3	1	C	2

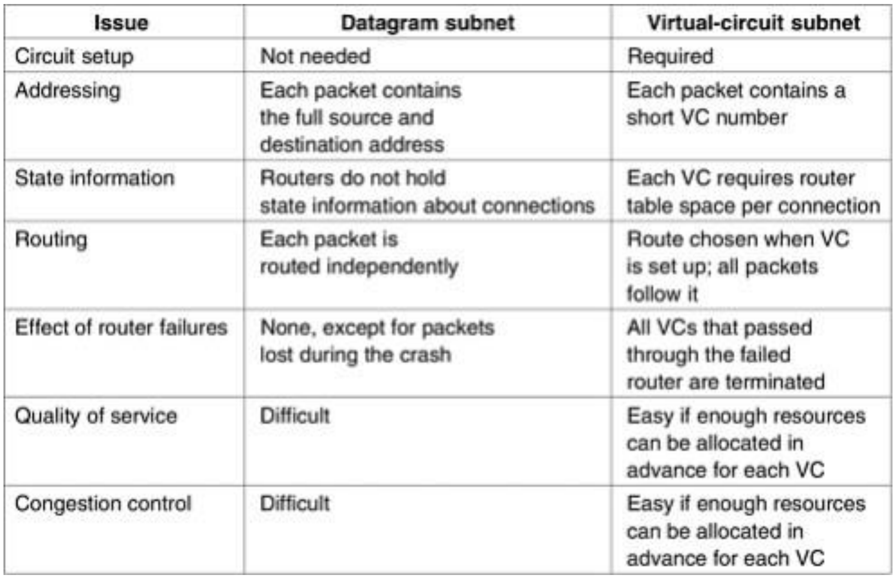
C's table			
A	1	E	1
A	2	E	2

E's table			
C	1	F	1
C	2	F	2

With connection-oriented service, each packet carries an identifier telling which virtual circuit it belongs to. As an example, consider the situation of Fig. 5-3. Here, host H1 has established connection 1 with host H2. It is remembered as the first entry in each of the routing tables. The first line of A's table says that if a packet bearing connection identifier 1 comes in from H1, it is to be sent to router C and given connection identifier 1. Similarly, the first entry at C routes the packet to E, also with connection identifier 1.

If H3 also wants to establish a connection to H2. It chooses connection identifier 1 (because it is initiating the connection and this is its only connection) and tells the subnet to establish the virtual circuit. This leads to the second row in the tables. Note that we have a conflict here because although A can easily distinguish connection 1 packets from H1 from connection 1 packets from H3, C cannot do this. For this reason, A assigns a different connection identifier to the outgoing traffic for the second connection. Avoiding conflicts of this kind is why routers need the ability to replace connection identifiers in outgoing packets, this is called **label switching**.

5. Comparison of Virtual-Circuit and Datagram Subnets:



Issue	Datagram subnet	Virtual-circuit subnet
Circuit setup	Not needed	Required
Addressing	Each packet contains the full source and destination address	Each packet contains a short VC number
State information	Routers do not hold state information about connections	Each VC requires router table space per connection
Routing	Each packet is routed independently	Route chosen when VC is set up; all packets follow it
Effect of router failures	None, except for packets lost during the crash	All VCs that passed through the failed router are terminated
Quality of service	Difficult	Easy if enough resources can be allocated in advance for each VC
Congestion control	Difficult	Easy if enough resources can be allocated in advance for each VC

ROUTING:

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

- If the subnet uses **datagrams** internally, this decision must be made anew for every arriving data packet since the best route may have changed since last time.
- If the subnet uses **virtual circuits** internally, routing decisions are made only when a new virtual circuit is being set up. Thereafter, data packets just follow the previously-established route. The latter case is sometimes called **session routing** because a route remains in force for an entire user session.

Routing - is making the decision which routes to use.
Forwarding - is what happens when a packet arrives.

Router - having two processes inside it. One of them handles each packet as it arrives, looking up the outgoing line to use for it in the routing tables. This process is **forwarding**. The other process is responsible for **filling in and updating the routing tables**. That is where the routing algorithm comes into play.
Properties are desirable in a routing algorithm: correctness, simplicity, robustness, stability, fairness, and optimality.
Routing algorithms can be grouped into two major classes: **nonadaptive and adaptive.**

- **Nonadaptive 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 (for all I and J) is computed in advance, off-line, and downloaded to the routers when the network is booted. This procedure is sometimes called **static routing.**

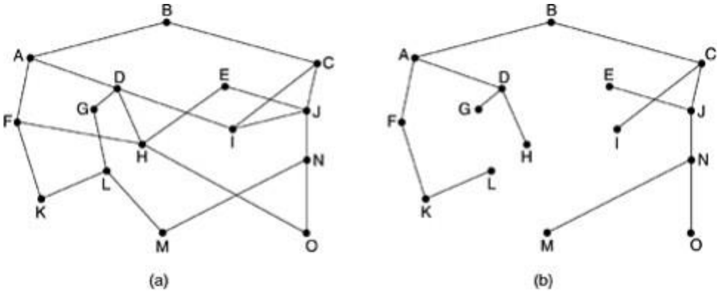
- **Adaptive algorithms,** change their routing decisions to reflect changes in the topology, and the traffic. Adaptive algorithms differ in where they get their information, when they change the routes, and what metric is used for optimization.

The Optimality Principle

A general statement about optimal routes without regard to network topology or traffic is known as the **optimality principle**. It states that if router J is on the optimal path from router I to router K, then the optimal path from J to K also falls along the same route.

The set of optimal routes from all sources to a given destination form a tree rooted at the destination. Such a tree is called a **sink tree** and is illustrated in Fig. 5-6, where the distance metric is the number of hops. Note that a sink tree is not necessarily unique; other trees with the same path lengths may exist. The goal of all routing algorithms is to discover and use the sink trees for all routers.

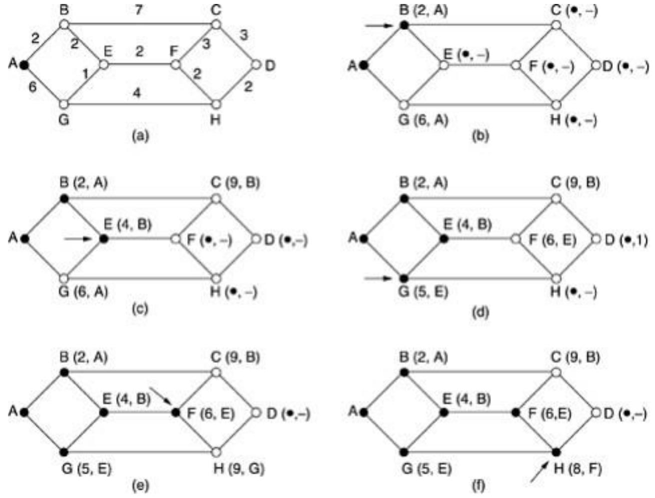
Figure 5-6. (a) A subnet. (b) A sink tree for router B



Shortest Path Routing

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 (often called a link). To choose a route between a given pair of routers, the algorithm just finds the shortest path between them on the graph.

Figure 5-7. The first five steps used in computing the shortest path from A to D. Thearrows indicate the working node.



- Measuring path length is the number of hops → the paths ABC and ABE in Fig. 5-7 are equally long.
- Another metric is the geographic distance → in kilometers, in which case ABC is clearly much longer than ABE.

Dijkstra's shortest path algorithm:

Compute the shortest path starting at the terminal node, t, rather than at the source node, s. Since the shortest path from t to s in an undirected graph is the same as the shortest path from s to t. The reason for searching backward is that each node is labeled with its predecessor rather than its successor. When the final path is copied into the output variable, path, the path is thus reversed.

FLOODING:

Flooding is a static algorithm, in which every incoming packet is sent out on every outgoing line except the one it arrived on. Flooding generates vast numbers of duplicate packets. One measure avoid this is to

- have **a hop counter** contained in the header of each packet, which is decremented at each hop, with the packet being discarded when the counter reaches zero. The hop counter should be initialized to the length of the path from source to destination. If the sender does not know how long the path is, it can initialize the counter to the worst case the full diameter of the subnet.
- keep track of which **packets have been flooded**, to avoid sending them out a second time. achieve this goal is to have the source router put a sequence number in each packet it receives from its hosts. Each router then needs a list per source router telling which sequence numbers originating at that source have already been seen. If an incoming packet is on the list, it is not flooded. To **prevent the list from growing** without bound, each list should be augmented by **a counter, k**, meaning that all sequence numbers through k have been seen. When a packet comes in, it is easy to check if the packet is a duplicate; if so, it is discarded. Furthermore, the full list below k is not needed, since k effectively summarizes it.
- **selective flooding -** the routers do not send every incoming packet out on every line, only on those lines that are going approximately in the right direction.

Uses of flooding:

- in military applications, where large numbers of routers may be blown to bits at any instant, the tremendous robustness of flooding is highly desirable.
- In distributed database applications, it is sometimes necessary to update all the databases concurrently, in which case flooding can be useful.
- In wireless networks, all messages transmitted by a station can be received by all other stations within its radio range, which is, in fact, flooding, and some algorithms utilize this property.
- It is as a metric against which other routing algorithms can be compared. Flooding always chooses the shortest path because it chooses every possible path in parallel.

Dynamic algorithms:

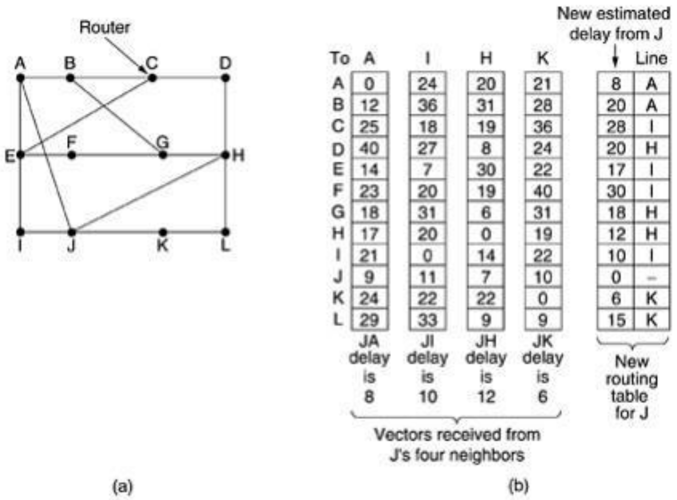
Take the current network load into account. Two dynamic algorithms in particular, distance vector routing and link state routing, are the most popular.

DISTANCE VECTOR ROUTING / BELLMAN-FORD ROUTING ALGORITHM /FORD-FULKERSON ALGORITHM:

Distance vector routing algorithms operate by having each router maintain a table (i.e, a vector) giving the best known distance to each destination and which line to use to get there. These tables are updated by exchanging information with the neighbors.
Each router maintains a **routing table** indexed by, and containing one entry for, each router in the subnet. This entry contains two parts: the preferred **outgoing line** to use for that destination and an **estimate of the time or distance** to that destination. The metric used might be number of hops, time delay in milliseconds, total number of packets queued along the path etc.

Imagine that one of these tables has just come in from neighbor X, with Xi being X's estimate of how long it takes to get to router i. If the router knows that the delay to X is m msec, it also knows that it can reach router i via X in Xi + m msec. By performing this calculation for each neighbor, a router can find out which estimate seems the best and use that estimate and the corresponding line in its new routing table. Note that the old routing table is not used in the calculation.
This updating process is illustrated in Fig. 5-9. Part (a) shows a subnet. The first four columns of part (b) show the delay vectors received from the neighbors of router J. A claims to have a 12-msec delay to B, a 25-msec delay to C, a 40-msec delay to D, etc. Suppose that J has measured or estimated its delay to its neighbors, A, I, H, and K as 8, 10, 12, and 6 msec, respectively.

Figure 5-9. (a) A subnet. (b) Input from A, I, H, K, and the new routing table for J.



To	A	I	H	K	New estimated delay from J	Line
A	0	24	20	21	8	A
B	12	36	31	28	20	A
C	25	18	19	36	28	I
D	40	27	8	24	20	H
E	14	7	30	22	17	I
F	23	20	19	40	30	I
G	18	31	6	31	18	H
H	17	20	0	19	12	H
I	21	0	14	22	10	I
J	9	11	7	10	0	–
K	24	22	22	0	6	K
L	29	33	9	9	15	K
	JA delay is 8	JI delay is 10	JH delay is 12	JK delay is 6	New routing table for J	

Vectors received from J's four neighbors

How J computes its new route to router G. It knows that it can get to A in 8 msec, and A claims to be able to get to G in 18 msec, so J knows it can count on a delay of 26 msec to G if it forwards packets bound for G to A. Similarly, it computes the delay to G via I, H, and K as 41 (31 + 10), 18 (6 + 12), and 37 (31 + 6) msec, respectively. The best of these values is 18, so it makes an entry in its routing table that the delay to G is 18 msec and that the route to use is via H. The same calculation is performed for all the other destinations, with the new routing table shown in the last column of the figure.

Drawback : The Count-to-Infinity Problem:

It converges to the correct answer, it may do so slowly. Consider the five-node (linear) subnet of Fig. 5-10, where the delay metric is the number of hops. Suppose A is down initially and all the other routers know this. In other words, they have all recorded the delay to A as infinity.

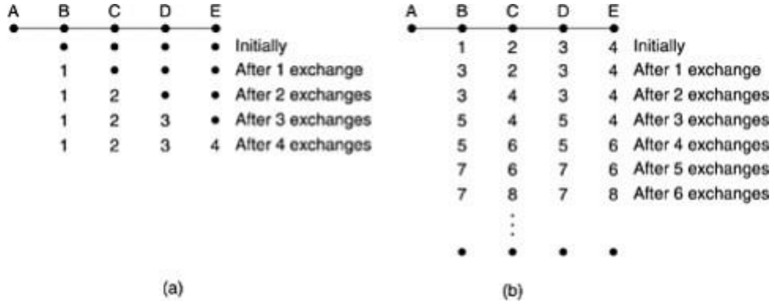

A B C D E
• • • • Initially
1 • • • After 1 exchange
1 2 • • After 2 exchanges
1 2 3 • After 3 exchanges
1 2 3 4 After 4 exchanges
(a)
A B C D E
1 2 3 4 Initially
3 2 3 4 After 1 exchange
3 4 3 4 After 2 exchanges
5 4 5 4 After 3 exchanges
5 6 5 6 After 4 exchanges
7 6 7 6 After 5 exchanges
7 8 7 8 After 6 exchanges
• • • •
(b)

fig 5-10

When A comes up, the other routers learn about it via the vector exchanges. At the time of the first exchange, B learns that its left neighbor has zero delay to A. B now makes an entry in its routing table that A is one hop away to the left. All the other routers still think that A is down. At this point, the routing table entries for A are as shown in the second row of Fig. 5-10(a). On the next exchange, C learns that B has a path of length 1 to A, so it updates its routing table to indicate a path of length 2, but D and E do not hear the good news until later. Clearly, the good news is spreading at the rate of one hop per exchange. In a subnet whose longest path is of length N hops, within N exchanges everyone will know about newly-revived lines and routers.

Fig. 5-10(b), in which all the lines and routers are initially up. Routers B, C, D, and E have distances to A of 1, 2, 3, and 4, respectively. Suddenly A goes down, or alternatively, the line between A and B is cut, which is effectively the same thing from B's point of view.

At the first packet exchange, B does not hear anything from A. Fortunately, C says: Do not worry; I have a path to A of length 2. Little does B know that C's path runs through B itself. For all B knows, C might have ten lines all with separate paths to A of length 2. As a result, B thinks it can reach A via C, with a path length of 3. D and E do not update their entries for A on the first exchange.

On the second exchange, C notices that each of its neighbors claims to have a path to A of length 3. It picks one of the them at random and makes its new distance to A 4, as shown in the third row of Fig. 5-10(b). Subsequent exchanges produce the history shown in the rest of Fig. 5-10(b).

From this figure, it should be clear why bad news travels slowly: no router ever has a value more than one higher than the minimum of all its neighbors. Gradually, all routers work their way up to infinity, but the number of exchanges required depends on the numerical value used for infinity. For this reason, it is wise to set infinity to the longest path plus 1. If the metric is time delay, there is no well-defined upper bound, so a high value is needed to prevent a path with a long delay from being treated as down. This problem is known as the **count-to-infinity problem**. The core of the problem is that when X tells Y that it has a path somewhere, Y has no way of knowing whether it itself is on the path

The delay metric was queue length, it did not take line bandwidth into account when choosing routes. The second problem the algorithm often took too long to converge (the count-to-infinity problem)

LINK STATE ROUTING

The idea behind link state routing is simple and can be stated as five parts. Each router must do the following:
1. Discover its neighbors and learn their network addresses.
2. Measure the delay or cost to each of its neighbors.
3. Construct a packet telling all it has just learned.
4. Send this packet to all other routers.
5. Compute the shortest path to every other router.

1. Discover its neighbors and learn their network addresses.

- Send a special HELLO packet on each point-to-point line.
- The router on the other end is expected to send back a reply telling who it is.
- When two or more routers are connected by a LAN.
- model the LAN as a node itself

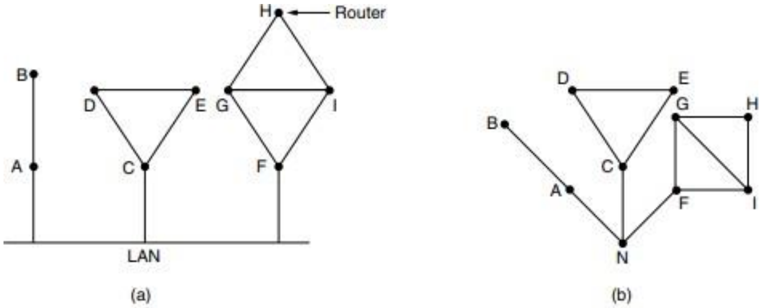


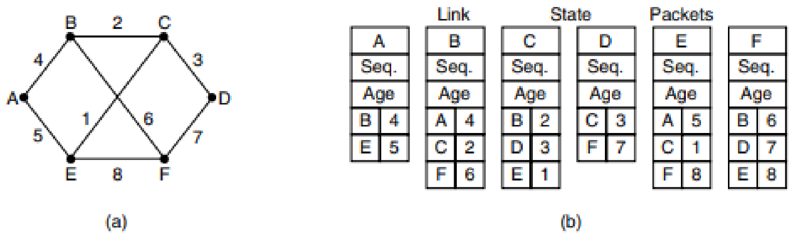
Figure 5-11. (a) Nine routers and a broadcast LAN. (b) A graph model of (a).

2. Measure the delay or cost to each of its neighbors - Setting Link Costs

- each link to have a distance or cost metric for finding shortest paths.
- make the cost inversely proportional to the bandwidth of the link.
- send over the line a special **ECHO packet** that the other side is required to send back immediately. By measuring the round-trip time and dividing it by two, the sending router can get a reasonable estimate of the delay.

3. Building Link State Packets

- each router to build a packet containing all the data.
- The packet starts with the identity of the sender, followed by a sequence number and age and a list of neighbors.



A	B	C	D	E	F
Seq.	Seq.	Seq.	Seq.	Seq.	Seq.
Age	Age	Age	Age	Age	Age
B 4	A 4	B 2	C 3	A 5	B 6
E 5	C 2	D 3	F 7	C 1	D 7
	F 6	E 1		F 8	E 8

Figure 5-12. (a) A network. (b) The link state packets for this network.

determining when to build them:

- build them periodically, that is, at regular intervals.
- build them when some significant event occurs, such as a line or neighbor going down or coming back up again or changing its properties appreciably

4. Distributing the Link State Packets

basic distribution algorithm - use flooding to distribute the link state packets. To keep the flood in check, each packet contains a sequence number that is incremented for each new packet sent. Routers keep track of all the (source router, sequence) pairs they see. When a new link state packet comes in, it is checked against the list of packets already seen. If it is new, it is forwarded on all lines except the one it arrived on. If it is a duplicate, it is discarded. If a packet with a sequence number lower than the highest one seen so far ever arrives, it is rejected as being obsolete since the router has more recent data.

When a link state packet comes in to a router for flooding, it is not queued for transmission immediately. Instead, it is put in a holding area to wait a short while in case more links are coming up or going down. If another link state packet from the same source comes in before the first packet is transmitted, their sequence numbers are compared. If they are equal, the duplicate is discarded. If they are different, the older one is thrown out. To guard against errors on the links, all link state packets are acknowledged.

Drawbacks of algorithm:

- if the sequence numbers wrap around,- The solution here is to use a 32-bit sequence number. With one link state packet per second.
- if a router ever crashes, it will lose track of its sequence number. If it starts again at 0, the next packet will be rejected as a duplicate.
- if a sequence number is ever corrupted.

The solution to all these problems is to include **the age** of each packet after the sequence number and decrement it once per second. When the age hits zero, the information from that router is discarded.

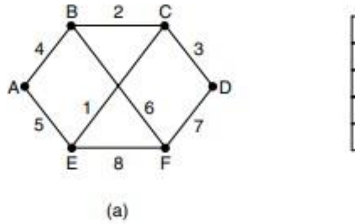
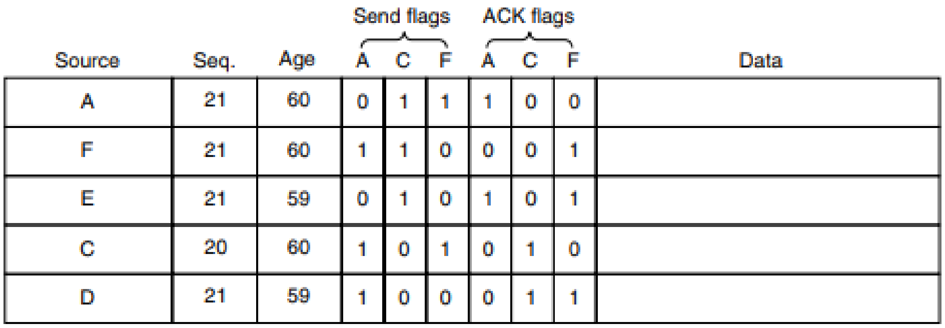


Figure 5-12. (a) A network



			Send flags			ACK flags			
Source	Seq.	Age	A	C	F	A	C	F	Data
A	21	60	0	1	1	1	0	0	
F	21	60	1	1	0	0	0	1	
E	21	59	0	1	0	1	0	1	
C	20	60	1	0	1	0	1	0	
D	21	59	1	0	0	0	1	1	

Figure 5-13. The packet buffer for router *B* in Fig. 5-12(a).

The data structure used by router B for the network shown in Fig. 5-12(a) is depicted in Fig. 5-13. Each row here corresponds to a recently arrived, but as yet not fully processed, link state packet. The table records where the packet originated, its sequence number and age, and the data. In addition, there are send and acknowledgement flags for each of B’s three links (to A, C, and F, respectively). The send flags mean that the packet must be sent on the indicated link. The acknowledgement flags mean that it must be acknowledged there.

The link state packet from A arrives directly, so it must be sent to C and F and acknowledged to A, as indicated by the flag bits. Similarly, the packet from F has to be forwarded to A and C and acknowledged to F. However, the situation with the third packet, from E, is different. It arrives twice, once via EAB and once via EFB.

Consequently, it has to be sent only to C but must be acknowledged to both A and F, as indicated by the bits. If a duplicate arrives while the original is still in the buffer, bits have to be changed. For example, if a copy of C’s state arrives from F before the fourth entry in the table has been forwarded, the six bits will be changed to 100011 to indicate that the packet must be acknowledged to F but not sent there.

5. Computing the New Routes

Once a router has accumulated a full set of link state packets, it can construct the entire network graph because every link is represented. Now Dijkstra’s algorithm can be run locally to construct the shortest paths to all possible destinations. This information is installed in the routing tables, and normal operation is resumed.

For a network with n routers, each of which has k neighbors, the memory required to store the input data is proportional to kn.

Any cast:

Delivery models in which a source sends to a single destination called **unicast,** to all destinations called **broadcast,** and to a group of destinations called **multicast**. Another delivery model, called **anycast** is sometimes also useful. In anycast, a packet is delivered to the nearest member of a group. Schemes that find these paths are called **anycast** routing.

Within each network, an **intradomain or interior gateway protocol** is used for routing. Across the networks that make up the internet, an **interdomain or exterior gateway protocol**is used. The networks may all use different intradomain protocols, but they must use the same interdomain protocol. In the Internet, the interdomain routing protocol is called **BGP (Border Gateway Protocol)**.

Each network is operated independently of all the others, it is often referred to as an **AS (Autonomous System).** Eg: an ISP network. Routing inside an autonomous system is called **Interior Routing (**RIP, OSPF)**.** Routing between AS is called as **Exterior Routing (BGP).**

ROUTING INFORMATION PROTOCOL:

It is an interior router protocol used inside an autonomous system. It is based on distance vector routing

Points to be remembered in RIP:

- in autonomous system we consider routers and networks. Routers have routing table while networks do not have.
- Destination of a routing table in a network which is the first column of routing table.
- Metric used is the distance which is the number of links to reach the destination called hop count.
- Infinity is defined as 16 means the maximum hop permitted is 15.
- Next hop column defines address of the router to which packet to be sent to reach the destination.

Destination	Hop count	Next hop	Other information

Routing table updation is done using RIP Response message.

RIP Routing Table Updation algorithm:

- RIP response message arrived
- Increment hop count by one for each advertised destination
- Repeat the following steps for each advertised destination
  - Add advertised information to the table if destination Is not present in the routing table.
  - Replace entry in the table with advertised one in next hop field is same
  - Replace entry in routing table if advertising hop count is smaller than the one in table.

RIP message format:

RIP message are of two types:

1. Message that deliver routing information
2. Message that request routing information

Command	version	All zeros
Address family identifier		All zeros
IP address		
All zeros		
All zeros		
Metric PC		
Repeat of last 20 bytes		

Command – indicate if it is request or response.

Version – version used

Zeros –provide backup compatibility with previous standards of RIP
AFI – specifies the address

IP address – ip address for the entry

Metric – number of internetwork hops encountered. Maximum of 15.

Disavantages:

- It only understands the shortest route to destination based on the count of hops
- Depends on routers for computing routing update
- Routing table need to be broadcasted every 30secnds
- Distance are based on hops not on cost

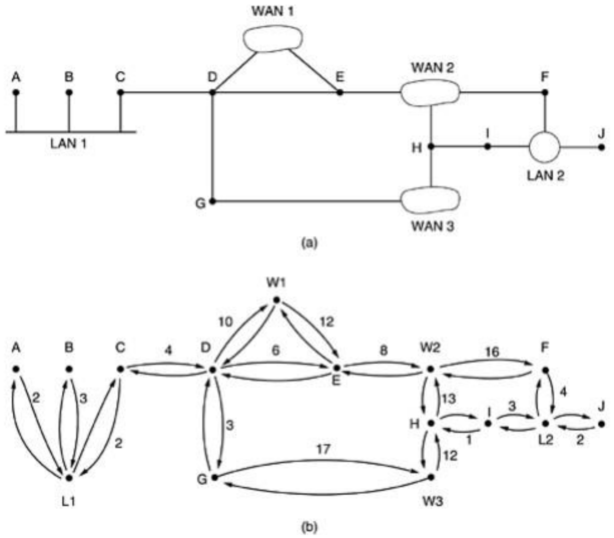
OSPF (OPEN SHORTEST PATH FIRST) - The Interior Gateway Routing Protocol

The Internet is made up of a large number of autonomous systems. Each AS is operated by a different organization and can use its own routing algorithm inside. For example, the internal networks of companies X, Y, and Z are usually seen as three ASes if all three are on the Internet. All three may use different routing algorithms internally. A routing algorithm within an AS is called an **interior gateway protocol**; an algorithm for routing between ASes is called an **exterior gateway protocol.**

OSPF supports three kinds of connections and networks:
1. Point-to-point lines between exactly two routers.
2. Multiaccess networks with broadcasting (e.g., most LANs).
3. Multiaccess networks without broadcasting (e.g., most packet-switched WANs).

A multiaccess network is one that can have multiple routers on it, each of which can directly communicate with all the others. All LANs and WANs have this property. Figure 5-64(a) shows an AS containing all three kinds of networks. Note that hosts do not generally play a role in OSPF.

Figure 5-64. (a) An autonomous system. (b) A graph representation of (a).



OSPF operates by abstracting the collection of actual networks, routers, and lines into a directed graph in which each arc is assigned a cost (distance, delay, etc.). It then computes the shortest path based on the weights on the arcs. A serial connection between two routers is represented by a pair of arcs, one in each direction. Their weights may be different. A multiaccess network is represented by a node for the network itself plus a node for each router. The arcs from the network node to the routers have weight 0 and are omitted from the graph.

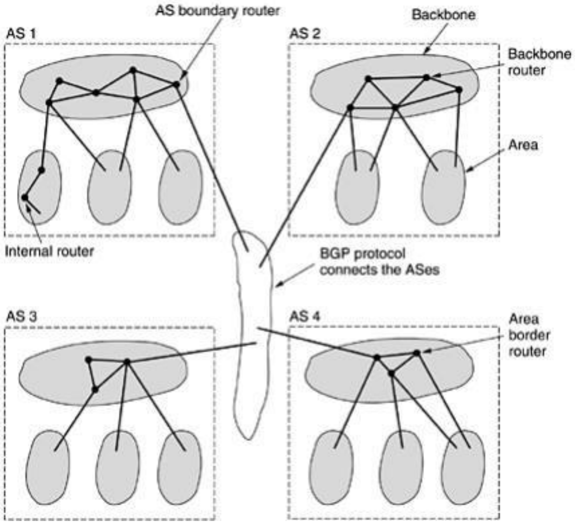
OSPF allows ASes to be divided into numbered **areas**, where an area is a network or a set of contiguous networks. Areas do not overlap but need not be exhaustive, that is, some routers may belong to no area. An area is a generalization of a subnet. Outside an area, its topology and details are not visible.

Every AS has a **backbone area,** called **area 0**. All areas are connected to the backbone, possibly by tunnels, so it is possible to go from any area in the AS to any other area in the AS via the backbone.

During normal operation, three kinds of routes may be needed:

1. intra-area - Intra-area routes are the easiest, since the source router already knows the shortest path to the destination router
2. interarea - Interarea routing always proceeds in three steps:
    a. go from the source to the backbone;
    b. go across the backbone to the destination area;
    c. go to the destination
3. inter-AS.

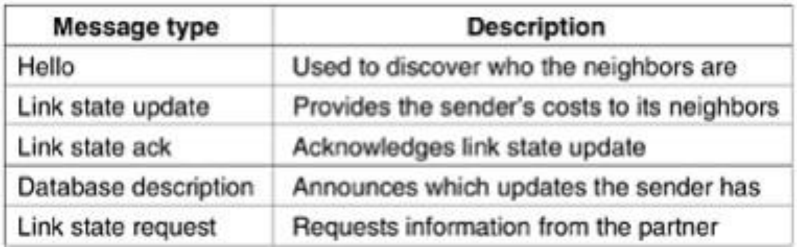
Fig 5-65 :The relation between ASes, backbones, and areas in OSPF.



OSPF distinguishes four classes of routers:

1. Internal routers are wholly within one area.
2. Area border routers connect two or more areas.
3. Backbone routers are on the backbone.
4. AS boundary routers talk to routers in other ASes.

Figure 5-66. The five types of OSPF messages.



Message type	Description
Hello	Used to discover who the neighbors are
Link state update	Provides the sender's costs to its neighbors
Link state ack	Acknowledges link state update
Database description	Announces which updates the sender has
Link state request	Requests information from the partner

When a router boots, it sends **HELLO** messages on all of its point-to-point lines and multicasts them on LANs to the group consisting of all the other routers.

One router is elected as the designated router. It is said to be adjacent to all the other routers on its LAN, and exchanges information with them. Neighboring routers that are not adjacent do not exchange information with each other. A backup designated router is always kept up to date to ease the transition should the primary designated router crash and need to replaced immediately.

Each router periodically floods **LINK STATE UPDATE** messages to each of its adjacent routers. This message gives its state and provides the costs used in the topological database. The flooding messages are acknowledged, to make them reliable. Each message has a sequence number, so a router can see whether an incoming LINK STATE UPDATE is older or newer than what it currently has. Routers also send these messages when a line goes up or down or its cost changes.

DATABASE DESCRIPTION messages give the sequence numbers of all the link state entries currently held by the sender. By comparing its own values with those of the sender, the receiver can determine who has the most recent values. These messages are used when a line is brought up.

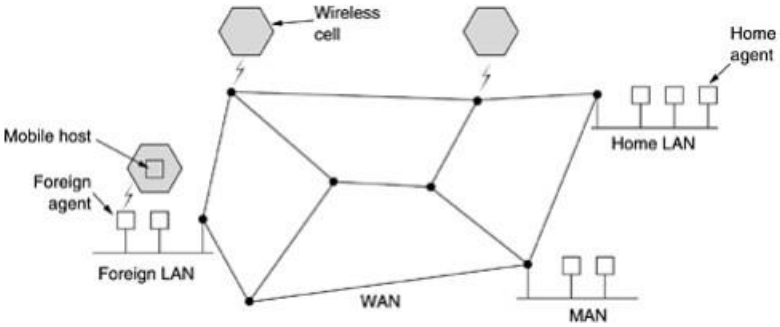
Either partner can request link state information from the other one by using **LINK STATE REQUEST** messages. The result of this algorithm is that each pair of adjacent routers checks to see who has the most recent data, and new information is spread throughout the area this way. All these messages are sent as raw IP packets.

Using flooding, each router informs all the other routers in its area of its neighbors and costs. This information allows each router and The backbone area to construct the graph for its area(s) and compute the shortest path. In addition, the backbone routers accept information from the area border routers in order to compute the best route from each backbone router to every other router. This information is propagated back to the area border routers, which advertise it within their areas. Using this information, a router about to send an interarea packet can select the best exit router to the backbone.

ROUTING FOR MOBILE HOSTS

The model of the world that network designers typically use is shown in Fig. 5-18.

Figure 5-18. A WAN to which LANs, MANs, and wireless cells are attached.



Hosts that never move are said to be **stationary.** They are connected to the network by copper wires or fiber optics. **Migratory hosts** are basically stationary hosts who move from one fixed site to another from time to time but use the network only when they are physically connected to it. **Roaming hosts** actually compute on the run and want to maintain their connections as they move around. We will use the term **mobile hosts** to mean either of the latter two categories, that is, all hosts that are away from home and still want to be connected.

All hosts are assumed to have a permanent home location that never changes. Hosts also have a permanent home address that can be used to determine their home locations, analogous to the way the telephone number 1-212-5551212 indicates the United States (country code 1) and Manhattan (212). The routing goal in systems with mobile hosts is to make it possible to send packets to mobile hosts using their home addresses and have the packets efficiently reach them wherever they may be.

the world is divided up (geographically) into small units. Called **areas**, where an area is typically a LAN or wireless cell. Each area has one or more foreign agents, which are processes that keep track of all mobile hosts visiting the area. In addition, each area has a home agent, which keeps track of hosts whose home is in the area, but who are currently visiting another area.

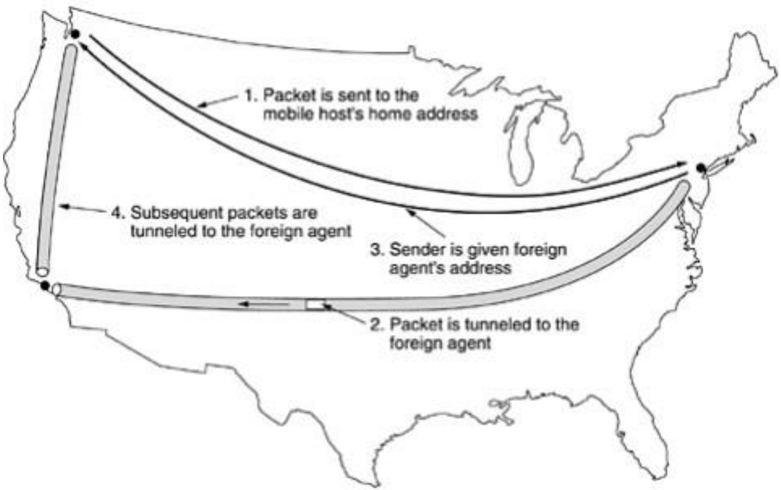
The registration procedure typically works like this:

1. Periodically, each foreign agent broadcasts a packet announcing its existence and address. A newly-arrived mobile host may wait for one of these messages, but if none arrives quickly enough, the mobile host can broadcast a packet.
2. The mobile host registers with the foreign agent, giving its home address, current datalink layer address, and some security information.
3. The foreign agent contacts the mobile host's home agent. It also includes the security information to convince the home agent that the mobile host is really there.
4. The home agent examines the security information, which contains a timestamp, toprove that it was generated within the past few seconds letting he foreign agent to proceed.
5. When the foreign agent gets the acknowledgement from the home agent, it makes anentry in its tables and informs the mobile host that it is now registered.

The home agent then does two things.

1. it encapsulates the packet in the payload field of an outer packet and sends the latter to the foreign agent (step 2 in Fig. 5-19). This mechanism is called **tunneling;** After getting the encapsulated packet, the foreign agent removes the original packet from thepayload field and sends it to the mobile host as a data link frame.
2. the home agent tells the sender to henceforth send packets to the mobile host by encapsulating them in the payload of packets explicitly addressed to the foreign agent instead of just sending them to the mobile host's home address (step 3). Subsequent packets can now be routed directly to the host via the foreign agent (step 4), bypassing the home location entirely.

Figure 5-19. Packet routing for mobile hosts.


1. Packet is sent to the mobile host's home address
4. Subsequent packets are tunneled to the foreign agent
3. Sender is given foreign agent's address
2. Packet is tunneled to the foreign agent

MODULE IV

CONGESTION

occur in a computer network when the resource demands exceed the capacity Packets may be lost due to too much queuing in the network. During Congestion the network throughput may drop and the path delay may become very high Congestion in a network may occur if users send data into the network at a rate greater than allowed by network resources. for example, Congestion may occur because the switched in a network have a limited buffer size to store arrived packets before processing

Causes of Congestion

1: Unpredictable statistical fluctuation of traffic flows

2: faults conditions within the network

3: Slow processor speed. if the router's CPU speed is very low and performing tasks like queuing buffer, tables updating etc. queries are built up, even though the line capacity is not fully utilized

4: Inefficient control policies

5: Bandwidth of the links is important in Congestion. The links to be used must be of high Bandwidth to avoid Congestion

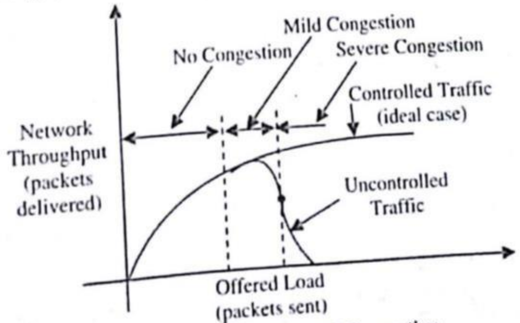


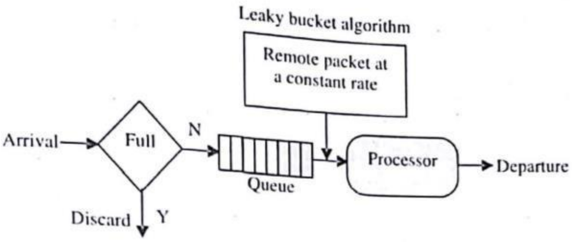
Figure 4.1: Effect of Congestion

TYPES OF CONGESTION CONTROL ALGORITHMS

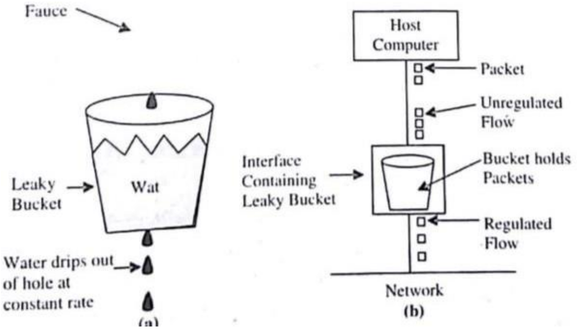
Congestion in a frame relay network is a problem that must be avoided because it decreases throughput and increases delay

1: Leaky Bucket Algorithm: If there is a hole at the bottom of a Bucket, then no matter at what rate the Bucket is filled up, the water leaks out drop by drop at a constant rate from the hole. Each host is connected by an interface that has finite queue acting like leaky Bucket.

When a packet comes to a host with the queue full, it is discarded. The host 1s allowed to put one packet per clock tick info the network. This can he enforced by the interface card or by the operating system. This converts an uneven flow of packets from the user process in an even flow of packers onto the network. Conceptually. each host is connected to the network by an interface Containing a leaky bucket, that is, a finite internal queue. If a packet arrives at the queue when it is full , the packet is discarded



In other word, if one or more processes within the host try to send a packet when the maximum number is already queued, the new packet is unceremoniously discarded. This arrangement can be built into the hardware interface or simulated by the host operating system. It was first proposed by Turner (1986) and is called the leaky bucket algorithm. In fact it is nothing other than a single-server queuing system with constant service time. This mechanism turns an uneven flow of packets from the user processes inside the host into an even flow of packets onto the network. Smoothing out bursts and greatly reducing the chances of congestion



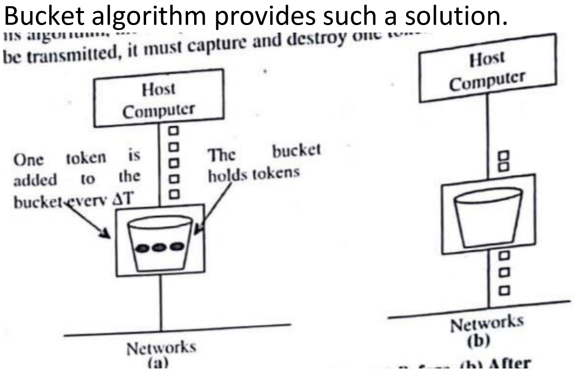
Implementing the original leaky bucket algorithm is easy. The leaky bucket consists of a finite queue. When a packet arrives, if there is room on the 4ueue it is appended to the queue otherwise, it is discarded. At every clock tick, one packet is transmitted (unless the queue is empty).

This arrangement can be simulated in the operating system or can be built into the hardware. Implementation of this algorithm is easy and consists of a finite queue. Whenever a packet arrives, if there is room in the queue it is queued up and if there is no room then the packet is discarded

2: Token Bucket Algorithm

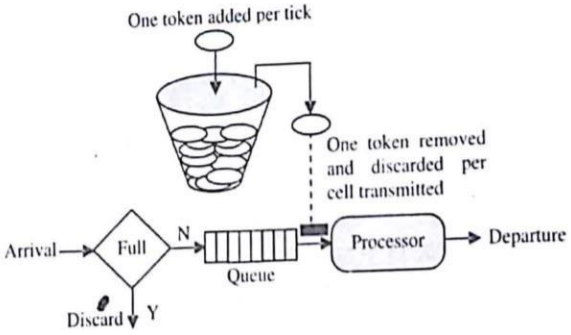
This algorithm allows bursts for short transmission while making sure that no data is lost. In contrast to the Leaky Bucket algorithm, not the data that is to be send but tokens are queued in a time-depended queue. One token is needed to send a single portion of data Implementations contain a token counter that is incremented on every time interval, so that the counter grows over time up until a maximum counter value is reached. The token counter is decremented by one for every data portion sent. When the token counter is zero no data can be transmitted.

For many applications it is better to allow the output to speed up somewhat when a larger burst arrives than to lose the data. Token Bucket algorithm provides such a solution.


be transmitted, it must capture and destroy one
Host Computer
One token is added to the bucket every ΔT
The bucket holds tokens
Networks
(a)
Host Computer
Networks
(b)
(b) After

Main steps of this algorithm can be described as follows:

* In regular intervals tokens are thrown into the bucket.
* The bucket has a maximum capacity.
* If there is a ready packet, a token is removed from the bucket, and the packet is send.
* If there is no token in the bucket, the packet cannot be send.



Token Bucket Algorithm

*Token dependent

*If bucket is full token are discarded, but not the packet.

*Packets can only transmitted when there are enough token.

*It allows large bursts to he sent at faster rate after that constant rate.

*It saves token to send large bursts.

Leaky Bucket Algorithm

*Token independent

*If bucket is full packet or data is discarded.

*Packets are transmitted continuously.

*It sends the packet at constant rate.

*It does not save token.

QUALITY OF SERVICE

is defined as something flow seeks to attain A Stream of packets from a source to destination is called flow. In a connection oriented network all packets belonging to a flow follow the same route, in a connection-less service they may follow different routes

The needs of each flow can be characterised by primary parameters vi, reliability, delay, jitter and bandwidth. together these determine the QoS(Quality of Service) the flow requires

QoS defines a set of attributes related to the performance of the connection for each connection the user can request a particular attributes.

flow Characteristics

1: Reliability: Reliability is a characteristic that flow needs. lack of reliability means losing a packet or acknowledgment which entails retransmission. However, the sensitivity of application programs to reliability is not the same

2: Delay: Source-to-destination delay is another characteristic Again application can tolerate delay in different degree. In this case, telephony, audio conferencing, video conferencing and remote login need minimum delay, while delay in file transfer or e-mail is lessimportant

3: Jitter: Jitter is the variation in delay for packets belonging to the same flow. For example, if four packets depart at times 0,1,2 and 3 and arrive at 20,21,22 and 23, all have thesame delay, 20 units of time. Jitter is defined as the variation in the packet delay. High jitter means the difference between delays is large; low jitter means the variation is small

4: Bandwidth: Different applications need different bandwidths. In video conferencing one need to send millions of bits per second to refresh a colour screen while the total number of bits in an e-mail may not reach even a million.

QoS Attributes

1: User Related Attributes: These are related to the end user in the sense that they define how fast a user wants to send/receive data

The Attributes are negotiated and defined at the time of the contact between the user and the network service provider

*Sustained Cell Rate(SCR): This is the average cell rate over a period of time, which could be more or less the actual transmission rates, as long as average is maintained

*Peak Cell Rate(PCR): This is the maximum transmission rate at a point of time

*Minimum Cell Rate(MCR): This is the minimum cell rate that the network guarantee a user

*Cell Variation Delay Tolerance(CVDT): This is a unit of measuring the changes in cell transmission times

2: Network Related Attributes: These Attributes defines the chara of a network

*Cell Loss Ratio(CLR): This Attribute define the fraction of cells lost/delivered too late during transmission

*Cell Transfer Delay(CTD): This is the average time required for a cell travel from the source to destination

*Cell Delay Variation(CDV): This is the difference between maximum and minimum values of CTD

*Cell Error Ratio(CER): This parameter defines the ratio of cells that contain errors

INTERNETWORKING

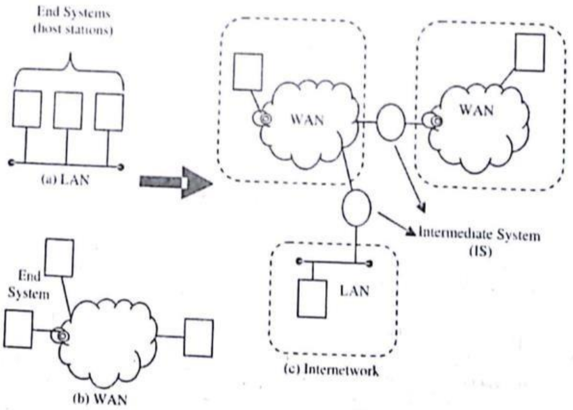
As the computer got smaller, cheaper and yet more powerful. more and more organisation, companies and people began having their own private network, even Internetworks in case of large organisations

Most of them wanted to join the rest of information world by further connecting to the internet. some of organisations used internet asa vehicle of communication between their remotely located private network/ Internetworks

All of these development saw the Internetworking technology to evolve as an important technology. an Internetworking is a collection of individual networks, connected by intermediate networking devices, that function as a single large network. there are various types of networks like LAN,WAN,MAN

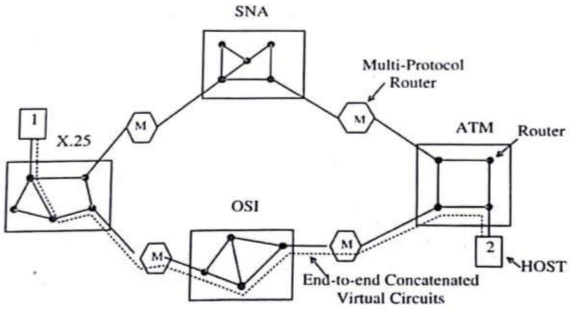
An Internetwork may also be defined as a network of computer communication network every authorised member of which could communication with every other authorised member directly or indirectly

Shows Interconnection Of 2 WAN and 1 LAN



Types or Internetworking

1: Concatenated Virtual Circuit: In the concatenated virtual circuit model, shown in figure connection to a host in a distant network is set up in a way similar to the way connectionsare normally established. The subnet sees that the destination is remote and builds a virtual circuit to the router nearest the destination network. Then it constructs a virtual circuit fromthat router to an external "gateway" (multiprotocol router). This gateway records theexistence of the virtual circuit in its tables and proceeds to build another virtual circuit to a router in the next subnet. This process continues until the destination host has been reached Once data packets begin flowing along the path, each gateway relays incoming packets.Converting between packet formats and virtual circuit numbers as needed. Clearly, all data packets must traverse the same sequence of gateways, and thus arrive in order

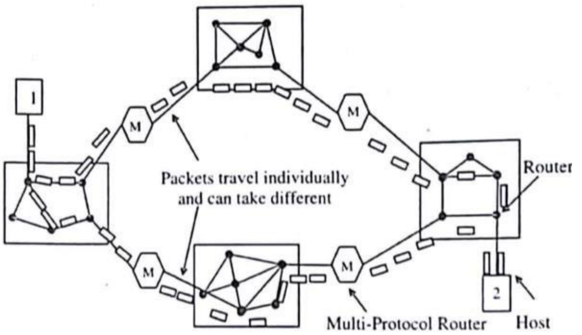


The essential feature of this approach is that a sequence of virtual circuits is set up from the source through one or more gateways to the destination. Each gateway maintains tables telling which virtual circuits pass through it.

2: Connectionless internetworking: The alternative internetwork model is a the datagram model, in this figure datagram from host 1 to host 2 are shown taking different routes through the internetwork

In this model, the only service t11e network layer offers to the transport layer is the ability to inject datagram into the subnet and hope forit will get to the destination. There is no notion of virtual circuit at all in the network layer. let alone a concatenation of them. This model doesnot require all packets belonging to one connection to traverse the same sequence or gateways.

A routing decision is made separately for each packet, possibly depending on the traffic at the moment the packet is sent. This strategy can use multiple routes and thus achieve a higher bandwidth than the concatenated virtual circuit model. On the other hand, there is no guarantee that the packets arrive at the destination in order, assuming that they arrive at all



Problems in Connectionless Internetworking

1: Conversion: If each network has its own network layer protocol, it is not possible for a packet from one network to transit another one. Multiprotocol routers tries to translate from one format to another. But such conversions will always be incomplete and often move to failure unless the two formats are close relatives with the same information fields For this reason conversion is rarely attempted.

2: Addressing: Imagine a simple case: a host on the Internet is trying to send an IP packet to a host on an adjoining OSI network. The OSI datagram protocol. Problem is that IP packets all carry the 32-bit Internet address of the destination host in a header field. OSI hosts do not have 32-bit Internet addresses. They use decimal addresses similar to telephone numbers

INTERNET PROTOCOL (IP)

IP is a datagram-oriented protocol, treating each packet independently. Also Internet Protocol makes no attempt to determine if packets reach their destination or to take corrective action if they do not. Internet Protocol provides the following functions: 1: Addressing, 2: Fragmentation

3: Packet timeouts

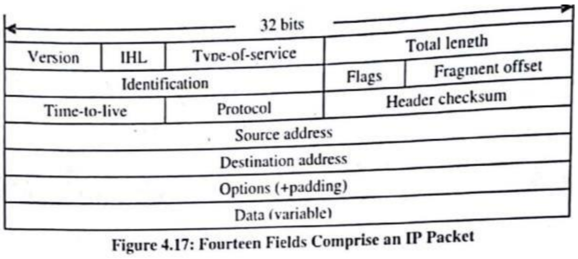
Internet Protocol (IP) is a network layer (Layer 3) protocol chat contains addressing information and some control information that enables packets to be routed. Along with the transmission Control Protocol(TCP), IP represents the heart of the Internet protocols.

IP has two primary responsibilities

1: Providing Connectionless, best effort delivery of datagram through an internetwork

2: Providing fragmentation and reassembly of datagram to support data links with different maximum transmission unit(MTU) sizes

IP Packet Format



32 bits

Version	IHL	Type-of-service	Total length
Identification		Flags	Fragment offset
Time-to-live		Protocol	Header checksum
Source address			
Destination address			
Options (+padding)			
Data (variable)			

Figure 4.17: Fourteen Fields Comprise an IP Packet

1: Version: Indicates the version of IP currently used.

2: IP Header Length (1HL): Indicates the datagram header length in 32-bit words.

3: Type-of-Service: Specifies how an upper-layer protocol would like a current datagram to be handled, and assigns datagram various levels of importance.

4: Total Length: Specifies the length, in bytes, of the entire IP packet, including the data and header.

5: Identification: Contains an integer that identifies the current datagram. This field is used to help piece together datagram fragments.

6: Flags: Consists of a 3-bit field of which the two tow-order (least-significant) bits control fragmentation.

7: Fragment Offset: Indicates the position of the fragment's data relative to the beginning of the data in the original datagram

8: Time-to-Live: Maintains a counter that gradually decrements down to zero, at which point the datagram is discarded.
9: Protocol: Indicates which upper-layer protocol receives incoming packets after IP processing is complete.
10: Header Checksum: Helps ensure IP header integrity.
11: Source Address: Specifies 1he sending node.
12: Destination Address: Specifies the receiving node.
13: Options: Allows IP to support various options, such as security.
14: Data: Contains upper-layer information.

IP ADDRESSING

As with any other network-layer protocol, the IP addressing scheme is integral to the process of routing IP datagram through internetwork. Each IP address has specific components and follows a basic format These IP addresses can be subdivided and used to create addresses for sub networks.

1: Network Number: It identifies a network and must be assigned by the internet network information centre(InterNIC) if the network is to be part of internet. an ISP can obtain block of network addresses from InterNIC and can itself assign address space as necessary

2: Host Number: It identifies a host on a network and assigned by the local network administrator

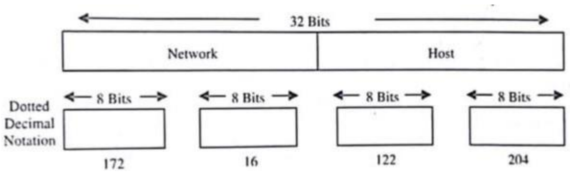
TYPES OF IP ADDRESSING

1: Classful Addressing: In the classful addressing system all the IP addresses that are available are divided into the five classes A.B,C,D and E. in which class A,B and C address are frequently used because class D is for Multicast and is rarely used and class E is reserved and is not currently used. Each of the IP address belongs to a particular class that's why they are classful addresses

IP Address Format

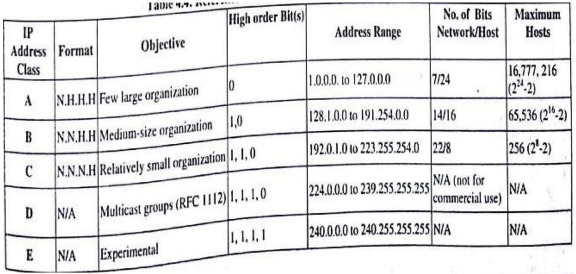
The 32 bits IP address is divided into four octets and each octet is written in eight bit decimal numbers. These four octet are separated by dots and ranges from O to 255

The binary weights of each bit in the octet is 128,64,32,16,8,4,2,1. The format of the 32-bit IP address is illustrated in the figure

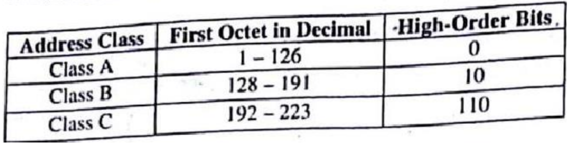


IP Address Classes

A class is used to recognise the part of network address and node address given in an IP address. There are 5 classes associated with IP addresses. A,B,C,D,E where A,B,C are used for commercial purpose



IP Address Class	Format	Objective	High order Bit(s)	Address Range	No. of Bits Network/Host	Maximum Hosts
A	N.H.H.H	Few large organization	0	1.0.0.0. to 127.0.0.0	7/24	16,777, 216 (2^{24}-2)
B	N.N.H.H	Medium-size organization	1,0	128.1.0.0 to 191.254.0.0	14/16	65,536 (2^{16}-2)
C	N.N.N.H	Relatively small organization	1, 1, 0	192.0.1.0 to 223.255.254.0	22/8	256 (2^{8}-2)
D	N/A	Multicast groups (RFC 1112)	1, 1, 1, 0	224.0.0.0 to 239.255.255.255	N/A (not for commercial use)	N/A
E	N/A	Experimental	1, 1, 1, 1	240.0.0.0 to 240.255.255.255	N/A	N/A



Address Class	First Octet in Decimal	·High-Order Bits.
Class A	1 – 126	0
Class B	128 – 191	10
Class C	192 – 223	110

2: Classless Addressing: There were certain problems with classful addressing such as address depletion and less organization access to internet. to overcome these problems, classful addressing is replaced with Classless addressing. as the name of Addressing schemes implies, the address are not divided into classes,however they are divided into blocks and the size of blocks varies according to the size of entity to which the addresses are to be allocated. IPv6 is classless addressing

The Internet authorities have enforced certain limitations on classing address blocks to make the handling of addresses easier. These limitations are as follows:

i) The addresses of a block must be contiguous.
ii) Each block must have a power of 2(that is 1,2,4,8...)number of addresses.
iii) The first address in a block must be evenly divisible by the total number of addresses in that block.

SUBNETTING

is a unique and powerful feature that is exclusive to the TCP/IP protocol and is one ofthe reasons TCP/IP offers great scalability.

Subnetting allows network address to be furtherdivided, apart from the already established classful boundaries, into smaller, more manageable networks. This division provides forunparalleled scalability and hierarchy, and gives a network administrator benefits such as reduced network traffic, less susceptibility to broadcast traffic, network optimisation. and greater ease of management. For example, if you were to borrow one bit from the host portion of a Class B network, your subnet mask would be 255.255.128.0

Subnet Mask

There are two parts to the IP address, the network portion and the host portion. Node assigned that IP address as well as other nodes that must communicate with it have no idea of the location of the line between host and network portions of the address. The subnet mask provides the answer to this dilemma. The subnet mask follows the IP address and details the line indicating where the network portion ofthe address ends and the host portion begin.

Like the IP address, the subnet mask is in a 4-octet, 32-byte format.

An example of a subnet mask is 255.0.0.0. a value of 255 means match all. Each of the threeconfigurable IP address

Classes has a default subnet mask:1: Class A 255.0.0.0

2: Class B 255.255.0.0

3: Class C 255.255.255.0

IPv4 (INTERNET PROTOCOL VERSION 4)

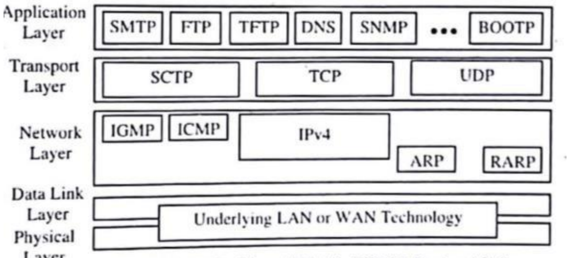
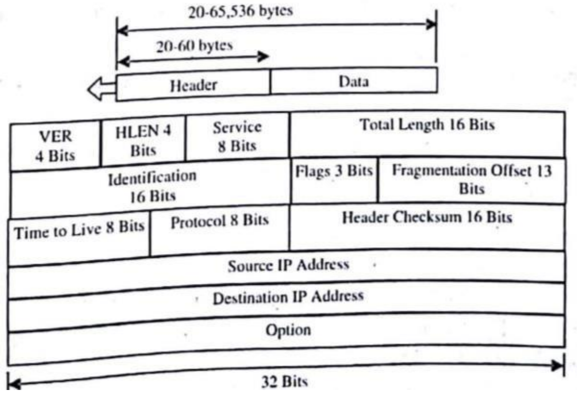


Figure 4.19: Position of IPv4 in TCP/IP Protocol Suite

IPv4 is an unreliable and connectionless datagram protocol a best-effort delivery service.The term best-effort means that IPv4 provides no error control or flow control (except for errordetection on the header). IPv4 assumes the unreliability of the underlying layers and does its best to get a transmission through to itsdestination. but with no guarantees.

IPv4 Datagram Format



Limitation of IPv4

1: The IP address relies on network layer address to identify end-points on networks, andeach networked device has a unique IP address 2: Uses a 32-bit addressing scheme, which givesit 4 billion possible addresses

3: If a network has slightly more number of hostthan a particular class, then it needs either two IP addresses of that class or the next class of IP address

4: Identified limitations of the IPv4 protocol areComplex host and router configuration, non- hierarchical addressing. difficulty in renumbering address large routing tables, non-trivial implementations in providing security.

QoS (Quality of Services) mobility and multi-homing, multicasting etc.

MODULE V

Internet Control Protocols – ICMP, ARP, RARP, BOOTP. Internet Multicasting – IGMP,Exterior Routing Protocols – BGP. IPv6 – Addressing – Issues, ICMPv6.

The Internet has several control protocols used in the network layer, including ICMP, ARP, RARP, BOOTP, and DHCP.

ICMP - The Internet Control Message Protocol

The operation of the Internet is monitored closely by the routers. When something unexpected occurs, the event is reported by the ICMP (Internet Control Message Protocol), which is also used to test the Internet. Each ICMP message type is encapsulated in an IP packet.

Types of messages:

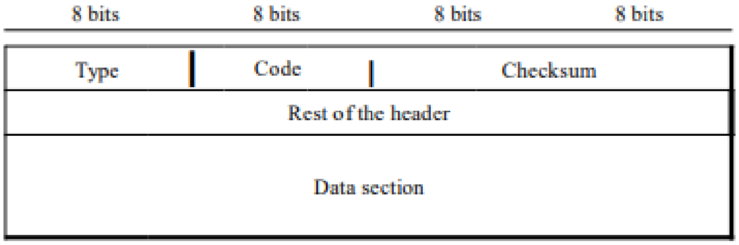
ICMP messages are divided into two broad categories: **error-reporting messages** and **query messages.**

- The **error-reporting messages** report problems that a router or a host (destination) may encounter when it processes an IP packet.
- The **query messages**, which occur in pairs, help a host or a network manager get specific information from a router or another host. For example, nodes can discover their neighbors

Message Format

An ICMP message has an 8-byte header and a variable-size data section. Although the general format of the header is different for each message type, the first 4 bytes are common to all.

Figure 21.8 *Generalformat of[CM? messages*



ICMP type, defines the type of the message.

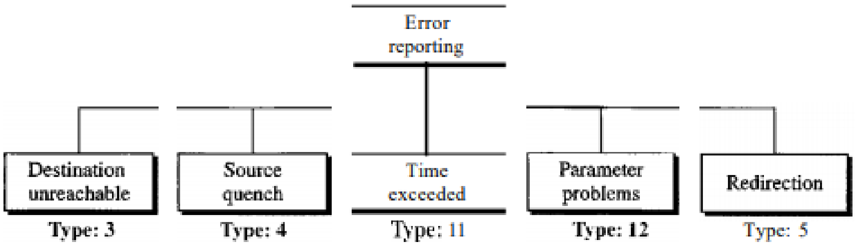
The **code field** specifies the reason for the particular message type.The last common field is the **checksum field** for error handling.

The **rest of the header** is specific for each message type.

The **data section** in error messages carries information for finding the original packet that had the error. In query messages, the data section carries extra information based on the type of the query.

One of the main responsibilities of ICMP is to report errors. ICMP always reports error messagesto the original source. Five types of errors are handled: destination unreachable, source quench, time exceeded, parameter problems, and redirection

Figure 21.9 *Error-reporting messages*



The following are **important points about ICMP error messages:**

- No ICMP error message will be generated in response to a datagram carrying an ICMP error message.
- No ICMP error message will be generated for a fragmented datagram that is not the first fragment.
- No ICMP error message will be generated for a datagram having a multicast address.
- No ICMP error message will be generated for a datagram having a special address such as 127.0.0.0 or 0.0.0.0.

IP header of the original datagram plus the first 8 bytes of data in that datagram. The original datagram header is added to give the original source, which receives the error message, informa tion about the datagram itself. The 8 bytes of data are included because on UDP and TCP protocols, the first 8 bytes provide information about the port numbers (UDP and TCP) and sequence number (TCP). This information is needed so the source can inform the protocols (TCPor UDP) about the error.

The **DESTINATION UNREACHABLE** message is used

- when the subnet or a router cannot locate the destination or
- when a packet with the DF bit cannot be delivered because a "small-packet" network stands in the way.

The **TIME EXCEEDED** message is sent

The time-exceeded message is generated in two cases:

- routers use routing tables to find the next hop (next router) that must receive the packet. If there are errors in one or more routing tables, a packet can travel in a loop or a cycle, going from one router to the next or visiting a series of routers endlessly. Each datagram contains a field called time to live that controls this situation. When a datagram visits a router, the value of this field is decremented by 1. When the time-to-live value reaches 0, after decrementing, the router discards the datagram. How ever, when the datagram is discarded, a time-exceeded message must be sent by the router to the original source. Second, a time-exceeded message is also generated when not all fragments that make up a message arrive at the destination host within a certain time limit.
- when a packet is dropped because its counter has reached zero.
- This event is a symptom that packets are looping, that there is enormous congestion, or

- The timer values are being set too low.
- If there are errors in one or more routing tables, a packet can travel in a loop or a cycle, going from one router to the next or visiting a series of routers endlessly. each datagram contains a field called time to live that controls this situation. When a datagram visits a router, the value of this field is decremented by 1. When the time-to-live value reaches 0, after decrementing, the router discards the datagram. However, when the datagram is discarded, a time-exceeded message must be sent by the router to the original source. Second, a time-exceeded message is also generated when not all fragments that make up a message arrive at the destination host within a certain time limit.

The **PARAMETER PROBLEM** message indicates

- an illegal value has been detected in a header field.
- A bug in the sending host's IP software or possibly in the software of a router transited.
- If a router or the destination host discovers an ambiguous or missing value in any field of the datagram, it discards the datagram and sends a parameter-problem message back to the source.
- Any ambiguity in the header part of a datagram can Create serious problems as the data gram travels through the Internet. If a router or the destination host discovers an ambig uous or missing value in any field of the datagram, it discards the datagram and sends a parameter-problem message back to the source.

The **SOURCE QUENCH** message was used

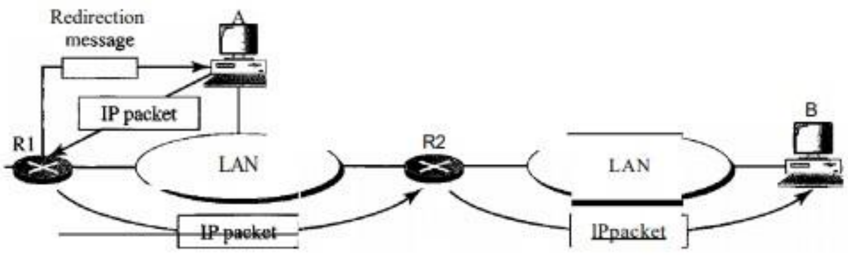
- to throttle hosts that were sending too many packets. When a host received this message, it was expected to slow down. It is rarely used any more because when congestion occurs, these packets tend to worsen it.
- The source-quench message in ICMP was designed to add a kind of flow control to the IP. When a router or host discards a datagram due to congestion, it sends a source-quench message to the sender of the datagram. This message has two purposes. First, it informs the source that the datagram has been discarded. Second, it warns the source that there is congestion somewhere in the path and that the source should slow down (quench) the sending process.
- The IP protocol is a connectionless protocol. There is no communication between the source host, which produces the datagram, the routers, which forward it, and the destination host, which processes it. One of the ramifications of this absence of communication is the lack of flow control. IP does not have a flow control mechanism embedded in the protocol. The lack of flow control can create a major problem in the operation of IP: congestion. The source host never knows if the routers or the destination host has been overwhelmed with datagrams. The source host never knows if it is producing datagrams faster than can be forwarded by routers or processed by the destination host. The lack of flow control can create congestion in routers or the destination host. A router or a host has a limited-size queue (buffer) for incoming datagrams waiting to be forwarded (in the case of a router) or to be processed (in the case of a host). If the datagrams are received much faster than they can be forwarded or processed, the queue may overflow. In this case, the router or the host has no choice but to discard some of the datagrams. The source-quench message in ICMP was designed to add a kind of flow control to the IP. When a router or host discards a datagram due to con gestion, it sends a source-quench message to the sender of the datagram. This

message has two purposes. First, it informs the source that the datagram has been discarded. Second, it warns the source that there is congestion somewhere in the path and that the source should slow down (quench) the sending process.

The **REDIRECT** message is used

- when a router notices that a packet seems to be routed wrong. It is used by the router to tell the sending host about the probable error.
- When a router needs to send a packet destined for another network, it must know the IP address of the next appropriate router. The same is true if the sender is a host. Both routers and hosts, then, must have a routing table to find the address of the router or the next router. Routing is dynamic.
- However, for efficiency, hosts do not take part in the routing update process because there are many more hosts in an internet than routerS. Updating the routing tables of hosts dynamically produces unacceptable traffic. The hosts usually use static routing. When a host comes up, its routing table has a limited number of entries. It usually knows the IP address of only one router, the default router. For this reason, the host may send a data gram, which is destined for another network, to the wrong router. In this case, the router that receives the datagram will forward the datagram to the correct router. However, to update the routing table of the host, it sends a redirection message to the host. This concept of redirection is shown in Figure 21.11. Host A wants to send a datagram to host B.

Figure **21.11** *Redirection concept*



- Router R2 is obviously the most efficient routing choice, but host A did not choose router R2. The datagram goes to R1 instead. Router R1, after consulting its table, finds that the packet should have gone to R2. It sends the packet to R2 and, at the same time, sends a redirection message to host A. Host A's routing table can now be updated.

Query

ICMP can diagnose some network problems. This is accomplished through the query messages, a group of four different pairs of messages, as shown in Figure 21.12. In this type of ICMP message, a node sends a message that is answered in a specific format by the destination node. Aquery message is encapsulated in an IP packet, which in tum is encapsulated in a data link layer frame.

Figure 21.12 *Query messages*

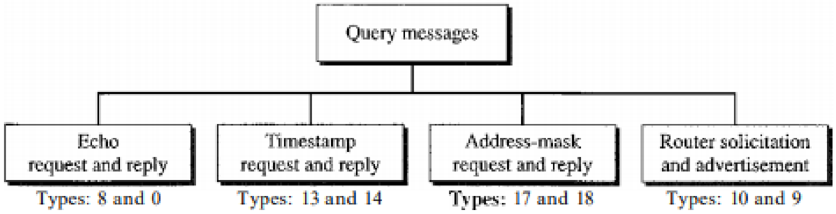
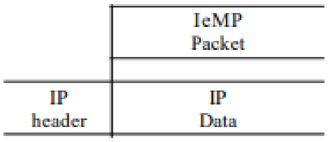


Figure 21.13 *Encapsulation ofICMP query messages*



The **ECHO and ECHO REPLY** messages are used to

- see if a given destination is reachable and alive.
- Upon receiving the ECHO message, the destination is expected to send an ECHO REPLY message back.
- The echo-request and echo-reply messages are designed for diagnostic purposes. Network managers and users utilize this pair of messages to identify network problems. The combination of echo-request and echo-reply messages determines whether two systems (hosts or routers) can communicate with each other. The echo-request and echo-reply messages can be used to determine if there is communication at the IP level. Because ICMP messages are encapsulated in IP datagrams, the receipt of an echo-reply message by the machine that sent the echo request is proof that the IP protocols in the sender and receiver are communicating with each other using the IP datagram. Also, it is proof that the intermediate routers are receiving, processing, and forwarding IP datagrams

The **TIMESTAMP REQUEST and TIMESTAMP REPLY** messages are similar,

- except that the arrival time of the message and the departure time of the reply are recorded in the reply.
- This facility is used to measure network performance.
- Two machines (hosts or routers) can use the timestamp request and timestamp reply messages to determine the round-trip time needed for an IP datagram to travel between them. It can also be used to synchronize the clocks in two machines.

ADDRESS-MASK REQUEST AND REPLY

A host may know its IP address, but it may not know the corresponding mask. For example, a host may know its IP address as 159.31.17.24, but it may not know that the corresponding mask is /24. To obtain its mask, a host sends an address-mask-request message to a router on the LAN.

If the host knows the address of the router, it sends the request directly to the router. If it does not know, it broadcasts the message. The router receiving the address-mask-request message responds with an address-mask-reply message, providing the necessary

ROUTER SOLICITATION AND ADVERTISEMENT

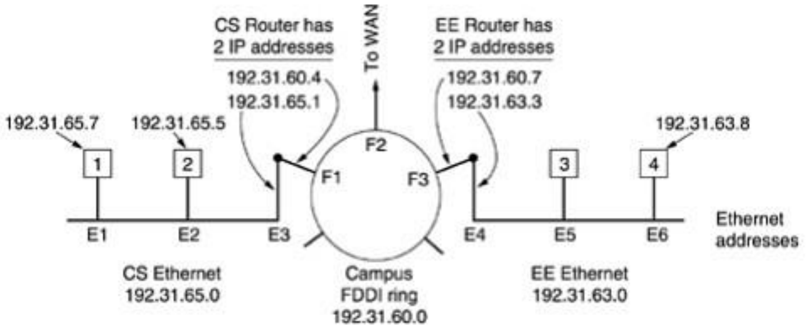
A host that wants to send data to a host on another network needs to know the address of routers connected to its own network. Also, the host must know if the routers are alive and functioning. The router-solicitation and router-advertisement messages can help in this situation. A host can broadcast (or multicast) a router-solicitation message. The router or routers that receive the solicitation message broadcast their routing information using the router-advertisement message. A router can also periodically send router-advertisement messages even if no host has solic ited.Note that when a router sends out an advertisement, it announces not only its own presence but also the presence of all routers on the network of which it is aware.

ARP—The Address Resolution Protocol - ARP solves the problem of finding out which Ethernet address corresponds to a given IP address.

Every machine on the Internet has one (or more) IP addresses, these cannot actually be used for sending packets because the data link layer hardware does not understand Internet addresses. Most hosts at organizations are attached to a LAN by an interface board that only understands LAN addresses. For example, every Ethernet board ever manufactured comes equipped with a unique 48-bit Ethernet address. The boards send and receive frames based on 48-bit Ethernet addresses.

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

Figure 5-62. Three interconnected /24 networks: two Ethernets and an FDDI ring.



Small university with several class C (now called /24) networks is illustrated. We have two Ethernets, one in the Computer Science Dept., with IP address 192.31.65.0 and one in Electrical Engineering, with IP address 192.31.63.0. These are connected by a campus backbone ring (e.g., FDDI) with IP address 192.31.60.0. Each machine on an Ethernet has a unique Ethernet address, labeled E1 through E6, and each machine on the FDDI ring has an FDDI address, labeled F1 through F3.

- **a user on host 1 sends a packet to a user on host 2:**

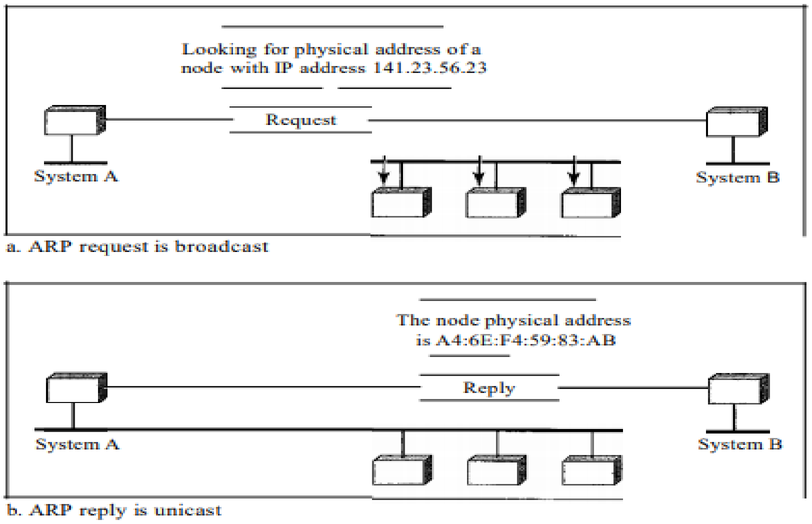
1. assume the sender knows the name of the intended receiver, eg:mary@eagle.cs.uni.edu
2. find the IP address for host 2, known as eagle.cs.uni.edu.

3. This lookup is performed by the Domain Name System, that DNS returns the IP address for host 2 (192.31.65.5).
4. The upper layer software on host 1 now builds a packet with 192.31.65.5 in the Destination address field and gives it to the IP software to transmit
5. The IP software can look at the address and see that the destination is on its own network, but it needs to find the destination's Ethernet address:

- One solution is to have a **configuration file/static mapping** somewhere in the system that maps IP addresses onto Ethernet addresses. While this solution is certainly possible, for organizations with thousands of machines, keeping all these files up to date is an error-prone, time-consuming job. Limitations of static mapping:
    - 1. A machine could change its NIC, resulting in a new physical address.
    - 2. In some LANs, such as LocalTalk, the physical address changes every time the computer is turned on.
    - 3. A mobile computer can move from one physical network to another, resulting in a change in its physical address.
    - To implement these changes, a static mapping table must be updated periodically. This overhead could affect network performance
- Host 1 to output a broadcast packet onto the Ethernet asking: Who owns IP address 192.31.65.5? The broadcast will arrive at every machine on Ethernet 192.31.65.0, and each one will check its IP address. Host 2 alone will respond with its Ethernet address (E2). The protocol used for asking this question and getting the reply is called **ARP (Address Resolution Protocol).** Almost every machine on the Internet runs it. ARP is defined in RFC 826.
- The advantage of using ARP over configuration files is the simplicity. The system manager does not have to do much except assign each machine an IP address and decide about subnet masks. ARP does the rest.

6. The IP software on host 1 builds an Ethernet frame addressed to E2, puts the IP packet (addressed to 192.31.65.5) in the payload field, and dumps it onto the Ethernet.
7. The Ethernet board of host 2 detects this frame, recognizes it as a frame for itself, scoops it up, and causes an interrupt.
8. The Ethernet driver extracts the IP packet from the payload and passes it to the IP software, which sees that it is correctly addressed and processes it.

Figure 21.1 *ARP operation*


Looking for physical address of a node with IP address 141.23.56.23
Request
System A
System B
The node physical address is A4:6E:F4:59:83:AB
Reply
System A
System B

a. ARP request is broadcast

b. ARP reply is unicast

Optimizations to make ARP work more efficiently:

- All machines on the Ethernet can enter this mapping into their ARP caches.

1. Once a machine has run ARP, it caches the result in case it needs to contact the same machine shortly. Next time it will find the mapping in its own cache, thus eliminating the need for a second broadcast.
2. Host 2 will need to send back a reply, forcing it, too, to run ARP to determine the sender's Ethernet address. This ARP broadcast can be avoided by having host 1 include its IP-to-Ethernet mapping in the ARP packet. When the ARP broadcast arrives at host 2, the pair (192.31.65.7, E1) is entered into host 2's ARP cache for future use.

- Every machine broadcast its mapping when it boots. This broadcast is generally done in the form of an ARP looking for its own IP address. There should not be a response, but a side effect of the broadcast is to make an entry in everyone's ARP cache. If a response does (unexpectedly) arrive, two machines have been assigned the same IP address. The new one should inform the system manager and not boot.
- every machine broadcast its mapping when it boots. This broadcast is generally done in the form of an ARP looking for its own IP address. There should not be a response, but a side effect of the broadcast is to make an entry in everyone's ARP cache. If a response does (unexpectedly) arrive, two machines have been assigned the same IP address. The new one should inform the system manager and not boot.

- **host 1 wants to send a packet to host 4 (192.31.63.8) / From host 1 to a distantnetwork over a WAN :**

Using ARP will fail because host 4 will not see the broadcast (routers do not forward Ethernet-level broadcasts). There are two solutions:

1. First, the CS router could be configured to respond to ARP requests for network 192.31.63.0 (and possibly other local networks). In this case, host 1 will make an ARP cache entry of (192.31.63.8, E3) and send all traffic for host 4 to the local router. This solution is called **proxy ARP**.

2. The second solution is to have host 1 immediately see that the destination is on a remote network and just send all such traffic to a default Ethernet address that handles all remote traffic, in this case E3. This solution does not require having the CS router know which remote networks it is serving.

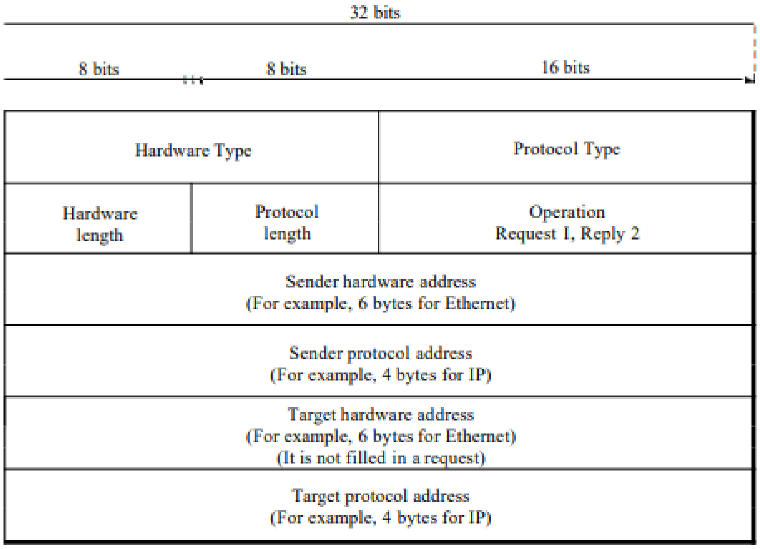
Either way, host 1 packs the IP packet into the payload field of an Ethernet frame addressed to E3. When the CS router gets the Ethernet frame, it removes the IP packet from the payload field and looks up the IP address in its routing tables. It discovers that packets for network 192.31.63.0are supposed to go to router 192.31.60.7. If it does not already know the FDDI address of 192.31.60.7, it broadcasts an ARP packet onto the ring and learns that its ring address is F3. It then inserts the packet into the payload field of an FDDI frame addressed to F3 and puts it on thering.

At the EE router, the FDDI driver removes the packet from the payload field and gives it to the IP software, which sees that it needs to send the packet to 192.31.63.8. If this IP address is not inits ARP cache, it broadcasts an ARP request on the EE Ethernet and learns that the destination address is E6,soit builds an Ethernet frame addressed to E6, puts the packet in the payload field, and sends it over the Ethernet. When the Ethernet frame arrives at host 4, the packet is extracted from the frame and passed to the IP software for processing.

From host 1 to a distant network over a WAN works essentially the same way, except that this time the CS router's tables tell it to use the WAN router whose FDDI address is F2.

ARP PACKET

Figure 21.2 *ARP packet*


32 bits
8 bits
8 bits
16 bits
Hardware Type
Protocol Type
Hardware length
Protocol length
Operation Request 1, Reply 2
Sender hardware address (For example, 6 bytes for Ethernet)
Sender protocol address (For example, 4 bytes for IP)
Target hardware address (For example, 6 bytes for Ethernet) (It is not filled in a request)
Target protocol address (For example, 4 bytes for IP)

Hardware type. This is a 16-bit field defining the type of the network on which ARP is running. Each LAN has been assigned an integer based on its type. For example, Ethernet is given type 1. **Protocol type**. This is a 16-bit field defining the protocol. For example, the value of this field for the IPv4 protocol is 080016.

Hardware length. This is an 8-bit field defining the length of the physical address in bytes. For example, for Ethernet the value is 6.

Protocol length. This is an 8-bit field defining the length of the logical address in bytes. For example, for the IPv4 protocol the value is 4.

Operation. This is a 16-bit field defining the type of packet. Two packet types are defined: ARP request (1) and ARP reply (2).

Sender hardware address. This is a variable-length field defining the physical address of the sender. For example, for Ethernet this field is 6 bytes long.

Sender protocol address. This is a variable-length field defining the logical (for example, IP) address of the sender. For the IP protocol, this field is 4 bytes long.

Target hardware address. This is a variable-length field defining the physical address of the target. For example, for Ethernet this field is 6 bytes long. For an ARP request message, this field is alI Os because the sender does not know the physical address of the target.

Target protocol address. This is a variable-length field defining the logical (for example, IP) address of the target. For the IPv4 protocol, this field is 4 bytes long.

Operation

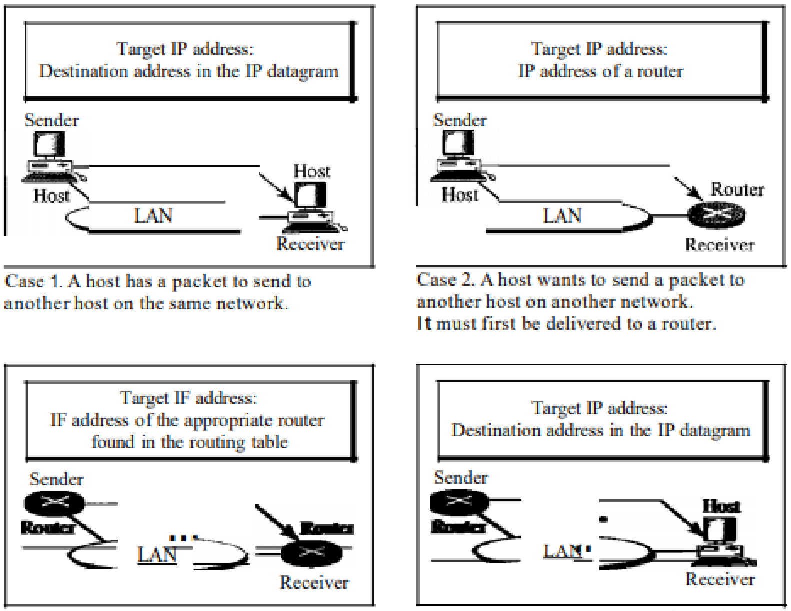
These are the steps involved in an ARP process:

1. The sender knows the IP address of the target.
2. IP asks ARP to create an ARP request message, filling in the sender physical address, the sender IP address, and the target IP address. The target physical address field is filled with Os.
3. The message is passed to the data link layer where it is encapsulated in a frame by using the physical address of the sender as the source address and the physical broadcast address as the destination address.
4. Every host or router receives the frame. Because the frame contains a broadcast destination address, all stations remove the message and pass it to ARP. All machines except the one targeted drop the packet. The target machine recognizes its IP address.
5. The target machine replies with an ARP reply message that contains its physical address. The message is unicast.
6. The sender receives the reply message. It now knows the physical address of the target machine.
7. The IP datagram, which carries data for the target machine, is now encapsulated in a frame and is unicast to the destination.

Four Different Cases

The following are four different cases in which the services of ARP can be used (see Figure 21.4).

Figure 21.4 *Four cases using ARP*



Case 1. A host has a packet to send to another host on the same network.

Case 2. A host wants to send a packet to another host on another network. It must first be delivered to a router.

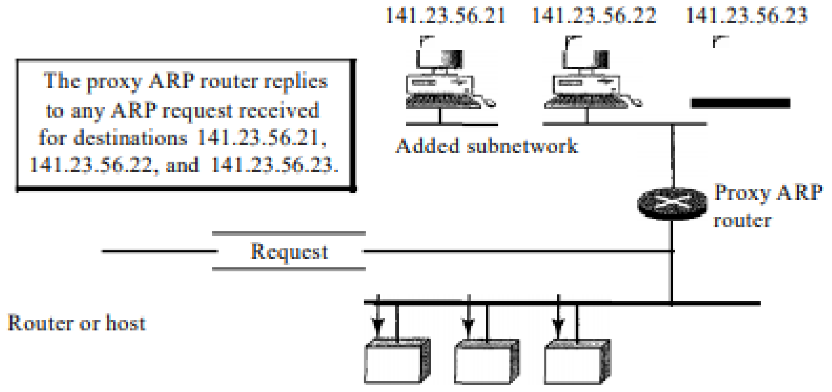
Case 3. A router receives a packet to be sent to a host on another network. It must first be delivered to the appropriate router.

Case 4. A router receives a packet to be sent to a host on the same network.

ProxyARP

A technique called proxy ARP is used to create a subnetting effect. A proxy ARP is an ARP that acts on behalf of a set of hosts. Whenever a router running a proxy ARP receives an ARP request looking for the IP address of one of these hosts, the router sends an ARP reply announcing its own hardware (physical) address. After the router receives the actual IP packet, it sends the packet to the appropriate host or router.

Figure 21.6 *Proxy ARP*


141.23.56.21
141.23.56.22
141.23.56.23
The proxy ARP router replies to any ARP request received for destinations 141.23.56.21, 141.23.56.22, and 141.23.56.23.
Added subnetwork
Proxy ARP router
Request
Router or host

In Figure 21.6 the ARP installed on the right-hand host will answer only to anARP request with a target IP address of 141.23.56.23.

The administrator may need to create a subnet without changing the whole system to recognize subnetted addresses. One solution is to add a router running a proxy ARP. In this case, the router acts on behalf of all the hosts installed on the subnet. When it receives an ARP request with a target IP address that matches the address of one of its proteges (141.23.56.21, 141.23.56.22, or 141.23.56.23), it sends an ARP reply and announces its hardware address as the target hardware address. When the router receives the IP packet, it sends the packet to the appropriate host.

Mapping Physical to Logical Address:

RARP, BOOTP, and DHCP There are occasions in which a host knows its physical address, but needs to know its logical address. This may happen in two cases:

1. A diskless station is just booted. The station can find its physical address by checking its interface, but it does not know its IP address.
2. An organization does not have enough IP addresses to assign to each station; it needs to assign IP addresses on demand. The station can send its physical address and ask for a short time lease

RARP - Reverse Address Resolution Protocol - (defined in RFC 903)

Reverse Address Resolution Protocol (RARP) finds the logical address for a machine that knows only its physical address. Each host or router is assigned one or more logical (IP) addresses, which are unique and independent of the physical (hardware) address of the machine.

To create an IP datagram, a host or a router needs to know its own IP address or addresses. The IP address of a machine is usually read from its configuration file stored on a disk file. However,a diskless machine is usually booted from ROM, which has minimum booting information. The ROM is installed by the manufacturer. It cannot include the IP address because the IP addresses on a network are assigned by the network administrator

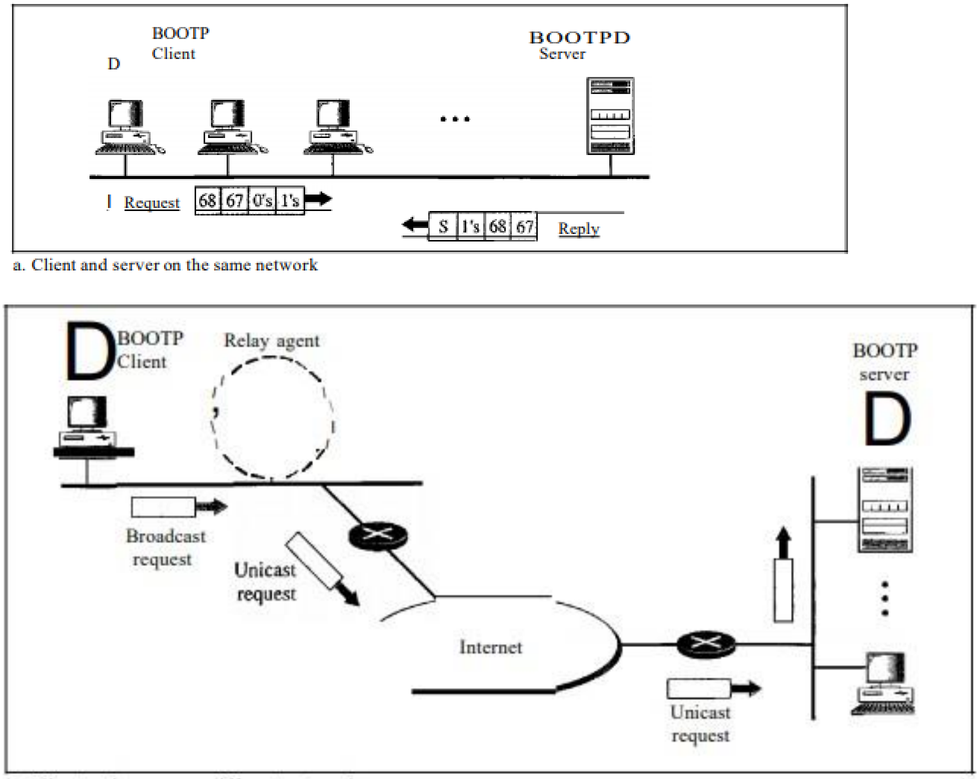
The machine can get its physical address (by reading its NIC, for example), which is unique locally. It can then use the physical address to get the logical address by using the RARP protocol. A RARP request is created and broadcast on the local network. Another machine on thelocal network that knows all the IP addresses will respond with a RARP reply. The requesting machine must be running a RARP client program; the responding machine must be running a RARP server program.

There is a serious problem with RARP: Broadcasting is done at the data link layer. The physical broadcast address, allis in the case of Ethernet, does not pass the boundaries of a network. This means that if an administrator has several networks or several subnets, it needs to assign a RARP server for each network or subnet. This is the reason that RARP is almost obsolete. Two protocols, BOOTP and DHCp, are replacing RARP.

BOOTP

The Bootstrap Protocol (BOOTP) is a client/server protocol designed to provide physical addressto logical address mapping. BOOTP is an application layer protocol. The administrator may put the client and the server on the same network or on different networks, as shown in Figure 21.7. BOOTP messages are encapsulated in a UDP packet, and the UDP packet itself is encapsulated in an IP packet.

Figure 21.7 *BOOTP client and server on the same and different network*



a. Client and server on the same network

b. Client and server on different networks

A client can send an IP datagram when it knows neither its own IP address (the source address) nor the server's IP address (the destination address) by using all 1 as as the source address and all1s as the destination address.

One of the advantages of BOOTP over RARP is that the client and server are application-layer processes.

One problem that must be solved. The BOOTP request is broadcast because the client does not know the IP address of the server. A broadcast IP datagram cannot pass through any router. To solve the problem, there is a need for an intermediary. One of the hosts (or a router that can be configured to operate at the application layer) can be used as a relay. The host in this case is called **a relay agent**. The relay agent knows the unicast address of a BOOTP server. When it receives this type of packet, it encapsulates the message in a unicast datagram and sends the request to the BOOTP server. The packet, carrying a unicast destination address, is routed by anyrouter and reaches the BOOTP server. The BOOTP server knows the message comes from a relay agent because one of the fields in the request message defines the IP address of the relay agent. The relay agent, after receiving the reply, sends it to the BOOTP client.

BOOTP is not a dynamic configuration protocol. When a client requests its IP address, the BOOTP server consults a table that matches the physical address of the client with its IP address.This implies that the binding between the physical address and the IP address of the client already exists. The binding is predetermined.

If a host moves from one physical network to another. If a host wants a temporary IP address. BOOTP cannot handle these situations because the binding between the physical and IP addresses is static and fixed in a table until changed by the administrator. BOOTP is a static configuration protocol

Internet Multicasting – IGMP:

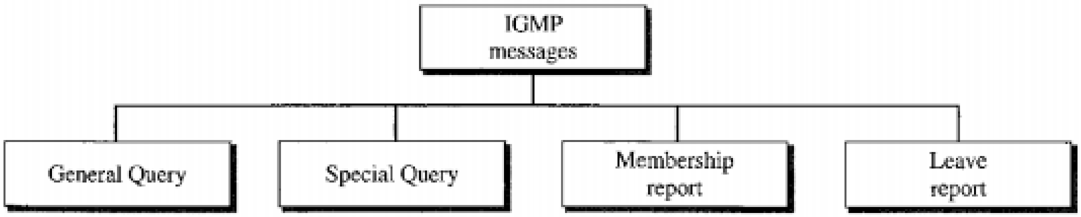
The IP protocol can be involved in two types of communication**: unicasting** and **multicasting**. **Unicasting** is the communication between one sender and one receiver. It is a one-to-one communication. Processes which send the same message to a large number of receivers simultaneously is called **multicasting**, which is a one-to-many communication. Examples are updating replicated, distributed databases, transmitting stock quotes to multiple brokers, and handling digital conference (i.e., multiparty) telephone calls.

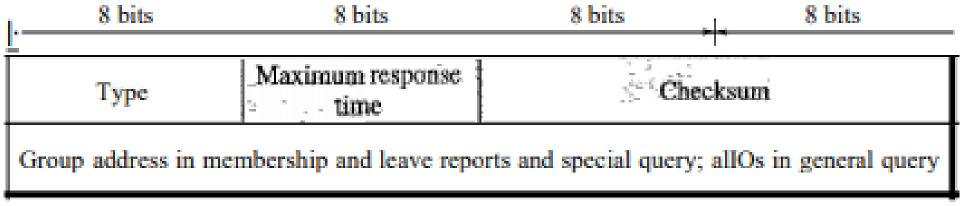
The Internet Group Management Protocol (IGMP) is one of the necessary, but not sufficient protocols that is involved in multicasting. Group Management For multicasting in the Internet we need routers that are able to route multicast packets. The routing tables of these routers must be updated by using one of the multicasting routing protocols.

IGMP is not a multicasting routing protocol; it is a protocol that **manages group membership**.In any network, there are one or more multicast routers that distribute multicast packets to hostsor other routers. The IGMP protocol gives the multicast routers information about the membership status of hosts (routers) connected to the network. IGMP is a group management protocol. It helps a multicast router create and update a list of loyal members related to each router interface.

IGMP Messages:

IGMPv2 has three types of messages: the query, the membership report, and the leave report. There are two types of query messages: general and special.




8 bits
8 bits
8 bits
8 bits
Type
Maximum response time
Checksum
Group address in membership and leave reports and special query; allOs in general query

Message Format:

Type. This 8-bit field defines the type of message, as shown in Table 21.1. The value of the typeis shown in both hexadecimal and binary notation.

Table 21.1 *IGMP type field*

Type	*Value*
General or special query	Ox11 or 00010001
Membership report	Ox16 or 00010110
Leave report	Ox17 or 00010111

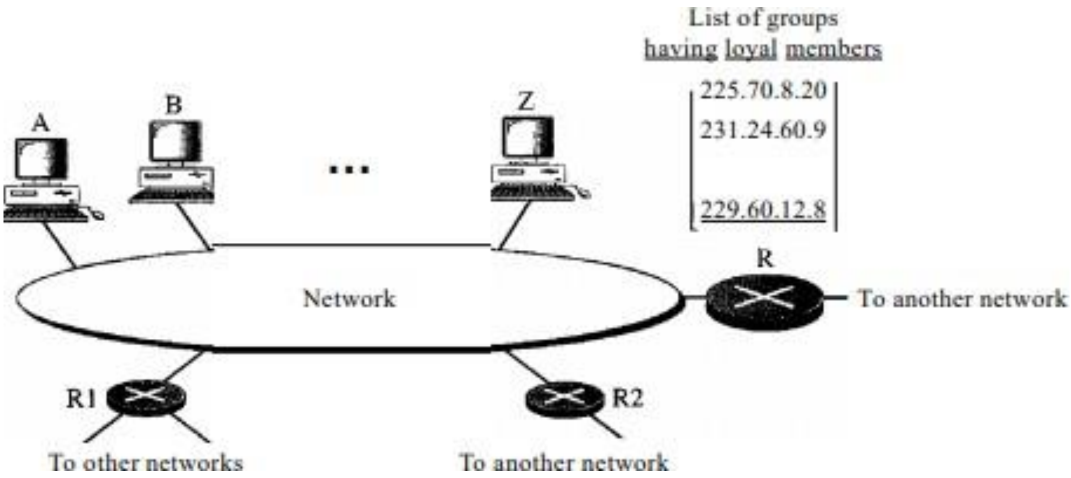
Maximum Response Time. This 8-bit field defines the amount of time in which a query must be answered. The value is in tenths of a second; for example, if the value is 100, it means 10 s. The value is nonzero in the query message; it is set to zero in the other two message types.

Checksum. This is a 16-bit field carrying the checksum. The checksum is calculated over the 8-byte message.

Group address. The value of this field is 0 for a general query message. The value defines the groupid (multicast address of the group) in the special query, the membership report, and the leave report messages.

IGMP Operation

IGMP operates locally. A multicast router connected to a network has a list of multicast addresses of the groups with at least one loyal member in that network.



For each group, there is one router that has the duty of distributing the multicast packets destined for that group. This means that if there are three multicast routers connected to a network, their lists of groupids are mutually exclusive. For example, in Figure only router R distributes packets with the multicast address of 225.70.8.20. A host or multicast router can have membership in a group. When a **host has membership**, it means that one of its processes (an application program) receives multicast packets from some group. When a **router has membership**, it means that a network connected to one of its other interfaces receives these multicast packets.
We say that the host or the router has **an interest in the group**. In both cases, the host and the router keep a list of groupids and relay their interest to the distributing router.

In the Figure, router R is the distributing router. There are two other multicast routers (R1 and R2) that, depending on the group list maintained by router R, could be the recipients of router Rin this network. Routers RI and R2 may be distributors for some of these groups in other networks, but not on this network.

1. Joining a Group

A host or a router can join a group. A host maintains a list of processes that have membership in a group. When a process wants to join a new group, it sends its request to the host. The host addsthe name of the process and the name of the requested group to its list. If this is the first entry forthis particular group, the host sends a membership report message. If this is not the first entry, there is no need to send the membership report. The protocol requires that the membership **report be sent twice**, one after the other within a few moments. In this way, if the first one is lost or damaged, the second one replaces it.

2. Leaving a Group

When a host sees that no process is interested in a specific group, it sends a leave report. Similarly, when a router sees that none of the networks connected to its interfaces is interested ina specific group, it sends a leave report about that group. When a multicast router receives a leave report, The router allows a specified time for any host or router to respond. If, during this time, no interest (membership report) is received, the router assumes that there are no loyal members in the network and purges the group from its list.

3. Monitoring membership

There is only one host interested in a group, but the host is shut down or removed from the system. The multicast router will never receive a leave report. The multicast router is responsiblefor monitoring all the hosts or routers in a LAN to see if they want to continue their membershipin a group. The router periodically (by default, every 125 s) sends a general query message. In this message, the group address field is set to 0.0.0.0. This means the query for membership continuation is for all groups in which a host is involved. The general query message does not define a particular group.

The query message has a maximum response time of 10. When a host or router receives the general query message, it responds with a membership report if it is interested in a group. However, if there is a common interest (two hosts, for example, are interested in the same group), only one response is sent for that group to prevent unnecessary traffic. This is called a **delayed response**. The query message must be sent by only one router called **the query router**,also to prevent unnecessary traffic.

4. Delayed Response

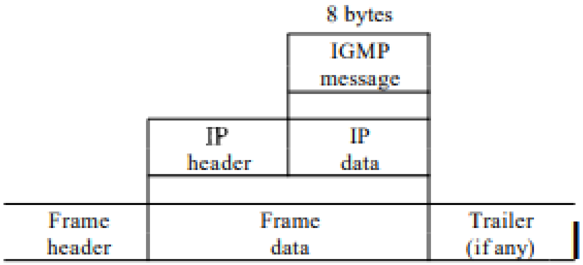
To prevent unnecessary traffic, IGMP uses a delayed response strategy. When a host or router receives a query message, it does not respond immediately; it delays the response. Each host or router uses a random number to create a timer, which expires between I and lOs. A timer is set for each group in the list. For example, the timer for the first group may expire in 2 s, but the timer for the third group may expire in 5 s. Each host or router waits until its timer has expired before sending a membership report message. During this waiting time, if the timer of another host or router, for the same group, expires earlier, that host or router sends a membership report.Because, the report is broadcast, the waiting host or router receives the report and knows that there is no need to send a duplicate report for this group; thus, the waiting station cancels its corresponding timer.

5. Query Router

Query messages may create a lot of responses. To prevent unnecessary traffic, IGMP designates one router as the query router for each network. Only this designated router sends the query message, and the other routers are passive.

Encapsulation

The IGMP message is encapsulated in an IP datagram, which is itself encapsulated in a frame.



Encapsulation at Network Layer

The value of the protocol field is 2 for the IGMP protocol. Every IP packet carrying this value inits protocol field has data delivered to the IGMP protocol. When the message is encapsulated in the IP datagram, the value of TTL must be 1. This is required because the domain of IGMP is theLAN. No IGMP message must travel beyond the LAN. A TTL value of 1 guarantees that the message does not leave the LAN since this value is decremented to 0 by the next router and, consequently, the packet is discarded.

Encapsulation at Data Link Layer

At the network layer, the IGMP message is encapsulated in an IP packet and is treated as an IP packet. However, because the IP packet has a multicast IP address, the ARP protocol cannot find the corresponding MAC (physical) address to forward the packet at the data link layer. What happens next depends on whether the underlying data link layer supports physical multicast addresses.

Physical Multicast Support Most LANs support physical multicast addressing. An Ethernet physical address (MAC address) is six octets (48 bits) long. If the first 25 bits in an Ethernet address identifies a physical multicast address for the TCP/IP protocol. The remaining 23 bits can be used to define a group. **To convert an IP multicast address into an Ethernet address**,the multicast router extracts the least significant 23 bits of a class D IP address and inserts theminto a multicast Ethernet physical address.

An Ethernet multicast physical address is in the range 01 :00:5E:00:OO:OO to01:00:5E:7F:FF:FF.

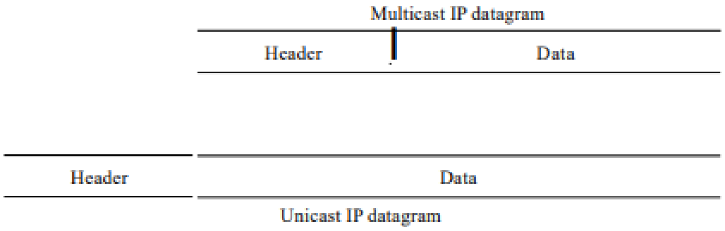
Change the multicast IP address 230.43.14.7 to an Ethernet multicast physical address.

Solution in two steps:
a. Write the rightmost 23 bits of the IP address in hexadecimal. This can be done by changing the rightmost 3 bytes to hexadecimal and then subtracting 8 from the leftmost digit if it is greater than or equal to 8. In our example, the result is 2B:OE:07.
b. We add the result of part a to the starting Ethernet multicast address, which is 01:00:5E:00:00:00. The result is 01:00:5E:2B:OE:07

Change the multicast IP address 238.212.24.9 to an Ethernet multicast address.

Solution
a. The rightmost 3 bytes in hexadecimal is D4: 18:09. We need to subtract 8 from the leftmost digit, resulting in 54:18:09.
b. We add the result of part a to the Ethernet multicast starting address. The result is

No Physical Multicast Support Most WANs do not support physical multicast addressing. To send a multicast packet through these networks, a process called **tunneling** is used. In tunneling, the multicast packet is encapsulated in a unicast packet and sent through the network, where it emerges from the other side as a multicast packet.



EXTERIOR ROUTING PROTOCOLS – BGP

Border Gateway Protocol (BGP) is an interdomain routing protocol using path vector routing. The Internet is divided into hierarchical domains called **autonomous systems.** Autonomous systems are divided into three categories: stub, multihomed, and transit.

Stub AS.

A stub AS has only one connection to another AS. The interdomain data traffic in a stub AS canbe either created or terminated in the AS. The hosts in the AS can send data traffic to other ASs.The hosts in the AS can receive data coming from hosts in other ASs. Data traffic, however, cannot pass through a stub AS. A stub AS is either a source or a sink. A good example of a stubAS is a small corporation or a small local ISP.

Multihomed AS.

A multihomed AS has more than one connection to other ASs, but it is still only a source or sink for data traffic. It can receive data traffic from more than one AS. It can send data traffic to more than one AS, but there is no transient traffic. It does not allow data coming from one AS and going to another AS to pass through. A good example of a multihomed AS is a large corporationthat is connected to more than one regional or national AS that does not allow transient traffic.

Transit AS.

A transit AS is a multihomed AS that also allows transient traffic. Good examples of transit ASs are national and international ISPs (Internet backbones).

Attributes are divided into two broad categories: well known and optional. **A well-known** attribute is one that every BGP router must recognize. An **optional attribute** is one that needs not be recognized by every router.

Well-known attributes are themselves divided into two categories: mandatory and discretionary.**A well-known mandatory attribute** is one that must appear in the description of a route. A well-known discretionary attribute is one that must be recognized by each router, but is not required to be included in every update message.
Examples:
ORIGIN. This defines the source of the routing information (RIP, OSPF, and so on).
AS_PATH. This defines the list of autonomous systems through which the destination can be reached.
NEXT-HOP, which defines the next router to which the data packet should be sent.

The optional attributes can also be subdivided into two categories: transitive and nontransitive.An **optional transitive attribute** is one that must be passed to the next router by the router that has not implemented this attribute.

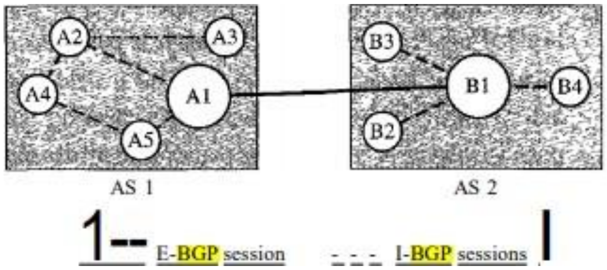
An **optional nontransitive attribute** is one that must be discarded if the receiving router has notimplemented it.

BGP Sessions

The exchange of routing information between two routers using BGP takes place in a session. A session is a connection that is established between two BGP routers only for the sake of exchanging routing information. To create a reliable environment, BGP uses the services of TCP. For this reason, BGP sessions are sometimes referred to as **semi-permanent connections**.

BGP can have two types of sessions: **external BGP (E-BGP)** and **internal BGP (I-BGP)**

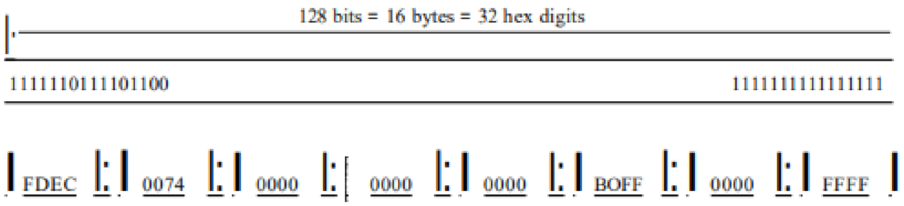
sessions.
The **E-BGP session** is used to exchange information between two speaker nodes belonging to two different autonomous systems.
The **I-BGP session**, on the other hand, is used to exchange routing information between two routers inside an autonomous systems.



IPV6 – ADDRESSING – ISSUES

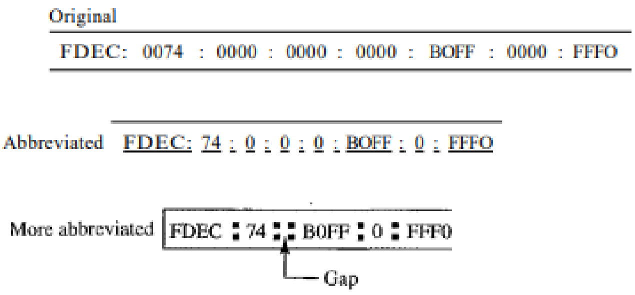
An IPv6 address consists of 16 bytes (octets); it is 128 bits long. IPv6 specifies hexadecimal colon notation. In this notation, 128 bits is divided into eight sections, each 2 bytes in length. Two bytes in hexadecimal notation requires four hexadecimal digits. Therefore, the address consists of 32 hexadecimal digits, with every four digits separated by a colon.

Figure 19.14 *IPv6 address in binary and hexadecimal colon notation*



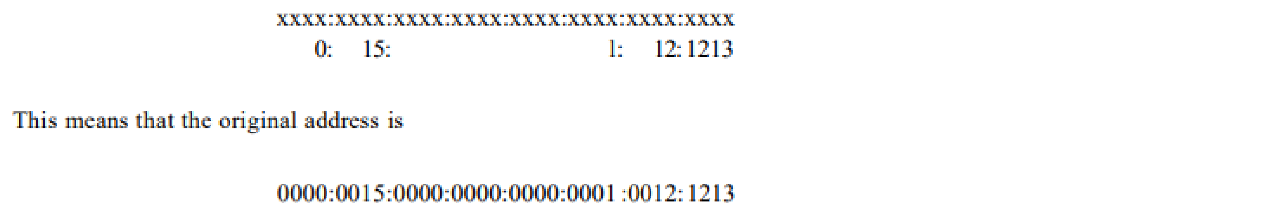
Although the IP address, even in hexadecimal format, is very long, many of the digits are zer9s.In this case, we can abbreviate the address. The leading zeros of a section (four digits between two colons) can be omitted. Only the leading zeros can be dropped, not the trailing zeros.

Figure 19.15 *Abbreviated IPv6 addresses*


Original
FDEC: 0074 : 0000 : 0000 : 0000 : BOFF : 0000 : FFFO
Abbreviated FDEC: 74 : 0 : 0 : 0 : BOFF : 0 : FFFO
More abbreviated FDEC : 74 :: B0FF : 0 : FFF0
Gap

0074 can be written as 74, OOOF as F, and 0000 as O. We can remove the zeros altogether and replace them with a double semicolon. Note that this type of abbreviation is allowed only once per address. If there are two runs of zero sections, only one of them can be abbreviated.

Expand the address 0:15::1:12:1213 to its original.



```
xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx
   0:   15:                   1:   12:1213
```

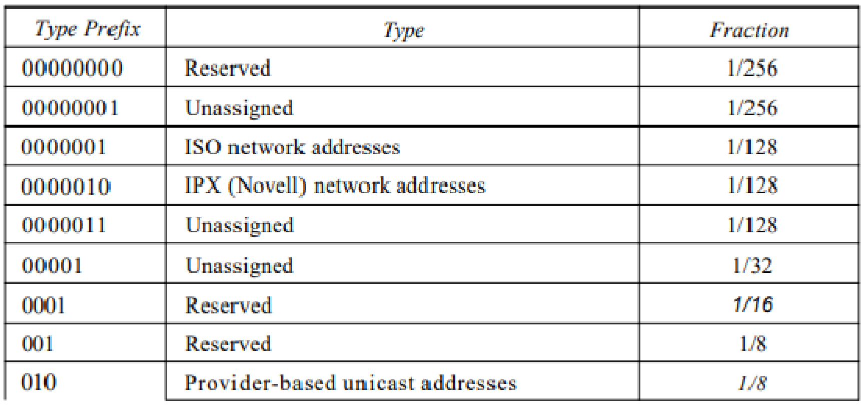
This means that the original address is

```
0000:0015:0000:0000:0000:0001 :0012: 1213
```

Address Space

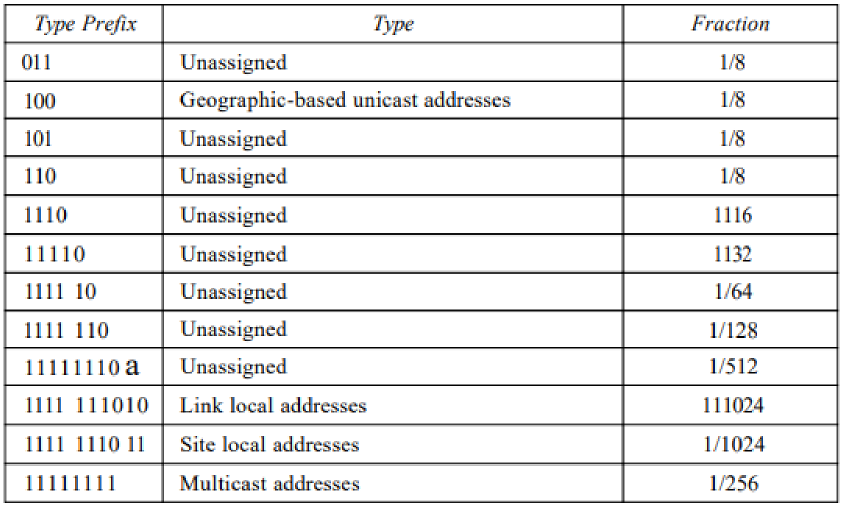
IPv6 has 2^{128} addresses available. IPv6 divided the address into several categories. A few leftmost bits, called **the type prefix**, in each address define its category. The type prefix is variable in length, but it is designed such that no code is identical to the first part of any other code. In this way, there is no ambiguity; when an address is given, the type prefix can easily be determined.

Table 19.5 *Type prefixes for IPv6 addresses*



Type Prefix	*Type*	*Fraction*
00000000	Reserved	1/256
00000001	Unassigned	1/256
0000001	ISO network addresses	1/128
0000010	IPX (Novell) network addresses	1/128
0000011	Unassigned	1/128
00001	Unassigned	1/32
0001	Reserved	1/16
001	Reserved	1/8
010	Provider-based unicast addresses	*1/8*

Table 19.5 *Type prefixes for IPv6 addresses (continued)*

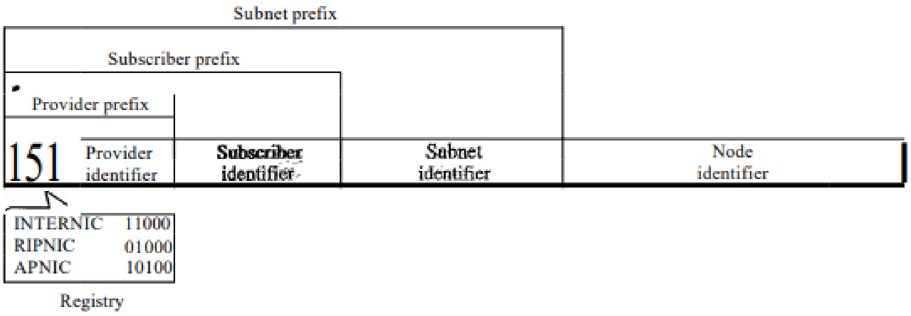


Type Prefix	*Type*	*Fraction*
011	Unassigned	1/8
100	Geographic-based unicast addresses	1/8
101	Unassigned	1/8
110	Unassigned	1/8
1110	Unassigned	1116
11110	Unassigned	1132
1111 10	Unassigned	1/64
1111 110	Unassigned	1/128
11111110 a	Unassigned	1/512
1111 111010	Link local addresses	111024
1111 1110 11	Site local addresses	1/1024
11111111	Multicast addresses	1/256

Unicast Addresses

A unicast address defines a single computer. The packet sent to a unicast address must be delivered to that specific computer. IPv6 defines two types of unicast addresses: **geographically based** and **provider-based**. The provider-based address is generally used by a normal host as a unicast address.

Figure 19.16 *Prefixes for provider-based unicast address*

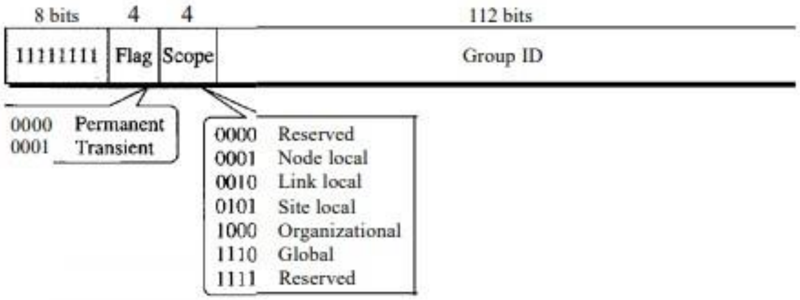


Type identifier. This 3-bit field defines the address as a provider-based address.
Registry identifier. This 5-bit field indicates the agency that has registered the address. **Provider identifier.** This variable-length field identifies the provider for Internet access (such asan ISP). A 16-bit length is recommended for this field.
Subscriber identifier. When an organization subscribes to the Internet through a provider, it is assigned a subscriber identification. A 24-bit length is recommended for this field.
Subnet identifier. Each subscriber can have many different subnetworks, and each subnetwork can have an identifier. The subnet identifier defines a specific subnetwork under the territory of the subscriber. A 32-bit length is recommended for this field.
Node identifier. The last field defines the identity of the node connected to a subnet. A length of 48 bits is recommended for this field to make it compatible with the 48-bit link (physical) address used by Ethernet. In the future, this link address will probably be the same as the node physical address.

Multicast Addresses

Multicast addresses are used to define a group of hosts instead of just one. A packet sent to a multicast address must be delivered to each member of the group.

Figure 19.17 *Multicast address in IPv6*



A flag that defines the group address as either permanent or transient. A **permanent group** address is defined by the Internet authorities and can be accessed at all times. A **transient group** address, on the other hand, is used only temporarily. Systems engaged in a teleconference, for example, can use a transient group address.
The third field defines **the scope** of the group address. Many different scopes are provided.

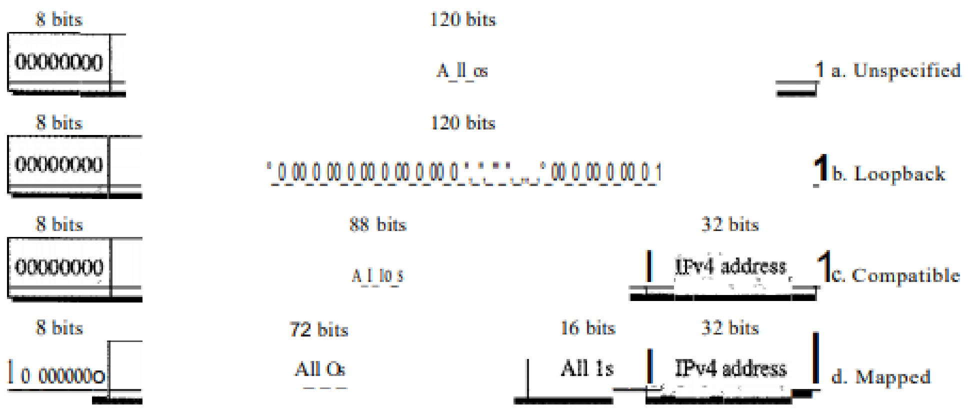
Anycast addresses.

An anycast address, like a multicast address, also defines a group of nodes. However, a packet destined for an anycast address is delivered to only one ofthe members ofthe anycast group, the nearest one (the one with the shortest route)

Reserved Addresses

These addresses start with eight Os (type prefix is 00000000). A few subcategories are defined in this category, as shown in Figure 19.18.

Figure 19.18 *Reserved addresses in IPv6*



An unspecified address is used when a host does not know its own address and sends an inquiryto find its address.

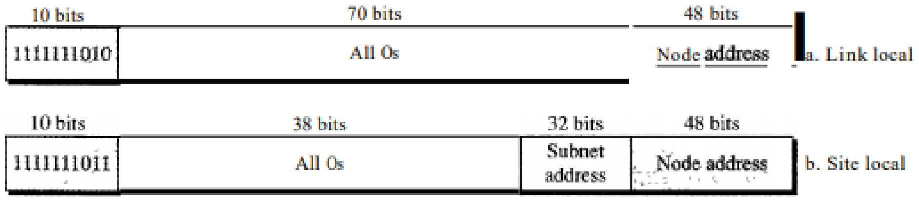
A loopback address is used by a host to test itself without going into the network.

A compatible address is used during the transition from IPv4 to IPv6. It is used when a computer using IPv6 wants to send a message to another computer using IPv6, but the message needs to pass through a part of the network that still operates in IPv4.
A mapped address is also used during transition. However, it is used when a computer that has migrated to IPv6 wants to send a packet to a computer still using IPv4.

Local Addresses

These addresses are used when an organization wants to use IPv6 protocol without being connected to the global Internet. In other words, they provide addressing for private networks. Nobody outside the organization can send a message to the nodes using these addresses. Two types of addresses are defined for this purpose:

Figure 19.19 *Local addresses in IPv6*



A link local address is used in an isolated subnet; a site local address is used in an isolated site with several subnets.

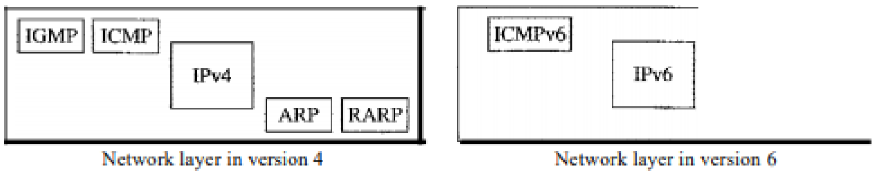
Advantages

The next-generation IP, or IPv6, has some advantages over IPv4 that can be summarized as follows:

- Larger address space. An IPv6 address is 128 bits long, as we discussed in Chapter 19. Compared with the 32-bit address of IPv4, this is a huge (296) increase in the address space. O
- Better header format. IPv6 uses a new header format in which options are separated from the base header and inserted, when needed, between the base header and the upper-layer data. This simplifies and speeds up the routing process because most of the options do not need to be checked by routers.
- New options. IPv6 has new options to allow for additional functionalities.
- Allowance for extension. IPv6 is designed to allow the extension of the protocol if required by new technologies or applications.
- Support for resource allocation. In IPv6, the type-of-service field has been removed, but a mechanism (calledjlow label) has been added to enable the source to request special handling of the packet. This mechanism can be used to support traffic such as real-time audio and video.
- Support for more security. The encryption and authentication options in IPv6 provide confidentiality and integrity of the packet.

ICMPv6.

Figure 21.23 *Comparison ofnetwork layers in version 4 and version 6*



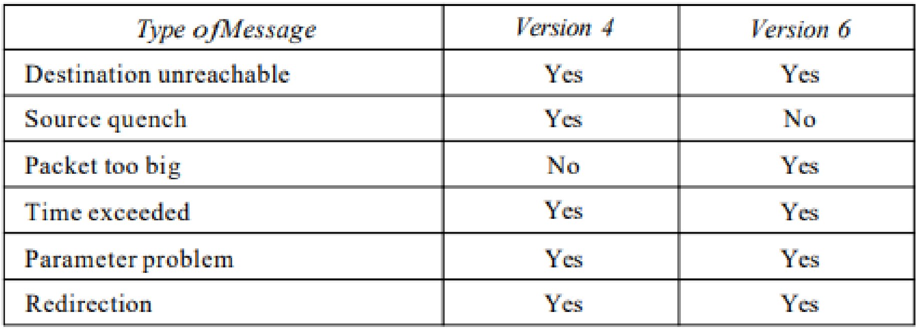
Comparison of the network layer of version 4 to version 6:

The ARP and IGMP protocols in version 4 are combined in ICMPv6. The RARP protocol is dropped from the suite because it was rarely used and BOOTP has the same functionality. Just asin ICMPv4, we divide the ICMP messages into two categories.

Error Reporting

As we saw in our discussion of version 4, one of the main responsibilities of ICMP is to report errors. Five types of errors are handled: destination unreachable, packet too big, time exceeded, parameter problems, and redirection. ICMPv6 forms an error packet, which is then encapsulatedin an IP datagram. This is delivered to the original source of the failed datagram.

Table 21.3 *Comparison oferror-reporting messages in ICMPv4 and ICMPv6*



Type ofMessage	*Version 4*	*Version 6*
Destination unreachable	Yes	Yes
Source quench	Yes	No
Packet too big	No	Yes
Time exceeded	Yes	Yes
Parameter problem	Yes	Yes
Redirection	Yes	Yes

The source-quench message is eliminated in version 6 because the priority and the flow label fields allow the router to control congestion and discard the least important messages. In this version, there is no need to inform the sender to slow down. The packet-too-big message is added because fragmentation is the responsibility of the sender in IPv6. Ifthe sender does not make the right packet size decision, the router has no choice but to drop the packet and send an error message to the sender.

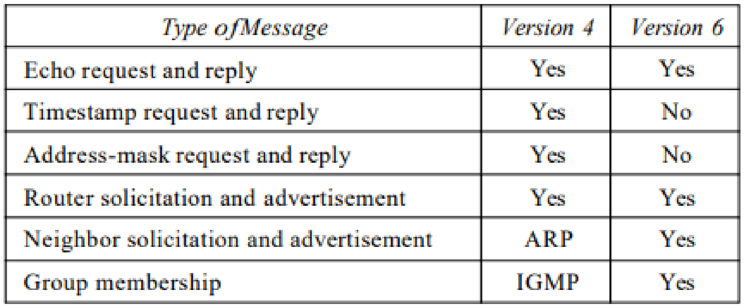
- **Packet Too Big**

This is a new type of message added to version 6. If a router receives a datagram that is larger than the maximum transmission unit (MTU) size of the network through which the datagram should pass, two things happen. First, the router discards the datagram and then an ICMP error packet-a packet-too-big message-is sent to the source.

Query

ICMP can diagnose some network problems. This is accomplished through the query messages. Four different groups of messages have been defined: echo request and reply, router solicitation and advertisement, neighbor solicitation and advertisement, and group membership.

Table 21.4 *Comparison ofquery messages in ICMPv4 and ICMPv6*



Type ofMessage	*Version 4*	*Version 6*
Echo request and reply	Yes	Yes
Timestamp request and reply	Yes	No
Address-mask request and reply	Yes	No
Router solicitation and advertisement	Yes	Yes
Neighbor solicitation and advertisement	ARP	Yes
Group membership	IGMP	Yes

Two sets of query messages are eliminated from ICMPv6: time-stamp request and reply-and address-mask request and reply. The timestamp request and reply messages **are eliminated** because they are implemented in other protocols such as TCP and because they were rarely usedin the past. The address-mask request and reply messages are eliminated in IPv6 because the subnet section of an address allows the subscriber to use up to 232 - 1 subnets. Therefore, subnetmasking, as defined in IPv4, is not needed here.

Neighbor Solicitation and Advertisement

the network layer in version 4 contains an independent protocol called Address Resolution Protocol (ARP). In version 6, this protocol is eliminated, and its duties are included in ICMPv6.The idea is exactly the same, but the format of the message has changed.

Group Membership

The network layer in version 4 contains an independent protocol called IGMP. In version 6, thisprotocol is eliminated, and its duties are included in ICMPv6. The purpose is exactly the same.