

Computer Networking

Teacher: Arabinda Sahoo

CSE 3034

Computer Networking

4

1

Physical Layer, Data Link, The Medium Access Control Sublayer, Network Layer, The Transport, Application Layer, Network Security

Communication: Connection Oriented Communication, Project: WWW Redirection, Connectionless Communication and Multicast, Project : Internet Radio, Project: Server Performance

Textbook

- Computer Networks by Tannenbaum, Pearson India
- UNIX Systems Programming: Communication, Concurrency and Threads by Robbins and Robbins, Pearson

Course Format: 3 Classes/week, 1hr/Class, 1 Lab/Week, 2hr/Lab, 1 credit = 4 Credits

Introduction

- During the 20th century, the key technology was information gathering, processing, and distribution.
- Among other developments, the installation of worldwide telephone networks, the invention of radio and television, the birth and growth of the computer industry, and the launching of communication satellites.
- The merging of computers and communications has had a profound influence on the way computer systems are organized.
- The old model of a single computer serving all of the organization's computational needs has been replaced by one in which a large number of separate but interconnected computers do the job. These systems are called computer networks.
- **Computer network : a collection of autonomous computers interconnected by a single technology.**

Introduction

- Computer network : a collection of autonomous computers interconnected by a single technology.
- Two computers are said to be interconnected if they are able to exchange information.
- Networks come in many sizes, shapes and forms. Although it may sound strange to some people, neither the Internet nor the World Wide Web is a computer network.
- The Internet is not a single network but a network of networks and the Web is a distributed system that runs on top of the Internet.

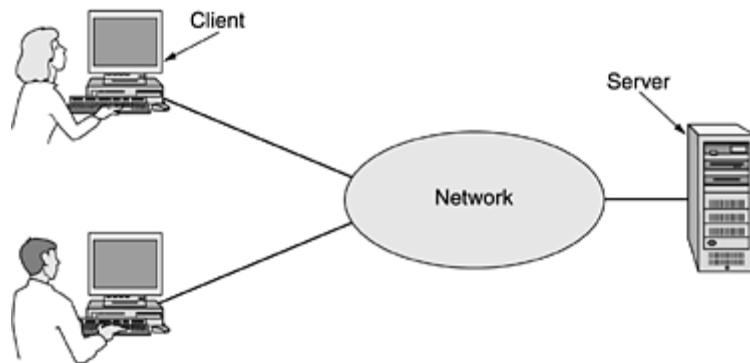
Computer network and a distributed system

- **Distributed system:** a collection of independent computers appears to its users as a single coherent system.
- A well-known example of a distributed system is the World Wide Web, in which everything looks like a document (Web page).
- In a **computer network**, this coherence, model, and software are absent.
- A distributed system is a software system built on top of a network. The software gives it a high degree of cohesiveness and transparency.
- The distinction between a network and a distributed system lies with the software (especially the operating system), rather than with the hardware.

Uses of Computer Networks

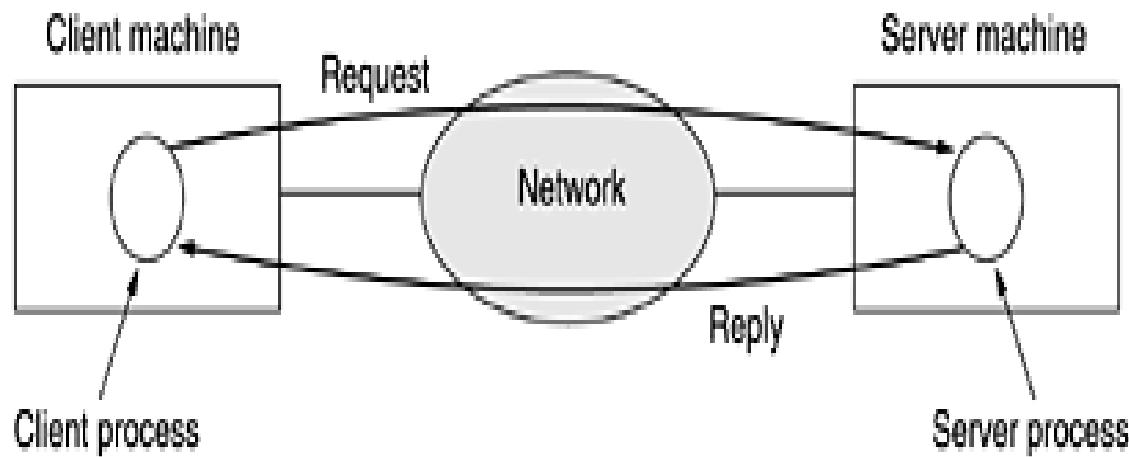
1. Business Applications

- Resource sharing
- Information sharing
- Client-server model
- In client-server model , two processes are involved, one on the client machine and one on the server machine. Communication takes the form of the client process sending a message over the network to the server process.



The client-server model involves requests and replies

- The client process then waits for a reply message. When the server process gets the request, it performs the requested work or looks up the requested data and sends back a reply.



Business Applications

- A second goal of setting up a computer network has to do with people rather than information or even computers. 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. It is called e-commerce (electronic commerce).

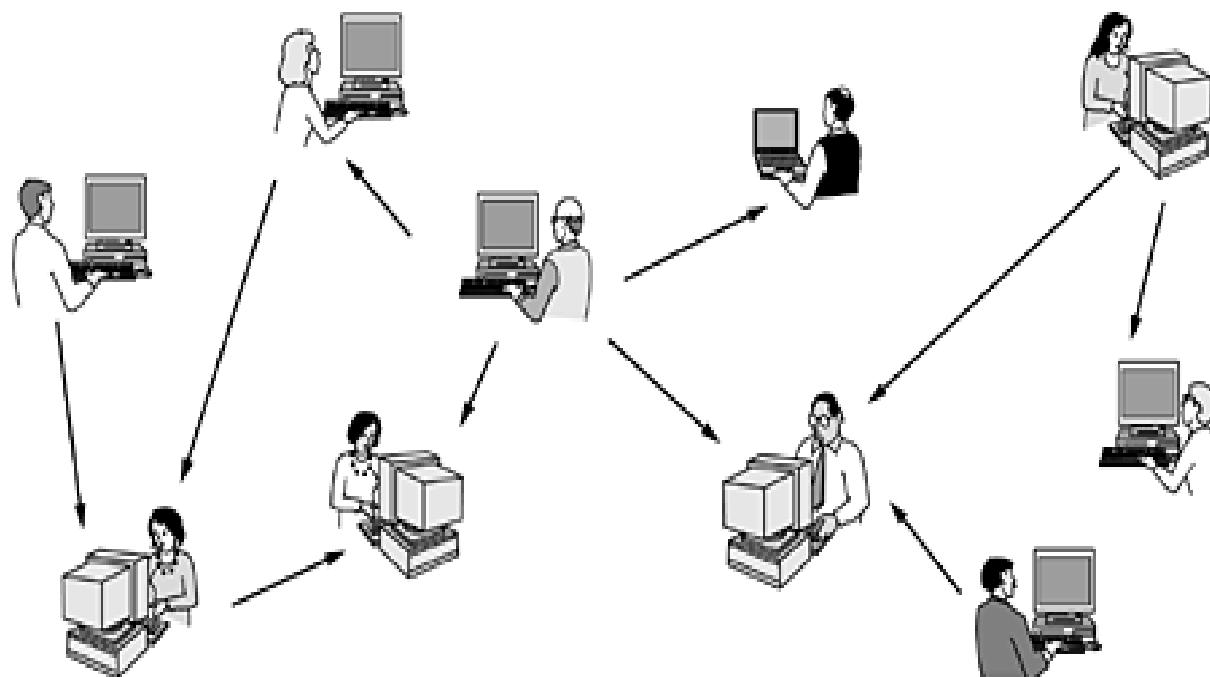
2. Home Applications

- Some of the more popular uses of the Internet for home users are as follows:
 - Access to remote information.
 - Person-to-person communication.
 - Interactive entertainment.
 - Electronic commerce.
- **Access to remote information** comes in many forms. It can be surfing the World Wide Web for information or just for fun.
- Information available includes the arts, business, cooking, government, health, history, hobbies, recreation, science, sports, travel, and many others.

Person-to-Person communication

- The second broad category of network use is person-to-person communication. E-mail is already used on a daily basis by millions of people all over the world and its use is growing rapidly.
- Another type of person-to-person communication often goes by the name of peer-to-peer communication, to distinguish it from the client-server model.
- In this form, individuals who form group can communicate with others in the group. Every person can, in principle, communicate with one or more other people.
- In a peer-to-peer system there are no fixed clients and servers

Peer-to-Peer system



- Third category is entertainment, which is a huge and growing industry.
- Our fourth category is electronic commerce in the broadest sense of the term
- Some forms of e-commerce.

Tag	Full name	Example
B2C	Business-to-consumer	Ordering books on-line
B2B	Business-to-business	Car manufacturer ordering tires from supplier
G2C	Government-to-consumer	Government distributing tax forms electronically
C2C	Consumer-to-consumer	Auctioning second-hand products on line
P2P	Peer-to-peer	File sharing

3. Mobile Users

- Mobile computers, such as notebook computers and personal digital assistants (PDAs), are one of the fastest-growing segments of the computer industry.
- Many owners of these computers have desktop machines back at the office and want to be connected to their home base even when away from home or en route. Since having a wired connection is impossible in cars and airplanes, there is a lot of interest in wireless networks.

Combinations of wireless networks and mobile computing

Wireless	Mobile	Applications
No	No	Desktop computers in offices
No	Yes	A notebook computer used in a hotel room
Yes	No	Networks in older, unwired buildings
Yes	Yes	Portable office; PDA for store inventory

Computer Networking

Network Hardware: Broadcast network, Pont to point network, LAN, MAN, WAN, Wireless network, Home network

Network Hardware

- There is no generally accepted taxonomy into which all computer networks fit, but two dimensions stand out as important: **transmission technology** and **scale**.
- Broadly speaking, there are two types of transmission technology that are in widespread use. They are as follows:
 - **Broadcast links.**
 - **Point-to-point links.**

Broadcast links

- Broadcast networks have a single communication channel that is shared by all the machines on the network.
- Short messages, called packets in certain contexts, sent by any machine are received by all the others. An address field within the packet specifies the intended recipient.
- Broadcast systems generally also allow the possibility of addressing a packet to all destinations by using a special code in the address field. When a packet with this code is transmitted, it is received and processed by every machine on the network. This mode of operation is called broadcasting.
- Some broadcast systems also support transmission to a subset of the machines, something known as multicasting. One possible scheme is to reserve one bit to indicate multicasting. The remaining $n - 1$ address bits can hold a group number.

Point-to-point links

- In contrast, 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.
- Often multiple routes, of different lengths, are possible, so finding good ones is important in point-to-point networks.
- As a general rule (although there are many exceptions), smaller, geographically localized networks tend to use broadcasting, whereas larger networks usually are point-to-point.

Classification by scale

- An alternative criterion for classifying networks is their scale.

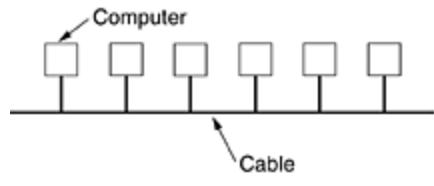
Interprocessor distance	Processors located in same	Example
1 m	Square meter	Personal area network
10 m	Room	
100 m	Building	Local area network
1 km	Campus	
10 km	City	Metropolitan area network
100 km	Country	
1000 km	Continent	Wide area network
10,000 km	Planet	The Internet

- Personal Area Network: networks that are meant for one person.
- Beyond the personal area networks come longer-range networks. These can be divided into
 - local
 - metropolitan
 - wide area networks.

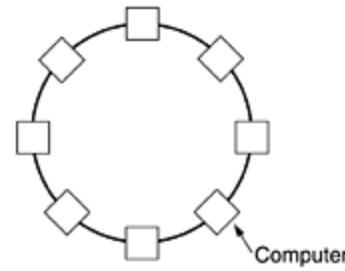
Local Area Networks (LANs)

- Local area networks, generally called LANs, are privately-owned networks within a single building or campus of up to a few kilometers in size.
- They are widely used to connect personal computers and workstations in company offices and factories to share resources (e.g., printers) and exchange information.
- LANs are distinguished from other kinds of networks by three characteristics:
 - (1) their size,
 - (2) their transmission technology
 - (3) their topology.

- Ethernet is a way of connecting computers together in a local area network or LAN.
- Various topologies are possible for broadcast LANs.(a) Bus. (b) Ring.



(a)

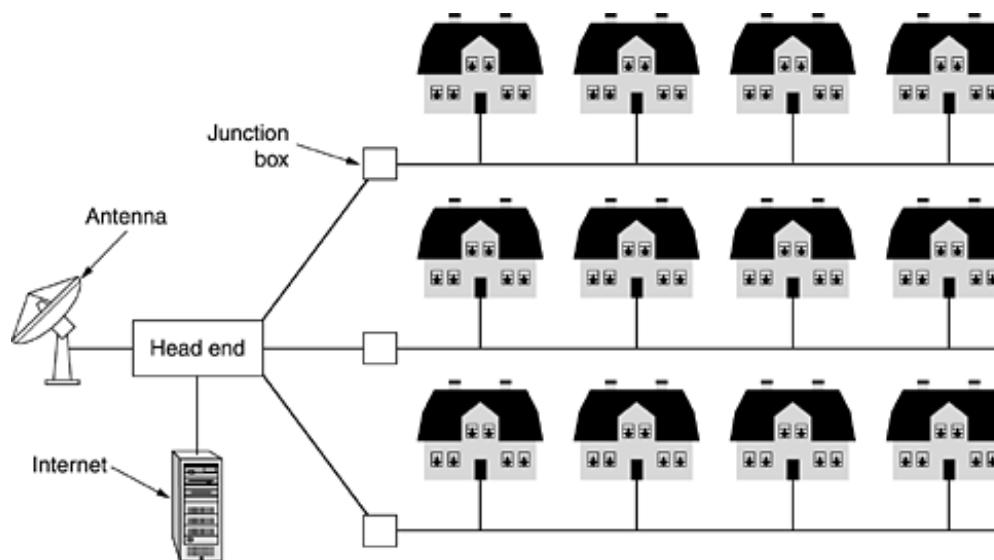


(b)

- Broadcast networks can be further divided into **static and dynamic**, depending on how the channel is allocated.
- A **typical static allocation** would be to divide time into discrete intervals and use a round-robin algorithm, allowing each machine to broadcast only when its time slot comes up.
- **Static allocation wastes channel capacity** when a machine has nothing to say during its allocated slot, so most systems attempt to allocate the channel dynamically (i.e., on demand).

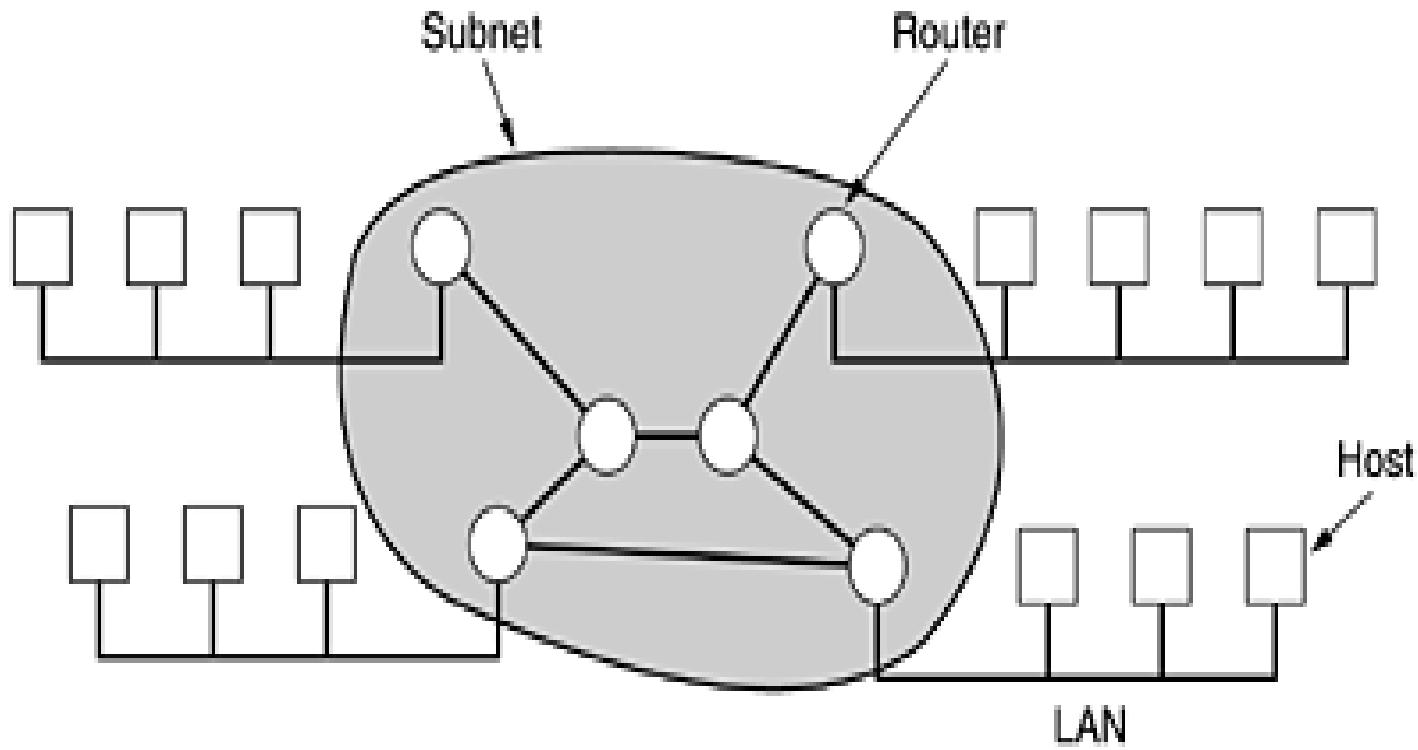
Metropolitan Area Networks (MAN)

- A metropolitan area network, or MAN, covers a city. The best-known example of a MAN is the **cable television network** available in many cities.



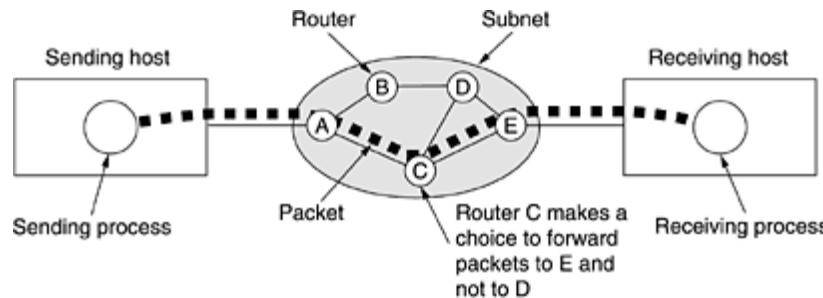
Wide Area Networks (WAN)

- A wide area network, or WAN, spans a large geographical area, a country or continent.
- It contains a collection of machines intended for running user (i.e., application) programs and call these machines hosts.
- The hosts are connected by a communication subnet, or just subnet for short. **The hosts are owned by the customers** (e.g., people's personal computers),
- The communication subnet is typically owned and operated by a telephone company or Internet service provider.
- The job of the subnet is to carry messages from host to host.
- In most wide area networks, the subnet consists of two distinct components: **transmission lines and switching elements**.



Packet-Switched WAN

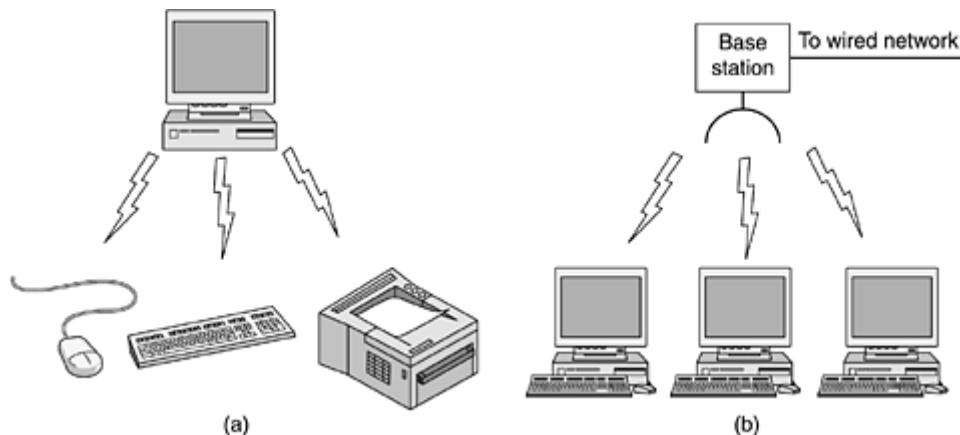
- In most WANs, the network contains numerous transmission lines, each one connecting a pair of routers. If two routers that do not share a transmission line wish to communicate, they must do this indirectly, via other routers.
- When a packet is sent from one router to another via one or more intermediate routers, the packet is received at each intermediate router, stored there until the required output line is free, and then forwarded.
- A subnet organized according to this principle is called a store-and-forward or packet-switched subnet.
- Nearly all wide area networks (except those using satellites) have store-and-forward subnets. When the packets are small and all the same size, they are called cells.



Wireless Networks

- wireless networks can be divided into three main categories:
 - System interconnection.
 - Wireless LANs.
 - Wireless WANs.

- System interconnection is all about interconnecting the components of a computer using short-range radio. Almost every computer has a monitor, keyboard, mouse, and printer connected to the main unit by cables.
- The next step up in wireless networking are the wireless LANs. These are systems in which every computer has a radio modem and antenna with which it can communicate with other systems.
- The third kind of wireless network is used in wide area systems. The radio network used for cellular telephones is an example of a low-bandwidth wireless system.
- This system has already gone through three generations. The first generation was analog and for voice only. The second generation was digital and for voice only. The third generation is digital and is for both voice and data.



Home Networks

- The fundamental idea is that in the future most homes will be set up for networking. Every device in the home will be capable of communicating with every other device, and all of them will be accessible over the Internet.
- Many devices are capable of being networked.
- Some of the categories (with examples) are as follows:
 - Computers (desktop PC, notebook PC, PDA, shared peripherals).
 - Entertainment (TV, DVD, VCR, camcorder, camera, stereo, MP3).
 - Telecommunications (telephone, mobile telephone, intercom, fax).
 - Appliances (microwave, refrigerator, clock, furnace, airco, lights).
 - Telemetry (utility meter, smoke/burglar alarm, thermostat, babycam).

Internetworks

- A collection of interconnected networks is called an internetwork or internet.
- A common form of internet is a collection of LANs connected by a WAN.
 - Subnets, networks, and internetworks are often confused.
 - subnet makes the most sense in the context of a wide area network, where it refers to the collection of routers and communication lines owned by the network operator.
 - The combination of a subnet and its hosts forms a network.
 - An internetwork is formed when distinct networks are interconnected.

QUIZ

Computer Networking

Network Software: Protocol Hierarchies, Design issues for the layers

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.

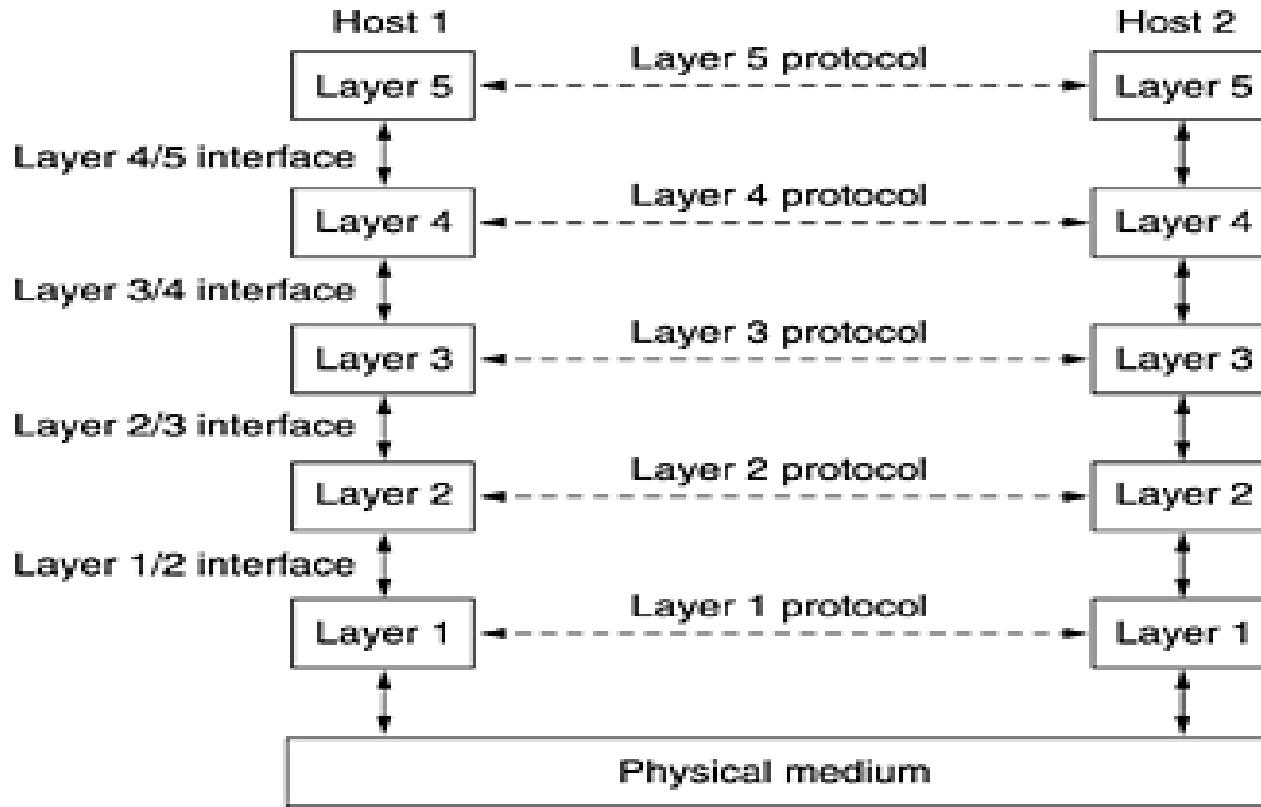
Protocol Hierarchies

- A protocol is a standard set of rules that allow electronic devices to communicate with each other.
- To reduce the 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.

Protocol Hierarchies

- 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.
- Basically, a protocol is an agreement between the communicating parties on how communication is to proceed. Violating the protocol will make communication more difficult

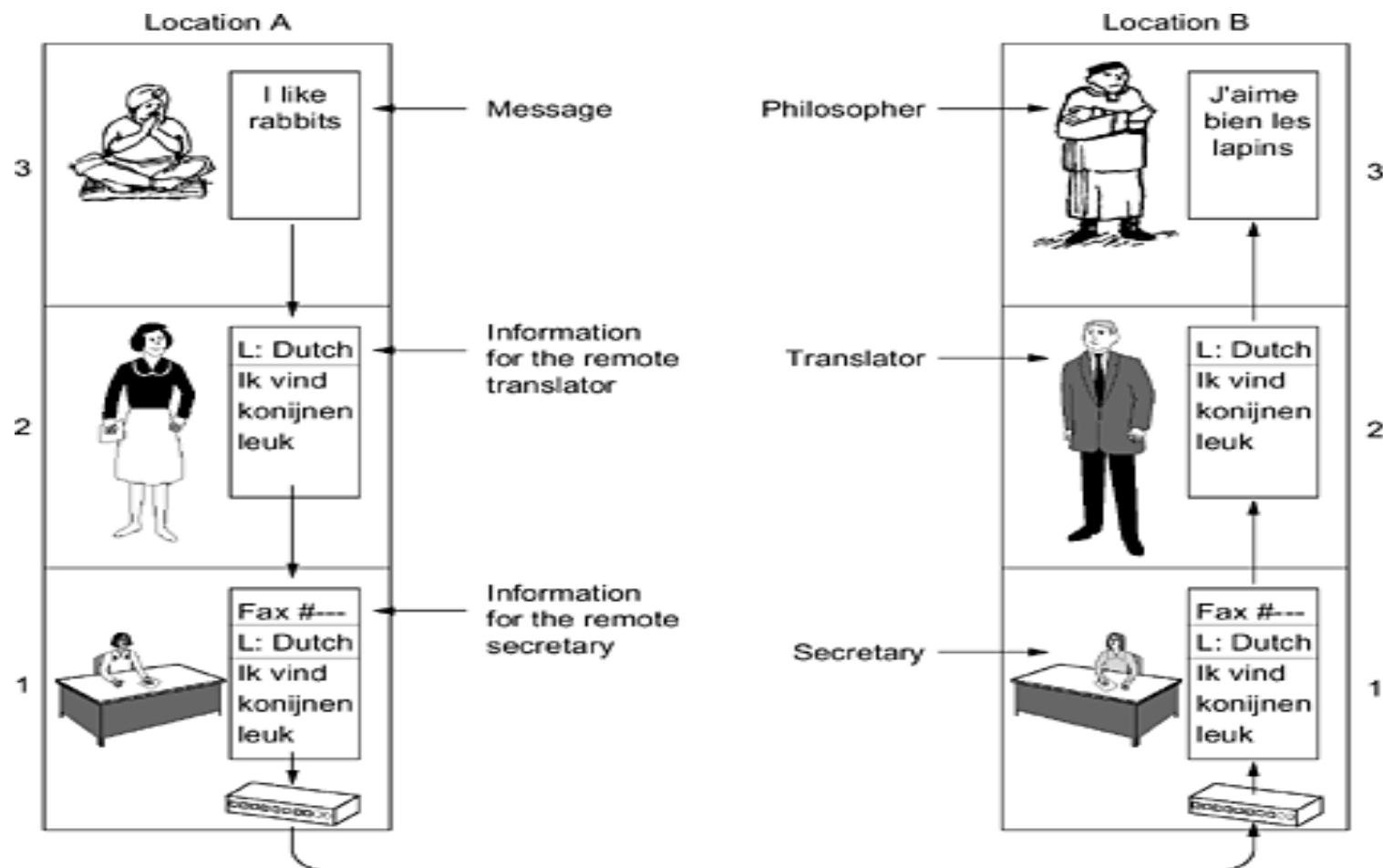
Figure: Layers, protocols, and interfaces.



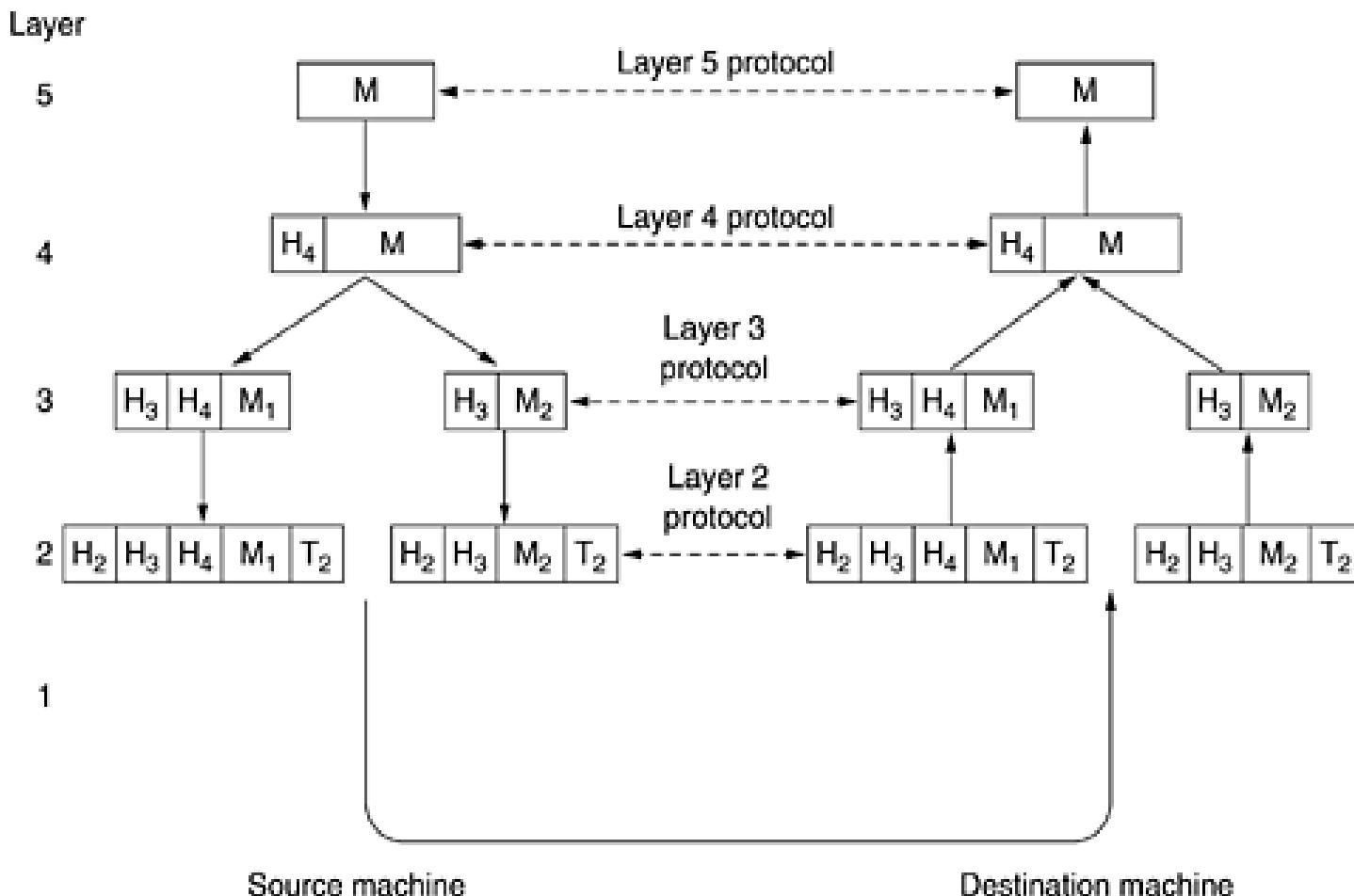
- A five-layer network is illustrated in figure. The entities comprising the corresponding layers on different machines are called peers.
- The peers may be processes, hardware devices, or even human beings. In other words, it is the peers that communicate by using the protocol.
- In reality, no data are directly transferred from layer n on one machine to layer n on another machine. Instead, each layer passes data and control information to the layer immediately below it, until the lowest layer is reached.
- Below layer 1 is the physical medium through which actual communication occurs.

- Between each pair of adjacent layers is an interface. The interface defines which primitive operations and services the lower layer makes available to the upper one.
- **A set of layers and protocols is called a network architecture.** The specification of an architecture must contain enough information to allow an implementer to write the program or build the hardware for each layer so that it will correctly obey the appropriate protocol.
- A list of protocols used by a certain system, one protocol per layer, is called a **protocol stack**.

- Idea of multilayer communication: The philosopher-translator-secretary architecture



- Example information flow supporting virtual communication in layer 5.



Design Issues for the Layers

- Every layer needs a mechanism for identifying senders and receivers.
- Another set of design decisions concerns the rules for data transfer.
- Error control is an important issue because physical communication circuits are not perfect.
- Not all communication channels preserve the order of messages sent on them.
- An issue that occurs at every level is how to keep a fast sender from swamping a slow receiver with data.
- Another problem that must be solved at several levels is the inability of all processes to accept arbitrarily long messages.
- When there are multiple paths between source and destination, a route must be chosen. Sometimes this decision must be split over two or more layers.

Computer Networking

**Network Software: Connection oriented and connection less services,
Service primitives, Relationship of services to protocols**

Connection-Oriented and Connectionless Services

- Layers can offer two different types of service to the layers above them: **connection-oriented** and **connectionless**.
- **Connection-oriented service** is modeled after the telephone system. To talk to someone, you pick up the phone, dial the number, talk, and then hang up.
- Similarly, to use a connection-oriented network service, the service user **first establishes a connection**, **uses the connection**, and then **releases the connection**.
- In some cases when a connection is established, the sender, receiver, and subnet conduct a negotiation about parameters to be used, such as **maximum message size**, **quality of service required**, and **other issues**.

Connection-Oriented and Connectionless Services

- In contrast, **connectionless service** is modeled after the postal system. Each message (letter) carries the full destination address, and each one is routed through the system independent of all the others.
- Normally, when two messages are sent to the same destination, the first one sent will be the first one to arrive. However, it is possible that the first one sent can be delayed so that the second one arrives first.
- Each service can be characterized by a quality of service. Some services are reliable in the sense that they never lose data.
- Usually, a reliable service is implemented by having the receiver acknowledge the receipt of each message so the sender is sure that it arrived. The acknowledgement process introduces overhead and delays, which are often worth it but are sometimes undesirable.

Connection-Oriented and Connectionless Services

- A typical situation in which a reliable connection-oriented service is appropriate is **file transfer**.
- Reliable connection-oriented service has two minor variations: **message sequences and byte streams**.
- For some applications, the transit delays introduced by acknowledgements are unacceptable.
- Not all applications require connections.
- Unreliable (meaning not acknowledged) connectionless service is often called datagram service, in analogy with telegram service, which also does not return an acknowledgement to the sender.

Connection-Oriented and Connectionless Services

- In other situations, the convenience of not having to establish a connection to send one short message is desired, but reliability is essential.
- The acknowledged datagram service can be provided for these applications.
- Another service is the request-reply service. In this service the sender transmits a single datagram containing a request; the reply contains the answer.
- Request-reply is commonly used to implement communication in the client-server model: the client issues a request and the server responds to it.

Six different types of service

	Service	Example
Connection-oriented	Reliable message stream	Sequence of pages
	Reliable byte stream	Remote login
Connection-less	Unreliable connection	Digitized voice
	Unreliable datagram	Electronic junk mail
	Acknowledged datagram	Registered mail
	Request-reply	Database query

- The concept of using unreliable communication may be confusing at first. After all, why would anyone actually prefer unreliable communication to reliable communication?
- First of all, reliable communication (in our sense, that is, acknowledged) may not be available. For example, Ethernet does not provide reliable communication. Packets can occasionally be damaged in transit.
- Second, the delays inherent in providing a reliable service may be unacceptable, especially in real-time applications such as multimedia.

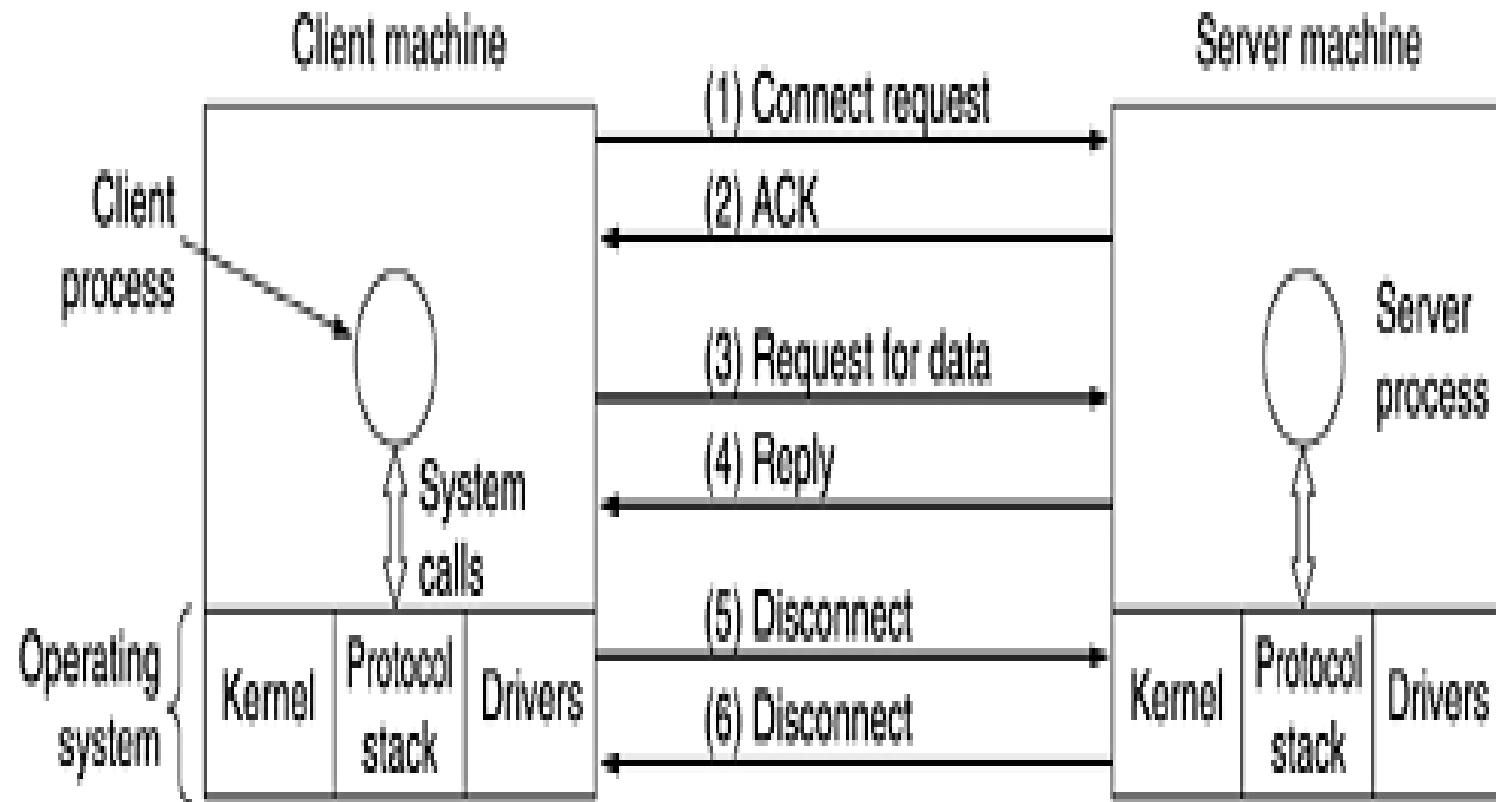
Service Primitives

- 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 entity.
- The set of primitives available depends on the nature of the service being provided.
- The primitives for connection-oriented service are different from those of connectionless service.

Five service primitives for implementing a simple connection-oriented service.

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

Packets sent in a simple client-server interaction on a connection-oriented network



These primitives might be used as follows.

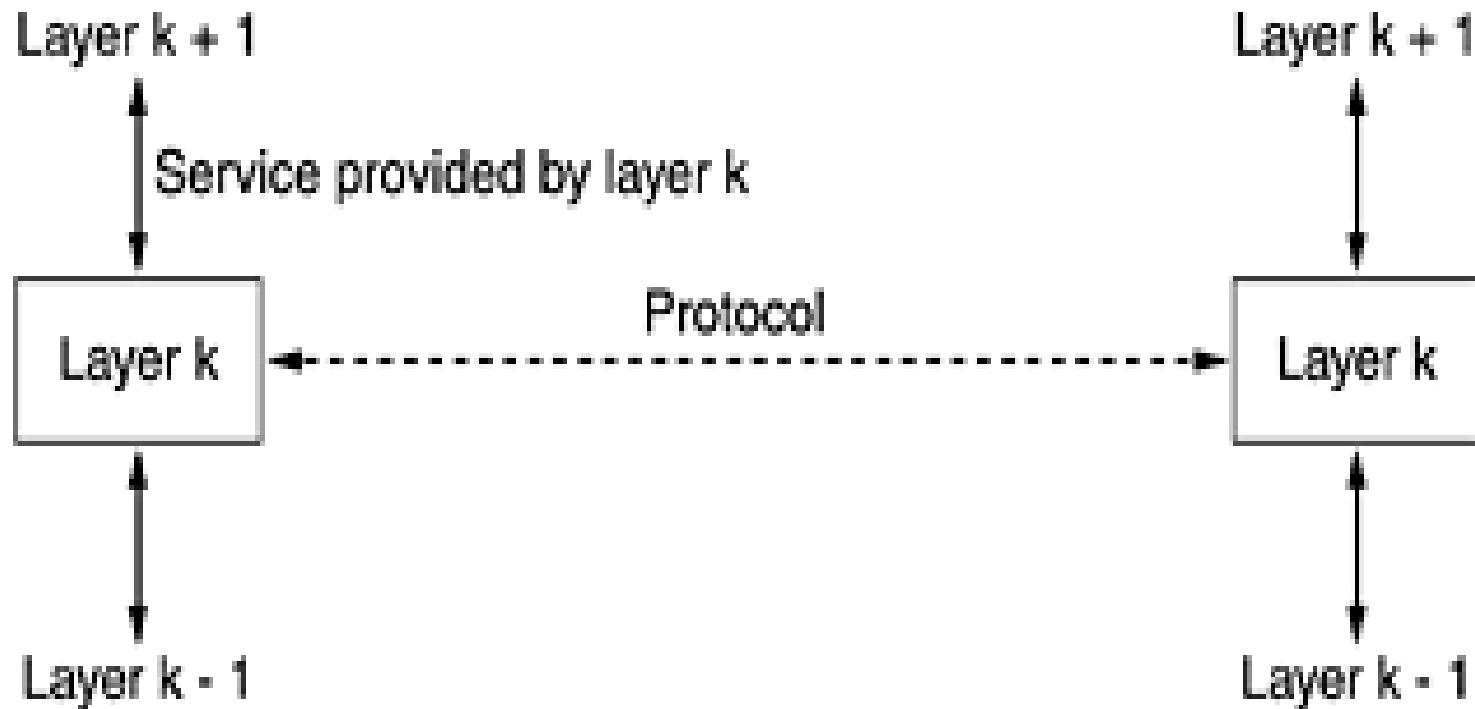
- First, the server executes LISTEN to indicate that it is prepared to accept incoming connections.
- Next, the client process executes CONNECT to establish a connection with the server.
- The CONNECT call needs to specify who to connect to, so it might have a parameter giving the server's address. The operating system then typically sends a packet to the peer asking it to connect.
- The client process is suspended until there is a response. When the packet arrives at the server, it is processed by the operating system there.
- When the system sees that the packet is requesting a connection, it checks to see if there is a listener. If so, it does two things: **unblocks the listener and sends back an acknowledgement (2)**.
- The arrival of this acknowledgement then releases the client. At this point the client and server are both running and they have a connection established.
- If a connection request arrives and there is no listener, the result is undefined. In some systems the packet may be queued for a short time in anticipation of a LISTEN.

- The next step is for the server to execute RECEIVE to prepare to accept the first request. Normally, the server does this immediately upon being released from the LISTEN, before the acknowledgement can get back to the client. The RECEIVE call blocks the server.
- Then the client executes SEND to transmit its request (3) followed by the execution of RECEIVE to get the reply.
- The arrival of the request packet at the server machine unblocks the server process so it can process the request. After it has done the work, it uses SEND to return the answer to the client (4).
- The arrival of this packet unblocks the client, which can now inspect the answer. If the client has additional requests, it can make them now. If it is done, it can use DISCONNECT to terminate the connection.
- Usually, an initial DISCONNECT is a blocking call, suspending the client and sending a packet to the server saying that the connection is no longer needed (5).
- When the server gets the packet, it also issues a DISCONNECT of its own, acknowledging the client and releasing the connection. When the server's packet (6) gets back to the client machine, the client process is released and the connection is broken.

The Relationship of Services to Protocols

- A service is a set of primitives (operations) that a layer provides to the layer above it. The service defines what operations the layer is prepared to perform on behalf of its users, but it says nothing at all about how these operations are implemented.
- A service relates to an interface between two layers, with the lower layer being the service provider and the upper layer being the service user.
- A protocol, in contrast, is a set of rules governing the format and meaning of the packets, or messages that are exchanged by the peer entities within a layer.
- Entities use protocols to implement their service definitions. They are free to change their protocols at will, provided they do not change the service visible to their users.
- In this way, the service and the protocol are completely decoupled.

The relationship between a service and a protocol



QUIZ TEST

Computer Networking

Reference Models

Reference Models

- Two important network architectures, the
 - ❑ OSI (Open Systems Interconnection) reference model
 - ❑ TCP/IP (Transmission Control Protocol/Internet Protocol) reference model
- The OSI model is a conceptual model that characterizes and standardizes the communication functions of a telecommunication or computing system without regard to its underlying internal structure and technology.
- The **OSI Model** is a conceptual framework used to describe the functions of a networking system.
- The **TCP/IP Model** is a suite of communication protocols used to interconnect network devices on the internet.

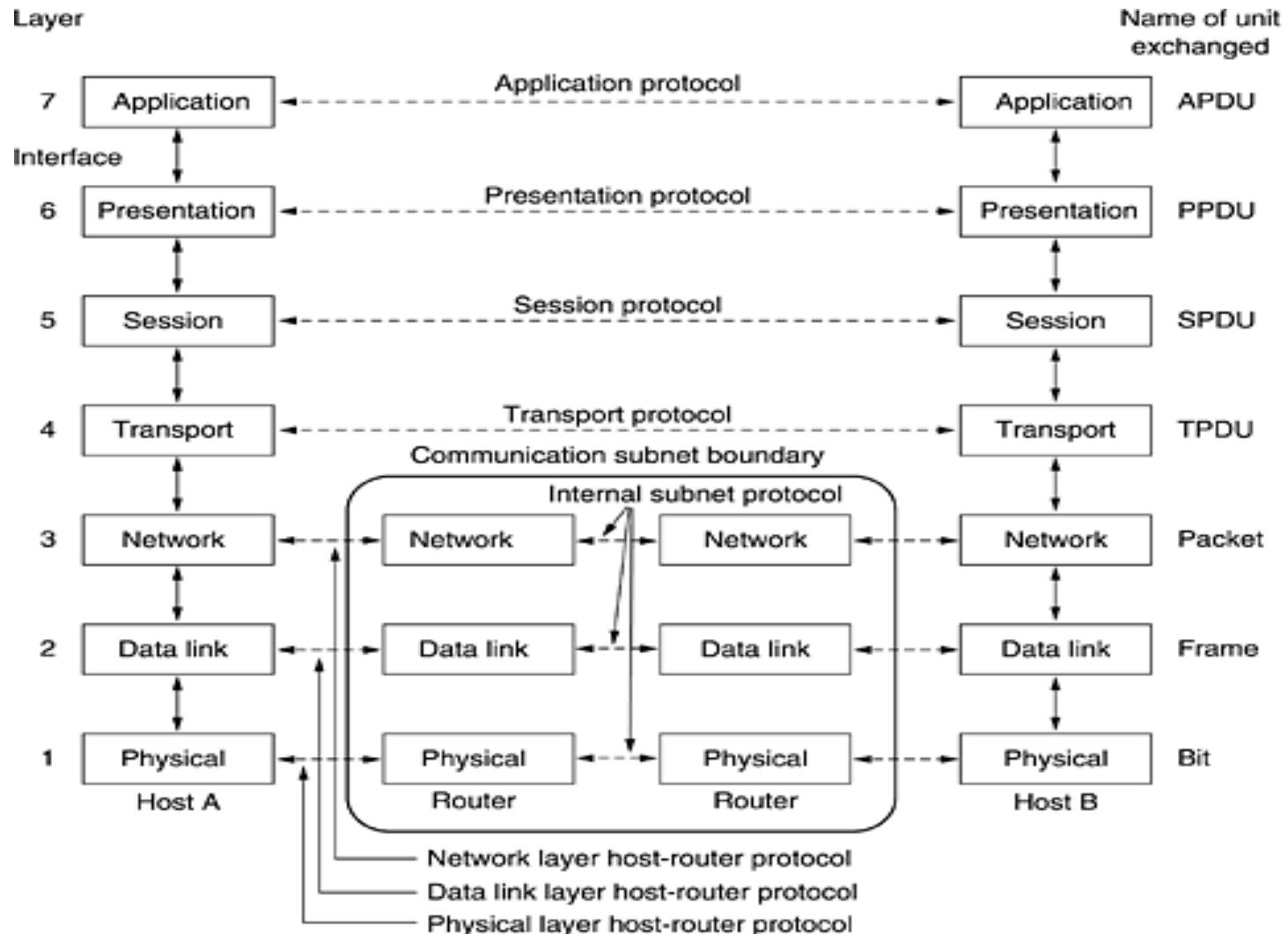
- OSI model gives guidelines on how communication needs to be done, while TCP/IP protocols layout standards on which the Internet was developed. So, TCP/IP is a more practical model.
- The OSI model is based on a proposal developed by the International Standards Organization (ISO) as a first step toward international standardization of the protocols used in the various layers. It was revised in 1995.
- The model is called the ISO OSI (Open Systems Interconnection) Reference Model because it deals with connecting open systems—that is, systems that are open for communication with other systems. We will just call it the OSI model for short.
- Although the protocols associated with the OSI model are rarely used any more, the model itself is actually quite general and still valid, and the features discussed at each layer are still very important.
- The TCP/IP model has the opposite properties: the model itself is not of much use but the protocols are widely used.

OSI Reference Model

- The OSI model has seven layers. The principles that were applied to arrive at the seven layers can be briefly summarized as follows:
 1. A layer should be created where a different abstraction is needed.
 2. Each layer should perform a well-defined function.
 3. The function of each layer should be chosen with an eye toward defining internationally standardized protocols.
 4. The layer boundaries should be chosen to minimize the information flow across the interfaces.
 5. The number of layers should be large enough that distinct functions need not be thrown together in the same layer out of necessity and small enough that the architecture does not become unwieldy.

- The OSI model has seven layers:
 - Physical Layer
 - Data Link Layer
 - Network Layer
 - Transport Layer
 - Session Layer
 - Presentation Layer
 - Application Layer
- It's easy to remember the sequence of OSI Model 7 Layers using this simple sentence:
 - “**P**lease **D**o **N**ot **T**hrow **S**eafod **P**izza **A**way”
 - “**A**ll **P**eople **S**eem **T**o **N**eed **D**ata **P**rocessing.”

The OSI reference model



The Physical Layer

- The physical layer is concerned with transmitting raw bits over a communication channel.
- Physical layer in the OSI model plays the role of interacting with actual hardware and signaling mechanism.
- Physical layer is the only layer of OSI network model which actually deals with the physical connectivity of two different stations.
- Converts data from the upper layers into 1s and 0s for transmission over media.
- Data-link layer hands over frames to physical layer. Physical layer converts them to electrical pulses, which represent binary data. The binary data is then sent over the wired or wireless media.
- Defined on this layer: Cable standards, wireless standards, and fiber optic standards.
- Copper wiring, fiber optic cable, radio frequencies, anything that can be used to transmit data is defined on the Physical layer of the OSI Model.

The Physical Layer

- The design issues have to do with making sure that when one side sends a 1 bit, it is received by the other side as a 1 bit, not as a 0 bit.
- Typical questions here are
 - how many volts should be used to represent a 1 and how many for a 0,
 - how many nanoseconds a bit lasts, whether transmission may proceed simultaneously in both directions,
 - how the initial connection is established and how it is torn down (disassemble or disintegrate) when both sides are finished, and
 - how many pins the network connector has and what each pin is used for.
- The design issues here largely deal with mechanical, electrical, and timing interfaces, and the physical transmission medium, which lies below the physical layer.

The Data Link Layer

- Data link layer is responsible for converting data stream to signals bit by bit and to send that over the underlying hardware.
- At the receiving end, Data link layer picks up data from hardware which are in the form of electrical signals, assembles them in a recognizable frame format, and hands over to upper layer.
- Data link layer Transforms the raw data bits to a data frame (few hundred/thousand bits).
- It accomplishes this task by having the sender break up the input data into data frames (typically a few hundred or a few thousand bytes) and transmit the frames sequentially.
- Is responsible for moving frames from node to node or computer to computer.
- Protocols defined include **Ethernet Protocol** and **Point-to-Point Protocol (PPP)**

The Data Link Layer

- Two sublayers: Logical Link Control (LLC) and the Media Access Control (MAC)
 - Logical Link Control (LLC)
 - Data Link layer addressing, flow control, address notification, error control
 - Media Access Control (MAC)
 - Determines which computer has access to the network media at any given time
 - Determines where one frame ends and the next one starts, called frame synchronization
- One issue that arises in the data link layer (and most of the higher layers as well) is how to keep a fast transmitter from drowning a slow receiver in data.
- Broadcast networks have an additional issue in the data link layer: how to control access to the shared channel. A special sublayer of the data link layer, the media access control sublayer, deals with this problem.

The Network Layer

- The network layer controls the operation of the subnet.
- Responsible for moving packets (data) from one end of the network to the other, called *end-to-end communications*
- Network layer manages options pertaining to host and network addressing, managing sub-networks, and internetworking.
- Handles the issues raised due to different physical addresses of machines belonging to different networks
- A key design issue is determining how packets are routed from source to destination.
- If too many packets are present in the subnet at the same time, they will get in one another's way, forming bottlenecks. The control of such congestion also belongs to the network layer.
- More generally, the quality of service provided (delay, transit time, etc.) is also a network layer issue.

The Network Layer

- When a packet has to travel from one network to another to get to its destination, many problems can arise.
- The addressing used by the second network may be different from the first one. The second one may not accept the packet at all because it is too large.
- The protocols may differ, and so on. It is up to the network layer to overcome all these problems to allow heterogeneous networks to be interconnected.
- Two different subnet may have different addressing schemes or non-compatible addressing types. Same with protocols, two different subnet may be operating on different protocols which are not compatible with each other.
- Network layer has the responsibility to route the packets from source to destination, mapping different addressing schemes and protocols.

The Transport Layer

- Accepts data from higher levels and splits it into smaller segments that can be sent to network layer.
- Transport layer takes data from upper layer (i.e. Application layer) and then breaks it into smaller size segments, numbers each byte, and hands over to lower layer (Network Layer) for delivery.
- This Layer is the first one which breaks the information data, supplied by Application layer in to smaller units called segments.
- Also, reassembles data segments into data for the use of higher layers
- Puts segments in correct order (called sequencing) so they can be reassembled in correct order at destination.
- This layer ensures that data must be received in the same sequence in which it was sent.
- Concerned with the reliability of the transport of sent data.
- May use a *connection-oriented protocol* such as TCP (Transmission Control Protocol) to ensure destination received segments
- May use a *connectionless protocol* such as UDP (User Datagram Protocol) to send segments without assurance of delivery

The Session Layer

- The session layer allows users on different machines to establish active communication sessions between them.
- It's main aim is to establish, manage, terminate, maintain and synchronize the interaction between communicating systems.
- Provides duplex, half-duplex, or simplex communications between devices
- Sessions offer various services, including
 - Dialog control
 - Token management
 - Synchronization

- **Dialog control** :keeping track of whose turn it is to transmit. This layer allows two systems to start communication with each other.
- **Token management:** preventing two parties from attempting the same critical operation at the same time.
- **Synchronization:** checkpointing long transmissions to allow them to continue from where they were after a crash. This layer allows a process to add checkpoints which are considered as synchronization points into stream of data.

The Presentation Layer

- The primary goal of this layer is to take care of the **syntax** and **semantics** of the information exchanged between two communicating systems.
- Presentation layer takes care that the data is sent in such a way that the receiver will understand the information (data) and will be able to use the data.
- Since different computer may deal with different data representations a standard encoding is done, thus handles three primary tasks:
 - Translation , –Compression , –Encryption

- **Translation:** Before being transmitted, information in the form of characters and numbers should be changed to bit streams. The presentation layer is responsible for interoperability between encoding methods as different computers use different encoding methods. It translates data between the formats the network requires and the format the computer.
- **Encryption:** It carries out encryption at the transmitter and decryption at the receiver.
- **Compression:** The primary role of Data compression is to reduce the number of bits to be transmitted.

The Application Layer

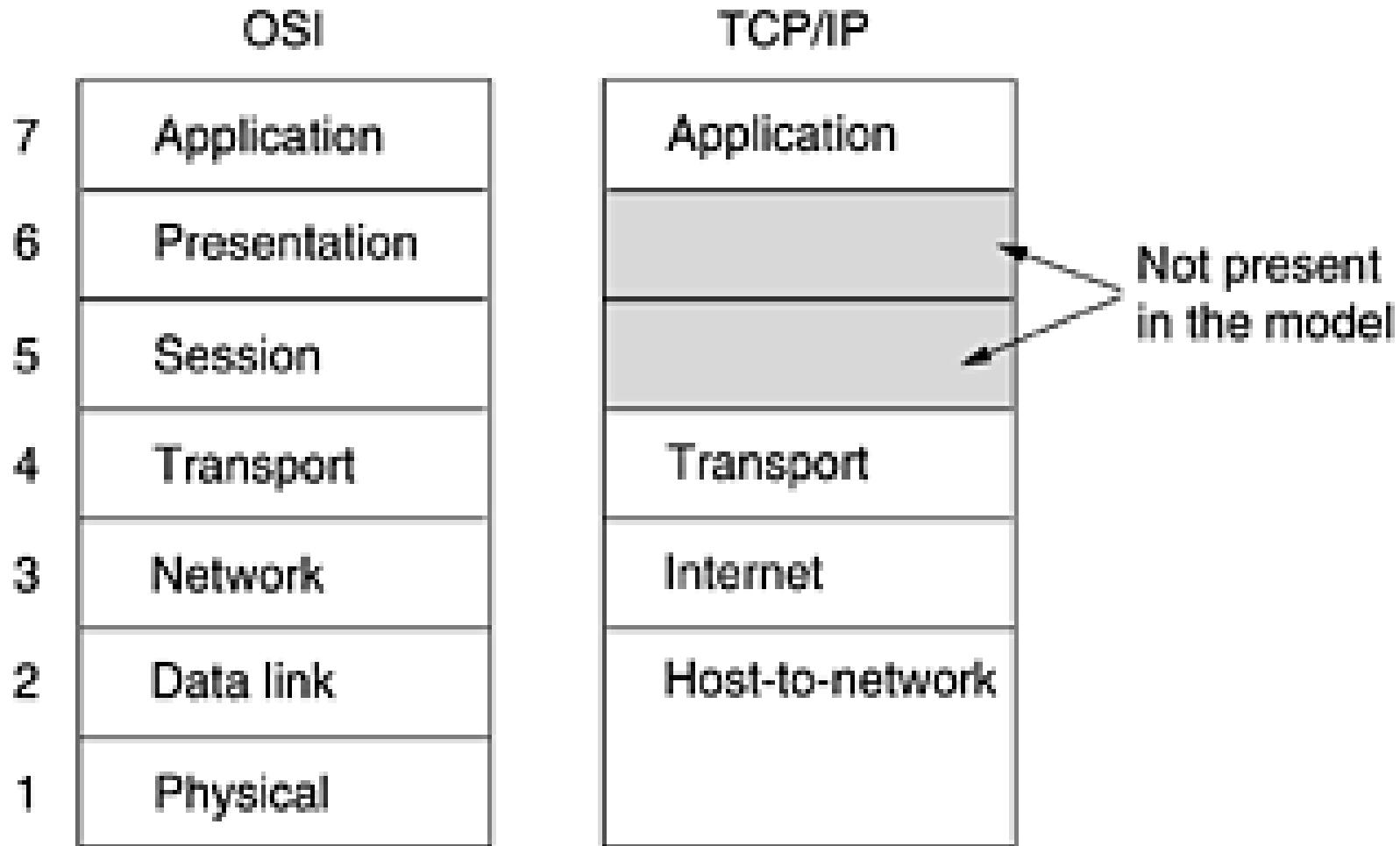
- The application layer contains a variety of protocols that are commonly needed by users. This layer which enables user or software to get access to the network. Some services provided by this layer includes: E-Mail, transferring files, distributing the results to user, directory services, network resources, etc.
- One widely-used application protocol is HTTP (HyperText Transfer Protocol), which is the basis for the World Wide Web.
- Other Application protocols that are used are: File Transfer Protocol (FTP), Trivial File Transfer Protocol (TFTP), Simple Mail Transfer Protocol (SMTP), TELNET, Domain Name System (DNS) etc.
- Functions of Application Layer
 - Mail Services
 - Network Virtual Terminal
 - Directory Services
 - File Transfer, Access and Management (FTAM)

- **Mail Services:** This layer provides the basis for E-mail forwarding and storage.
- **Network Virtual Terminal:** It allows a user to log on to a remote host.
- **Directory Services:** This layer provides access for global information about various services.
- **File Transfer, Access and Management (FTAM):** It is a standard mechanism to access files and manages it. Users can access files in a remote computer and manage it. They can also retrieve files from a remote computer.

The TCP/IP Reference Model

- This model was proposed earlier to OSI model.
- The main aim behind the development of this protocol suite is to support/interconnect different types of network (e.g. interconnection of radio network and computer network).
- Another major goal is connections to remain intact as long as the source and destination machines were functioning, even if some of the machines or transmission lines in between were suddenly put out of operation.
- The original TCP/IP protocol suite was defined having four protocol layers: **Host-to-network, internet, transport and application.**
- When TCP/IP is compared to OSI it can be seen that the host-to-network layer is equivalent to the combination of physical and data link layer. Also, the internet layer is equivalent to the network layer, and the application layer is roughly doing the job of the session, presentation, and application layers.

The TCP/IP reference model.



The Host-to-Network Layer

- The TCP/IP reference model does not really say much about what happens here, except to point out that the host has to connect to the network using some protocol so it can send IP packets to it.
- This protocol is not defined and varies from host to host and network to network. Books and papers about the TCP/IP model rarely discuss it.
- At the physical and data link layers, *TCP/IP* does not define any specific protocol. It supports all the standard and proprietary protocols.

The Internet Layer

- All these requirements led to the choice of a packet-switching network based on a connectionless internetwork layer. This layer, called the internet layer.
- The job of this layer is to permit hosts to inject packets into any network and have them 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.

The Internet Layer

- 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.

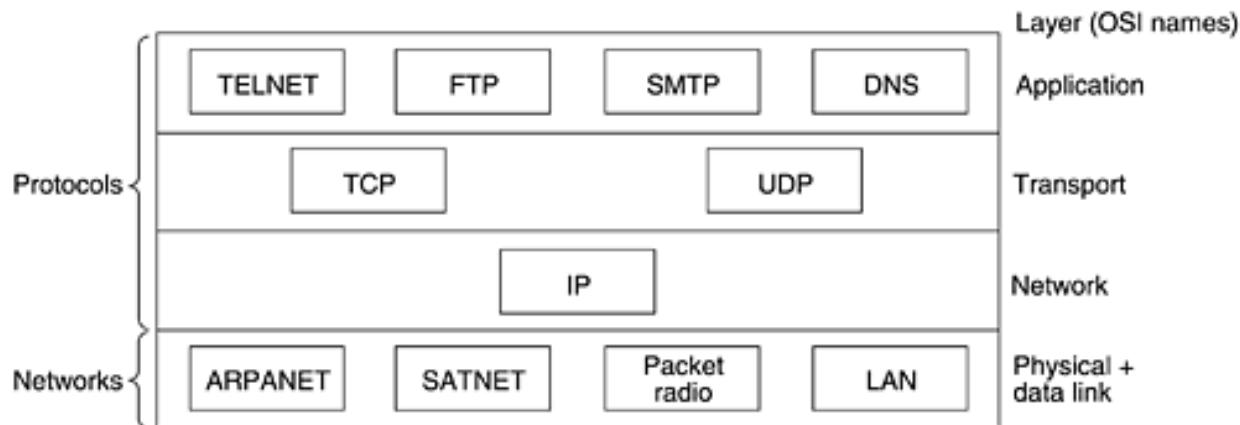
Transport Layer

- Allow peer entities on the source and destination hosts to carry on a conversation, just as in the OSI transport layer. Makes the use of either of the two types of transport protocol **TCP(Transmission Control Protocol)** , **UDP(User Datagram Protocol)**.
- **TCP(Transmission Control Protocol) :**
 - A reliable connection-oriented protocol.
 - 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 to make sure a fast sender cannot swamp a slow receiver with more messages than it can handle.

Transport Layer

- **UDP(User Datagram Protocol):**
 - An unreliable, connectionless protocol for applications that do not want sequencing or flow control and wish to provide their own.
 - Also widely used for one-shot, client-server-type request-reply queries and applications in which prompt delivery is more important than accurate delivery, such as transmitting speech or video.
 - The relation of IP, TCP, and UDP is shown in figure.

(Fig: Protocols and networks in the TCP/IP model initially)



The Application Layer

- Application layer protocols define the rules when implementing specific network applications
- Rely on the underlying layers to provide accurate and efficient data delivery
- Typical protocols:
 - FTP – File Transfer Protocol (For file transfer)
 - Telnet – Remote terminal protocol (For remote login on any other computer on the network)
 - SMTP – Simple Mail Transfer Protocol (For mail transfer)
 - HTTP – Hypertext Transfer Protocol (For Web browsing)
 - NNTP-Network News Transfer Protocol (For transfer newsgroup articles between systems over the Internet)

A Comparison of the OSI and TCP/IP Reference Models

- Since both OSI and TCP/IP reference models have developed looking at the operation of communication between users of a computer network they support some similar characteristics as listed below.

Similarities :

- Both lie on the concept of a stack of independent protocols.
- Functionality of the layers is roughly similar

Differences :

OSI reference model	TCP/IP reference model
uses 7 different layers.	Uses 4 different layers.
Supports both connectionless & connection oriented service in the network layer but only connection oriented service in transport layer.	Supports only connectionless service in the network layer but both connectionless & connection oriented service in transport layer.
Clearly distincts service, interface & protocol.	Doesn't clearly distinguish service, interface & protocol.
Protocols are better hidden and can be replaced relatively easily as the technology changes.	Protocols are not hidden and can not be replaced easily as the technology changes (e.g. Replacing IP with a different protocol is virtually impossible).
The reference model was devised before the corresponding protocols were invented.	The protocols came first, and the model was really just a description of the existing protocols since the protocols fit perfectly.

Example Networks

Example Networks

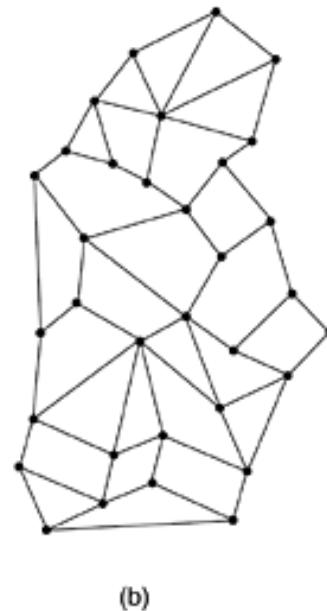
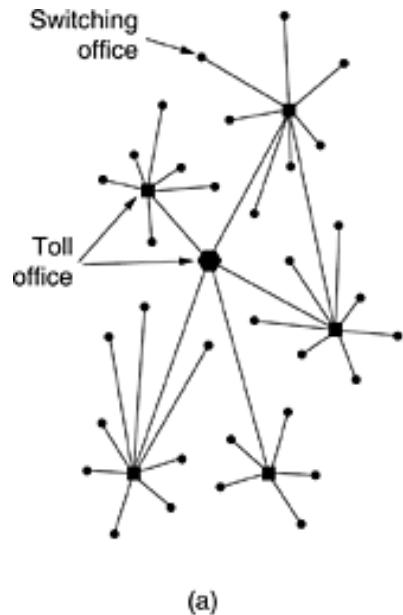
- The subject of computer networking covers many different kinds of networks, large and small, well known and less well known. They have different goals, scales, and technologies.
- The computer networks that are functioning in the current scenario are associated with so many attributes like size, technology, goals etc.
- Some examples : Internet, ATM (Asynchronous Transfer Mode), Ethernet, IEEE 802.11

The Internet

- To introduce the internet the first statement which holds true is that internet has revolutionized many aspects of our daily lives. People use internet for various reasons, if accounted it may be more than our knowledge.
- The Internet is not a network at all, but a vast collection of different networks that use certain common protocols and provide certain common services.
- It is an unusual system in that it was not planned by anyone and is not controlled by anyone.

The ARPANET (Advanced Research Projects Agency Network)

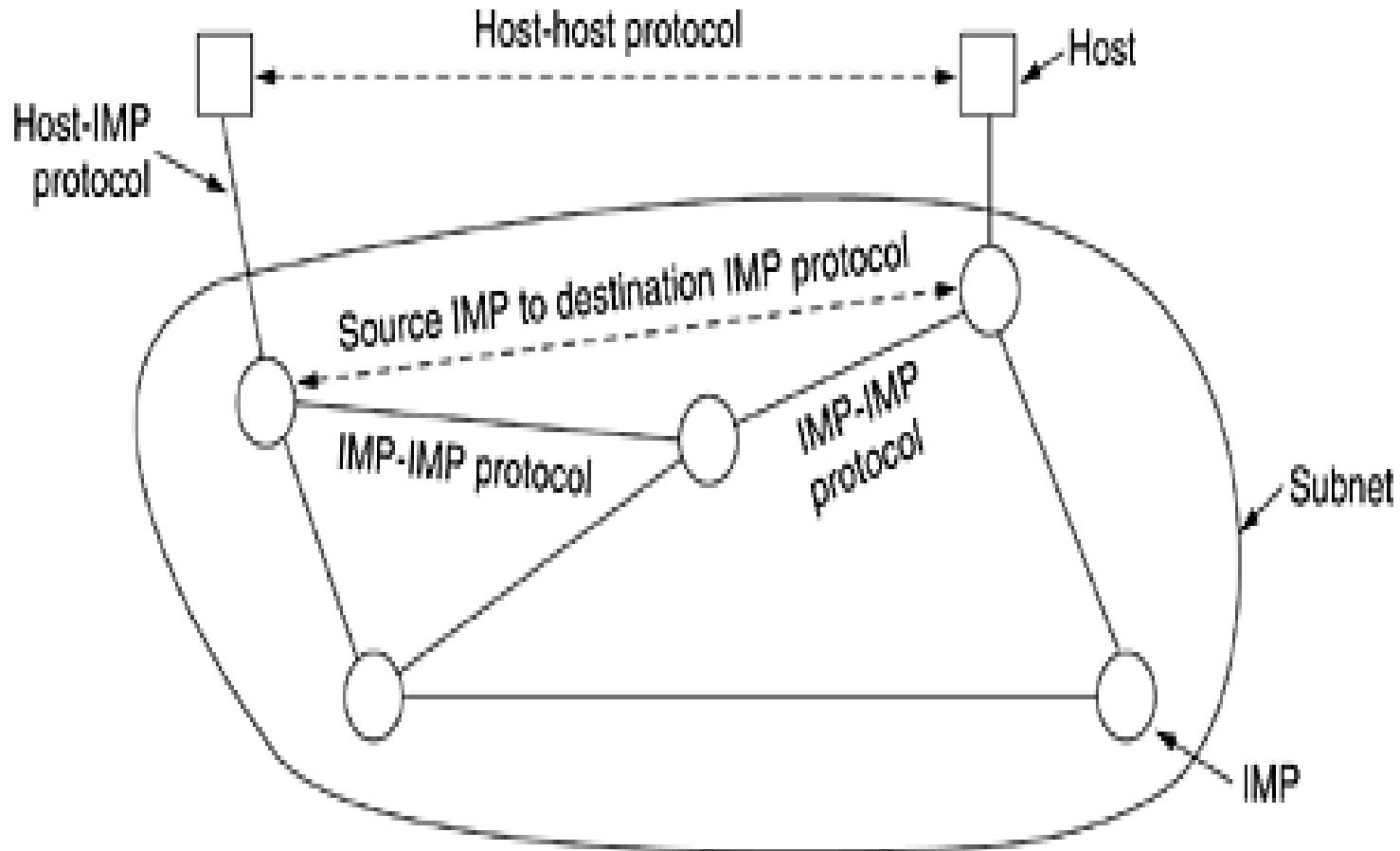
- The history started with the want from DoD (**DEPARTMENT OF DEFENSE**) in late 1950 .
 - The work started using the base of existing public telephone network.
 - The vulnerability of the system was that the destruction of a few key toll offices could fragment the system into many isolated islands.
- (a) Structure of the telephone system. (b) Baran's proposed distributed switching system.



The ARPANET

- Around 1960, the DoD awarded a contract to the RAND Corporation to find a solution. One of its employees, Paul Baran, proposed the incorporation of digital packet switching technology in a highly distributed and fault tolerant system
- However due to lack of support from the biggest and richest corporation AT&T, the idea was dismissed.
- Following several years, with the interest of U.S., ARPA (Advanced Research Projects Agency) is created to find the solution related to design of the command and control network.
- In 1967, Larry Roberts, director of ARPA bought the idea suggested by Wesley Clark related to building of a packet switched subnet, where each host has its own router: with a name **ARPANET**

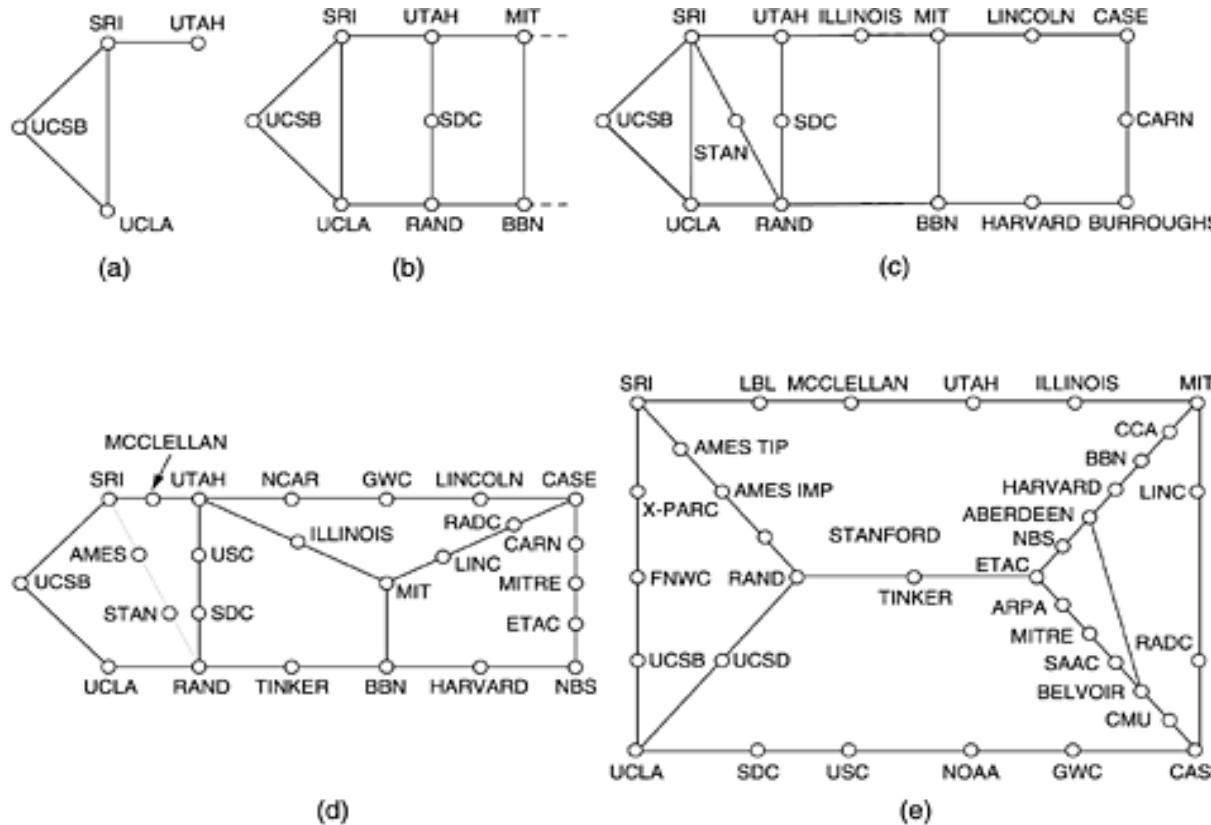
The original ARPANET design



- In ARPANET, the subnet consists of minicomputers called **IMPs (Interface Message Processors)** connected by 56-kbps transmission lines.
- For high reliability, each IMP connected to at least two other IMPs.
- The subnet was to be a datagram subnet, so if some lines and IMPs were destroyed, messages could be automatically rerouted along alternative paths. Each node of the network was to consist of an IMP and a host, in the same room, connected by a short wire.
- A host could send messages of up to 8063 bits to its IMP, which would then break these up into packets of at most 1008 bits and forward them independently toward the destination. Each packet was received in its entirety before being forwarded, so the subnet was the first electronic store-and-forward packet-switching network.
- The IMPs were interconnected by 56-kbps lines. The 56kbps lines were also leased from telephone companies.

- The software was split into two parts: subnet and host.
- The subnet software consisted of the IMP end of the host-IMP connection, the IMP-IMP protocol, and a source IMP to destination IMP protocol designed to improve reliability.
- Outside the subnet, software was also needed, namely, the host end of the host-IMP connection, the host-host protocol, and the application software.
- An experimental network went on the air in December 1969 with four nodes: at UCLA, UCSB, SRI, and the University of Utah.
- The network grew quickly as more IMPs were delivered and installed; it soon spanned the United States. Further, with installation of more IMPs the network grew quickly.

Growth of the ARPANET. (a) December 1969. (b) July 1970. (c) March 1971. (d) April 1972. (e) September 1972.

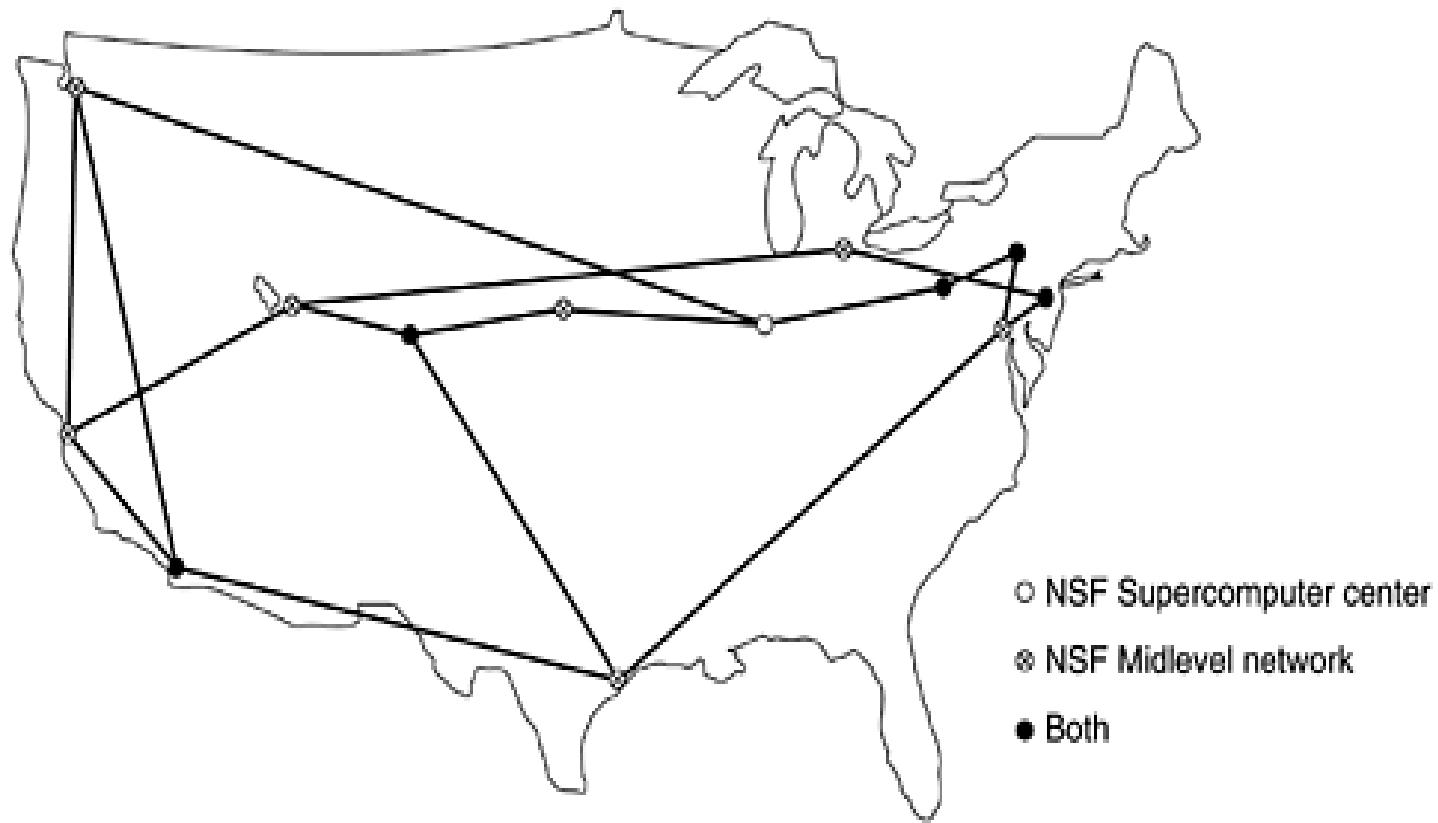


- During that period it was found that ARPANET protocols were not suitable for running over multiple networks. Further research on this leads to the invention of TCP/IP protocol which was specifically designed to handle communication over internetworks.
- Furthermore, with TCP/IP, it was easy for the LANs to connect to the ARPANET. During the 1980s, many additional networks, especially LANs, were connected to the ARPANET.
- However, As the scale increased, finding hosts became increasingly expensive, so **DNS (Domain Name System)** was created to organize machines into domains and map host names onto IP addresses.

NSFNET

- In the late 1970's to facilitate the research on network issues, NSF (the U.S. National Science Foundation) had taken a response to design a successor to the ARPANET that would be open to all university research groups.
- NSF decided to build a backbone network to connect its six supercomputer centers where in each supercomputer was attached with a microcomputer called a **fuzzball**.
- The fuzzballs were connected with 56-kbps leased lines and formed the subnet.
- The software technology was different however: the fuzzballs spoke TCP/IP right from the start, making it the first TCP/IP WAN.
- NSF also funded some regional networks that connected to the backbone to allow users at thousands of universities, research labs, libraries, and museums to access any of the supercomputers and to communicate with one another.
- The complete network, including the backbone and the regional networks, was called **NSFNET** . Further, It connected to the ARPANET through a link between an IMP and a fuzzball.

The NSFNET backbone in 1988



- Following to instantaneous success of NSFNET, the NSF started using the fiber optic channels at 448 kbps to provide the version 2 backbone. Further, by 1990, the second backbone was upgraded to 1.5 Mbps.
- Later on, looking at the inability in financial support from government, NSF encouraged few non government organizations (one of which is IBM) to form a non profit corporation, **ANS (Advanced Networks and Services)**, as the first step along the road to commercialization.
- In 1990, ANS took over NSFNET and upgraded the 1.5-Mbps links to 45 Mbps to form **ANSNET**, which after a running 5 years sold to America Online. But by then, various companies were offering commercial IP service.
- In the due course of time, to ease the transition and make sure every regional network could communicate with every other regional network, NSF awarded contracts to four different network operators to establish a **NAP (Network Access Point)**.

- Every network operator had provided backbone service to the NSF regional networks as well as to connect to all the NAPs.
- In terms of technical it means that a packet originating on any regional network had a choice of backbone carriers to get from its NAP to the destination's NAP.
- During the 1990s, many other countries and regions also built national research networks, often patterned on the ARPANET and NSFNET.
- These included EuropaNET and EBONE in Europe, which started out with 2-Mbps lines and then upgraded to 34-Mbps lines.

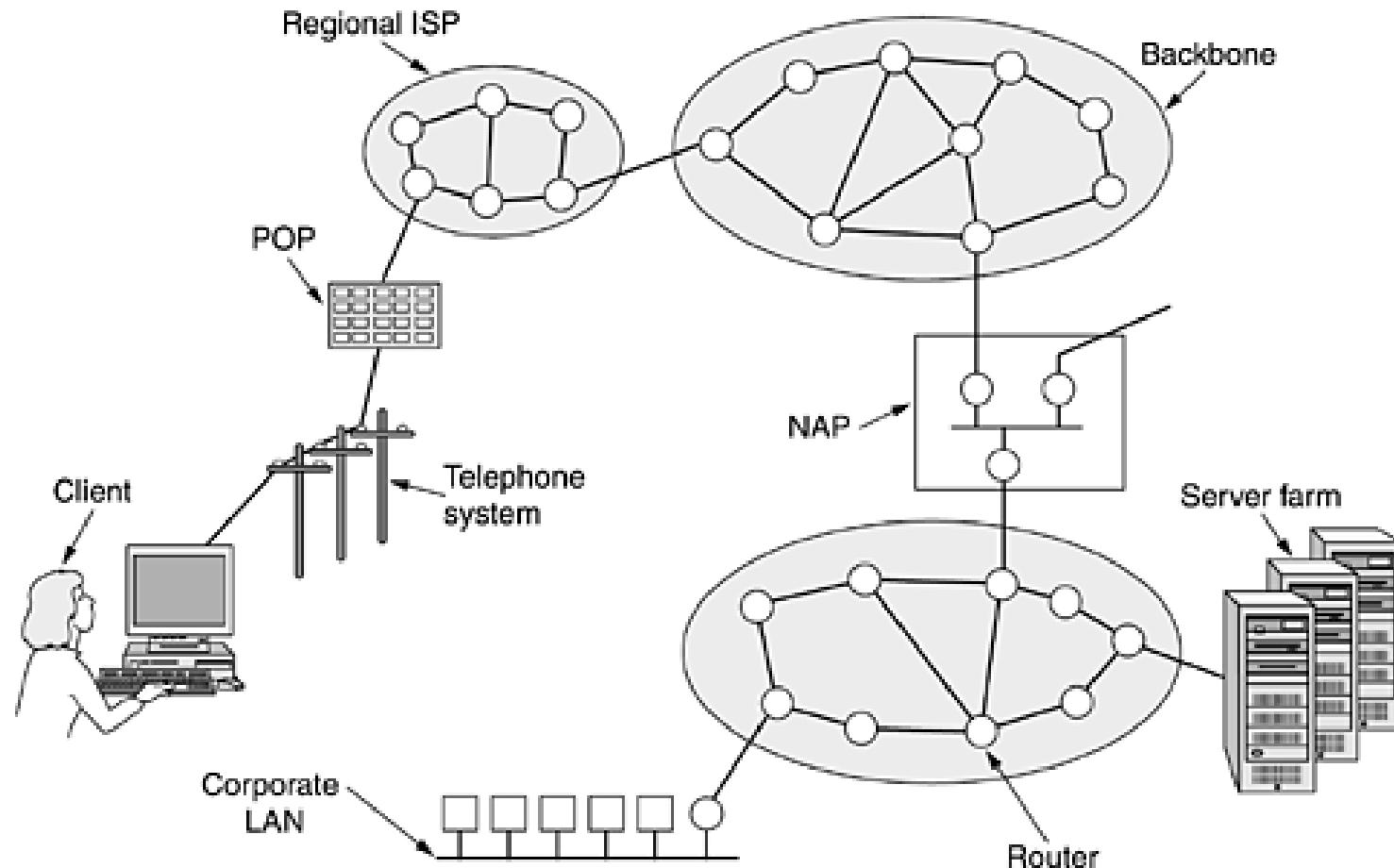
Internet Usage:

- The number of networks, machines, and users connected to the ARPANET grew rapidly after TCP/IP became the only official protocol on January 1, 1983.
- When NSFNET and the ARPANET were interconnected, the growth became exponential. Many regional networks joined up, and connections were made to networks in Canada, Europe, and the Pacific.
- Sometime in the mid-1980s, people began viewing the collection of networks as an internet, and later as the Internet.
- The glue that holds the Internet together is the TCP/IP reference model and TCP/IP protocol stack. TCP/IP makes universal service possible.
- **What does it actually mean to be on the Internet?** The definition is that a machine is on the Internet if it runs the TCP/IP protocol stack, has an IP address, and can send IP packets to all the other machines on the Internet.

- However, the issue is clouded somewhat by the fact that millions of personal computers can call up an Internet service provider using a modem, be assigned a temporary IP address, and send IP packets to other Internet hosts.
- Traditionally (meaning 1970 to about 1990), the Internet and its predecessors had four main applications: **E-mail. News. Remote login. File transfer.**
- Up until the early 1990s, the Internet was largely populated by academic, government, and industrial researchers. One new application, the **WWW (World Wide Web)** changed all that and brought millions of new, nonacademic users to the net.
- In the due course of time, numerous other kinds of pages have come into existence through WWW
- During 1990s the ISPs (Internet Service Providers) made the facility available.

Architecture of the Internet:

- **Figure. Overview of the Internet.**



- Let us assume client calls his or her ISP over a dial-up telephone line.
- The modem is a card within the PC that converts the digital signals the computer produces to analog signals that can pass unhindered over the telephone system.
- These signals are transferred to the ISP's **POP (Point of Presence)**, where they are removed from the telephone system and injected into the ISP's regional network.
- The ISP's regional network consists of interconnected routers in the various cities the ISP serves. If the packet is destined for a host served directly by the ISP, the packet is delivered to the host. Otherwise, it is handed over to the **ISP's backbone operator**.
- If a packet given to the backbone is destined for an ISP or company served by the backbone, it is sent to the closest router and handed off there.
- To allow packets to hop between backbones, all the major backbones connect at the **NAPs (Network Access Point)**.
- A Network Access Point was a public network exchange facility where Internet service providers connected with one another in peering arrangements.

Connection-Oriented Networks: ATM (Asynchronous Transfer Mode)

- In connectionless design every packet is routed independently of every other packet. As a consequence, if some routers go down during a session, no harm is done as long as the system can reconfigure itself dynamically so that subsequent packets can find some route to the destination, even if it is different from that which previous packets used.
- The connection-oriented camp comes from the world of telephone companies. In the telephone system, a caller must dial the called party's number and wait for a connection before talking or sending data. This connection setup establishes a route through the telephone system that is maintained until the call is terminated. All words or packets follow the same route.
- If a line or switch on the path goes down, the call is aborted.
- Why do the telephone companies like it then? There are two reasons:
 - Quality of service.
 - Billing.

- connection-oriented networks
 - X.25 and Frame Relay
 - ATM (Asynchronous Transfer Mode)
- The word synchronous refers to things happening at the same time. Synchronous transmission is defined as the process by which data or a signal is transferred from one application system or device to another at constant periods or intervals, usually monitored by a clock. This means that the transmitting and receiving systems send and receive data at the same rate or speed.
- In the case of asynchronous transmission, the data or signals being transmitted and received are not done in synchronization. The time interval between the sending and receiving devices enable transmission and reception at their own pace. This means the data sending transmitter may not be at the same rate as the data receptor. This mode of transmission is not monitored by the same rate and the transmission is said to be asynchronous.
- Transmission in synchronous mode closely tied to a clock and ATM is not.

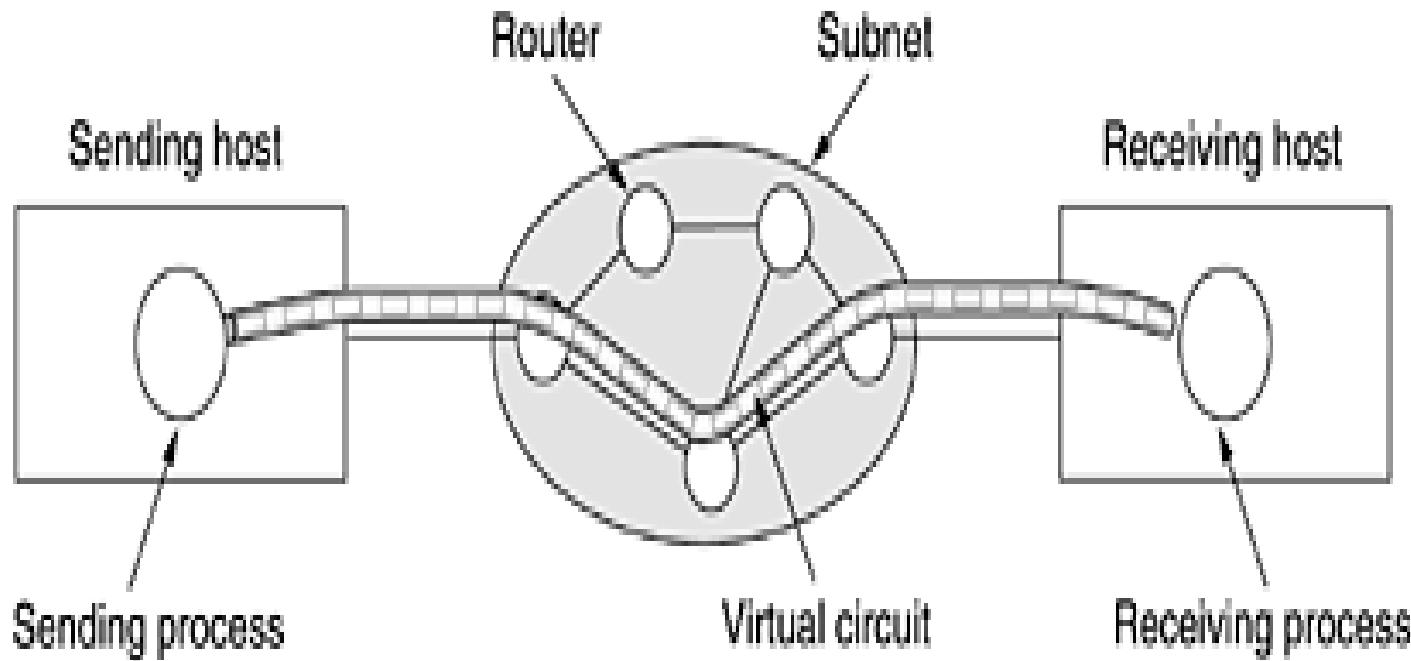
X.25 and Frame Relay

- The initial connection-oriented networks implemented in 1970s (i.e. **X.25**) and later on in 1980s (i.e **frame delay**) usually works with synchronous transmsion characteristics.
- **To use X.25**, a computer first established a connection to the remote computer, that is, placed a telephone call. This connection was given a connection number to be used in data transfer packets.
- **Frame relay** : it is a connection-oriented network with no error control and no flow control. Because it was connection-oriented, packets were delivered in order

Asynchronous Transfer Mode (ATM)

- connection-oriented network developed to work with an asynchronous transmission system. This network is named as **ATM (Asynchronous Transfer Mode)** network.
- ATM was designed in the early 1990s. The main aim of this network is to solve all the world's networking and telecommunications problems by merging voice, data, cable television, telex, telegraph, connected by strings and everything else into a single integrated system that could do everything for everyone.
- However, this was not happened at that time due to bad timing, technology and implementation. Later on, it was found to be successful and being widely used within the telephone system for moving IP packets.
- ATM was much more successful than OSI, and it is now widely used deep within the telephone system, often for moving IP packets.

ATM Virtual Circuits:



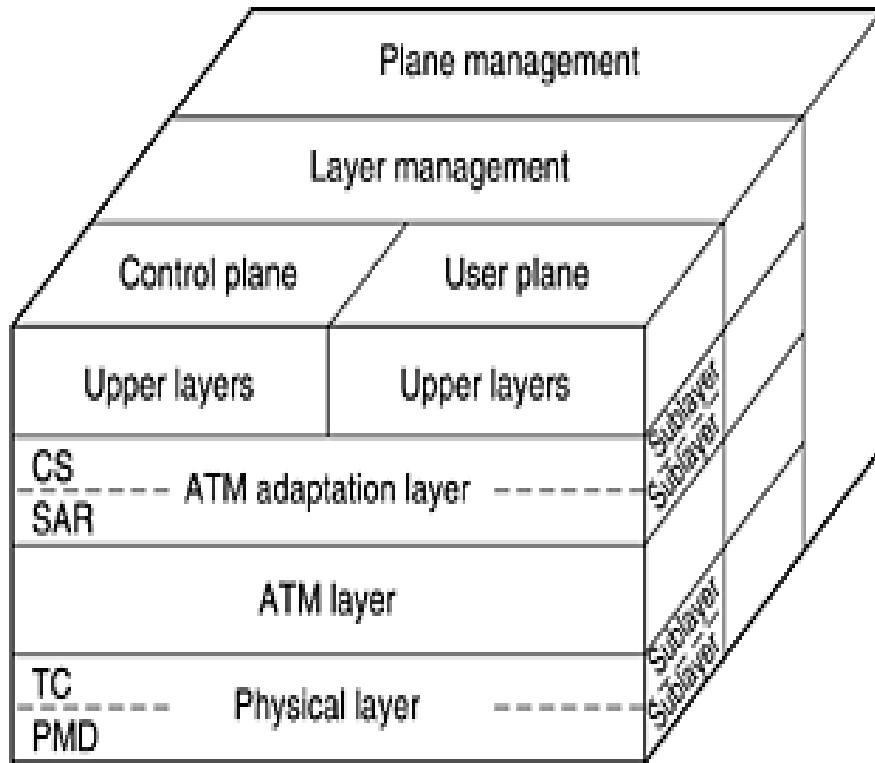
- Since ATM networks are connection-oriented, sending data requires first sending a packet to set up the connection.
- Connections are often called virtual circuits, in analogy with the physical circuits used within the telephone system.
- Most ATM networks also support permanent virtual circuits, which are permanent connections between two (distant) hosts. They are similar to leased lines in the telephone world.
- Each connection, temporary or permanent, has a unique connection identifier.
- Once a connection has been established, either side can begin transmitting data. The basic idea behind ATM is to transmit all information in small, fixed-size packets called cells.
- ATM Cell



- The cells are 53 bytes long, of which 5 bytes are header and 48 bytes are user data.
- Part of the header is the connection identifier, so the sending and receiving hosts and all the intermediate routers can tell which cells belong to which connections. This information allows each router to know how to route each incoming cell.
- Cell routing is done in hardware, at high speed. In fact, the main argument for having fixed-size cells is that it is easy to build hardware routers to handle short, fixed-length cells.
- Variable-length IP packets have to be routed by software, which is a slower process.

- Another plus of ATM is that the hardware can be set up to copy one incoming cell to multiple output lines, a property that is required for handling a television program that is being broadcast to many receivers. Finally, small cells do not block any line for very long, which makes guaranteeing quality of service easier.
- All cells follow the same route to the destination. Cell delivery is not guaranteed, but their order is. If cells 1 and 2 are sent in that order, then if both arrive, they will arrive in that order, never first 2 then 1. But either or both of them can be lost along the way.
- It is up to higher protocol levels to recover from lost cells.
- ATM networks are organized like traditional WANs, with lines and switches (routers). The most common speeds for ATM networks are 155 Mbps and 622 Mbps, although higher speeds are also supported.

The ATM Reference Model:



CS: Convergence sublayer

SAR: Segmentation and
reassembly sublayer

TC: Transmission convergence
sublayer

PMD: Physical medium
dependent sublayer

- ATM has its own reference model, different from the OSI model and also different from the TCP/IP model.
- It consists of three layers, the physical, ATM, and ATM adaptation layers, plus whatever users want to put on top of that (layers with a flexibility for an user defined upper layer(s) above that).

Physical layer :

- Deals with the physical medium: voltages, bit timing, and various other issues.
- No specific rules for the cells regarding the choice of transmission medium.
- ATM cells can be sent on a wire or fiber by themselves.

ATM layer :

- Deals with cells and cell transport.
- Defines the layout of a cell and tells what the header fields mean.
- Deals with establishment and release of virtual circuits.
- Handles congestion control issues.

ATM adaption layer (AAL) :

- Allow users to send packets larger than a cell.
- The ATM interface segments these packets, transmits to lower layer.
- Reassembles the segments (if any) at the other end.

User defined upper layer :

- User plane deals with data transport, flow control, error correction, and other user functions.
- Control plane is concerned with connection management.
- Layer and plane management functions relate to resource management and interlayer coordination.

The ATM layers and sublayers, and their functions.

OSI layer	ATM layer	ATM sublayer	Functionality
3/4	AAL	CS	Providing the standard interface (convergence)
		SAR	Segmentation and reassembly
2/3	ATM		Flow control Cell header generation/extraction Virtual circuit/path management Cell multiplexing/demultiplexing
2	Physical	TC	Cell rate decoupling Header checksum generation and verification Cell generation Packing/unpacking cells from the enclosing envelope Frame generation
			Bit timing Physical network access

PMD (Physical Medium Dependent) sublayer:

- Make the bits on and off to move through transmission medium (say cable)/carrier.
- Handles the bit timing.
- For different carriers and cables, this layer will be different.

TC (Transmission Convergence) sublayer:

- Converts the cells into bitstream in transmitting end and the reverse in receiving end.
- Handles all the issues related to telling where cells begin and end in the bitstream.

SAR (Segmentation And Reassembly) sublayer:

- Breaks up packets into cells on the transmission side and puts them back together again at the destination.

CS (Convergence Sublayer):

- handles different kinds of services to different applications (e.g., file transfer and video on demand have different requirements concerning error handling, timing, etc.).

Ethernet

- Both the Internet and ATM were designed for wide area networking. However, many companies, universities, and other organizations have large numbers of computers that must be connected. This need gave rise to the local area network.
- It is the most popular local area network which was developed and implemented in Xerox PARC (Palo alto Research center) in 1976. Prior to this the concept of interconnecting the computers was available, where the communication was taking place through short range radio devices. Such a network was implemented by University of Hawaii in 1970s and was named as ALOHANET.
- In ALOHANET a number of user terminals were connected to a central computer. The communication between each user terminal and central computer was taking place in two frequencies: **upstream (to the central computer)** and **downstream (from the central computer)**.

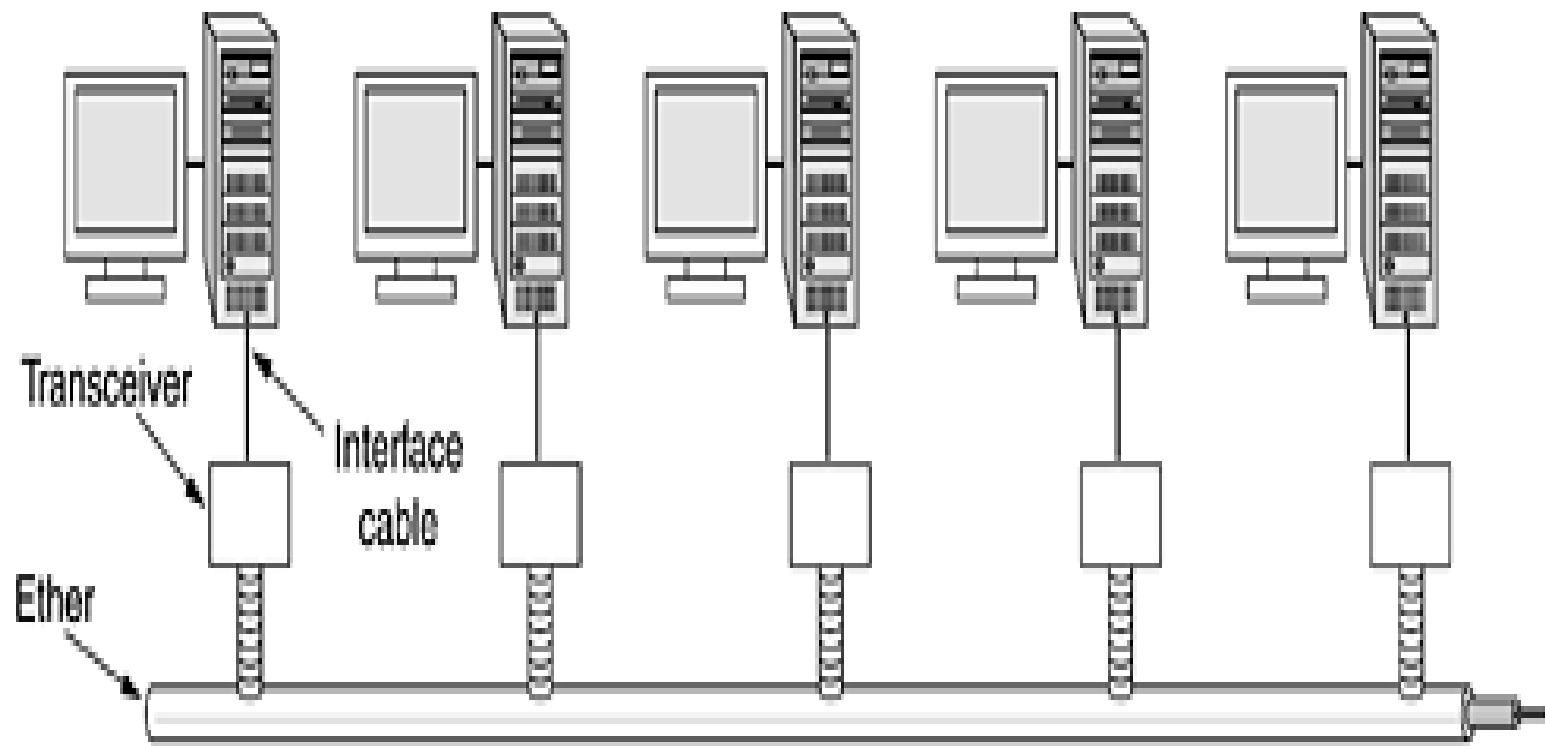
Ethernet

- When the user wanted to contact the computer, it just transmitted a packet containing the data in the upstream channel. If no one else was transmitting at that instant, the packet probably got through and was acknowledged on the downstream channel.
- This system worked fairly well under conditions of low traffic but bogged down badly when the upstream traffic was heavy because contention in upstream channel.

- **Ethernet** was named after the *luminiferous ether*, through which electromagnetic radiation was once thought to propagate. This network was developed after ALOHANET with a difference that in **Ethernet the transmission medium was not the vacuum but a thick coaxial cable (the ether)**.
- The coaxial cable is up to 2.5 k long (with repeaters at every 500meters).
- Up to 256 machines could be attached to the system via transceivers screwed onto the cable. A cable with multiple machines attached to it in parallel is called a multidrop cable. The system ran at 2.94 Mbps.

- Ethernet had a major improvement over ALOHANET: before transmitting, a computer first listened to the cable to see if someone else was already transmitting. If so, the computer held back until the current transmission finished. Doing so avoided interfering with existing transmissions, giving a much higher efficiency.
- ALOHANET did not work like this because it was impossible for a terminal on one island to sense the transmission of a terminal on a distant island. With a single cable, this problem does not exist.

Architecture of the original Ethernet



- Despite the computer listening before transmitting, a problem still arises: what happens if two or more computers all wait until the current transmission completes and then all start at once?
- The Xerox Ethernet was so successful that DEC, Intel, and Xerox drew up a standard in 1978 for a 10-Mbps Ethernet, called the DIX standard. With two minor changes, the DIX standard became the **IEEE 802.3** standard in 1983.
- Ethernet continued to develop and is still developing. New versions at 100 Mbps, 1000 Mbps, and still higher have come out. Also the cabling has improved, and switching and other features have been added.
- Ethernet (**IEEE 802.3**) is not the only LAN standard. The committee also standardized a token bus (**802.4**) and a token ring (**802.5**).

- The term **token** is nothing but a short packet and is used to make a turn for a computer being allowable for transmission of its data. Thus it was taken that a computer could only send if it possessed the token, thus avoiding collisions.
- Similarly, IBM had its own favorite: its proprietary token ring. The token was passed around the ring and whichever computer held the token was allowed to transmit before putting the token back on the ring.
- However, in due course of time 802.4 has vanished from sight but 802.5 had its existence and still in use at some IBM site (popular in the name **IBM token ring**).
- However in the war of LAN, Ethernet has taken the highest utility in compare to others like token bus and token ring.

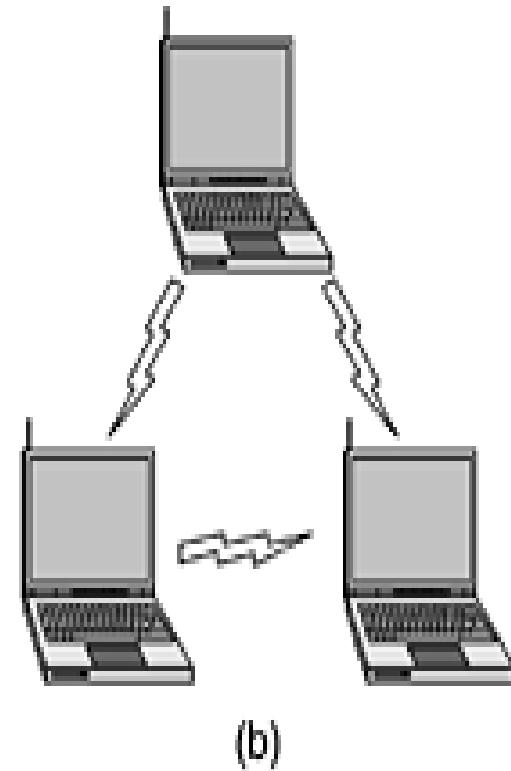
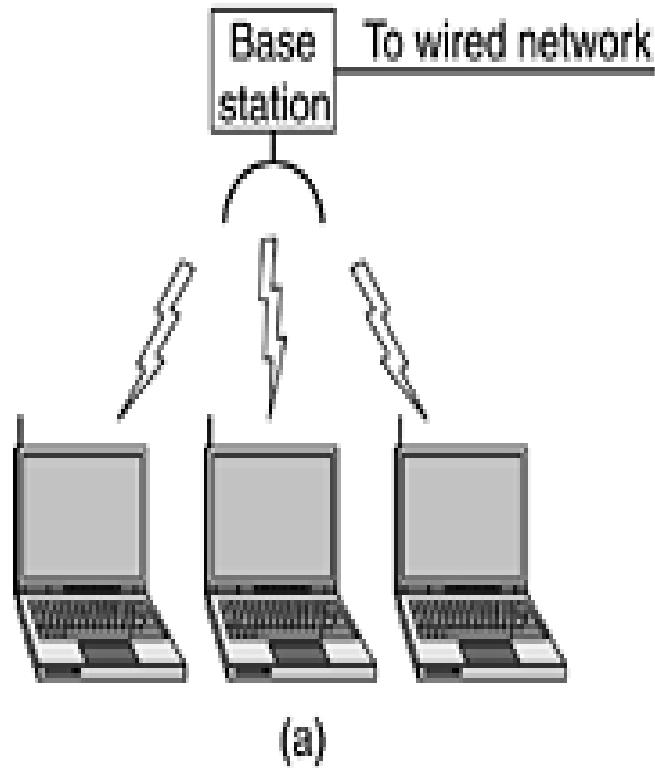
Wireless LANs: 802.11

- The idea of Wireless LAN was developed when it was thought of to equip both the office and the notebook computers with short-range radio transmitters and receivers and to allow them to communicate.
- The most practical approach is to equip both the office and the notebook computers with short-range radio transmitters and receivers to allow them to communicate.
- But during its implementation some systems faces problem because technical incompatibility between devices. For example a computer equipped with a brand X radio could not work in a room equipped with a brand Y base station.
- To short out this issue, the industry decided that a wireless LAN standard might be a good idea, so the IEEE committee that standardized the wired LANs was given the task of drawing up a wireless LAN standard.

- The standard it came up with was named 802.11. A common name for it is WiFi.
- The proposed standard had to work in two modes:
 1. **In the presence of a base station:** all communication was to go through the base station, called an **access point** in 802.11 terminology.
 2. **In the absence of a base station :** the computers would just send to one another directly. This mode is now sometimes called **adhoc networking**.

The two modes are illustrated in fig.

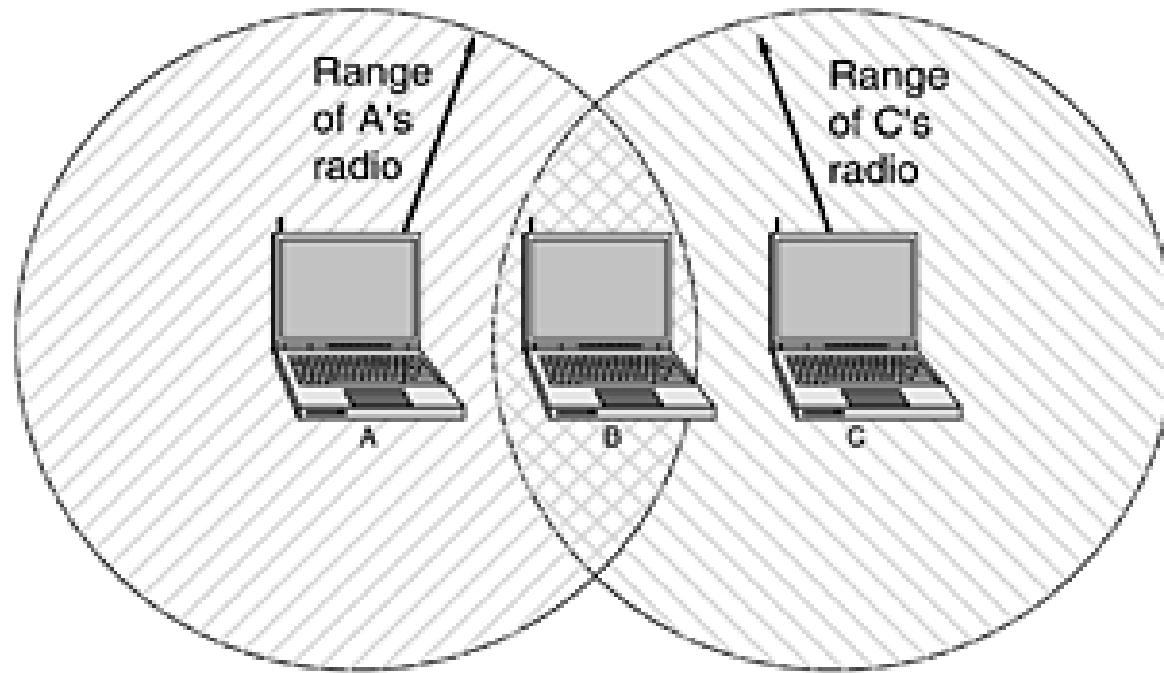
Figure. (a) Wireless networking with a base station. (b) Ad hoc networking.



- In particular, some of the many challenges that had to be met were:
 - finding a suitable frequency band that was available, preferably worldwide;
 - dealing with the fact that radio signals have a finite range;
 - ensuring that users' privacy was maintained;
 - taking limited battery life into account;
 - worrying about human safety (do radio waves cause cancer?);
 - understanding the implications of computer mobility;
 - building a system with enough bandwidth to be economically viable.

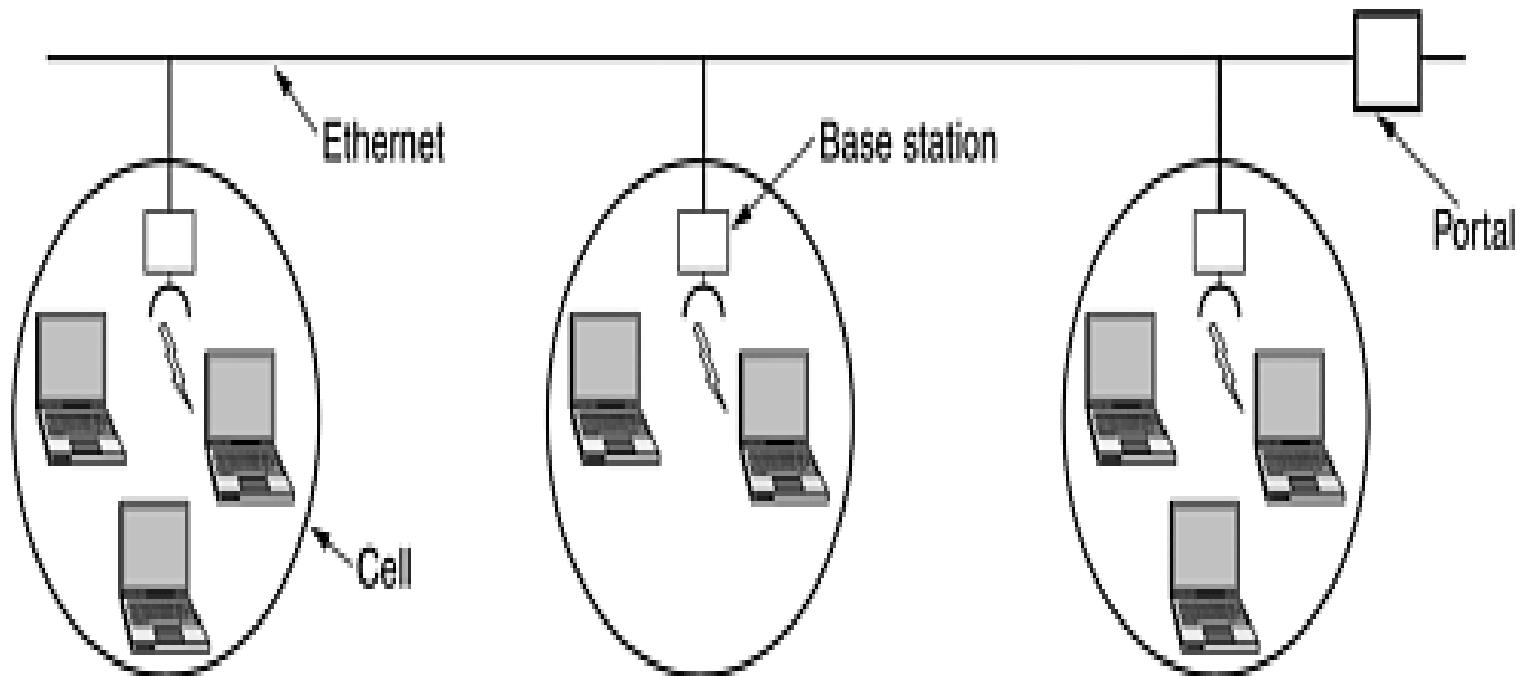
- At the time the standardization process started (mid-1990s), Ethernet had already come to dominate local area networking, so the committee decided to make 802.11 compatible with Ethernet above the data link layer.
- In particular, it should be possible to send an IP packet over the wireless LAN the same way a wired computer sent an IP packet over Ethernet. Nevertheless, in the physical and data link layers, several inherent differences with Ethernet exist and had to be dealt with by the standard.
- Though Wireless LAN had come up as a substitute of Ethernet but suffered from certain problems found out after installation.
- First, a computer on Ethernet always listens to the ether before transmitting. Only if the ether is idle does the computer begin transmitting. With wireless LANs, that idea does not work so well.

Figure. The range of a single radio may not cover the entire system.



- The second problem that had to be solved is that a radio signal can be reflected off solid objects, so it may be received multiple times (along multiple paths). This interference results in what is called **multipath fading**.
- The third problem is that a great deal of software is not aware of mobility.
- The fourth problem is that if a notebook computer is moved away from the ceiling-mounted base station it is using and into the range of a different base station, some way of handing it off is needed.
- Although this problem occurs with cellular telephones, it does not occur with Ethernet and needed to be solved.
- The network consists of multiple cells, each with its own base station, but with the base stations connected by Ethernet.
- From the outside, the entire system should look like a single Ethernet. The connection between the 802.11 system and the outside world is called a portal.

A multicell 802.11 network.



- After some work, the committee came up with a standard in 1997 that addressed these and other concerns. The wireless LAN it described ran at either **1 Mbps or 2 Mbps**.
- A split developed within the committee, resulting in two new standards in 1999.
 - The **802.11a** standard uses a wider frequency band and runs at speeds up to **54 Mbps**.
 - The **802.11b** standard uses the same frequency band as 802.11, but uses a different modulation technique to achieve **11 Mbps**.
- In the current scenario, the 802.11 is being widely used in organizations like airports, railway stations, hotels, shopping malls, and universities so far as the computational ability and internet access is concerned.

QUIZ

Problem

- A client-server system uses a satellite network, with the satellite at a height of 40,000 km. What is the best-case delay in response to a request?

- A client-server system uses a satellite network, with the satellite at a height of 40,000 km. What is the best-case delay in response to a request?

Answer:

The request has to go up and down, and the response has to go up and down. The total path length traversed is thus 160,000 km. The speed of light in air and vacuum is 300,000 km/sec, so the propagation delay alone is $160,000/300,000$ sec or about 533 msec.

- A system has an n-layer protocol hierarchy. Applications generate messages of length M bytes. At each of the layers, an h -byte header is added. What fraction of the network bandwidth is filled with headers?

- A system has an n -layer protocol hierarchy. Applications generate messages of length M bytes. At each of the layers, an h -byte header is added. What fraction of the network bandwidth is filled with headers?

Answer:

With n layers and h bytes added per layer, the total number of header bytes per message is hn , so the space wasted on headers is hn . The total message size is $M + nh$, so the fraction of bandwidth wasted on headers is $hn / (M + hn)$.

An image is 1024 x 768 pixels with 3 bytes/pixel. Assume the image is uncompressed. How long does it take to transmit it over a 56-kbps modem channel? Over a 1-Mbps cable modem? Over a 10-Mbps Ethernet? Over 100-Mbps Ethernet?

- An image is 1024 x 768 pixels with 3 bytes/pixel. Assume the image is uncompressed. How long does it take to transmit it over a 56-kbps modem channel? Over a 1-Mbps cable modem? Over a 10-Mbps Ethernet? Over 100-Mbps Ethernet?

Answer:

- The image is $1024 \times 768 \times 3$ bytes or 2,359,296 bytes. This is 18,874,368 bits. At 56,000 bits/sec, it takes about 337.042 sec. At 1,000,000 bits/sec, it takes about 18.874 sec. At 10,000,000 bits/sec, it takes about 1.887 sec. At 100,000,000 bits/sec, it takes about 0.189 sec.

- Assume 6 devices are arranged in a mesh topology. How many cables are needed? How many ports are needed for each device?

- Assume 6 devices are arranged in a mesh topology. How many cables are needed? How many ports are needed for each device?

Answer:

$$\text{No. of cables} = n*(n-1)/2 = 15 \text{ cables}$$

$$\text{No. of devices connected per device} = n-1 = 5$$

$$\text{No. of ports per device} = 5$$

- How long was a bit on the original IEEE 802.3 standard in meters? Use a transmission speed of 10 Mbps and assume the propagation speed in coax is $2/3$ the speed of light in vacuum.

- How long was a bit on the original 802.3 standard in meters? Use a transmission speed of 10 Mbps and assume the propagation speed in coax is 2/3 the speed of light in vacuum.
- Answer:

The speed of light in coax is about 200,000 km/sec, which is 200 meters/ μ sec. At 10 Mbps, it takes 0.1 μ sec to transmit a bit. Thus, the bit lasts 0.1 μ sec in time, during which it propagates 20 meters. Thus, a bit is 20 meters long here.

- A factor in the delay of a store-and-forward packet-switching system is how long it takes to store and forward a packet through a switch. If switching time is 10 μ sec, is this likely to be a major factor in the response of a client-server system where the client is in New York and the server is in California? Assume the propagation speed in copper and fiber to be 2/3 the speed of light in vacuum.

- A factor in the delay of a store-and-forward packet-switching system is how long it takes to store and forward a packet through a switch. If switching time is 10 μ sec, is this likely to be a major factor in the response of a client-server system where the client is in New York and the server is in California? Assume the propagation speed in copper and fiber to be 2/3 the speed of light in vacuum.

Answer:

- No. The speed of propagation is 200,000 km/sec or 200 meters/ μ sec. In 10 μ sec the signal travels 2 km. Thus, each switch adds the equivalent of 2 km of extra cable. If the client and server are separated by 5000 km, traversing even 50 switches adds only 100 km to the total path, which is only 2%. Thus, switching delay is not a major factor under these circumstances.

- A collection of five routers is to be connected in a point-to-point subnet. Between each pair of routers, the designers may put a high-speed line, a medium-speed line, a low-speed line, or no line. If it takes 100 ms of computer time to generate and inspect each topology, how long will it take to inspect all of them?

- A collection of five routers is to be connected in a point-to-point subnet. Between each pair of routers, the designers may put a high-speed line, a medium-speed line, a low-speed line, or no line. If it takes 100 ms of computer time to generate and inspect each topology, how long will it take to inspect all of them?
- Answer:

Call the routers A, B, C, D, and E. There are ten potential lines: AB, AC, AD, AE, BC, BD, BE, CD, CE, and DE. Each of these has four possibilities (three speeds or no line), so the total number of topologies is $4^{10} = 1,048,576$. At 100 ms each, it takes 104,857.6 sec, or slightly more than 29 hours to inspect them all.

Chapter 2. The Physical Layer

The Physical Layer

- It defines the mechanical, electrical, and timing interfaces to the network.
- Three kinds of transmission media:
 - Guided (copper wire and fiber optics),
 - Wireless (terrestrial radio),
 - Satellite.
- Three examples of communication systems used in practice for wide area computer networks: **the (fixed) telephone system, the mobile phone system, and the cable television system.**

The Theoretical Basis for Data Communication

- Information can be transmitted on wires by varying some physical property such as voltage or current.
- By representing the value of this voltage or current as a single-valued function of time, $f(t)$, we can model the behavior of the signal and analyze it mathematically.

Fourier Analysis

- In the early 19th century, the French mathematician Jean-Baptiste Fourier proved that any reasonably behaved periodic function, $g(t)$ with period T can be constructed as the sum of a (possibly infinite) number of sines and cosines:

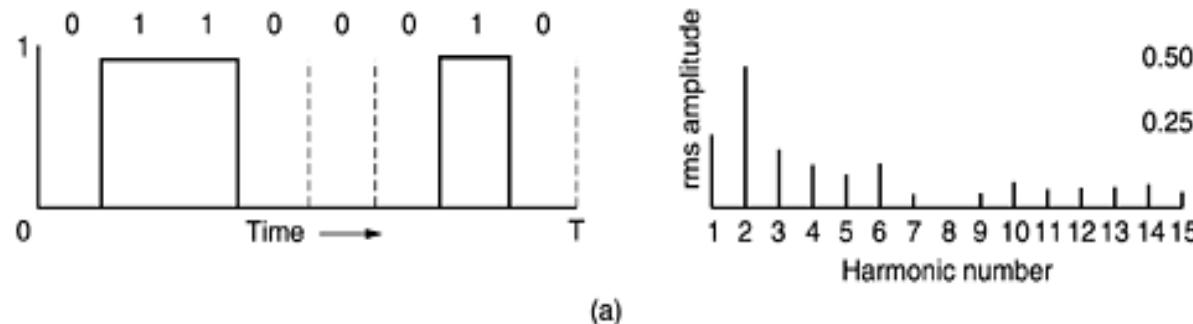
$$g(t) = \frac{1}{2}c + \sum_{n=1}^{\infty} a_n \sin(2\pi nft) + \sum_{n=1}^{\infty} b_n \cos(2\pi nft)$$

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

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

Bandwidth-Limited Signals

- let us consider a specific example: the transmission of the ASCII character "b" encoded in an 8-bit byte. The bit pattern that is to be transmitted is 01100010.



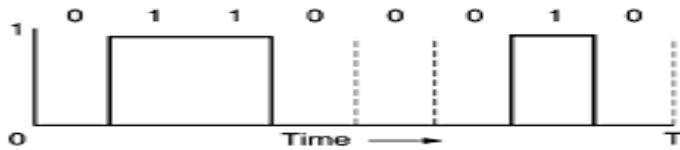
(a) A binary signal and its root-mean-square Fourier amplitudes.

$$a_n = \frac{1}{\pi n} [\cos(\pi n/4) - \cos(3\pi n/4) + \cos(6\pi n/4) - \cos(7\pi n/4)]$$

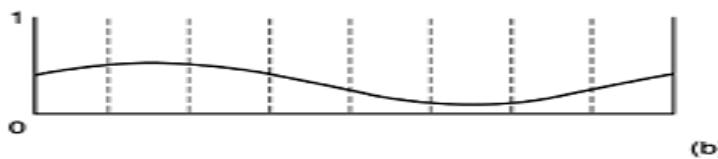
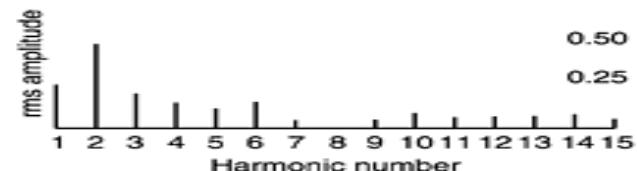
$$b_n = \frac{1}{\pi n} [\sin(3\pi n/4) - \sin(\pi n/4) + \sin(7\pi n/4) - \sin(6\pi n/4)]$$

$$c = 3/4$$

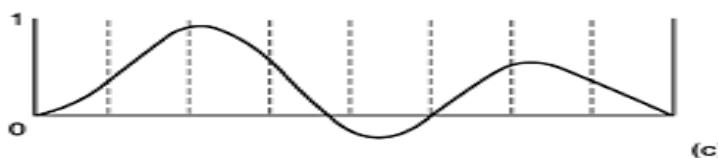
(b)-(e) Successive approximations to the original signal.



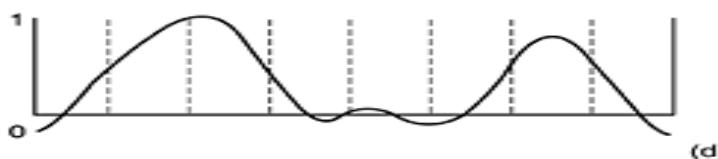
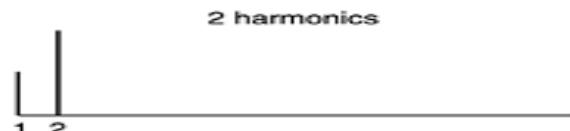
(a)



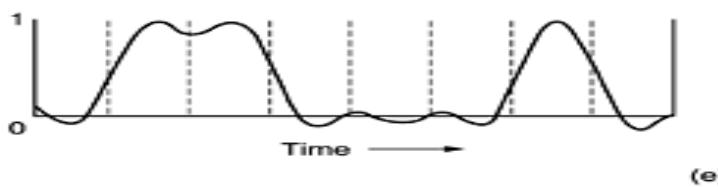
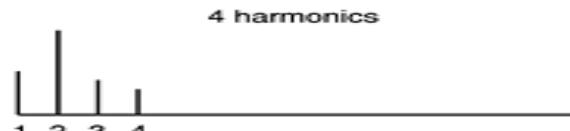
(b)



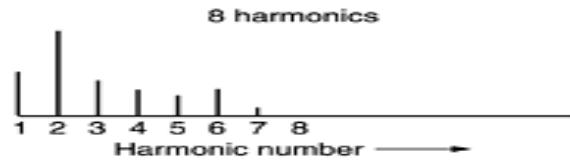
(c)



(d)



(e)



- No transmission facility can transmit signals without losing some power in the process.
- If all the Fourier components were equally diminished, the resulting signal would be reduced in amplitude but not distorted.
- Unfortunately, all transmission facilities diminish different Fourier components by different amounts, thus introducing distortion.
- Usually, the amplitudes are transmitted undiminished from 0 up to some frequency f_c [measured in cycles/sec or Hertz (Hz)] with all frequencies above this cutoff frequency attenuated.
- **The range of frequencies transmitted without being strongly attenuated is called the bandwidth.**
- In practice, the cutoff is not really sharp, so often the quoted bandwidth is from 0 to the frequency at which half the power gets through.
- **The bandwidth is a physical property of the transmission medium and usually depends on the construction, thickness, and length of the medium.**
- In some cases a filter is introduced into the circuit to limit the amount of bandwidth available to each customer.
- Figure 2-1(b) shows the signal that results from a channel that allows only the first harmonic (the fundamental, f) to pass through. Similarly, Fig. (c)-(e) show the spectra and reconstructed functions for higher-bandwidth channels.

Relation between data rate and harmonics

Example: Assume you want to send 8 bits at 9600 bps over an ordinary phone line, BW = 3000Hz

- $b=9600 \text{ bps}$

The time to send 8 bits is $T= 8/b=8/9600 = 0.83 \text{ msec.}$

- The frequency of the first harmonic is $1 / (8/b) = b/8=9600/8 =1/0.83 \text{ msec} = 1200 \text{ Hz}$ (periods per second)
- Ordinary phone lines have an artificial cut-off bandwidth of 3000Hz.
 $BW = 3000\text{Hz}$
- Thus the highest harmonic passed through is $= BW / (b/8) = 8 BW/ b = 2.5 \Rightarrow$ highest harmonic is 2
- The signal received would be tricky to reconstruct
=> limiting the bandwidth limits the data rate

Example: Assume you want to send 8 bits at 300 bps over an ordinary phone line. Find out the highest harmonic ?

Example: Assume you want to send 8 bits at 300 bps over an ordinary phone line

- $b=300 \text{ bps}$

The time to send 8 bits is $8/b=8/300 = 26.67 \text{ msec.}$

- The frequency of the first harmonic is $1 / (8/b) = b/8=300/8 =1/26.67 \text{ msec}= 37.5 \text{ Hz}$ (periods per second)
- Ordinary phone lines have an artificial cut-off bandwidth of 3000Hz.

$BW = 3000\text{Hz}$

- Thus the highest harmonic passed through is $= BW / (b/8) = 8 \text{ BW/ b} = 8*3000/300 => \text{highest harmonic is } 80$
- The signal received would be easy to reconstruct

Example: Assume you want to send 8 bits at 38400 bps over an ordinary phone line

- $b=38400$ bps

The time to send 8 bits is $8/b=8/38400 = 0.21$ msec.

- The frequency of the first harmonic is $1 / (8/b) = b/8=38400/8 =1/0.21$ msec= 4800 Hz (periods per second)
- Ordinary phone lines have an artificial cut-off bandwidth of 3000Hz.

$BW = 3000\text{Hz}$

- Thus the highest harmonic passed through is $= BW / (b/8) = 8 BW/ b = 8*3000/38400 =>$ highest harmonic is 0
- The signal received not possible to reconstruct
=> limiting the bandwidth limits the data rate

(Relation between data rate and harmonics for send a constant of 8 bits over a 3KHz channel.)

Bps	T (msec)	First harmonic (Hz)	# Harmonics sent
300	26.67	37.5	80
600	13.33	75	40
1200	6.67	150	20
2400	3.33	300	10
4800	1.67	600	5
9600	0.83	1200	2
19200	0.42	2400	1
38400	0.21	4800	0

The Maximum Data Rate of a Channel

- As early as 1924, an AT&T engineer, Henry Nyquist, realized that even a perfect channel has a finite transmission capacity.
- He derived an equation expressing the maximum data rate for a **finite bandwidth noiseless channel**.
- Nyquist proved that if an arbitrary signal has been run through a low-pass filter of bandwidth **H**, the filtered signal can be completely reconstructed by making only **2H** (exact) samples per second.
- Sampling the line faster than $2H$ times per second is pointless because the higher frequency components that such sampling could recover have already been filtered out.
- If the signal consists of V discrete levels, **Nyquist's theorem** states:

$$\text{maximum data rate} = 2H \log_2 V \text{ bits/sec}$$

- **For example, a noiseless 3-kHz channel cannot transmit binary (i.e., two-level) signals at a rate exceeding 6000 bps.**

- **What is the maximum data rate in a noiseless 6-kHz channel transmitting 16 bit signals .**
- Answer:

$$\text{maximum data rate} = 2H \log_2 V \text{ bits/sec}$$

- $H=6 \text{ kHz}= 6000, V= 16$
- Maximum data rate = $2 * 6000 * 4 = 48000 \text{ bps}$

- So far we have considered only noiseless channels. If random noise is present, the situation deteriorates rapidly. And there is always random (thermal) noise present due to the motion of the molecules in the system.
- In 1948, **Claude Shannon** carried Nyquist's work further and extended it to the case of a channel subject to random (that is, thermodynamic) noise.
- The amount of thermal noise present is measured by the ratio of the signal power to the noise power, called the **signal-to-noise ratio**.
- If we denote the signal power by S and the noise power by N, the signal-to-noise ratio is **S/N**.
- signal-to-noise ratio is S/N is also known as **SNR**.
- Usually, the ratio itself is not quoted; instead, the quantity $10 \log_{10} (S/N)$ is given. These units are called decibels (dB).
- **$SNR \text{ in dB} = 10 \log_{10} (S/N) \text{ dB}$**
- **Example:**
 - An S/N ratio of 10 is 10 dB,
 - The manufacturers of stereo amplifiers often characterize the bandwidth (frequency range) over which their product is linear by giving the 3-dB frequency on each end.
 - These are the points at which the amplification factor has been approximately halved (because $\log_{10} 3 = 0.5$).

Calculate the SNR in dB

- Signal to noise ratio of 100,
- Signal to noise ratio of 1000.

Answer:

$$\text{SNR in dB} = 10 \log_{10} (100) \text{ dB} = 10 * 2 = 20 \text{ dB}$$

$$\text{SNR in dB} = 10 \log_{10} (1000) \text{ dB} = 10 * 3 = 30 \text{ dB}$$

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

$$\text{maximum number of bits/sec} = H \log_2 (1 + S/N)$$

- **For example,** a channel of 3000-Hz bandwidth with a signal to thermal noise ratio of 30 dB (typical parameters of the analog part of the telephone system), $S/N=1000$, can never transmit much more than 30,000 bps, no matter how many or how few signal levels are used and no matter how often or how infrequently samples are taken.
- Shannon's result was derived from information-theory arguments and applies to any channel subject to thermal noise.

2. A noiseless 4-kHz channel is sampled every 1 msec. What is the maximum data rate?

2. A noiseless 4-kHz channel is sampled every 1 msec. What is the maximum data rate?

ANS:

A noiseless channel can carry an arbitrary large amount of information, no matter how often it is sampled.

Just send a lot of data per sample.

For 4-KHz channel, make 1000 samples/sec.

If each sample is 16 bits, the channel can send 16 Kbps.

If each sample is 1024 bits, the channel can send

$1000 \text{ samples/sec} * 1024 \text{ bits} = 1024 \text{ Mbps}$

The key word here is “noiseless”. With a normal 4 KHz channel, Shannon limit would not allow this.

For the 4 KHz channel we can make 8000 samples/sec.

In this case if each sample is 1024 bits this channel can send 8.2 Mbps.

3. Television channels are 6 MHz wide. How many bits/sec can be sent if four-level digital signals are used? Assume a noiseless channel.

3. Television channels are 6 MHz wide. How many bits/sec can be sent if four-level digital signals are used? Assume a noiseless channel.

- Using the **Nyquist theorem**,
- Max. data rate = $2H \log_2 V$ bits/sec",
we can sample = $2 (6\text{MHz}) \log_2 (4) = 24$
Mbps.

4. If a binary signal is sent over a 3-kHz channel whose signal-to-noise ratio is 20 dB, what is the maximum achievable data rate?

4. If a binary signal is sent over a 3-kHz channel whose signal-to-noise ratio is 20 dB, what is the maximum achievable data rate?

$$10\log_{10} S / N = 20 \text{ dB}$$

$$S / N = 100$$

Using Shannon theorem,

$$\text{Maximum number of bits/sec} = H * \log_2(1 + S / N) = 3 \text{ KHz} * \log_2(1 + 100)$$

$$\Rightarrow \log_2(101) \neq 6,658 \Rightarrow$$

$$= 3 * 6,658 = 19,975 \text{ Kbps}$$

Using Nyquist theorem,

$$\text{Maximum data rate in bits/sec} = 2 * H * \log_2 V = 2 * 3 * \log_2 2 = 6 \text{ Kbp}$$

The bottleneck is therefore the Nyquist limit, giving a maximum channel capacity of 6 Kbps.

1. Compute the Fourier coefficients for the function $f(t) = t$ ($0 \leq t \leq 1$).

1. Compute the Fourier coefficients for the function $f(t) = t$ ($0 \leq t \leq 1$).

ANS:

$$a_n = \frac{2}{T} \int_0^T g(t) \sin(2\pi n ft) dt = 2 \int_0^1 t \sin(2\pi n ft) dt \Rightarrow \text{Assume that } 2\pi n t = a \Rightarrow$$

$$2 \int_0^1 \frac{a}{2\pi n} \sin(a) \frac{da}{2\pi n} = \frac{1}{2\pi^2 n^2} \int_0^1 a \sin a da = \frac{1}{2\pi^2 n^2} \int_0^1 x \sin x dx$$

$$x = u \Rightarrow dx = du$$

$$\sin x dx = dv \Rightarrow -\cos x = v$$

$$\frac{1}{2\pi^2 n^2} \int_0^1 (x(-\cos x) - \int_0^1 -\cos x dx) dx = \frac{1}{2\pi^2 n^2} \int_0^1 (-\cos x * x + \sin x) dx$$

$$= \frac{1}{2\pi^2 n^2} \int_0^1 (-\cos(2\pi n t) * 2\pi n t + \sin(2\pi n t)) dt = \frac{-2\pi n}{2\pi^2 n^2} t \Big|_0^1 = -\frac{1}{\pi n}$$

$$b_n = \frac{2}{T} \int_0^T g(t) \cos(2\pi n ft) dt = 2 \int_0^1 t \cos(2\pi n ft) dt = 0$$

$$c = \frac{2}{T} \int_0^T g(t) dt = 2 \int_0^1 t dt = 2 \frac{t^2}{2} \Big|_0^1 = 1$$

2.2 Guided Transmission Media

Guided Transmission Media

- Media are roughly grouped into
- **Guided media**, such as copper wire and fiber optics
- **Unguided media**, such as radio and lasers through the air
- Guided Transmission media
 - Magnetic Media
 - Twisted Pair
 - Coaxial Cable
 - **Fiber Optics**

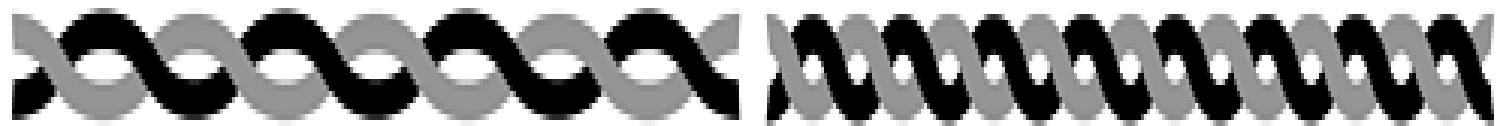
Magnetic Media

- One of the most common ways to transport data from one computer to another is to write them onto magnetic tape or removable media (e.g., recordable DVDs), physically transport the tape or disks to the destination machine, and read them back in again.
- Although this method is not as sophisticated as using a geosynchronous communication satellite, it is often more cost effective, especially for applications in which high bandwidth or cost per bit transported is the key factor.

Twisted Pair

- Although the bandwidth characteristics of magnetic tape are excellent, the delay characteristics are poor. Transmission time is measured in minutes or hours, not milliseconds.
- For many applications an on-line connection is needed.
- One of the oldest and still most common transmission media is twisted pair.
- A twisted pair consists of two insulated copper wires, typically about 1 mm thick. The wires are twisted together in a helical form, just like a DNA molecule.
- Twisting is done because two parallel wires constitute a fine antenna. When the wires are twisted, **the waves from different twists cancel out, so the wire radiates less effectively.**

Figure. (a) Category 3 UTP. (b) Category 5 UTP.



(a)

(b)

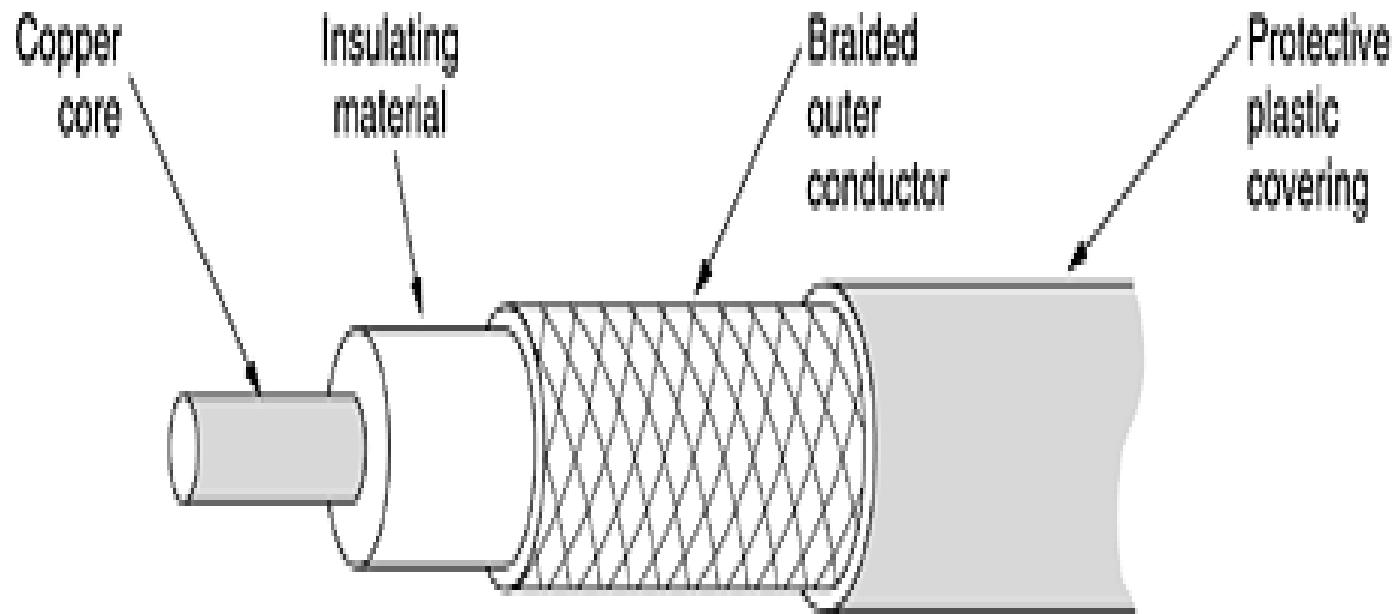
- The most common application of the twisted pair is the **telephone system**. Nearly all telephones are connected to the telephone company (telco) office by a twisted pair.
- Twisted pairs can run several kilometers without amplification, but for longer distances, repeaters are needed.
- When many twisted pairs run in parallel for a substantial distance, such as all the wires coming from an apartment building to the telephone company office, they are bundled together and encased in a protective sheath.
- The pairs in these bundles would interfere with one another if it were not for the twisting.
- **Twisted pairs can be used for transmitting either analog or digital signals.**
- **The bandwidth depends on the thickness of the wire and the distance traveled**, but several megabits/sec can be achieved for a few kilometers in many cases.
- Due to their adequate performance and low cost, twisted pairs are widely used.

- Twisted pair cabling comes in several varieties, two of which are important for **computer networks**. All of these wiring types are often referred to as **UTP (Unshielded Twisted Pair)**
- **Category 3 UTP - up to 16 MHz bandwidth.**
- **Category 5 UTP - up to 100 MHz bandwidth.**
- Category 3 twisted pairs consist of two insulated wires gently twisted together. Four such pairs are typically grouped in a plastic sheath to protect the wires and keep them together.
- This scheme allowed up to four regular telephones or two multiline telephones in each office to connect to the telephone company equipment in the wiring closet.
- Starting around 1988, the more advanced category 5 twisted pairs were introduced.
- They are similar to category 3 pairs, but with **more twists per centimeter**, which results in **less crosstalk** and a **better-quality signal over longer distances**, making them more suitable for high-speed computer communication.
- Up-and-coming categories are **6** and **7**, which are capable of handling signals with bandwidths of **250 MHz** and **600 MHz**, respectively.

Coaxial Cable

- Another common transmission medium is the coaxial cable (known as just "coax" and pronounced "co-ax").
- **It has better shielding than twisted pairs, so it can span longer distances at higher speeds.**
- Two kinds of coaxial cable are widely used. One kind, **50-ohm cable**, is commonly used when it is intended for **digital transmission from the start**.
- The other kind, **75-ohm cable**, is commonly used for **analog transmission** and **cable television** but is becoming more important with the **Internet over cable**.
- **A coaxial cable consists of a stiff copper wire as the core, surrounded by an insulating material. The insulator is encased by a cylindrical conductor, often as a closely-woven braided mesh. The outer conductor is covered in a protective plastic sheath.**

- **Figure. Cutaway view of a coaxial cable**



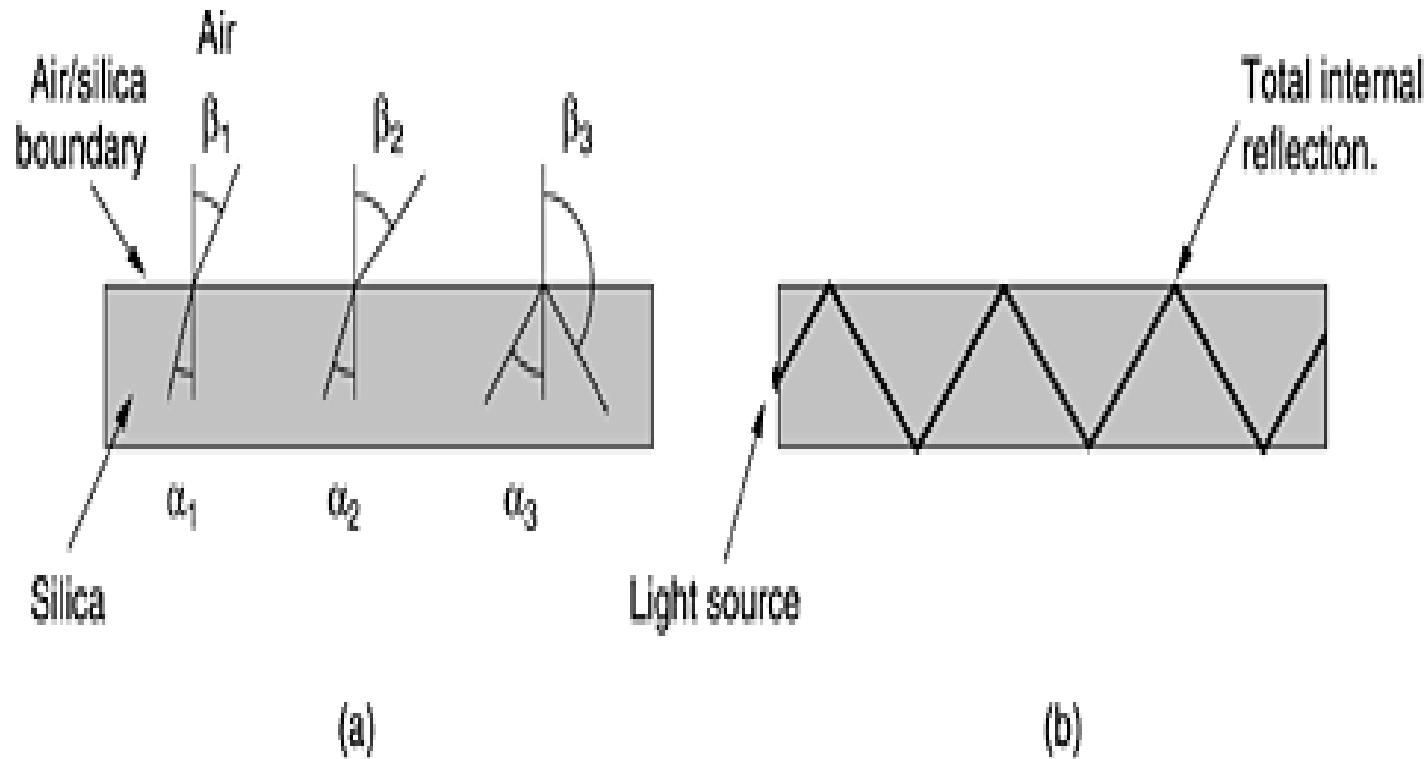
- The construction and shielding of the coaxial cable give it a good combination of **high bandwidth** and **excellent noise immunity**.
- The **bandwidth possible depends on the cable quality, length, and signal-to-noise ratio of the data signal**.
- Modern cables have a bandwidth of close to **1 GHz**.
- Coaxial cables used to be widely used within the **telephone system for long-distance lines** but have now largely been replaced by fiber optics on long-haul routes.
- Coax is still widely used for cable television and metropolitan area networks, however.

Fiber Optics

- In the last 20 years computing speed has increased by a factor of 20 for each decade (IBM PC in 81 ran at 4.77 MHz => 2GHz in 2001).
- In the same period, wide area data communication went from 56 kbps (the ARPANET) to 1 Gbps (modern optical communication), a gain of more than a factor of 125 per decade.
- Moreover the error rate has gone from 10^{-5} per bit to \sim zero in optical networks.
- Semiconductors are close to their physical limit.
- In contrast, **with current fiber technology, the achievable bandwidth is certainly in excess of 50,000 Gbps (50 Tbps)** .
- The current practical signaling limit of about 10 Gbps is due to our inability to convert between electrical and optical signals any faster, although in the laboratory, 100 Gbps has been achieved on a single fiber.

- An optical transmission system has three key components: **the light source, the transmission medium, and the detector.**
- Conventionally, a pulse of light indicates a 1 bit and the absence of light indicates a 0 bit. The transmission medium is an ultra-thin fiber of glass.
- The detector generates an electrical pulse when light falls on it.
- By attaching a light source to one end of an optical fiber and a detector to the other, we have a unidirectional data transmission system that accepts an electrical signal, converts and transmits it by light pulses, and then reconverts the output to an electrical signal at the receiving end.

- Figure 2-5. (a) Three examples of a light ray from inside a silica fiber impinging on the air/silica boundary at different angles. (b) Light trapped by total internal reflection.



- This transmission system would leak light and be useless in practice except for an interesting principle of physics.
- When a light ray passes from one medium to another, for example, from fused silica to air, the ray is refracted (bent) at the silica/air boundary, as shown in Fig (a).
- Here we see a light ray incident on the boundary at an angle a_1 emerging at an angle b_1 . **The amount of refraction depends on the properties of the two media (in particular, their indices of refraction).**
- **For angles of incidence above a certain critical value, the light is refracted back into the silica; none of it escapes into the air.**
- **Thus, a light ray incident at or above the critical angle is trapped inside the fiber, as shown in Fig. (b), and can propagate for many kilometers with virtually no loss.**

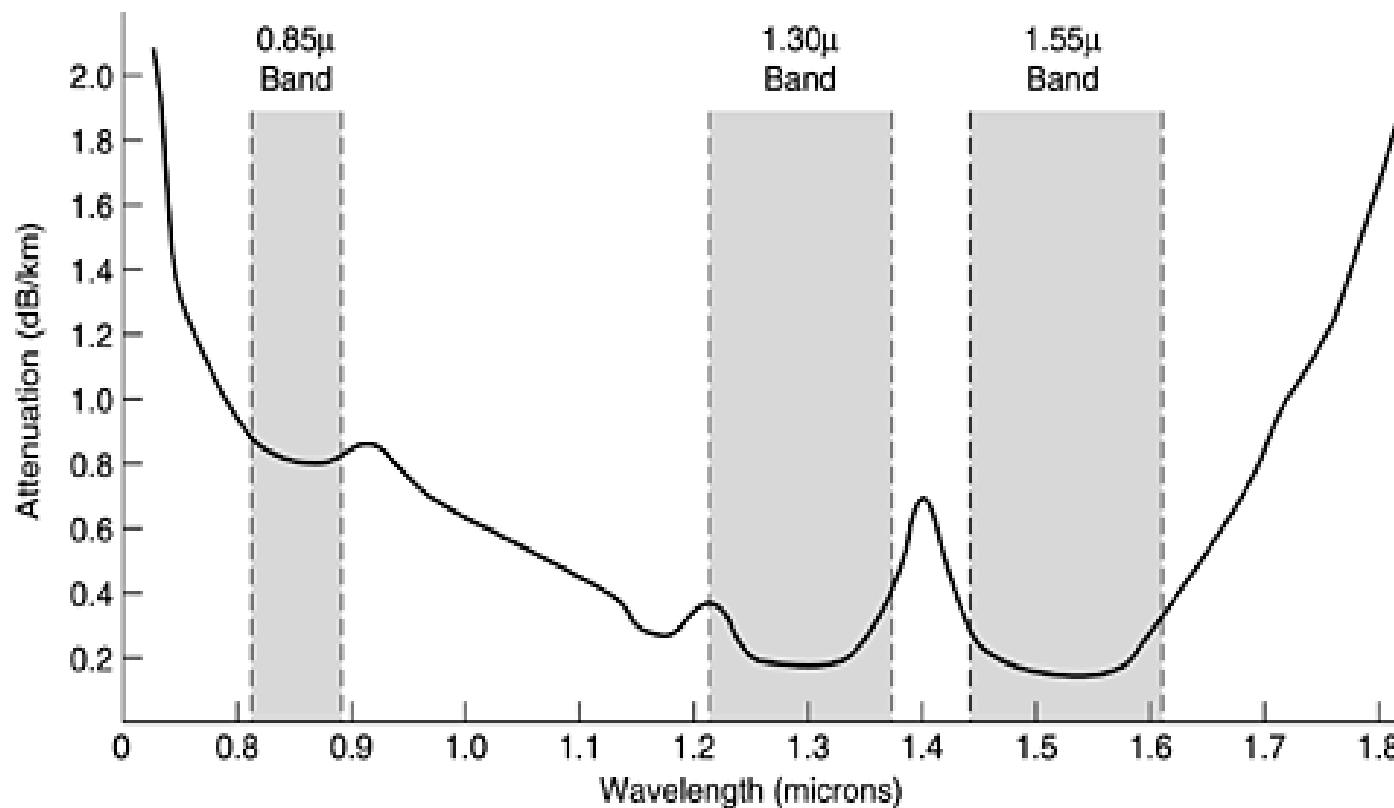
- The sketch of Fig. (b) shows only one trapped ray, but since any light ray incident on the boundary above the critical angle will be reflected internally, many different rays will be bouncing around at different angles.
- Each ray is said to have a different mode, so a fiber having this property is called a **multimode** fiber.
- However, if the fiber's diameter is reduced to a few wavelengths of light, the fiber acts like a wave guide, and the light can propagate only in a straight line, without bouncing, yielding a **single-mode** fiber.
- Single-mode fibers are more expensive but are widely used for longer distances.
- Currently available single-mode fibers can transmit data at **50 Gbps for 100 km without amplification**. Even higher data rates have been achieved in the laboratory for shorter distances.

Transmission of Light through Fiber

- Optical fibers are made of glass, which, in turn, is made from sand, an inexpensive raw material available in unlimited amounts.
- The attenuation of light through glass depends on the wavelength of the light (as well as on some physical properties of the glass).

$$\text{Attenuation in decibels} = 10 \log_{10} \frac{\text{transmitted power}}{\text{received power}}$$

Attenuation of light through fiber in the infrared region.



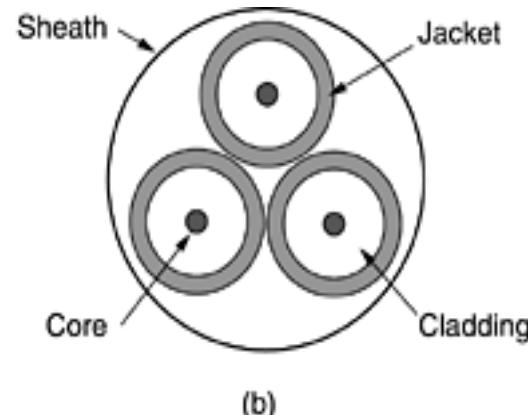
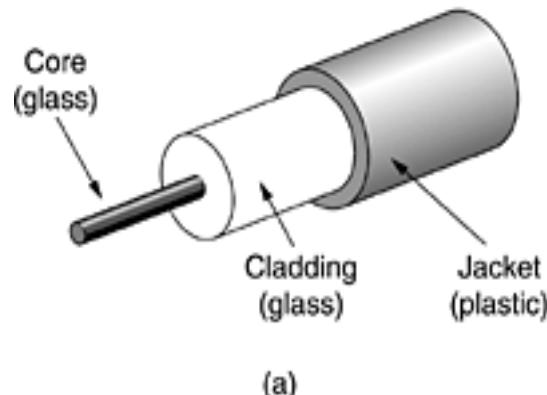
- Visible light has slightly shorter wavelengths, from 0.4 to 0.7 microns (1 micron is 10^{-6} meters).
- **Three wavelength bands are used for optical communication. They are centered at 0.85, 1.30, and 1.55 microns, respectively.**
- The last two have good attenuation properties (less than 5 percent loss per kilometer).
- The 0.85 micron band has higher attenuation, but at that wavelength the lasers and electronics can be made from the same material (gallium arsenide).
- All three bands are 25,000 to 30,000 GHz wide.

- Light pulses sent down a fiber spread out in length as they propagate. This spreading is called **chromatic dispersion**.
- The amount of it is wavelength dependent. One way to keep these spread-out pulses from overlapping is to increase the distance between them, but this can be done only by reducing the signaling rate.
- Fortunately, it has been discovered that by making the pulses in a special shape related to the **reciprocal of the hyperbolic cosine**, nearly all the dispersion effects cancel out, and it is possible to send pulses for thousands of kilometers without appreciable shape distortion.
- These pulses are called **solitons**.
- A considerable amount of research is going on to take solitons out of the lab and into the field.

Fiber Cables

- Fiber optic cables are similar to coax, except without the braid.
- At the center is the glass core through which the light propagates. In multimode fibers, the core is typically **50 microns in diameter**, about the thickness of a human hair.
- In single-mode fibers, the core is **8 to 10 microns**.
- The **core** is surrounded by a glass **cladding** with a lower index of refraction than the core, to keep all the light in the core. Next comes a thin **plastic jacket** to protect the cladding.
- Fibers are typically grouped in bundles, protected by an outer sheath. Figure 2-7(b) shows a sheath with three fibers.

(a) Side view of a single fiber. (b) End view of a sheath with three fibers.



- Fibers can be connected in three different ways.
- First, they can terminate in connectors and be **plugged into fiber sockets**. Connectors lose about **10 to 20** percent of the light, but they make it easy to reconfigure systems.
- Second, they can be **spliced mechanically**. Mechanical splices take trained personnel about 5 minutes and result in a **10 percent light loss**.
- Third, **two pieces of fiber can be fused (melted)** to form a solid connection. A fusion splice is almost as good as a single drawn fiber, but even here, a **small amount of attenuation occurs**.

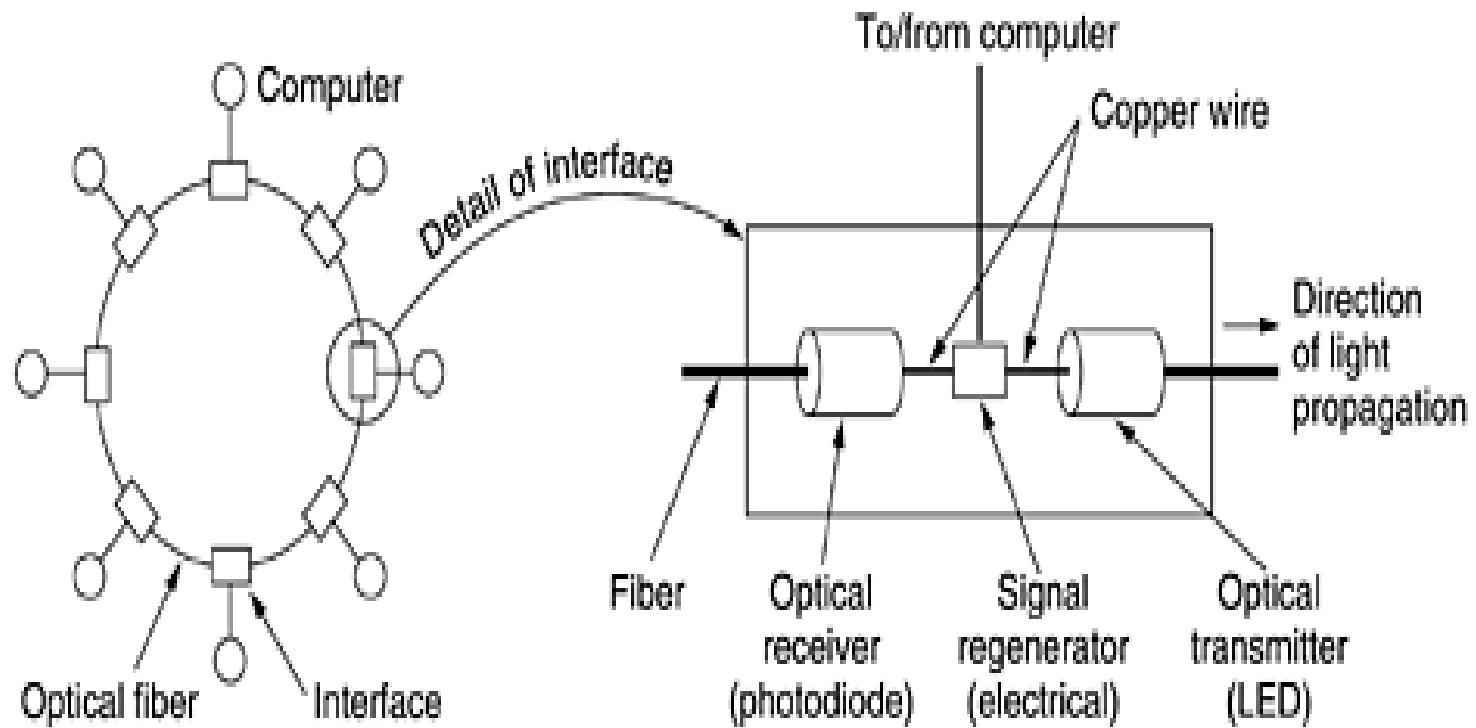
- For all three kinds of splices, reflections can occur at the point of the splice, and the reflected energy can interfere with the signal.
- Two kinds of light sources are typically used to do the signaling, **LEDs (Light Emitting Diodes) and semiconductor lasers**.
- They can be tuned in wavelength by inserting **Fabry-Perot or Mach-Zehnder interferometers** between the source and the fiber.
 - A comparison of semiconductor lasers and LEDs as light sources.

Item	LED	Semiconductor laser
Data rate	Low	High
Fiber type	Multimode	Multimode or single mode
Distance	Short	Long
Lifetime	Long life	Short life
Temperature sensitivity	Minor	Substantial
Cost	Low cost	Expensive

- The receiving end of an optical fiber consists of a **photodiode**, which gives off an electrical pulse when struck by light.
- **The typical response time of a photodiode is 1 nsec, which limits data rates to about 1 Gbps.**
- Thermal noise is also an issue, so a pulse of light must carry enough energy to be detected.
- By making the pulses powerful enough, the error rate can be made arbitrarily small.

Fiber Optic Networks

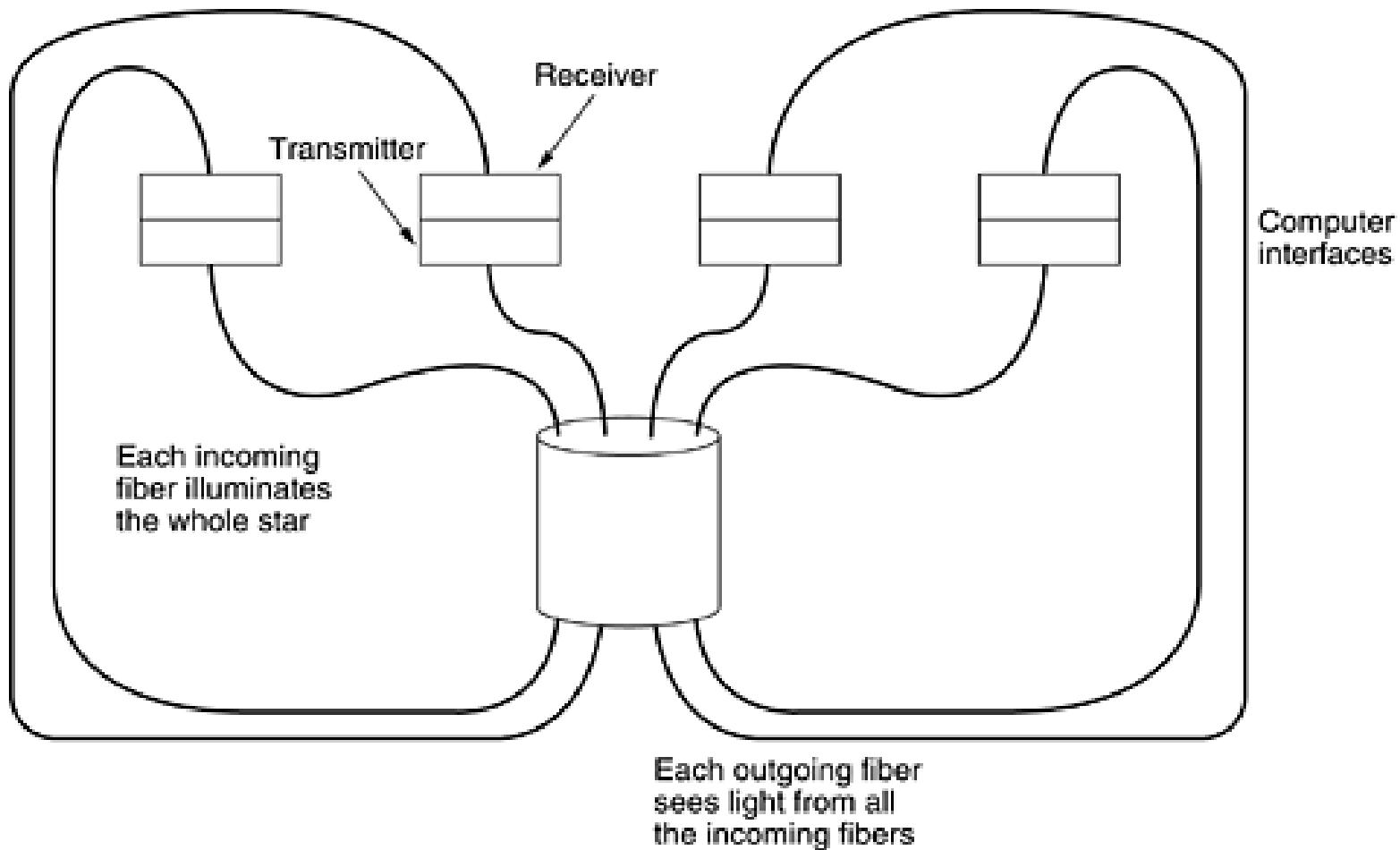
- A fiber optic ring with active repeaters.



- Fiber optics can be used for LANs as well as for long-haul transmission.
- One way around the problem is to realize that a ring network is really just a collection of point-to-point links.
- The interface at each computer passes the light pulse stream through to the next link and also serves as a T junction to allow the computer to send and accept messages.
- Two types of interfaces are used.
 - A **passive interface** consists of two taps fused onto the main fiber. One tap has an LED or laser diode at the end of it (for transmitting), and the other has a photodiode (for receiving).
 - The tap itself is completely passive and is thus **extremely reliable** because a broken LED or photodiode does not break the ring. It just takes one computer off-line.
- The other interface type, is the **active repeater**. The incoming light is converted to an electrical signal, regenerated to full strength if it has been weakened, and retransmitted as light.
 - The interface with the computer is an ordinary copper wire that comes into the signal regenerator.
 - Purely optical repeaters are now being used, too. These devices do not require the optical to electrical to optical conversions.

- If an active repeater fails, the ring is broken and the network goes down.
- The passive interfaces lose light at each junction, so the number of computers and total ring length are greatly restricted.
- A ring topology is not the only way to build a LAN using fiber optics.
- It is also possible to have hardware broadcasting by using the **passive star construction**.
- In this design, each interface has a fiber running from its transmitter to a silica cylinder, with the incoming fibers fused to one end of the cylinder.
- Similarly, fibers fused to the other end of the cylinder are run to each of the receivers.
- Whenever an interface emits a light pulse, it is diffused inside the passive star to illuminate all the receivers, thus achieving broadcast.
- In effect, the passive star combines all the incoming signals and transmits the merged result on all lines.
- Since the incoming energy is divided among all the outgoing lines, the number of nodes in the network is limited by the sensitivity of the photodiodes.

Figure. A passive star connection in a fiber optics network.



Comparison of Fiber Optics and Copper Wire

- Fiber optics can handle much higher bandwidths than copper.
- Due to the low attenuation in fiber optics , repeaters are needed only about every 50 km on long lines, versus about every 5 km for copper, a substantial cost saving.
- Fiber also has the advantage of not being affected by power surges, electromagnetic interference, or power failures. Nor is it affected by corrosive chemicals in the air, making it ideal for harsh factory environments.
- Fiber optic is thin and lightweight.

- Also, fiber is much lighter than copper, which greatly reduces the need for expensive mechanical support systems that must be maintained.
- For new routes, fiber wins hands down due to its much lower installation cost.
- Fibers do not leak light and are quite difficult to tap. These properties give fiber excellent security.
- **On the downside**, fiber is a less familiar technology requiring skills not all engineers have, and fibers can be damaged easily by being bent too much.
- Since optical transmission is inherently unidirectional, two-way communication requires either two fibers or two frequency bands on one fiber.
- Finally, fiber interfaces cost more than electrical interfaces.

2.3 Wireless Transmission

Wireless Transmission

- For mobile users, wireless transmission is vital.
- Some people believe that the future holds only two kinds of communication: fiber and wireless.
- All fixed (i.e., nonmobile) computers, telephones, faxes, and so on will use fiber, and all mobile ones will use wireless.
- Wireless has advantages for even fixed devices in some circumstances.

2.3.1 The Electromagnetic Spectrum

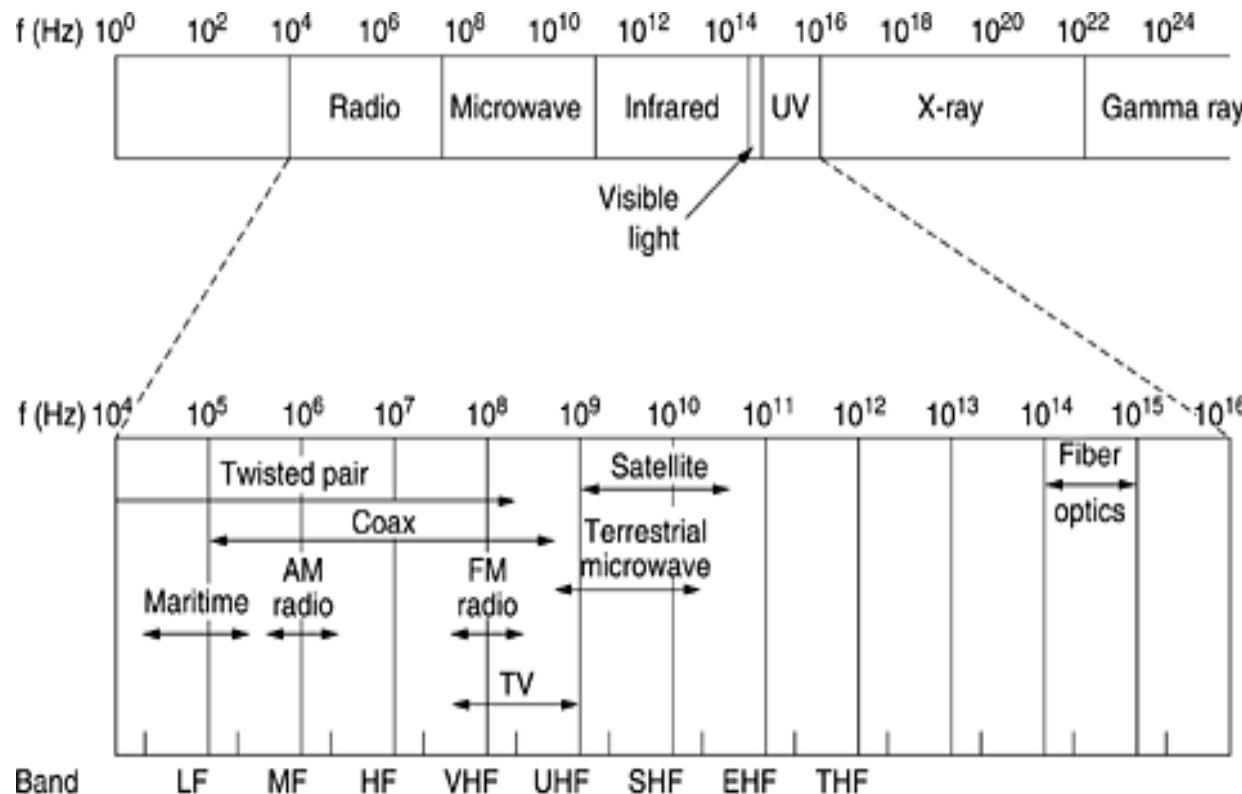
- When electrons move, they create electromagnetic waves that can propagate through space (even in a vacuum).
- The number of oscillations per second of a wave is called its frequency, f , and is measured in Hz.
- The distance between two consecutive maxima (or minima) is called the wavelength, which is universally designated by the Greek letter λ (lambda).
- When an antenna of the appropriate size is attached to an electrical circuit, the electromagnetic waves can be broadcast efficiently and received by a receiver some distance away.
- All wireless communication is based on this principle.

- In vacuum, all electromagnetic waves travel at the same speed, no matter what their frequency. This speed, usually called the speed of light, c , is approximately 3×10^8 m/sec, or about 1 foot (30 cm) per nanosecond.
- In copper or fiber the speed slows to about 2/3 of this value and becomes slightly frequency dependent. The speed of light is the ultimate speed limit. No object or signal can ever move faster than it.
- The fundamental relation between f , λ , and c (in vacuum) is

$$\lambda f = c$$

- Since c is a constant, if we know f , we can find λ , and vice versa.
- For example, 100-MHz waves are about 3 meters long, 1000-MHz waves are 0.3-meters long, and 0.1-meter waves have a frequency of 3000 MHz.

- **Figure: The electromagnetic spectrum and its uses for communication.**



- The radio, microwave, infrared, and visible light portions of the spectrum can all be used for transmitting information by modulating the amplitude, frequency, or phase of the waves.
- Ultraviolet light, X-rays, and gamma rays would be even better, due to their higher frequencies, but they are hard to produce and modulate, do not propagate well through buildings, and are dangerous to living things.
- The bands listed at the bottom of figure are the official ITU (International Telecommunication Union) names and are based on the wavelengths, so the LF band goes from 1 km to 10 km (approximately 30 kHz to 300 kHz).
- The terms LF, MF, and HF refer to low, medium, and high frequency, respectively.
- The higher bands were later named the Very, Ultra, Super, Extremely, and Tremendously High Frequency bands.
- Beyond that there are no names, but Incredibly, Astonishingly, and Prodigiously high frequency (IHF, AHF, and PHF) are used.

- The amount of information that an electromagnetic wave can carry is related to its bandwidth.
- It should now be obvious why networking people like fiber optics so much.
- If we solve $\lambda f = c$ for f and differentiate with respect to λ , and consider the absolute value

$$\Delta f = \frac{c \Delta \lambda}{\lambda^2}$$

- Thus, given the width of a wavelength band, $\Delta\lambda$, we can compute the corresponding frequency band, Δf , and from that the data rate the band can produce.

- The wider the band, the higher the data rate.
- As an example, consider the 1.30-micron band. Here we have $\lambda=1.3 \times 10^{-6}$ and $\Delta\lambda = 0.17 \times 10^{-6}$, so Δf is about 30 THz. At, say, 8 bits/Hz, we get 240 Tbps. (1 Tb= 10^{24})
- Most transmissions use a narrow frequency band (i.e., $\Delta f/f \ll 1$) to get the best reception (many watts/Hz).
- However, in some cases, a wide band is used.

2.3.2 Radio Transmission

- Radio waves are easy to generate, can travel long distances, and can penetrate buildings easily, so they are widely used for communication, both indoors and outdoors.
- Radio waves also are omnidirectional, meaning that they travel in all directions from the source, so the transmitter and receiver do not have to be carefully aligned physically.
- The properties of radio waves are frequency dependent. At low frequencies, radio waves pass through obstacles well, but the power falls off sharply with distance from the source, roughly as $1/r^2$ in air.
- At high frequencies, radio waves tend to travel in straight lines and bounce off obstacles. They are also absorbed by rain.
- At all frequencies, radio waves are subject to interference from motors and other electrical equipment.

2.3.3 Microwave Transmission

- Above 100 MHz, the waves travel in nearly straight lines and can therefore be narrowly focused.
- Concentrating all the energy into a small beam by means of a parabolic antenna (like the familiar satellite TV dish) gives a much higher signal-to-noise ratio, but the transmitting and receiving antennas must be accurately aligned with each other.
- Unlike radio waves at lower frequencies, microwaves do not pass through buildings well.
- Microwave communication is so widely used for long-distance telephone communication, mobile phones, television distribution, and other uses.
- Microwave is also relatively inexpensive.

2.3.4 Infrared and Millimeter Waves

- Unguided infrared and millimeter waves are widely used for short-range communication.
- The remote controls used on televisions, VCRs, and stereos all use infrared communication.
- They are relatively directional, cheap, and easy to build but have a major drawback: they do not pass through solid objects.
- On the other hand, the fact that infrared waves do not pass through solid walls well is also a plus. It means that an infrared system in one room of a building will not interfere with a similar system in adjacent rooms or buildings.
- Security of infrared systems is better than that of radio systems precisely for this reason.
- Therefore, no government license is needed to operate an infrared system, in contrast to radio systems.
- Infrared communication has a limited use on the desktop, for example, connecting notebook computers and printers, but it is not a major player in the communication game.

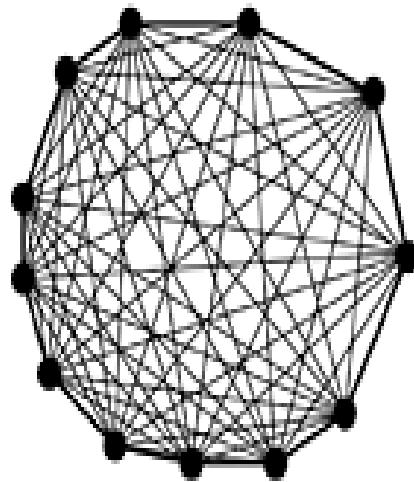
2.5 The Public Switched Telephone Network

The Public Switched Telephone Network

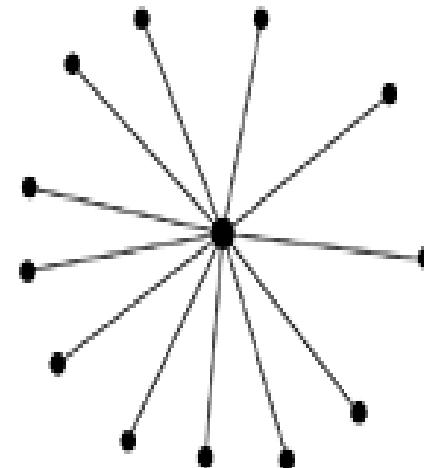
- When two computers owned by the same company or organization and located close to each other need to communicate, it is often easiest just to run a cable between them. LANs work this way.
- However, when the distances are large or there are many computers or the cables have to pass through a public road or other public right of way, the costs of running private cables are usually prohibitive.
- Consequently, the network designers must rely on the existing telecommunication facilities.
- These facilities, especially the PSTN (Public Switched Telephone Network).
- PSTN were usually designed, with a completely different goal in mind: transmitting the human voice in a more-or-less recognizable form.

2.5.1 Structure of the Telephone System

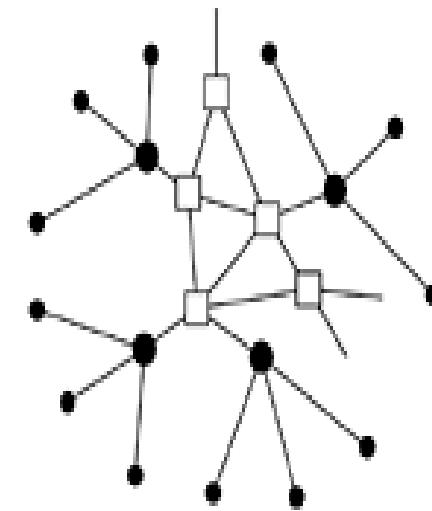
- Figure 2-20. (a) Fully-interconnected network. (b) Centralized switch. (c) Two-level hierarchy.



(a)



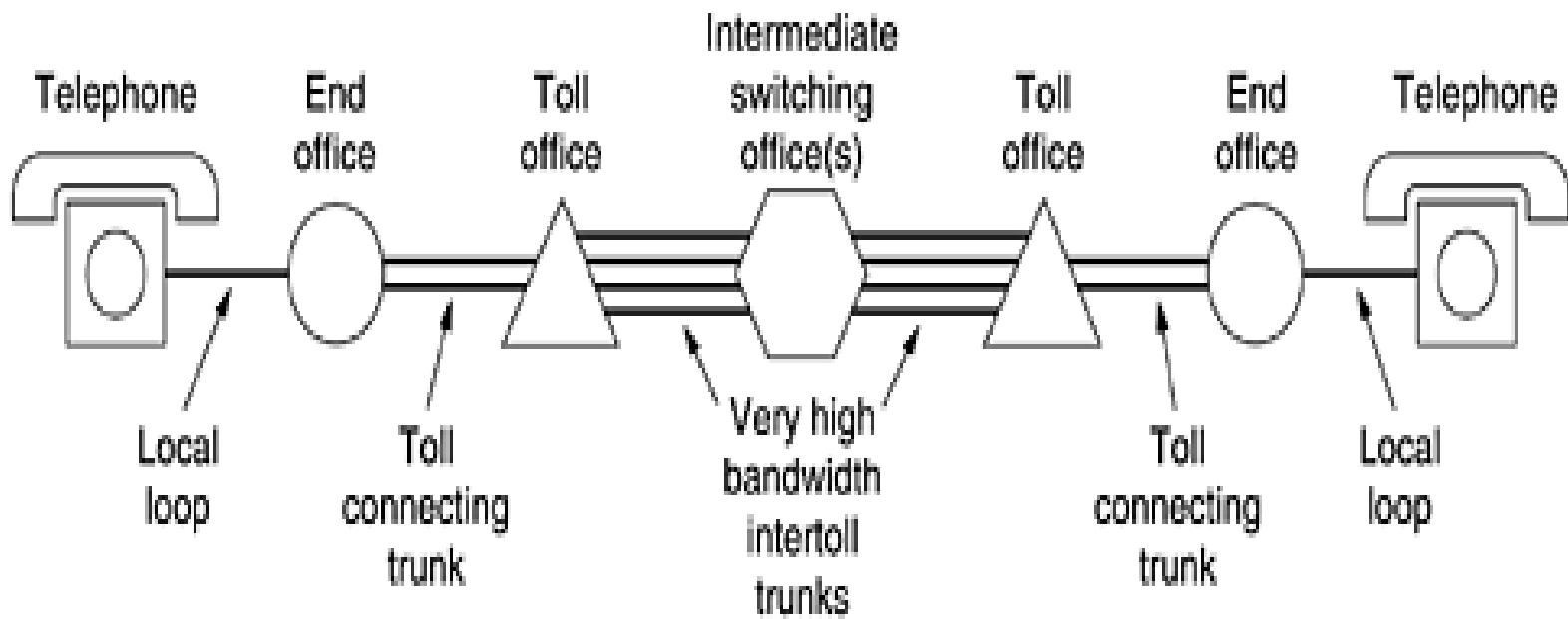
(b)



(c)

- The three major parts of the telephone system are:
 - ❑ The switching offices,
 - ❑ The wires between the customers and the switching offices
 - ❑ The long-distance connections between the switching offices.

- **Figure 2-21. A typical circuit route for a medium-distance call.**



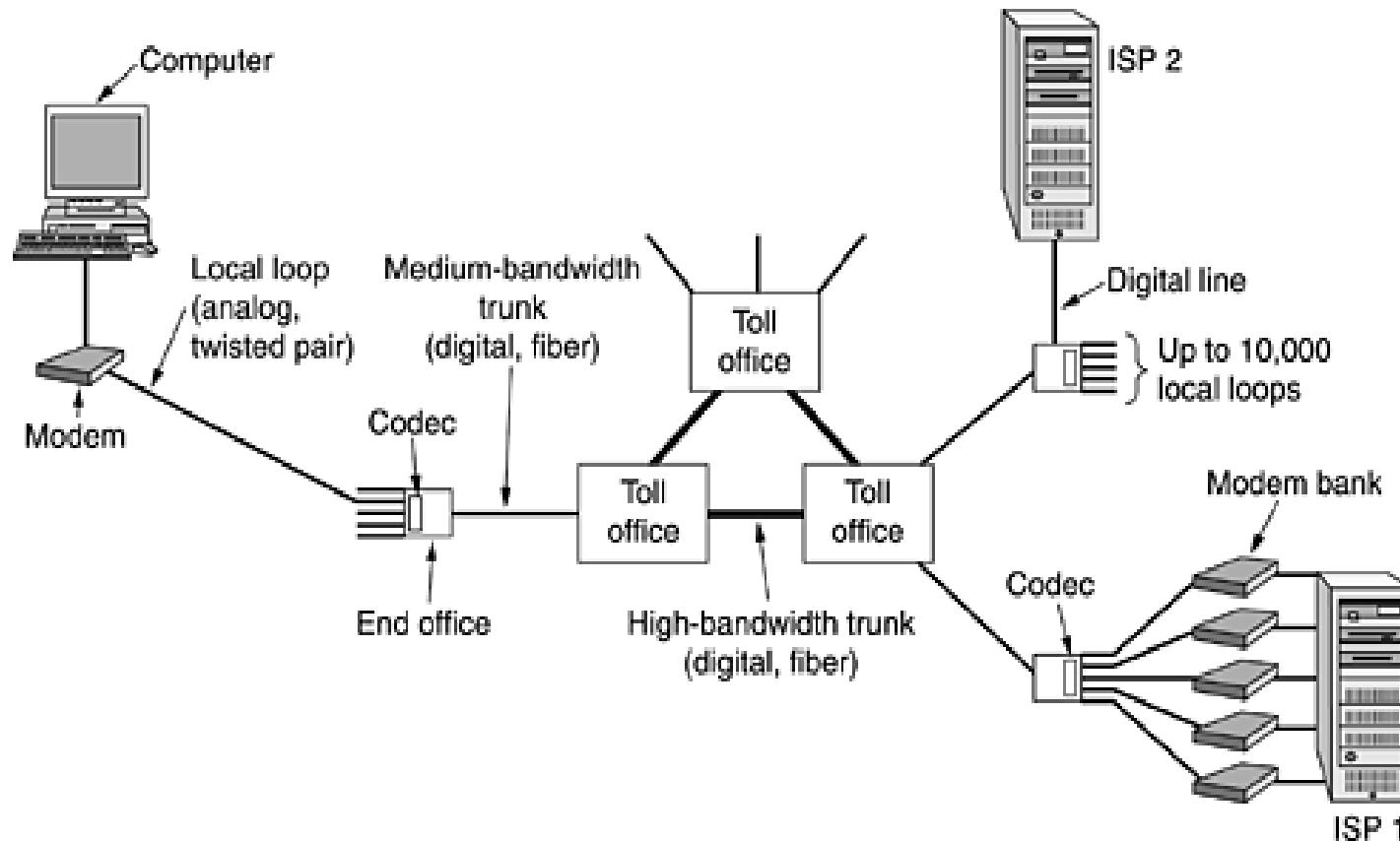
- Each telephone has two copper wires coming out of it that go directly to the telephone company's nearest end office (also called a local central office).
- The two-wire connections between each subscriber's telephone and the end office are known in the trade as the **local loop**.
- If a subscriber attached to a given end office calls another subscriber attached to the same end office, the switching mechanism within the office sets up a direct electrical connection between the two local loops.
- If the called telephone is attached to another end office, a different procedure has to be used. Each end office has a number of outgoing lines to one or more nearby switching centers, called toll offices.
- These lines are called toll connecting trunks. If both the caller's and callee's end offices happen to have a toll connecting trunk to the same toll office (a likely occurrence if they are relatively close by), the connection may be established within the toll office.

- If the caller and callee do not have a toll office in common, the path will have to be established somewhere higher up in the hierarchy.
- Primary, sectional, and regional offices form a network by which the toll offices are connected.
- The toll, primary, sectional, and regional exchanges communicate with each other via high- bandwidth intertoll trunks (also called interoffice trunks).
- A variety of transmission media are used for telecommunication. Local loops consist of twisted pairs nowadays.
- Between switching offices, coaxial cables, microwaves, and especially fiber optics are widely used.
- In the past, transmission throughout the telephone system was analog, with the actual voice signal being transmitted as an electrical voltage from source to destination.
- With the advent of fiber optics, digital electronics, and computers, all the trunks and switches are now digital.

- In summary, the telephone system consists of three major components:
 - Local loops** (analog twisted pairs going into houses and businesses).
 - Trunks** (digital fiber optics connecting the switching offices).
 - Switching offices** (where calls are moved from one trunk to another).
- The local loops provide everyone access to the whole system.
- For the long-haul trunks, the main issue is how to collect multiple calls together and send them out over the same fiber.

2.5.3 The Local Loop: Modems, ADSL, and Wireless

- Figure 2-23. The use of both analog and digital transmission for a computer to computer call. Conversion is done by the modems and codecs.



- An end office has up to 10,000 local loops.
- The area code + exchange indicated the end office.
- Example:(212) 601-xxxx was a specific end office with 10,000 subscribers, numbered 0000 through 9999.
- The two-wire local loop coming from a telephone company end office into houses and small businesses.
- The local loop is also frequently referred to as the "last mile," although the length can be up to several miles.
- It has used analog signaling for over 100 years and is likely to continue doing so for some years to come, due to the high cost of converting to digital.
- When a computer wishes to send digital data over an analog dial-up line, the data must first be converted to analog form for transmission over the local loop. This conversion is done by a device called a modem.

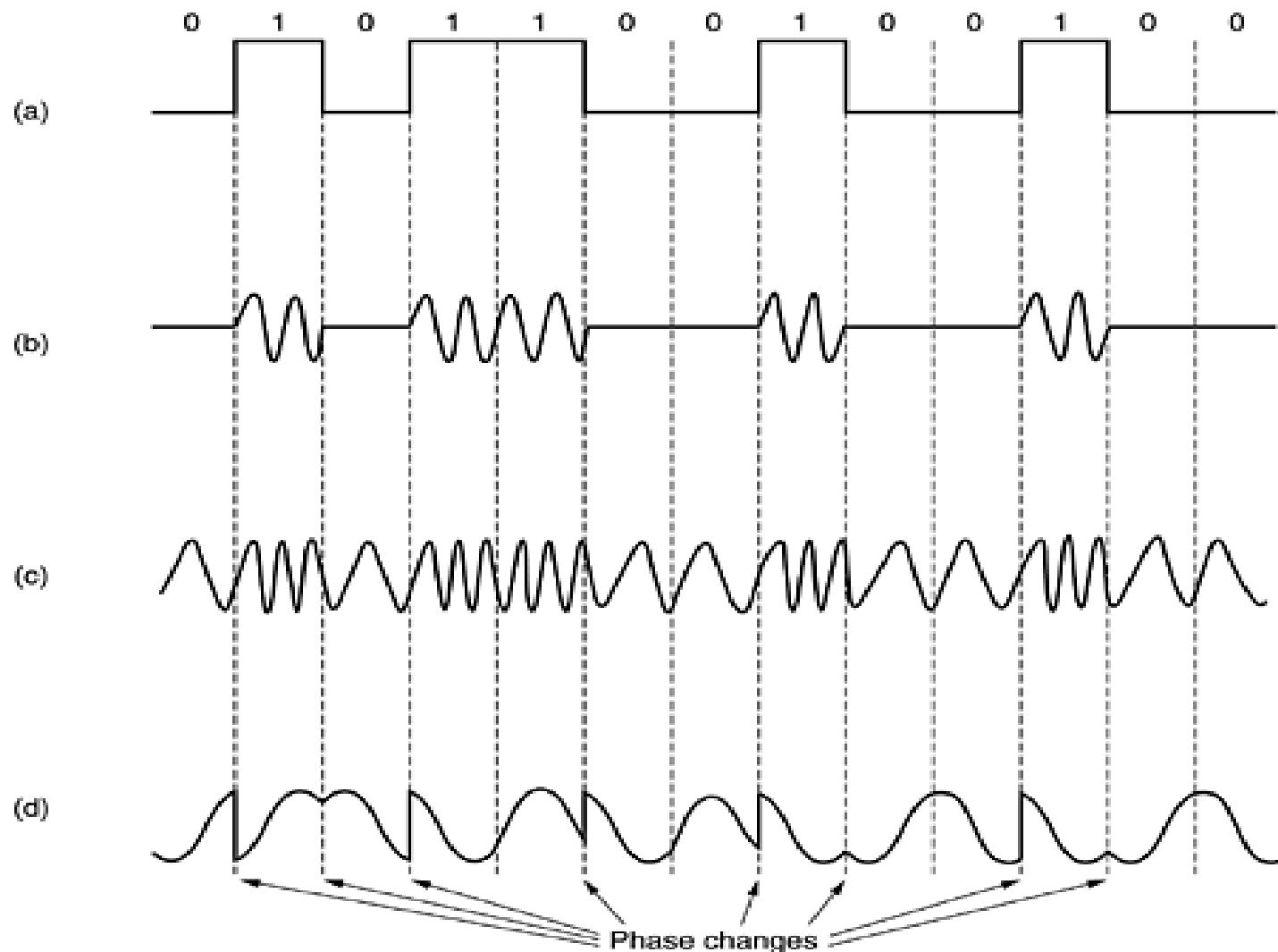
- At the telephone company end office the data are converted to digital form for transmission over the long-haul trunks.
- If the other end is a computer with a modem, the reverse conversion—digital to analog—is needed to traverse the local loop at the destination.
- ISP 1 (Internet Service Provider), which has a bank of modems, each connected to a different local loop.
- This ISP can handle as many connections as it has modems. (assuming its server or servers have enough computing power).
- Analog signaling consists of varying a voltage with time to represent an information stream.
- If transmission media were perfect, the receiver would receive exactly the same signal that the transmitter sent.
- Unfortunately, media are not perfect, so the received signal is not the same as the transmitted signal.
- For digital data, this difference can lead to errors.

- Transmission lines suffer from three major problems: **attenuation, delay distortion, and noise.**
- Attenuation is the loss of energy as the signal propagates outward. The loss is expressed in decibels per kilometer. The amount of energy lost depends on the frequency.
- To see the effect of this frequency dependence, imagine a signal not as a simple waveform, but as a series of Fourier components.
- Each component is attenuated by a different amount, which results in a different Fourier spectrum at the receiver.
- The different Fourier components also propagate at different speeds in the wire. This speed difference leads to distortion of the signal received at the other end.
- Another problem is noise, which is unwanted energy from sources other than the transmitter.
- Thermal noise is caused by the random motion of the electrons in a wire and is unavoidable.
- Crosstalk is caused by inductive coupling between two wires that are close to each other.
- Finally, there is impulse noise, caused by spikes on the power line or other causes. For digital data, impulse noise can wipe out one or more bits.

Modems

- Both attenuation and propagation speed are frequency dependent, it is undesirable to have a wide range of frequencies in the signal.
- Unfortunately, the square waves used in digital signals have a wide frequency spectrum and thus are subject to strong attenuation and delay distortion.
- These effects make baseband (DC) signaling unsuitable except at slow speeds and over short distances.
- To get around the problems associated with DC signaling, especially on telephone lines, AC signaling is used.
- A continuous tone in the 1000 to 2000-Hz range, called a sine wave carrier, is introduced.

- **Figure 2-24. (a) A binary signal. (b) Amplitude modulation. (c) Frequency modulation. (d) Phase modulation.**

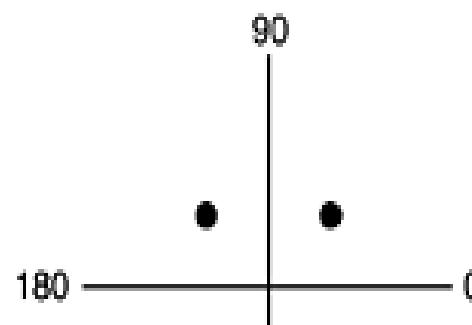


- Its amplitude, frequency, or phase can be modulated to transmit information.
- In amplitude modulation, two different amplitudes are used to represent 0 and 1, respectively.
- In frequency modulation, also known as frequency shift keying, two (or more) different tones are used. (The term keying is also widely used in the industry as a synonym for modulation.)
- In the simplest form of phase modulation, the carrier wave is systematically shifted 0 or 180 degrees at uniformly spaced intervals.
- A device that accepts a serial stream of bits as input and produces a carrier modulated by one (or more) of these methods (or vice versa) is called a modem (for modulator-demodulator).
- The modem is inserted between the (digital) computer and the (analog) telephone system.

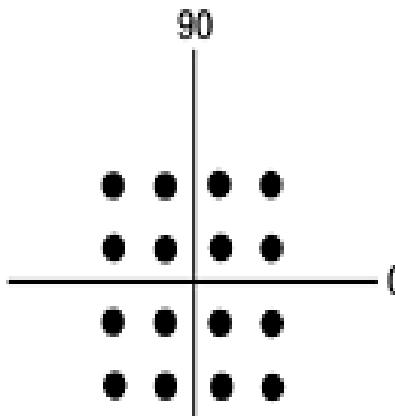
- To go to higher and higher speeds, it is not possible to just keep increasing the sampling rate.
- The **Nyquist theorem** says that even with a perfect 3000-Hz line , there is no point in sampling faster than 6000 Hz.
- **In practice, most modems sample 2400 times/sec and focus on getting more bits per sample.**
- **The number of samples per second is measured in baud.**
- In telecommunication and electronics, baud is a common unit of measurement of symbol rate,
- Baud is one of the components that determine the speed of communication over a data channel.
- During each baud, one symbol is sent. Thus, an n- baud line transmits n symbols/sec.
- For example, a 2400-baud line sends one symbol about every 416.667 μ sec.
- If the symbol consists of 0 volts for a logical 0 and 1 volt for a logical 1, the bit rate is 2400 bps.
- If, however, the voltages 0, 1, 2, and 3 volts are used, every symbol consists of 2 bits, so a 2400-baud line can transmit 2400 symbols/sec at a data rate of 4800 bps.

- Similarly, with four possible phase shifts, there are also 2 bits/symbol, so again here the bit rate is twice the baud rate. This technique is widely used and called **QPSK (Quadrature Phase Shift Keying)**.
- The concepts of **bandwidth**, **baud**, **symbol**, and **bit rate** are commonly confused.
- The **bandwidth** of a medium is the range of frequencies that pass through it with minimum attenuation. It is a physical property of the medium (usually from 0 to some maximum frequency) and measured in Hz.
- The **baud rate** is the number of samples/sec made. Each sample sends one piece of information, that is, **one symbol**. The baud rate and symbol rate are thus the same.
- The modulation technique (e.g., QPSK) determines the number of bits/symbol.
- The bit rate is the amount of information sent over the channel and is equal to the number of symbols/sec times the number of bits/symbol.
- All advanced modems use a combination of modulation techniques to transmit multiple bits per baud.
- Often multiple amplitudes and multiple phase shifts are combined to transmit several bits/symbol.

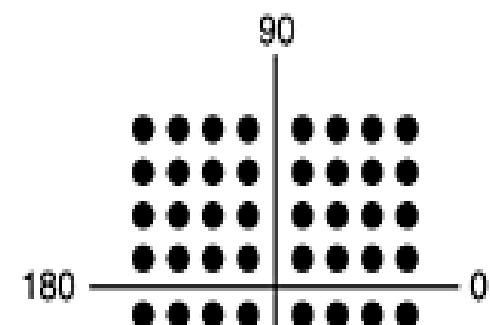
- **Figure 2-25. (a) QPSK. (b) QAM-16. (c) QAM-64.**



(a)



(b)

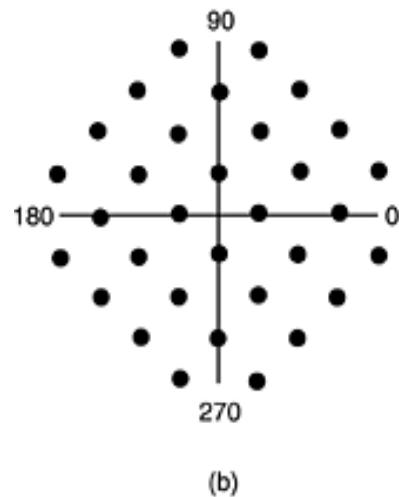


(c)

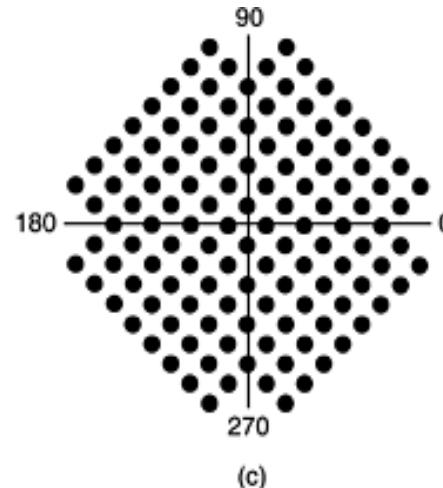
- Fig. 2-25(a) has four valid combinations and can be used to transmit 2 bits per symbol. It is QPSK.
- Fig. 2-25(b) has four amplitudes and four phases are used, for a total of 16 different combinations. This modulation scheme can be used to transmit 4 bits per symbol. It is called QAM-16 (Quadrature Amplitude Modulation).
- Figure 2-25(c) allows 64 different combinations, so 6 bits can be transmitted per symbol. It is called QAM-64. Higher-order QAMs also are used.
- Diagrams such as those of Fig. 2-25, which show the legal combinations of amplitude and phase, are called **constellation diagrams**.

- Each high-speed modem standard has its own constellation pattern and can talk only to other modems that use the same one (although most modems can emulate all the slower ones).
- With many points in the constellation pattern, even a small amount of noise in the detected amplitude or phase can result in an error and, potentially, many bad bits.
- To reduce the chance of an error, standards for the higher speeds modems do error correction by adding extra bits to each sample. The schemes are known as TCM (Trellis Coded Modulation).
- The V.32 modem standard uses 32 constellation points to transmit 4 data bits and 1 parity bit per symbol at 2400 baud to achieve 9600 bps with error correction.
- Its constellation pattern is shown in Fig. 2-26(a). The decision to "rotate" around the origin by 45 degrees was done for engineering reasons; the rotated and unrotated constellations have the same information capacity.
- The next step above 9600 bps is 14,400 bps. It is called V.32 bis. This speed is achieved by transmitting 6 data bits and 1 parity bit per sample at 2400 baud.

- **Figure 2-26. (a) V.32 for 9600 bps. (b) V32 bis for 14,400 bps.**



(b)



(c)

- Its constellation pattern has 128 points when QAM-128 is used and is shown in Fig. 2-26(b).
- Fax modems use this speed to transmit pages. QAM-256 is not used in any standard telephone modems, but it is used on cable networks.
- The next telephone modem after V.32 bis is V.34, which runs at 28,800 bps at 2400 baud with 12 data bits/symbol.
- The final modem in this series is V.34 bis which uses 14 data bits/symbol at 2400 baud to achieve 33,600 bps.

- To increase the effective data rate further, many modems compress the data before transmitting it, to get an effective data rate higher than 33,600 bps.
- All modern modems allow traffic in both directions at the same time (by using different frequencies for different directions).
- A connection that allows traffic in both directions simultaneously is called full duplex. A two-lane road is full duplex.
- A connection that allows traffic either way, but only one way at a time is called half duplex. A single railroad track is half duplex.
- A connection that allows traffic only one way is called simplex. A one-way street is simplex. Another example of a simplex connection is an optical fiber with a laser on one end and a light detector on the other end.

- The reason that standard modems stop at 33,600 is that the Shannon limit for the telephone system is about 35 kbps, so going faster than this would violate the laws of physics.
- whether 56-kbps modems are theoretically possible?
- But why is the theoretical limit 35 kbps? It has to do with the average length of the local loops and the quality of these lines. The 35 kbps is determined by the average length of the local loops.
- A call originating at the computer on the left and terminating at ISP1 goes over two local loops as an analog signal, once at the source and once at the destination. Each of these adds noise to the signal. If we could get rid of one of these local loops, the maximum rate would be doubled.

- ISP 2 does precisely that. It has a pure digital feed from the nearest end office. The digital signal used on the trunks is fed directly to ISP 2, eliminating the codecs, modems, and analog transmission on its end.
- Thus, when one end of the connection is purely digital, as it is with most ISPs now, the maximum data rate can be as high as 70 kbps. Between two home users with modems and analog lines, the maximum is 33.6 kbps.
- The reason that 56 kbps modems are in use has to do with the Nyquist theorem.
- The telephone channel is about 4000 Hz wide (including the guard bands). The maximum number of independent samples per second is thus 8000. The number of bits per sample in the U.S. is 8, one of which is used for control purposes, allowing 56,000 bit/sec of user data.

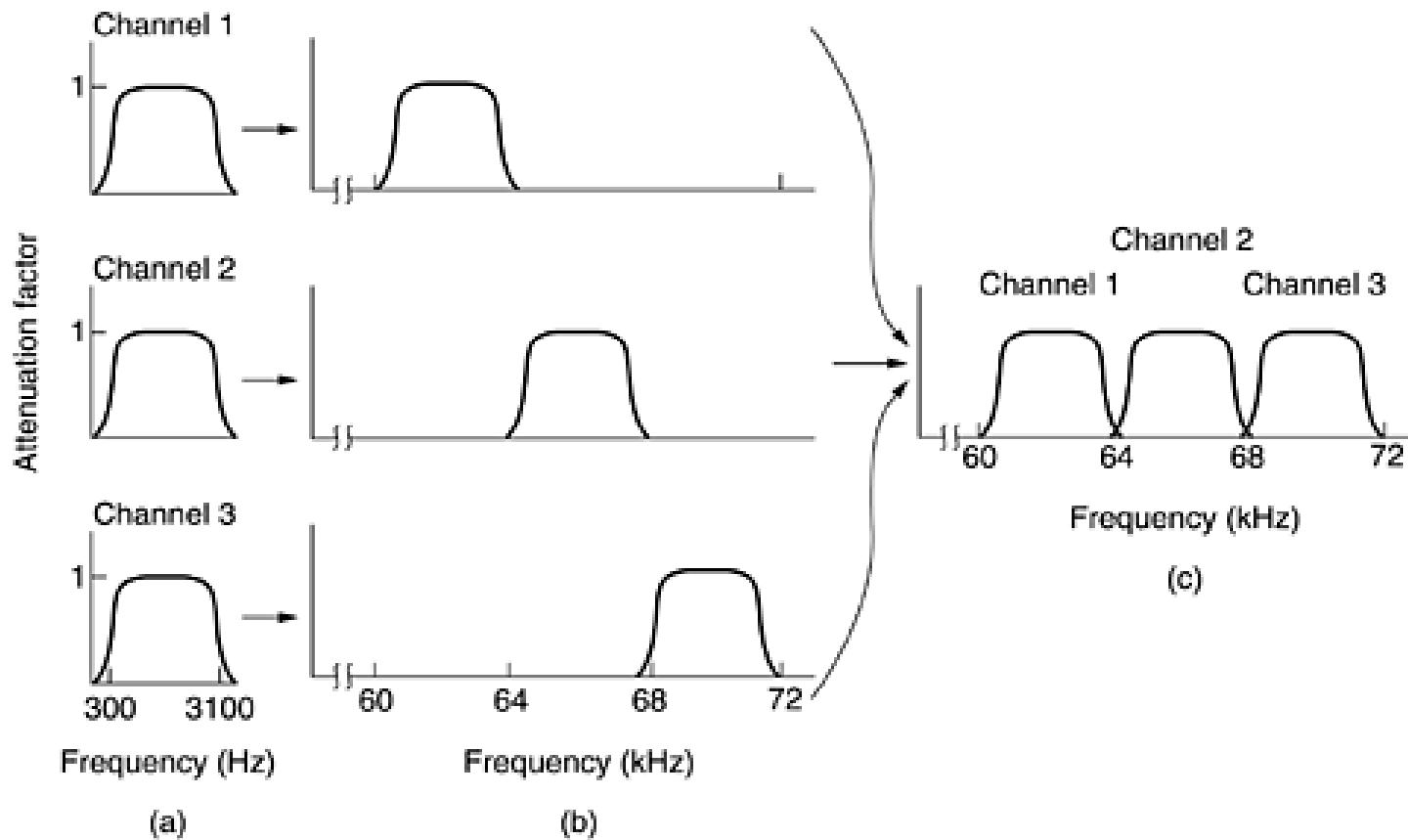
- In Europe, all 8 bits are available to users, so 64,000-bit/sec modems could have been used, but to get international agreement on a standard, 56,000 was chosen.
- This modem standard is called V.90. It provides for a 33.6-kbps upstream channel (user to ISP), but a 56 kbps downstream channel (ISP to user) because there is usually more data transport from the ISP to the user than the other way (e.g., requesting a Web page takes only a few bytes, but the actual page could be megabytes).
- In theory, an upstream channel wider than 33.6 kbps would have been possible, but since many local loops are too noisy for even 33.6 kbps, it was decided to allocate more of the bandwidth to the downstream channel to increase the chances of it actually working at 56 kbps.
- The next step beyond V.90 is V.92. These modems are capable of 48 kbps on the upstream channel if the line can handle it.

Trunks and Multiplexing

- Many conversations over a single physical trunk.
- Multiplexing schemes can be divided into two basic categories: FDM (Frequency Division Multiplexing) and TDM (Time Division Multiplexing).
- In FDM, the frequency spectrum is divided into frequency bands, with each user having exclusive possession of some band.
- In TDM, the users take turns (in a round-robin fashion), each one periodically getting the entire bandwidth for a little burst of time.
- AM radio broadcasting provides illustrations of both kinds of multiplexing.

Frequency Division Multiplexing

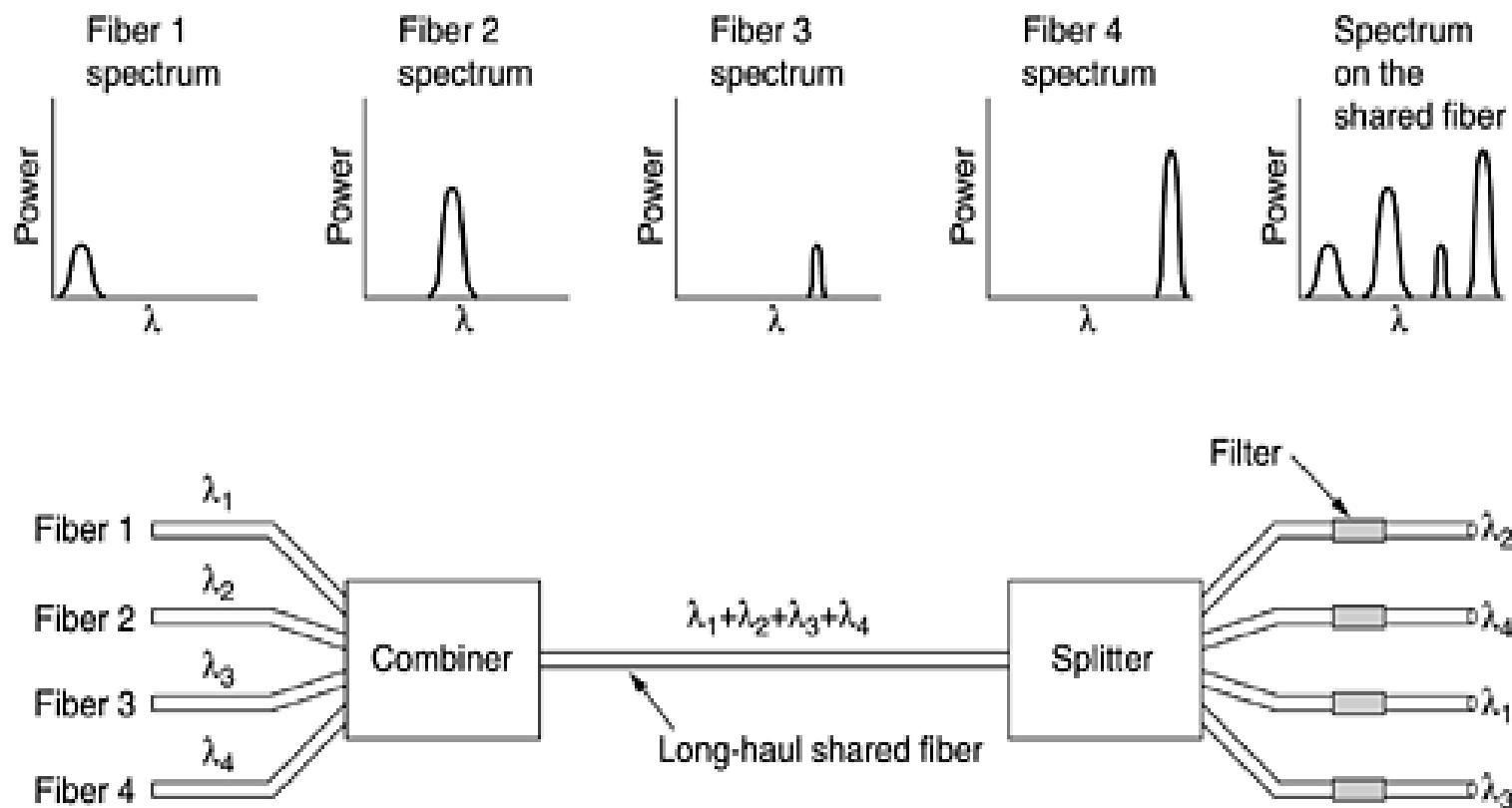
- Figure 2-31. Frequency division multiplexing. (a) The original bandwidths. (b) The bandwidths raised in frequency. (c) The multiplexed channel.



- Three voice-grade telephone channels are multiplexed using FDM. Filters limit the usable bandwidth to about 3100 Hz per voice-grade channel.
- When many channels are multiplexed together, 4000 Hz is allocated to each channel to keep them well separated.
- First the voice channels are raised in frequency, each by a different amount. Then they can be combined because no two channels now occupy the same portion of the spectrum.
- The FDM schemes used around the world are to some degree standardized. A widespread standard is twelve 4000-Hz voice channels multiplexed into the 60 to 108 kHz band. This unit is called a group.
- The 12-kHz to 60- kHz band is sometimes used for another group. Many carriers offer a 48- to 56-kbps leased line service to customers, based on the group.
- Five groups (60 voice channels) can be multiplexed to form a supergroup.
- The next unit is the mastergroup, which is five supergroups (CCITT standard) or ten supergroups (Bell system). Other standards of up to 230,000 voice channels also exist.

Wavelength Division Multiplexing

- Figure 2-32. Wavelength division multiplexing.



- For fiber optic channels, a variation of frequency division multiplexing is used. It is called WDM (Wavelength Division Multiplexing).
- Many fibers come together at an optical combiner, each with its energy present at a different wavelength.
- Many beams are combined onto a single shared fiber for transmission to a distant destination.
- At the far end, the beam is split up over as many fibers as there were on the input side.
- Each output fiber contains a short, specially-constructed core that filters out all but one wavelength.
- The resulting signals can be routed to their destination or recombined in different ways for additional multiplexed transport.

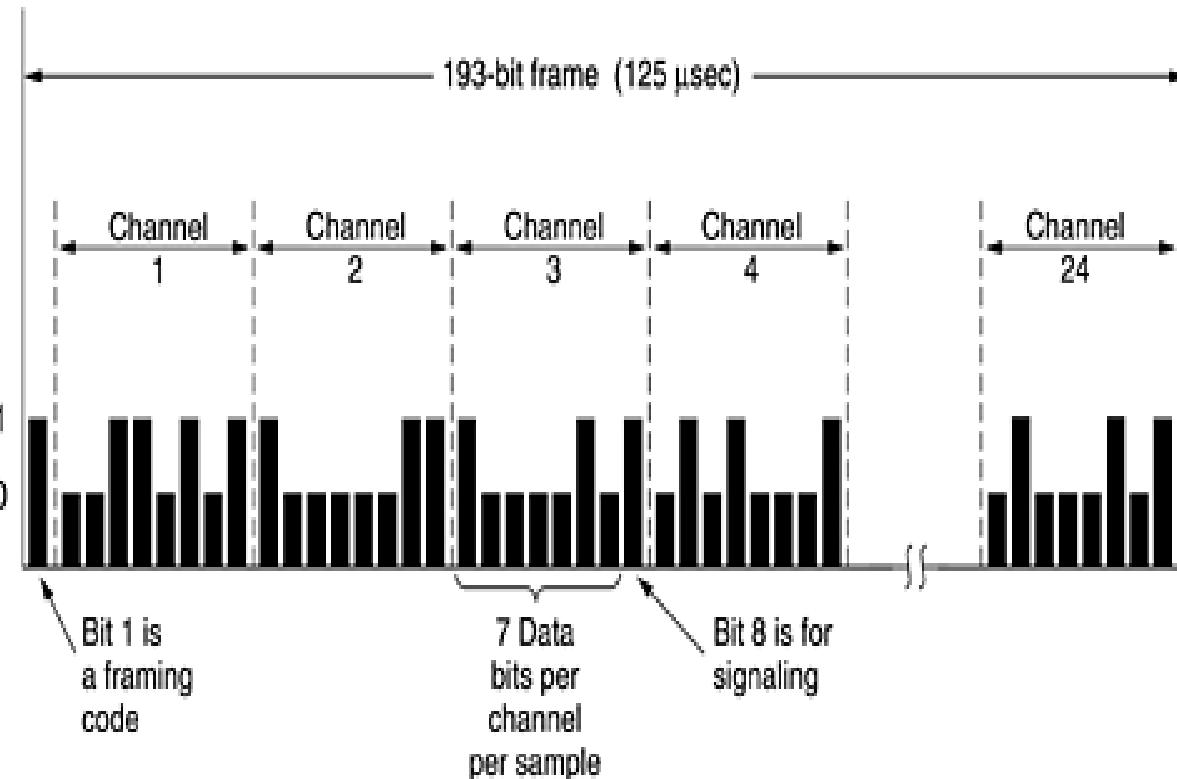
Time Division Multiplexing

- Although FDM is still used over copper wires or microwave channels, it requires analog circuitry and is not suitable for digital data.
- In contrast, TDM can be handled entirely by digital electronics, so it has become far more widespread in recent years. Unfortunately, it can only be used for digital data.
- Since the local loops produce analog signals, a conversion is needed from analog to digital in the end office, where all the individual local loops come together to be combined onto outgoing trunks..
- The analog signals are digitized in the end office by a device called a codec (coder-decoder), producing a series of 8- bit numbers.
- The codec makes 8000 samples per second (125 μ sec/sample) because the Nyquist theorem says that this is sufficient to capture all the information from the 4-kHz telephone channel bandwidth.
- At a lower sampling rate, information would be lost; at a higher one, no extra information would be gained. This technique is called PCM (Pulse Code Modulation).

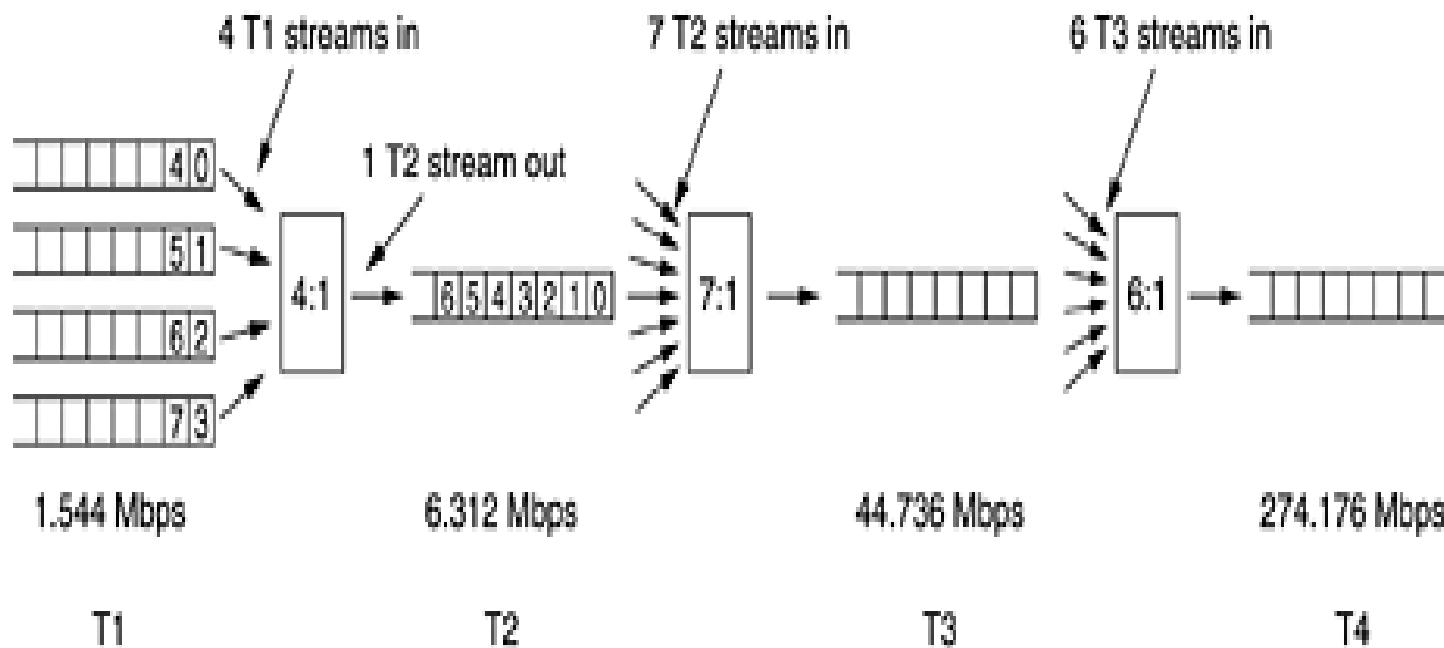
- PCM forms the heart of the modern telephone system. As a consequence, virtually all time intervals within the telephone system are multiples of 125 μ sec.
- No international standard for PCM, A variety of incompatible schemes are now in use in different countries around the world.
- The method used in North America and Japan is the T1 carrier.
- The T1 carrier consists of 24 voice channels multiplexed together.
- Usually, the analog signals are sampled on a round-robin basis with the resulting analog stream being fed to the codec rather than having 24 separate codecs and then merging the digital output.

- Each of the 24 channels, in turn, gets to insert 8 bits into the output stream. Seven bits are data and one is for control, yielding $7 \times 8000 = 56,000$ bps of data, and $1 \times 8000 = 8000$ bps of signaling information per channel.
- A frame consists of $24 \times 8 = 192$ bits plus one extra bit for framing, yielding 193 bits every 125 μ sec. This gives a gross data rate of $193/125 \mu$ sec = 1.544 Mbps.
- The 193rd bit is used for frame synchronization.
- It takes on the pattern 0101010101 Normally, the receiver keeps checking this bit to make sure that it has not lost synchronization.

- **Figure 2-33. The T1 carrier (1.544 Mbps).**



- **Figure 2-35. Multiplexing T1 streams onto higher carriers.**



- Time division multiplexing allows multiple T1 carriers to be multiplexed into higher-order carriers.
- Four T1 channels being multiplexed onto one T2 channel.
- Four T1 streams at 1.544 Mbps should generate 6.176 Mbps, but T2 is actually 6.312 Mbps. The extra bits are used for framing and recovery in case the carrier slips.
- T1 and T3 are widely used by customers, whereas T2 and T4 are only used within the telephone system itself, so they are not well known.
- At the next level, seven T2 streams are combined to form a T3 stream. Then six T3 streams are joined to form a T4 stream.
- At each step a small amount of overhead is added for framing and recovery in case the synchronization between sender and receiver is lost.

SONET/SDH

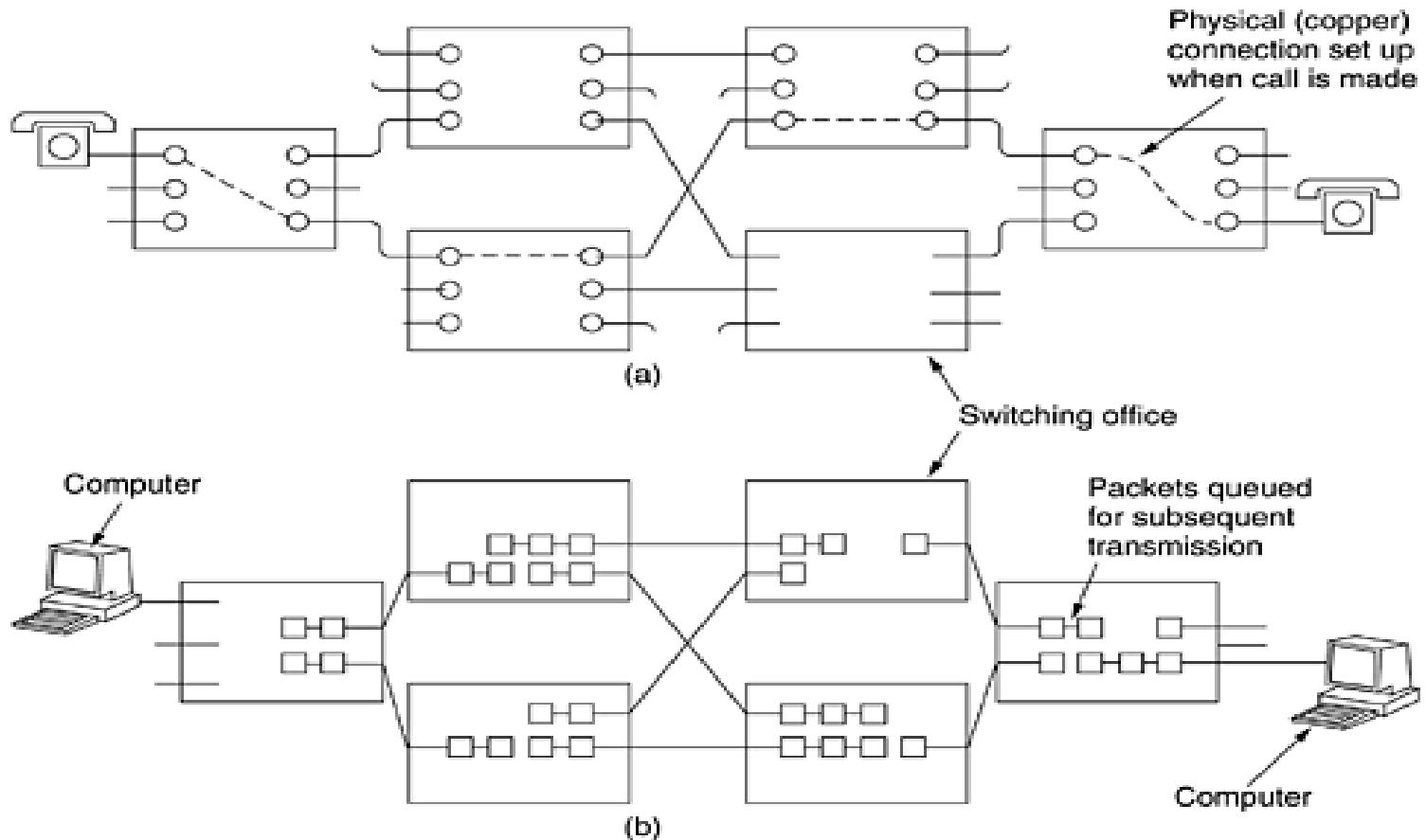
- In the early days of fiber optics, every telephone company had its own proprietary optical TDM system.
 - After AT&T was broken up in 1984, local telephone companies had to connect to multiple long-distance carriers, all with different optical TDM systems, so the need for standardization became obvious.
- SONET (Synchronous Optical NETwork) - 1985
 - SDH (Synchronous Digital Hierarchy) -1989
- Virtually all the long-distance telephone traffic in the United States, and much of it elsewhere, now uses trunks running SONET in the physical layer.

Switching

- The phone system is divided into two principal parts:
 - ❑ outside plant (the local loops and trunks, since they are physically outside the switching offices)
 - ❑ inside plant (the switches), which are inside the switching offices.
- Two different switching techniques are used nowadays: circuit switching and packet switching.

Circuit Switching

Figure 2-38. (a) Circuit switching. (b) Packet switching.



- When you or your computer places a telephone call, the switching equipment within the telephone system seeks out a physical path all the way from your telephone to the receiver's telephone. This technique is called circuit switching.
- Each of the six rectangles represents a carrier switching office (end office, toll office, etc.).
- When a call passes through a switching office, a physical connection is established between the line on which the call came in and one of the output lines.
- In the early days of the telephone, the connection was made by the operator plugging a jumper cable into the input and output sockets.
- Now it is replaced by automatic telephone switching equipment.
- The alternative to circuit switching is packet switching.
- With this technology, individual packets are sent as need be, with no dedicated path being set up in advance. It is up to each packet to find its way to the destination on its own
- An important property of circuit switching is the need to set up an end-to-end path before any data can be sent.
- Once a call has been set up, a dedicated path between both ends exists and will continue to exist until the call is finished.

Message Switching

- An alternative switching strategy is message switching.
- When this form of switching is used, no physical path is established in advance between sender and receiver. Instead, when the sender has a block of data to be sent, it is stored in the first switching office (i.e., router) and then forwarded later, one hop at a time.
- Each block is received in its entirety, inspected for errors, and then retransmitted. A network using this technique is called a store-and-forward network.
- The first electromechanical telecommunication systems used message switching, namely, for telegrams.

Packet Switching

- With message switching, there is no limit at all on block size, which means that routers (in a modern system) must have disks to buffer long blocks.
- It also means that a single block can tie up a router-router line for minutes, rendering message switching useless for interactive traffic.
- To get around these problems, packet switching was invented.
- Packet-switching networks place a tight upper limit on block size, allowing packets to be buffered in router main memory instead of on disk.
- By making sure that no user can monopolize any transmission line very long (milliseconds), packet-switching networks are well suited for handling interactive traffic.
- A further advantage of packet switching over message switching is: the first packet of a multipacket message can be forwarded before the second one has fully arrived, reducing delay and improving throughput.
- For these reasons, computer networks are usually packet switched, occasionally circuit switched, but never message switched.

A comparison of circuit-switched and packet-switched networks.

Item	Circuit switched	Packet switched
Call setup	Required	Not needed
Dedicated physical path	Yes	No
Each packet follows the same route	Yes	No
Packets arrive in order	Yes	No
Is a switch crash fatal	Yes	No
Bandwidth available	Fixed	Dynamic
Time of possible congestion	At setup time	On every packet
Potentially wasted bandwidth	Yes	No
Store-and-forward transmission	No	Yes
Transparency	Yes	No
Charging	Per minute	Per packet

- circuit switching requires that a circuit be set up end to end before communication begins. Packet switching does not require any advance setup. The first packet can just be sent as soon as it is available.
- The result of the connection setup with circuit switching is the reservation of bandwidth all the way from the sender to the receiver. All packets follow this path. Among other properties, having all packets follow the same path means that they cannot arrive out of order. With packet switching there is no path, so different packets can follow different paths, depending on network conditions at the time they are sent. They may arrive out of order.
- Packet switching is more fault tolerant than circuit switching.
- Setting up a path in advance also opens up the possibility of reserving bandwidth in advance. If bandwidth is reserved, then when a packet arrives, it can be sent out immediately over the reserved bandwidth. With packet switching, no bandwidth is reserved, so packets may have to wait their turn to be forwarded.
- congestion can occur at different times with circuit switching (at setup time) and packet switching (when packets are sent).
- If a circuit has been reserved for a particular user and there is no traffic to send, the bandwidth of that circuit is wasted. It cannot be used for other traffic. Packet switching does not waste bandwidth and thus is more efficient from a system-wide perspective.

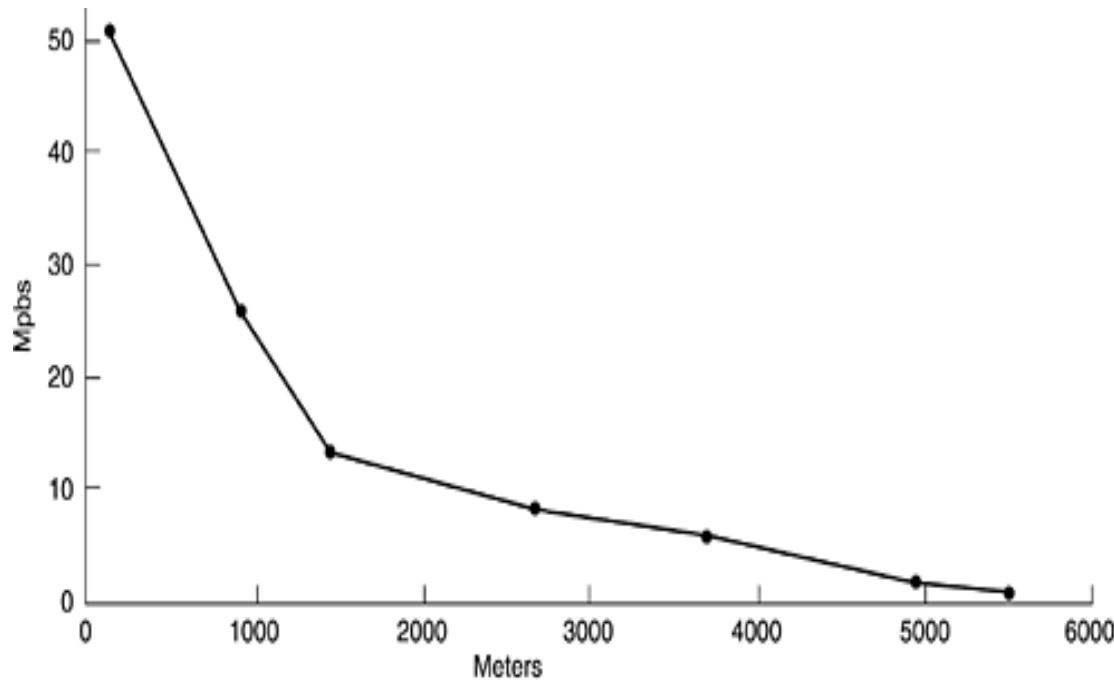
- Packet switching uses store-and-forward transmission. A packet is accumulated in a router's memory, then sent on to the next router. With circuit switching, the bits just flow through the wire continuously. The store-and-forward technique adds delay.
- Another difference is that circuit switching is completely transparent. The sender and receiver can use any bit rate, format, or framing method they want to. The carrier does not know or care. With packet switching, the carrier determines the basic parameters.
- A final difference between circuit and packet switching is the charging algorithm. With circuit switching, charging has historically been based on distance and time.

Digital Subscriber Lines

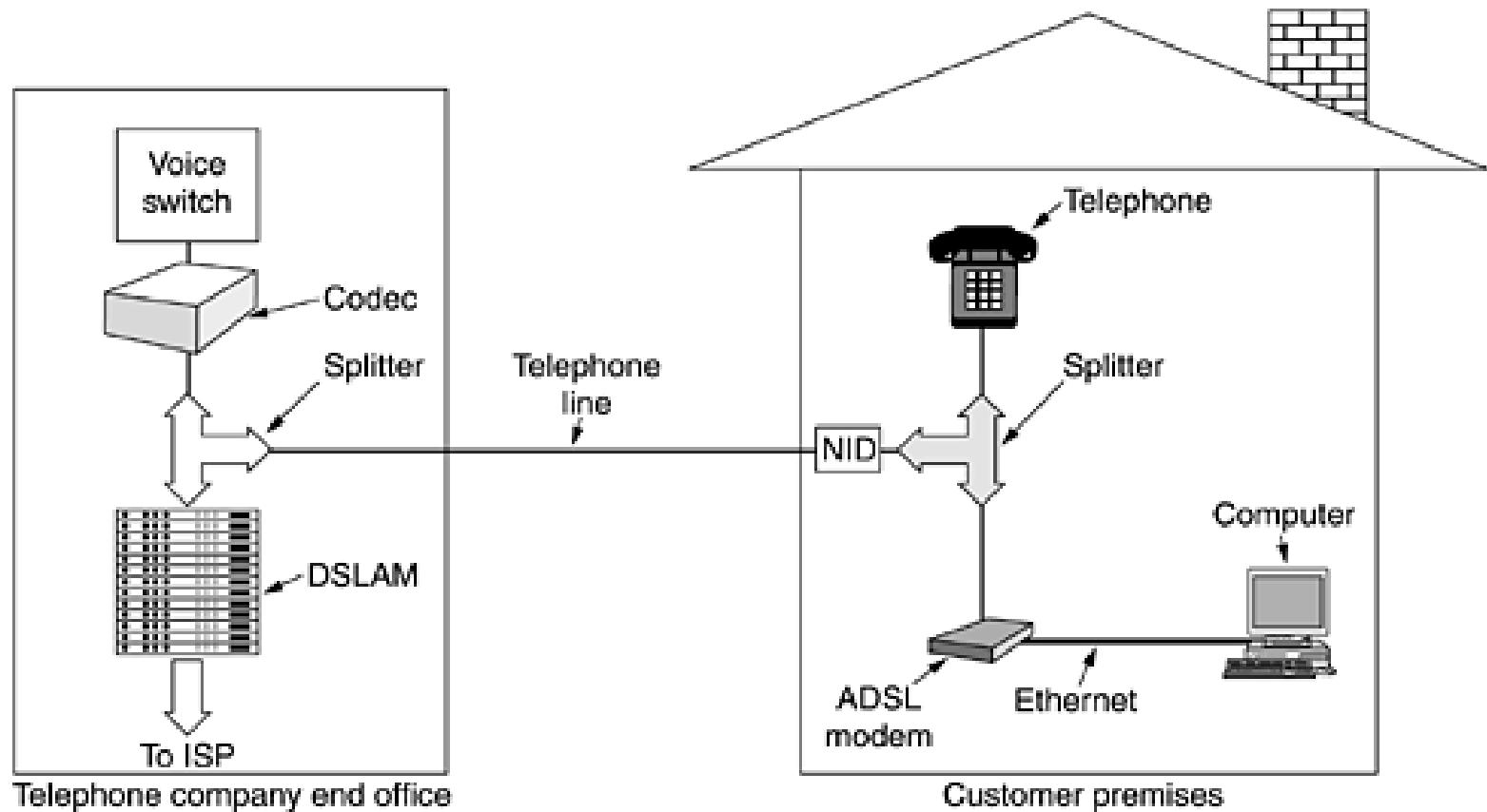
- When the telephone industry finally got to 56 kbps. Meanwhile, the cable TV industry was offering speeds up to 10 Mbps on shared cables, and satellite companies were planning to offer upward of 50 Mbps.
- As Internet access became an increasingly important part of their business, the telephone companies (LECs) began to realize they needed a more competitive product. Their answer was to start offering new digital services over the local loop.
- Services with more bandwidth than standard telephone service are sometimes called broadband, although the term really is more of a marketing concept than a specific technical concept.
- Initially, there were many overlapping offerings, all under the general name of xDSL (Digital Subscriber Line), for various x.
- The most popular of these services, ADSL (Asymmetric DSL).

- In telephone system, At the point where each local loop terminates in the end office, the wire runs through a filter that attenuates all frequencies below 300 Hz and above 3400 Hz.
- The cutoff is not sharp—300 Hz and 3400 Hz are the 3 dB points—so the bandwidth is usually quoted as 4000 Hz even though the distance between the 3 dB points is 3100 Hz.
- Data are thus also restricted to this narrow band.
- The trick that makes xDSL work is that when a customer subscribes to it, the incoming line is connected to a different kind of switch, one that does not have this filter, thus making the entire capacity of the local loop available.
- Unfortunately, the capacity of the local loop depends on several factors, including its length, thickness, and general quality.
- A plot of the potential bandwidth as a function of distance is given in Fig.

- The xDSL services have all been designed with certain goals in mind.
 - First, the services must work over the existing category 3 twisted pair local loops.
 - Second, they must not affect customers' existing telephones and fax machines.
 - Third, they must be much faster than 56 kbps.
 - Fourth, they should be always on, with just a monthly charge but no per-minute charge.



A typical ADSL equipment configuration.

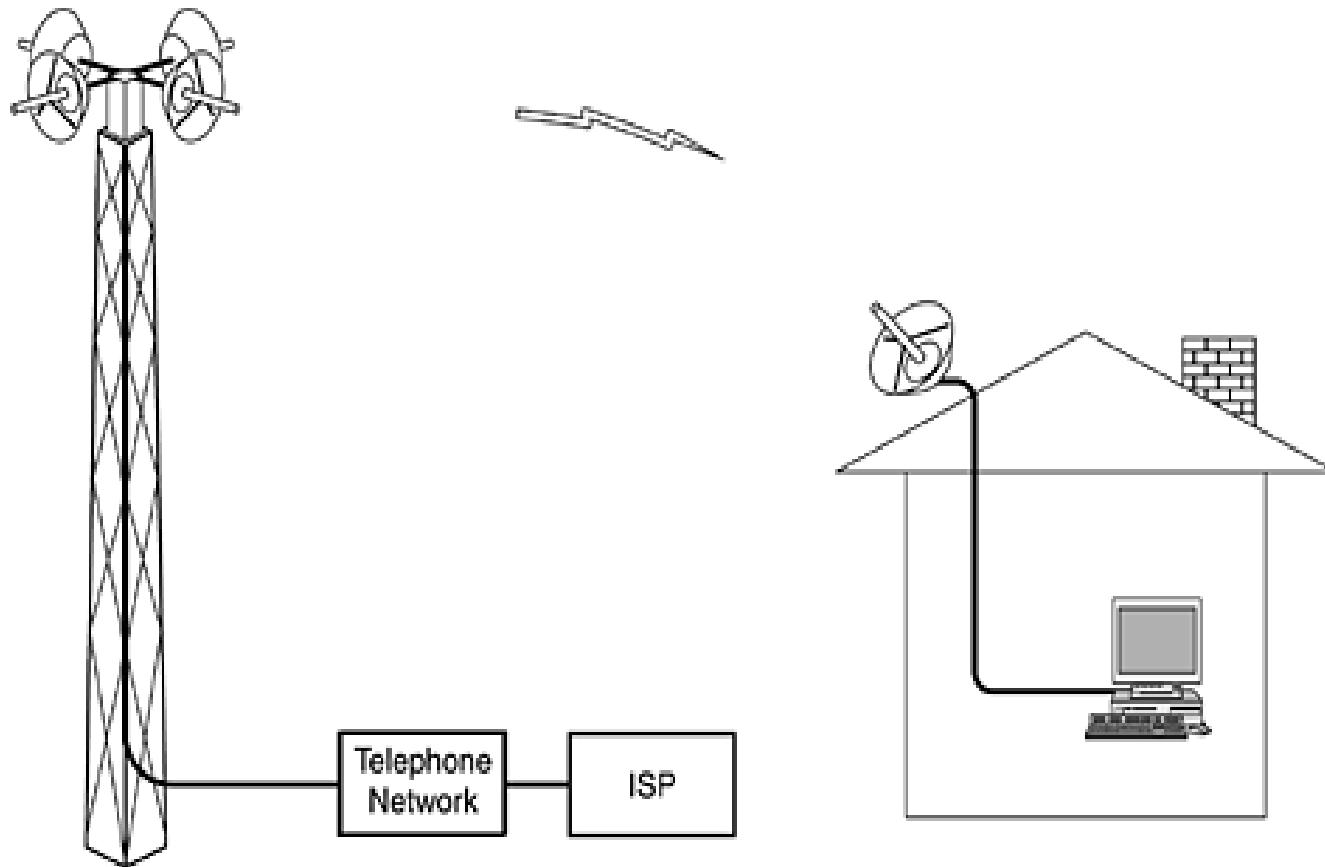


- A typical ADSL arrangement is shown in Fig. 2-29.
- In this scheme, a telephone company technician must install a NID (Network Interface Device) on the customer's premises. Close to the NID (or sometimes combined with it) is a splitter, an analog filter that separates the 0-4000 Hz band.
- The signal is routed to the existing telephone or fax machine, and the data signal is routed to an ADSL modem.
- The ADSL modem is actually a digital signal processor that has been set up to act as 250 QAM modems operating in parallel at different frequencies.
- At the other end of the wire, on the end office side, a corresponding splitter is installed. Here the voice portion of the signal is filtered out and sent to the normal voice switch.
- The signal above 26 kHz is routed to a new kind of device called a DSLAM (Digital Subscriber Line Access Multiplexer), which contains the same kind of digital signal processor as the ADSL modem.
- Once the digital signal has been recovered into a bit stream, packets are formed and sent off to the ISP.

Wireless Local Loops

- Local loop is a circuit line from a subscriber's phone to the local end office.
- But the implementation of local loop of wires is risky for the operators, especially in rural and remote areas due to less number of users and increased cost of installation.
- Hence, the solution for it is the usage of wireless local loop (WLL) which uses wireless links rather than copper wires to connect subscribers to the local central office.

Architecture of an LMDS (Local Multipoint Distribution Service) system.



- Tower with multiple antennas on it, each pointing in a different direction.
- Since millimeter waves are highly directional, each antenna defines a sector, independent of the other ones.
- At this frequency, the range is 2–5 km, which means that many towers are needed to cover a city.
- With current technology, each sector can have 36 Gbps downstream and 1 Mbps upstream, shared among all the users in that sector.

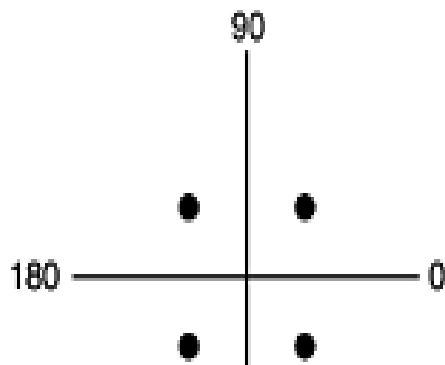
Q 20. Is an oil pipeline a simplex system, a half-duplex system, a full-duplex system, or none of the above?

Q 20. Is an oil pipeline a simplex system, a half-duplex system, a full-duplex system, or none of the above?

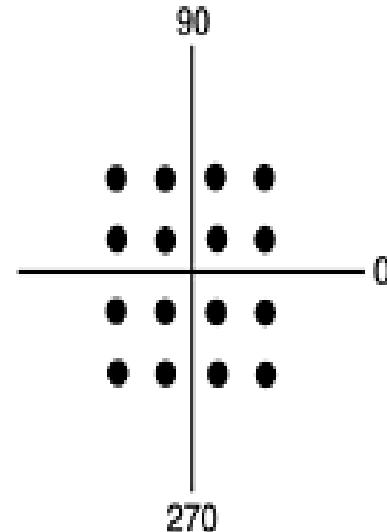
Answer:

- Like a single railroad track, it is half duplex. Oil can flow in either direction, but not both ways at once.

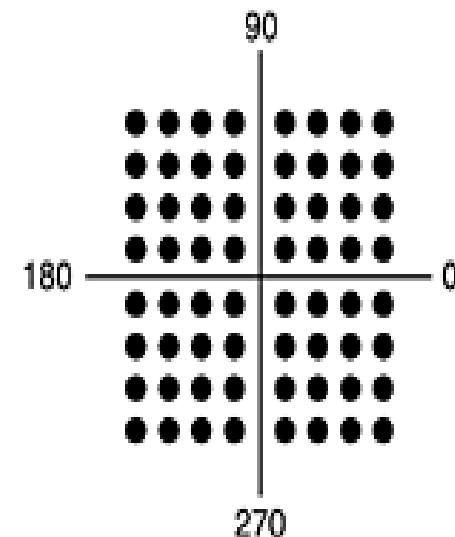
- Q 22. A modem constellation diagram similar to Fig. 2-25 has data points at the following coordinates: $(1, 1)$, $(1, -1)$, $(-1, 1)$, and $(-1, -1)$. How many bps can a modem with these parameters achieve at 1200 baud?



(a)



(b)



(c)

Answer:

- There are four legal values per baud, so the bit rate is twice the baud rate. At 1200 baud, the data rate is 2400 bps.

- Q. 23 A modem constellation diagram similar to Fig. 2-25 has data points at $(0, 1)$ and $(0, 2)$. Does the modem use phase modulation or amplitude modulation?

- Answer:

The phase shift is always 0, but two amplitudes are used, so this is straight amplitude modulation.

- Q 24:

In a constellation diagram, all the points lie on a circle centered on the origin. What kind of modulation is being used?

- Answer:
- If all the points are equidistant from the origin, they all have the same amplitude, so amplitude modulation is not being used. Frequency modulation is never used in constellation diagrams, so the encoding is pure phase shift keying.

- How many frequencies does a full-duplex QAM-64 modem use?

Answer:

- Two, one for upstream and one for downstream. The modulation scheme itself just uses amplitude and phase. The frequency is not modulated.

- Q 26. An ADSL system using DMT allocates 3/4 of the available data channels to the downstream link. It uses QAM-64 modulation on each channel. What is the capacity of the downstream link?

Q 26. An ADSL system using DMT allocates 3/4 of the available data channels to the downstream link. It uses QAM-64 modulation on each channel. What is the capacity of the downstream link?

Answer:

In Discrete multitone (**DMT**) there are 256 channels in all, minus 6 for POTS (Plain old telephone service) and 2 for control, leaving 248 for data.

If 3/4 of these are for downstream, that gives 186 channels for downstream. ADSL modulation is at 4000 baud, so with QAM-64 (6 bits/baud) we have 24,000 bps in each of the 186 channels.

The total bandwidth is then 4.464 Mbps downstream.

- Q 28. Ten signals, each requiring 4000 Hz, are multiplexed on to a single channel using FDM. How much minimum bandwidth is required for the multiplexed channel? Assume that the guard bands are 400 Hz wide.

Q 28. Ten signals, each requiring 4000 Hz, are multiplexed on to a single channel using FDM. How much minimum bandwidth is required for the multiplexed channel? Assume that the guard bands are 400 Hz wide.

Answer:

There are ten 4000 Hz signals.

We need nine guard bands to avoid any interference.

The minimum bandwidth required is $4000 \times 10 + 400 \times 9 = 43,600$ Hz.

Q 29 Why has the PCM sampling time been set at 125 μ sec?

Q 29 Why has the PCM sampling time been set at 125 μ sec?

Answer:

A sampling time of 125 μ sec corresponds to 8000 samples per second. According to the Nyquist theorem, this is the sampling frequency needed to capture all the information in a 4 kHz channel, such as a telephone channel. (Actually the nominal bandwidth is somewhat less, but the cutoff is not sharp.)

- Q 30 What is the percent overhead on a T1 carrier; that is, what percent of the 1.544 Mbps are not delivered to the end user?

Q 30 What is the percent overhead on a T1 carrier; that is, what percent of the 1.544 Mbps are not delivered to the end user?

Answer:

The end users get $7 \times 24 = 168$ of the 193 bits in a frame.

The overhead is therefore $25/193 = 13\%$.

**Numerical on Physical Layer
Computer Networking(CSE3034)**

1. A noiseless 4-kHz channel is sampled every 1 msec. What is the maximum data rate?

Sol:

1. The key word here is “noiseless”. With a normal 4 KHz channel, Shannon limit would not allow this. For the 4 KHz channel we can make 8000 samples/sec. In this case if each sample is 1024 bits this channel can send 8.2 Mbps.

2. Television channels are 6 MHz wide. How many bits/sec can be sent if four-level digital signals are used? Assume a noiseless channel.

Sol:

2. Using the Nyquist theorem, which is "Max. data rate = $2B \log_2 V$ bits/sec", we can sample = $2 (6\text{MHz}) \log_2 (4) = 24$ million times/sec. Therefore, using four level signals total data rate will be of 24 Mbps.

3. If a binary signal is sent over a 3-kHz channel whose signal-to-noise ratio is 20 dB, what is the maximum achievable data rate?

Sol:

3. Signal-to-Noise ratio (S/N) = 20 dB, which implies that, $10 \log_{10}(S/N) = 20$
 $\Rightarrow \log_{10}(S/N) = 2$
 $\Rightarrow S/N = 10^2 = 100.$

Now, from Shannon's theorem we know,

$$\begin{aligned}\text{Max. data rate} &= B \log_2(1 + S/N) \text{ bits/sec} \\ &= (3000 \text{ Hz}) \log_2(1 + 100) \text{ bits/sec} \\ &= (3000 \text{ Hz}) \log_2(101) \text{ bits/sec} \\ &= (3000) (6.643) \text{ bits/sec} \\ &= 19.92 \text{ kbps.}\end{aligned}$$

4. What signal-to-noise ratio is needed to put a T1 carrier on a 50-kHz line?

Sol:

4. $B = 50,000 \text{ Hz}$.

Now based on, Shannon's theorem, $B \log_2 (1 + S/N) \text{ bits/sec} = T1's \text{ data-rate}$.

$$50,000 \log_2(1 + S/N) = 1.544 \times 10^6 \log_2(1 + S/N) = 30.88$$

$$S/N = (2^{30.88}) - 1$$

$$\text{In dB, } S/N = 10 \log_{10} (S/N) = 10 \log_{10} ((2^{30.88}) - 1) = 92.95 \text{ dB.}$$

Therefore, the signal-to-noise ratio needs to be 92.95dB.

5. It is desired to send a sequence of computer screen images over an optical fiber. The screen is 480 x 640 pixels, each pixel being 24 bits. There are 60 screen images per second. How much bandwidth is needed, and how many microns of wavelength are needed for this band at 1.30 microns?

Sol:

5. Bandwidth needed is $480 * 640 * 24 * 60 = 442\,368\,000 \text{ bits}$

6. Radio antennas often work best when the diameter of the antenna is equal to the wavelength of the radio wave. Reasonable antennas range from 1 cm to 5 meters in diameter. What frequency range does this cover?

Sol:

6. $\text{Freq} = C/W$ where C: speed of light. W: wavelength. \ Convert 1 cm to m $\rightarrow 1 \text{ cm} = 0.01 \text{ m}$

For diameter of 1 cm: \ $\text{Freq} = (3 \times 10^8)/0.01 \ \text{GHz}$ \ $\text{Freq} = 3 \times 10^10 = 30 \text{ GHz}$

For diameter of 5 m: \ $\text{Freq} = (3 \times 10^8)/5 \ \text{MHz}$ \ $\text{Freq} = 6 \times 10^7 = 60 \text{ MHz}$

The cover range is from 60 MHz to 30 GHz.

7. A modem constellation diagram similar to Fig. 2-25 has data points at the following coordinates: (1, 1), (1, -1), (-1, 1), and (-1, -1). How many bps can a modem with these parameters achieve at 1200 baud?

Answer:

7. QPSK encodes 2 bits/symbol. rate = baud * bits/symbol = $1200 * 2 = 2400 \text{ bps}$

8. A modem constellation diagram similar to Fig. 2-25 has data points at (0, 1) and (0, 2). Does the modem use phase modulation or amplitude modulation?

Answer:

This is amplitude modulation because both points are the same angle from the positive x axis but are different distances away from origin (0,0).

9. How many frequencies does a full-duplex QAM-64 modem use?

Answer:

9. Two, one for upstream and one for downstream. The modulation scheme itself just uses amplitude and phase. The frequency is not modulated.

10. Ten signals, each requiring 4000 Hz, are multiplexed on to a single channel using FDM. How much minimum bandwidth is required for the multiplexed channel? Assume that the guard bands are 400 Hz wide.

Answer:

10. There are ten 4000 Hz signals.

Therefore, we need nine guard bands to avoid any interference.

Altogether, the minimum bandwidth required is $(4000 \times 10 + 400 \times 9)$ or, 43,600 Hz.

11. Why has the PCM sampling time been set at 125 μ sec?

Answer:

11. A sampling time of 125 μ sec (micro-sec) corresponds to 8000 samples per second, because, in 1 second (or, in 106 μ sec) we sample $(106/125)$ or, 8000. According to the Nyquist theorem (which is "Max. data rate = $2B \log_2 V$ bits/sec"), this is the sampling frequency needed to capture all the information in a ($B=$) 4-kHz channel, such as a telephone channel (Actually the nominal bandwidth is less, but the cutoff is not sharp.). Note:(assuming two signal level or, possible symbol), Max. data rate = $2B \log_2 V$ bits/sec = $2 \times 4,000 \times \log_2(2)$ bits/sec = $2 \times 4000 \times 1$ = 8,000.

12. What is the percent overhead on a T1 carrier; that is, what percent of the 1.544 Mbps are not delivered to the end user?

Answer:

12. With a modern T1 line, the end users get $8 \times 24 = 192$ of the 193 bits in a frame. The overhead is therefore $1 / 193 = 0.5\%$. Therefore, at least 0.5% of the 1.544 Mbps are not delivered to the end user.

13. Compare the maximum data rate of a noiseless 4-kHz channel using
 (a) Analog encoding (e.g., QPSK) with 2 bits per sample.
 (b) The T1 PCM system.

Answer:

13. In both cases 8000 samples/sec are possible. With dabit encoding, two bits are sent per sample. With T1, 7 bits are sent per period. The respective data rates are 16 kbps and 56 kbps.

14. If a binary signal is sent over a 3kHz bandwidth channel whose signal to noise ratio is 20dB, what is the maximum achievable data rate?

Answer:

14. From Shannon's theorem:

$$\text{Max Data Rate} = W \log_2(1+S/N)$$

Note that the signal to noise ratio (SNR) given here is a power ratio, yet we are given the SNR in decibels. We therefore need to convert back to a power ratio:

$$\text{SNR in Db} = 10 \log_2(1+S/N)$$

$$S/N=100$$

$$\text{Max Data Rate} = W \log_2(1+S/N) = 20 \text{ kbps}$$

The Nyquist limit for binary signalling over a 3kHz channel is

$$\text{Max Data Rate} = 2W \log_2 M = 6 \text{ kbps}$$

Therefore, the maximum achievable data rate is 6kbps. (To achieve higher rates than this (up to the Shannon limit), one would have to use a different signalling method.)

15. The loss in a cable is usually defined in decibels per kilometer (dB/km). If the signal at the beginning of a cable with -0.3 dB/km has a power of 2 mW, what is the power of the signal at 5 km?

Answer:

15. The loss in the cable in decibels is $5 \times (-0.3) = -1.5$ dB.

We can calculate the power as

$$dB = 10 \log_{10} \frac{P_2}{P_1} = -1.5$$

Thus $P_2 = 1.4 \text{ mW}$

16. A TV channel has a bandwidth of 6 MHz. If we send a digital signal using one channel, what are the data rates if we use one harmonic, three harmonics, and five harmonics?

Solution:

16. Using the first harmonic, data rate = $2 \times 6 \text{ MHz} = 12 \text{ Mbps}$

Using three harmonics, data rate = $(2 \times 6 \text{ MHz}) / 3 = 4 \text{ Mbps}$

Using five harmonics, data rate = $(2 \times 6 \text{ MHz}) / 5 = 2.4 \text{ Mbps}$

17. A signal travels from point A to point B. At point A, the signal power is 100 W. At point B, the power is 90 W. What is the attenuation in decibels?

Solution:

17. $\text{dB} = 10 \log_{10} (90 / 100) = -0.46 \text{ dB}$

18. If the bandwidth of the channel is 5 Kbps, how long does it take to send a frame of 100,000 bits out of this device?

Answer:

18. $100,000 \text{ bits} / 5 \text{ Kbps} = 20 \text{ s}$

19. A file contains 2 million bytes. How long does it take to download this file using a 56-Kbps channel? 1-Mbps channel?

Answer:

19. The file contains $2,000,000 \times 8 = 16,000,000$ bits. With a 56-Kbps channel, it takes $16,000,000 / 56,000 = 289 \text{ s}$. With a 1-Mbps channel, it takes 16 s

20. A signal with 200 milliwatts power passes through 10 devices, each with an average noise of 2 microwatts. What is the SNR? What is the SNR_{dB} ?

Answer:

20. We have $\text{SNR} = (200 \text{ mW}) / (10 \times 2 \times \mu\text{W}) = 10,000$

We then have $\text{SNR}_{\text{dB}} = 10 \log_{10} \text{SNR} = 40$

21. Calculate the baud rate for the given bit rate and type of modulation.

- a. 2000 bps, FSK
- b. 4000 bps, ASK
- c. 6000 bps, QPSK
- d. 36,000 bps, 64-QAM

Answer:

21. We use the formula

$$S = (1/r) \times N, \text{ but first we need to calculate the value of } r \text{ for each case.}$$

- a. $r = \log_2 2 = 1 \rightarrow S = (1/1) \times (2000 \text{ bps}) = 2000 \text{ baud}$
- b. $r = \log_2 2 = 1 \rightarrow S = (1/1) \times (4000 \text{ bps}) = 4000 \text{ baud}$
- c. $r = \log_2 4 = 2 \rightarrow S = (1/2) \times (6000 \text{ bps}) = 3000 \text{ baud}$
- d. $r = \log_2 64 = 6 \rightarrow S = (1/6) \times (36,000 \text{ bps}) = 6000 \text{ baud}$

22. Calculate the bit rate for the given baud rate and type of modulation.

- a. 1000 baud, FSK
- b. 1000 baud, ASK
- c. 1000 baud, BPSK
- d. 1000 baud, 16-QAM

Answer:

22. We use the formula $N = r \times S$, but first we need to calculate the value of r for each case.

- a. $r = \log_2 1 = 1 \rightarrow N = (1) \times (1000 \text{ bps}) = 1000 \text{ bps}$
- b. $r = \log_2 1 = 1 \rightarrow N = (1) \times (1000 \text{ bps}) = 1000 \text{ bps}$
- c. $r = \log_2 1 = 1 \rightarrow N = (1) \times (1000 \text{ bps}) = 1000 \text{ bps}$
- d. $r = \log_2 4 = 2 \rightarrow N = (2) \times (1000 \text{ bps}) = 2000 \text{ bps}$

23. A cable company uses one of the cable TV channels (with a bandwidth of 6 MHz) to provide digital communication for each resident. What is the available data rate for each resident if the company uses a 64-QAM technique?

Answer:

23. 36Mbps

24. We have 14 sources, each creating 500 8-bit characters per second. Since only some of these sources are active at any moment, we use statistical TDM to combine these sources using character interleaving. Each frame carries 6 slots at a time, but we need to add 4-bit addresses to each slot. Answer the following questions:

- a. What is the size of an output frame in bits?
- b. What is the output frame rate?
- c. What is the duration of an output frame?
- d. What is the output data rate?

Answer:

24. Frame size = (# of slots) x (character size + slot address) = 6 x (8 bits+ 4 bits) =72 bits

We can assume that we have only 6 input lines. Each frame needs to carry one character from each of these lines. This means that the link needs to send 500 frames/s

Frame duration = 1 / (frame rate) = 1 / 500 = 2ms

Data rate = (500 frames/s) x (72 bits/frame) = 36 Kbps

25. Two channels, one with a bit rate of 190 kbps and another with a bit rate of 180 kbps, are to be multiplexed using pulse-stuffing TDM with no synchronization bits. Answer the following questions:

- a. What is the size of a frame in bits?
- b. What is the frame rate?
- c. What is the duration of a frame?
- d. What is the data rate?

Answer:

25. We need to add extra bits to the second source to make both bit rates = 190Kbps. Now we have two sources, each of 190 Kbps. Since the data unit was not specified, assume that it is one bit. Frame size = 2 bits.

Frame rate = 190k frames/s

Frame duration = 1/frame rate = 1/190k = 5.26μs

Data rate = 190k*2 = 380kbps

Chapter 3: Data Link Layer Numericals

Q 1.

The following character encoding is used in a data link protocol:

A: 01000111; B: 11100011; FLAG:01111110; ESC: 11100000

Show the bit sequence transmitted (in binary) for the four-character frame: A B ESC FLAG when each of the following framing methods are used:

- a. Character count.
- b. Flag bytes with byte stuffing.
- c. Starting and ending flag bytes, with bit stuffing.

A.1.

- a. 00000100 01000111 11100011 11100000 01111110
- b. 01111110 01000111 11100011 11100000 11100000 11100000 01111110 01111110
- c. 01111110 01000111 110100011 111000000 011111010 01111110

Q 2.

The following data fragment occurs in the middle of a data stream for which the byte-stuffing algorithm described in the text is used:

A B ESC C ESC FLAG FLAG D. What is the output after stuffing?

A.2.

After stuffing the output is

A B ESC ESC C ESC ESC ESC FLAG ESC FLAG D

Q.3.

What is the maximum overhead in byte-stuffing algorithm?

A.3.

The maximum overhead in byte stuffing algorithm is 100%(i.e. when the payload contains only ESC and Flag bytress).

Q 4.

A bit string, 011110111110111110, needs to be transmitted at the data link layer. What is the string actually transmitted after bit stuffing?

A.4.

The actual transmitted bit string after bit stuffing is

011110111110011111010

Q.5.

Let us assume that $m = 3$ and $n = 4$. Find the list of valid datawords and codewords assuming the check bit is used to indicate even parity in the code word.

A.5.

Valid datawords : 000, 001, 010, 011, 100, 101, 110, 111

Valid codewords : 0000, 0011, 0101, 0110, 1001, 1010, 1100, 1111

Q.6.

What is the Hamming distance for each of the following codewords:

- a. (10000, 00000)
- b. (10101, 10000)
- c. (11111, 11111)
- d. (000, 000)

A.6.

- a. 1
- b. 2
- c. 0
- d. 0

Q.7.

Given the codeword of size 4 bit. If the size of dataword is 3 bit. What is the value of hamming distance for the codeword?

A. 7.

Hamming distance = 2

Q 8.

To provide more reliability than a single parity bit can give, an error-detecting coding scheme uses one parity bit for checking all the odd-numbered bits and a second parity bit for all the even-numbered bits.

What is the Hamming distance of this code?

A.8.

Making one change to any valid character cannot generate another valid character due to the nature of parity bits. Making two changes to even bits or two changes to odd bits will give another valid character, so the distance is 2.

Q. 9.

Find the minimum Hamming distance to be implemented in codeword for the following cases:

- Detection of two errors.
- Correction of two errors.
- Detection of 3 errors or correction of 2 errors.
- Detection of 6 errors or correction of 2 errors.

A.9.

- For error detection → Hamming distance = $d + 1 = 2 + 1 = 3$
- For error correction → Hamming distance = $2d + 1 = 2 \times 2 + 1 = 5$
- For error detection → Hamming distance = $d + 1 = 3 + 1 = 4$

For error correction → Hamming distance = $2d + 1 = 2 \times 2 + 1 = 5$

Therefore minimum Hamming distance should be **5**.

- For error detection → Hamming distance = $d + 1 = 6 + 1 = 7$

For error correction → Hamming distance = $2d + 1 = 2 \times 2 + 1 = 5$

Therefore minimum Hamming distance should be **7**.

Q.10.

Given in the table a set of valid dataword and codeword.

Dataword	Codeword
00	00000
01	01011
10	10101
11	11110

What is the dataword transmitted for the following codewords received assuming there is 1 bit error?

- a. 01010
- b. 11010

A.10.

- a. 01
- b. 11

Q 11.

Sixteen-bit messages are transmitted using a Hamming code. How many check bits are needed to ensure that the receiver can detect and correct single bit errors? Show the bit pattern transmitted for the message 1101001100110101. Assume that even parity is used in the Hamming code.

A.11.

5 check bits are needed at positions 1, 2, 4, 8, and 16.

The bit pattern transmitted for the message 1101001100110101 is 011010110011001110101

Q.12.

An 8 bit message using even-parity Hamming code is received as **101001001111**. Find the 8 bit message after getting decoded assuming no error during transmission?

A.12.

The 8 bit message after decoding is 10101111.

Q.13.

A 12-bit Hamming code whose hexadecimal value is 0xE4F arrives at a receiver. What was the original value in hexadecimal? Assume that not more than 1 bit is in error.

A.13.

If we number the bits from left to right starting at bit 1, in this example bit 2 (a parity bit) is incorrect. The 12-bit value transmitted (after Hamming encoding) was 0xA4F. The original 8-bit data value was 0xAF.

Q.14.

Suppose that data are transmitted in blocks of sizes 1000 bits. What is the maximum error rate under which error detection and retransmission mechanism (1 parity bit per block) is better than using Hamming code? Assume that bit errors are independent of one another and no bit error occurs during retransmission.

A.14.

From Eq. $(m+r+1) \leq 2^r$, we know that 10 check bits are needed for each block in case of using Hamming code. Total bits transmitted per block are 1010 bits. In case of error detection mechanism, one parity bit is transmitted per block (i.e. 1001). Suppose error rate is x per bit. Thus, a block may encounter a bit error $1000x$ times. Every time an error is encountered, 1001 bits have to be retransmitted. So, total bits transmitted per block are $1001 + 1000x \times 1001$ bits. For error detection and retransmission to be better, $1001 + 1000x \times 1001 < 1010$. So, the error rate must be less than 9×10^{-6} .

Q.15.

What is the remainder obtained by dividing x^7+x^5+1 by the generator polynomial x^3+1 ?

A.15.

The remainder is x^2+x+1 .

Q.16.

Given the dataword 101001111 and the divisor 10111. Show the generation of the CRC codeword at the sender site (using binary division).

A.16.

The codeword at the sender site is 1010011110001

Q.17.

A bit stream 10101010 is transmitted using the standard CRC method. The generator polynomial is x^3+x^2+1 . Show the actual bit string transmitted. Suppose the second bit from the left is inverted during transmission. Show that this error is detected at the receiver's end.

A.17.

The frame is 10101010. The generator is 1101. We must append 3 zeros to the message (i.e. 10101010000). The remainder after dividing 10101010000 by 1101 is 110. So actual bit string transmitted is 10101010110. Since the second bit from left is inverted during transmission, the bits received are 11101010110. Dividing this by 1101 doesn't give remainder 0. So the received bits contain error.

Q.18.

A bit stream 10011101 is transmitted using the standard CRC method. The generator polynomial is x^3+1 . Show the actual bit string transmitted. Suppose that the third bit from the left is inverted during transmission.

A.18.

The frame is 10011101. The generator is 1001. The message after appending three zeros is 10011101000. The remainder on dividing

1001110100 by 1001 is 100 . So, the actual bit string transmitted is 10011101100 . The received bit stream with an error in the third bit from the left is 10111101100 . Dividing this by 1001 produces a remainder 100 , not 0 . So the received bits contain error and needs retransmission.

Q.19.

A channel has a bit rate of 4 kbps and a propagation delay of 20 msec. For what range of frame sizes does stop-and-wait give an efficiency of at least 50%?

A.19.

Efficiency will be 50% when the time required to transmit the frame equals the round-trip propagation delay. At a transmission rate of 4 bits/msec, 160 bits takes 40 msec. For frame sizes above 160 bits, stop-and-wait is reasonably efficient.

Q.20.

A 3000-km-long T1 trunk is used to transmit 64-byte frames using protocol 5. If the propagation speed is $6 \mu\text{sec}/\text{km}$, how many bits should the sequence numbers be?

A.20.

To operate efficiently, the sequence space (actually, the send window size) must be large enough to allow the transmitter to keep transmitting until the first acknowledgement has been received. The propagation time is 18 ms. At T1 speed, which is 1.536 Mbps (excluding the 1 header bit), a 64-byte frame takes 0.300 msec. Therefore, the first frame fully arrives 18.3 msec after its transmission was started. The acknowledgement takes another 18 msec to get back, plus a small (negligible) time for the acknowledgement to arrive fully. In all, this time is 36.3 msec. The transmitter must have enough

window space to keep going for 36.3 msec. A frame takes 0.3 ms, so it takes 121 frames to fill the pipe. Seven-bit sequence numbers are needed.

Chapter 4. The Medium Access Control Sublayer

The Medium Access Control Sublayer

- Networks can be divided into two categories: those using point to-point connections and those using broadcast channels.
- In any broadcast network, the key issue is how to determine who gets to use the channel when there is competition for it.
- When only a single channel is available, determining who should go next is much harder.
- Broadcast channels are sometimes referred to as **multiaccess channels** or **random access channels**.

- The protocols used to determine who goes next on a multiaccess channel belong to a sublayer of the data link layer called the **MAC (Medium Access Control)** sublayer.
- The MAC sublayer is especially important in LANs, many of which use a multiaccess channel as the basis for communication.
- WANs, in contrast, use point-to-point links, except for satellite networks.
- Technically, the MAC sublayer is the bottom part of the data link layer.

4.1 The Channel Allocation Problem

- **Static Channel Allocation (Fixed Channel allocation)**
- **Dynamic Channel Allocation**

Static Channel Allocation in LANs and MANs

- The traditional way of allocating a single channel, such as a telephone trunk, among multiple competing users is Frequency Division Multiplexing (FDM).
- If there are N users, *the bandwidth* is divided into N equal-sized portions, each user being assigned one portion.
- Since each user has a private frequency band, there is no interference between users.
- When there is only a small and constant number of users, FDM is a simple and efficient allocation mechanism.

- **Disadvantage:** when the number of senders is large and continuously varying, FDM presents some problems.
- If the spectrum is cut up into N regions and fewer than N users are currently interested in communicating, a large piece of valuable spectrum will be wasted.
- If more than N users want to communicate, some of them will be denied permission for lack of bandwidth, even if some of the users who have been assigned a frequency band hardly ever transmit or receive anything.
- In TDM based allocation Each user is statically allocated every N th time slot. If a user does not use the allocated slot, it is just wastage of bandwidth.

Dynamic Channel Allocation in LANs and MANs

- In dynamic channel allocation scheme, frequency bands are not permanently assigned to the users.
- Instead channels are allotted to users dynamically as needed.
- This allocation scheme optimizes bandwidth usage and results in faster transmissions.
- The allocation is done considering a number of parameters so that transmission interference is minimized.

- All the work done in this area are based on five key assumptions:
- **1. Station Model:** The model consists of N *independent stations* (e.g., *computers, telephones, or personal communicators*), each with a program or user that generates frames for transmission. Stations are sometimes called **terminals**.
- Once a frame has been generated, the station is blocked and does nothing until the frame has been successfully transmitted.

- **2. Single Channel Assumption:** A single channel is available for all communication. All stations can transmit on it and all can receive from it.
- **3. Collision Assumption:** If two frames are transmitted simultaneously, they overlap in time and the resulting signal is garbled. This event is called a **collision**. All stations can detect collisions.
- A collided frame must be transmitted again later. There are no errors other than those generated by collisions.

- **4a. Continuous Time:** Frame transmission can begin at any instant. There is no master clock dividing time into discrete intervals.
- **4b. Slotted Time:** Time is divided into discrete intervals (slots). Frame transmissions always begin at the start of a slot.
- **5a. Carrier Sense:** Stations can tell if the channel is in use before trying to use it. If the channel is sensed as busy, no station will attempt to use it until it goes idle.
- **5b. No Carrier Sense.** Stations cannot sense the channel before trying to use it. They just go ahead and transmit. Only later can they determine whether the transmission was successful.

Difference

	FCA	DCA
Channel Allocation	Fixed number of channels or voice channels are allotted.	Channels to be allotted are not fixed initially.
Blockage	If all channels are occupied, then user call is blocked.	If all channels are blocked, then Base Station(BS) requests more channels from Mobile Station Center(MSC).
Algorithm	No need to complex algorithm.	Algorithm to determine efficient channel availability is quite complex in DCA.
Cost	FCA is cheaper than DCA.	DCA is costly as real time computation needed.
Cell Allocation	Once call is complete, channel remains with the cell.	Once call is complete, channel is returned back to Mobile Station Center.

4.2 Multiple Access Protocols

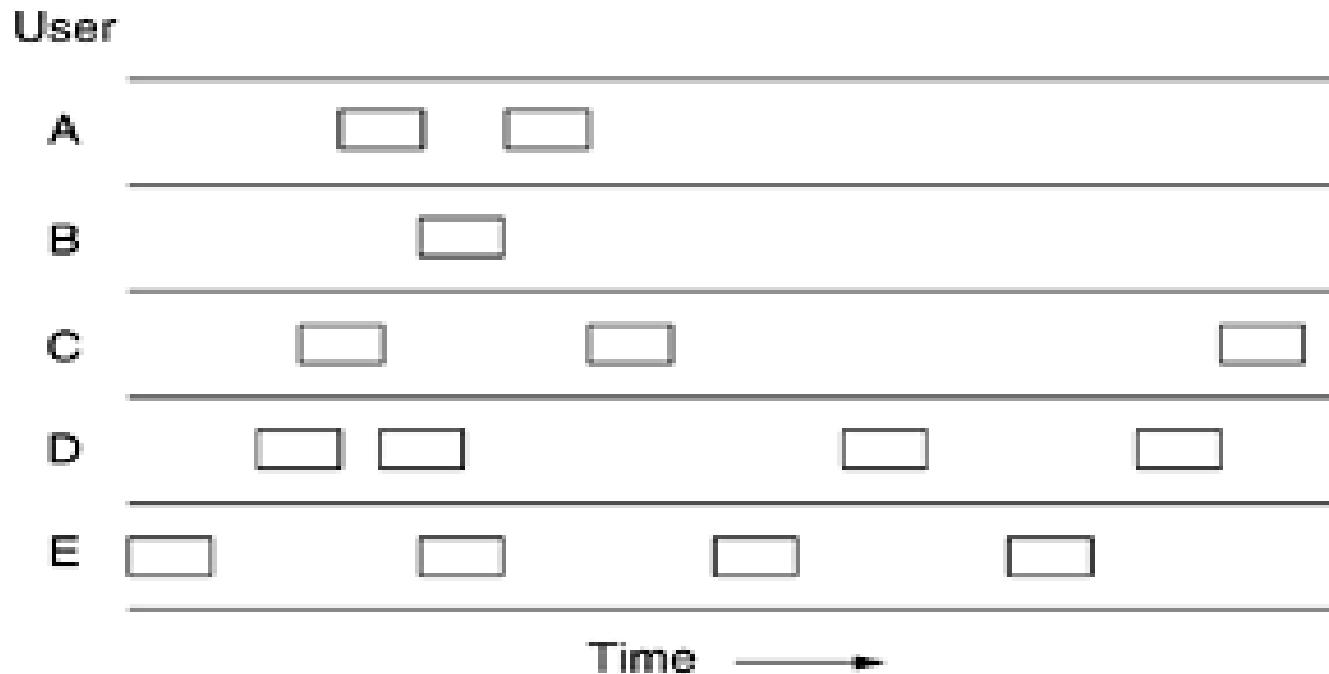
ALOHA

- In the 1970s, Norman Abramson developed this method to solve the channel allocation problem.
- Two versions of ALOHA here: **pure and slotted**.
- They differ with respect to whether time is divided into discrete slots into which all frames must fit.
- Pure ALOHA does not require global time synchronization; slotted ALOHA does.

Pure ALOHA

- The basic idea of an ALOHA system: Users transmit whenever they have data to be sent. There will be collisions, of course, and the colliding frames will be damaged.
- However, due to the feedback property of broadcasting, a sender can always find out whether its frame was destroyed by listening to the channel.
- With a LAN, the feedback is immediate; with a satellite, there is a delay of 270 msec before the sender knows if the transmission was successful.
- If listening while transmitting is not possible for some reason, acknowledgements are needed.
- If the frame was destroyed, the sender just waits a random amount of time and sends it again.
- Systems in which multiple users share a common channel in a way that can lead to conflicts are widely known as **contention** systems.

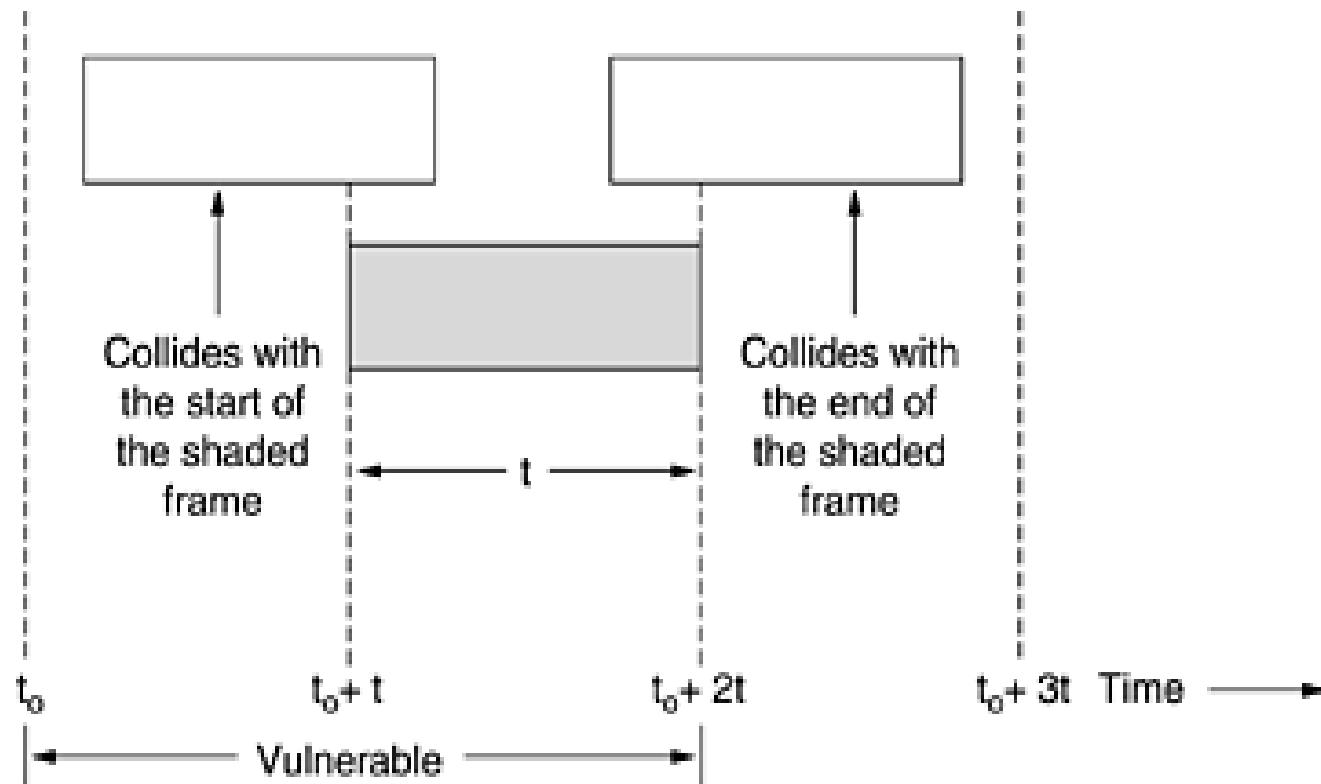
Figure 4-1. In pure ALOHA, frames are transmitted at completely arbitrary times.



- Whenever two frames try to occupy the channel at the same time, there will be a collision and both will be garbled.
- The checksum cannot (and should not) distinguish between a total loss and a near miss.
- A user is always in one of two states: **typing or waiting**.
- Initially, all users are in the typing state. When a line is finished, the user stops typing, waiting for a response.

- The station then transmits a frame containing the line and checks the channel to see if it was successful. If so, the user sees the reply and goes back to typing.
- If not, the user continues to wait and the frame is retransmitted over and over until it has been successfully sent.
- Let the "**frame time**" denote the amount of time needed to transmit the standard, fixed-length frame (i.e., the **frame length divided by the bit rate**).

Figure 4-2. Vulnerable period for the shaded frame.



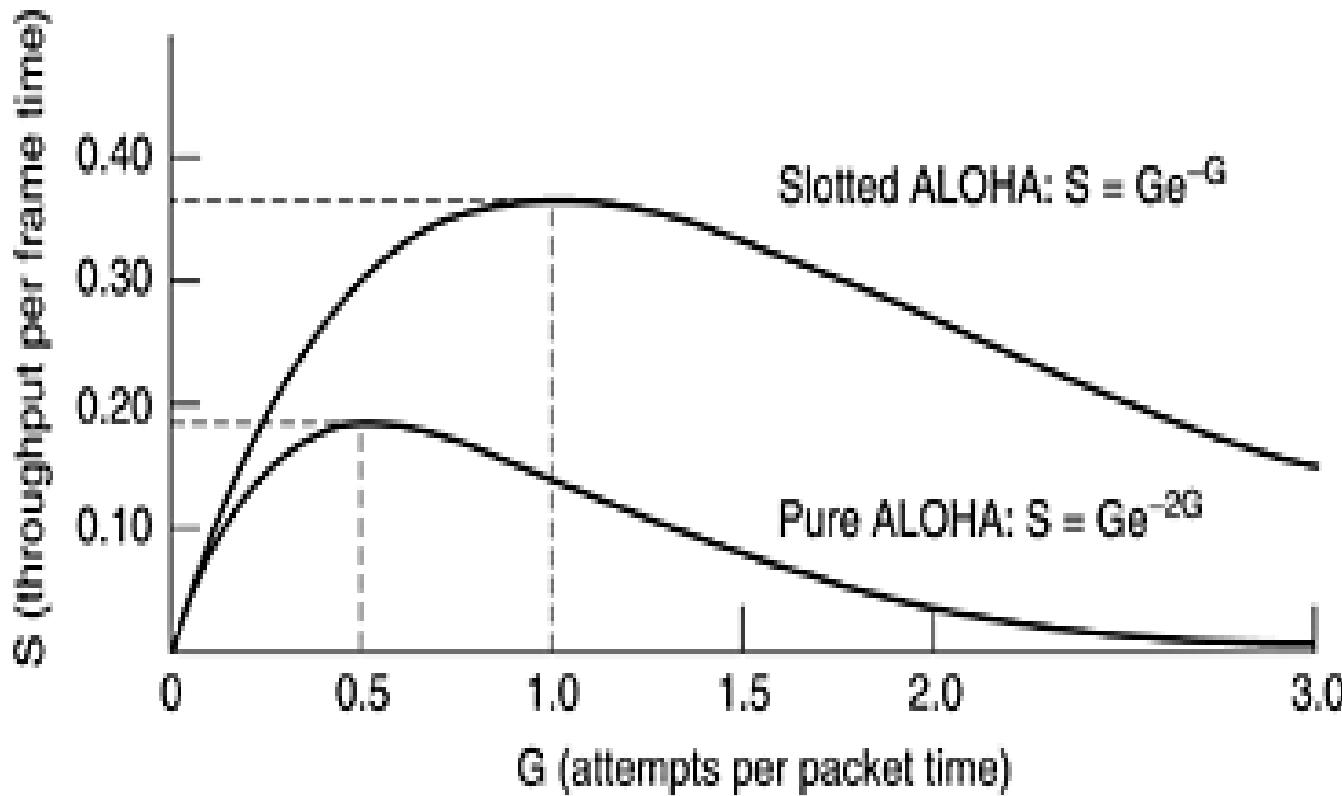
- A frame will not suffer a collision if no other frames are sent within one frame time of its start.
- Under what conditions will the shaded frame arrive undamaged?
- Let t be the time required to send a frame. If any other user has generated a frame between time t_0 and $t_0 + t$, the end of that frame will collide with the beginning of the shaded one.
- In pure ALOHA a station does not listen to the channel before transmitting, it has no way of knowing that another frame was already underway.
- Similarly, any other frame started between $t_0 + t$ and $t_0 + 2t$ will bump into the end of the shaded frame.
- An interesting question is: What is the efficiency of an ALOHA channel? what fraction of all transmitted frames escape collisions under these chaotic circumstances?

- The probability that k frames are generated during a given frame time is given by the Poisson distribution:

$$\Pr[k] = \frac{G^k e^{-G}}{k!}$$

- so the probability of zero frames is just e^{-G} .
- Throughput $S = Ge^{-2G}$
- The relation between the offered traffic and the throughput is shown in Fig. 4-3.
- The **maximum throughput** occurs at $G = 0.5$, with $S = 1/2e$, which is about **0.184**.
- In other words, the best we can hope for is a channel utilization of 18 percent.

Figure 4-3. Throughput versus offered traffic for ALOHA systems



Slotted ALOHA

- In 1972, Roberts published a method for doubling the capacity of an ALOHA system.
- His proposal was to divide time into discrete intervals, each interval corresponding to one frame.
- In Roberts' method, which has come to be known as slotted ALOHA, in contrast to Abramson's pure ALOHA, a computer is not permitted to send whenever a carriage return is typed.
- Instead, it wait for the beginning of the next slot. Thus, the continuous pure ALOHA is turned into a discrete one.
- Since the vulnerable period is now halved, the probability of no other traffic during the same slot as our test frame is e^{-G} which leads to

$$S = Ge^{-G}$$

- Slotted ALOHA peaks at $G = 1$, with a throughput of $S = 1/e$ or about 0.368, twice that of pure ALOHA.
- If the system is operating at $G = 1$, the probability of an empty slot is 0.368.
- The best we can hope for using slotted ALOHA is 37 percent of the slots empty, 37 percent successes, and 26 percent collisions.
- Operating at higher values of G reduces the number of empties but increases the number of collisions exponentially.
- When Internet access over the cable was invented, all of a sudden there was a problem of how to allocate a shared channel among multiple competing users.

Carrier Sense Multiple Access Protocols

Carrier Sense Multiple Access Protocols

- Carrier Sense Multiple Access (CSMA) is a network protocol for carrier transmission that operates in the Medium Access Control (MAC) layer.
- With slotted ALOHA the best channel utilization that can be achieved is $1/e$.
- In local area networks, however, it is possible for stations to detect what other stations are doing, and adapt their behavior accordingly. These networks can achieve a much better utilization than $1/e$.
- Protocols in which stations listen for a carrier (i.e., a transmission) and act accordingly are called **carrier sense protocols**.
- It senses or listens whether the shared channel for transmission is busy or not, and transmits if the channel is not busy.
- Using CSMA protocols, more than one users or nodes send and receive data through a shared medium that may be a single cable or optical fiber connecting multiple nodes.

Working Principle

- When a station has frames to transmit, it attempts to detect presence of the carrier signal from the other nodes connected to the shared channel.
- If a carrier signal is detected, it implies that a transmission is in progress.
- The station waits till the ongoing transmission executes to completion, and then initiates its own transmission.
- Generally, transmissions by the node are received by all other nodes connected to the channel.
- Since, the nodes detect for a transmission before sending their own frames, collision of frames is reduced.
- However, if two nodes detect an idle channel at the same time, they may simultaneously initiate transmission. This would cause the frames to garble resulting in a collision.

Persistent and Nonpersistent CSMA

1-persistent CSMA

- The first carrier sense protocol is called **1-persistent CSMA** (Carrier Sense Multiple Access).
- 1-persistent CSMA is a Carrier Sense Multiple Access (CMSA) protocol that operates in the Medium Access Control (MAC) layer.
- In 1-persistent CSMA, when a transmitting station has a frame to send and it senses a busy channel, it waits for the end of the transmission, and transmits immediately.
- Since, it sends with a probability 1, the name 1 – persistent CSMA is given.

Algorithm

The algorithm of 1-persistent CSMA is:

- When a frame is ready, the transmitting station checks whether the channel is idle or busy.
- If the channel is busy, the station waits and continually checks until the channel becomes idle.
- If the channel is idle then it transmits the frame immediately, with a probability 1.
- A collision may occur if two or more channels transmit simultaneously.
- If collision occurs, the station waits for a random period of time and restarts the algorithm all over again.

Disadvantages of 1-persistent CSMA

There are chances of collisions in the following situations:

- The propagation delay has an important effect on the performance of the protocol.
- The longer the propagation delay, the more important this effect becomes, and the worse the performance of the protocol.
- **Situation 1:** Suppose that a station A has transmitted a frame, which has not yet reached another station B due to propagation delay. Station B assumes that the channel is idle and transmits its frame. Thus a collision occurs.
- Even if the propagation delay is zero, there will still be collisions. If two stations become ready in the middle of a third station's transmission, both will wait politely until the transmission ends and then both will begin transmitting exactly simultaneously, resulting in a collision.
- **Situation 2:** Suppose that a station A is transmitting while stations B and C are waiting for the transmission to complete. At the instance station A completes transmission, both stations B and C start transmitting simultaneously at the same time. This results in collision.

Advantage of 1-persistent CSMA

- It has better throughput than ALOHA protocols.
- This protocol is far better than pure ALOHA because both stations have the decency to stop from interfering with the third station's frame.
- This approach will lead to a higher performance than pure ALOHA. Exactly the same holds for slotted ALOHA.

Non-persistent CSMA protocol

- A second carrier sense protocol is **nonpersistent CSMA**.
- In this protocol, a conscious attempt is made to be less greedy than in the previous one.

Algorithm

The algorithm of non-persistent CSMA is

- When a frame is ready, the transmitting station checks whether the channel is idle or busy.
- If the channel is idle then it transmits the frame immediately.
- If the channel is busy, the station waits for a random time period during which it does not check whether the channel is idle or busy.
- At the end of the waiting time period, it again checks the status of the channel and restarts the algorithm.
- Consequently, this algorithm leads to better channel utilization but longer delays than 1-persistent CSMA.

Advantage of non-persistent CSMA

- Its rate of collision is much less than 1-persistent CSMA. This is because each station waits for a random amount of time before attempting retransmission.
- The probability that multiple stations will wait for same amount of time is extremely low. So, collision between contending stations is greatly reduced.

Disadvantage of non-persistent CSMA

- It reduces the bandwidth usage of network. This is because the channel remains idle even if there are stations who have frames to transmit.
- This occurs since each station wait for a random time before attempting retransmission. There may be multiple stations who are waiting while the channel is idle.

p-persistent CSMA

- The last protocol is **p-persistent CSMA**.
- P-persistent CSMA is an approach of Carrier Sense Multiple Access (CSMA) protocol that combines the advantages of 1-persistent CSMA and non-persistent CSMA.
- In p-persistent CSMA, when a transmitting station has a frame to send and it senses a busy channel, it waits for the end of the transmission, and then transmits with a probability p .
- Since, it sends with a probability p , the name p – persistent CSMA is given.

Algorithm

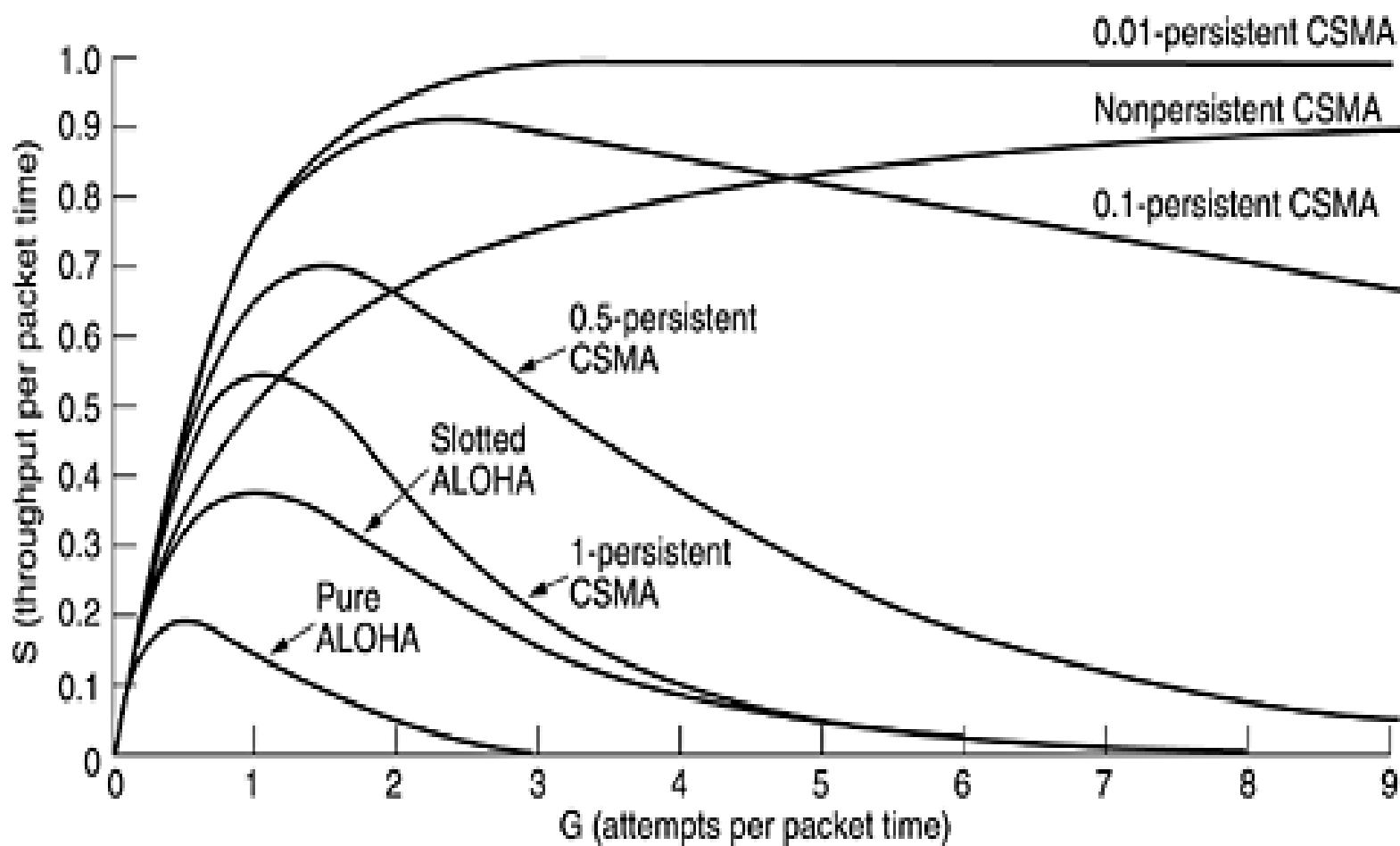
The algorithm of p-persistent CMSA is:

- When a frame is ready, the transmitting station checks whether the channel is idle or busy.
- If the channel is idle then it transmits the frame immediately.
- If the channel is busy, the station waits and continually checks until the channel becomes idle.
- When the channel becomes idle, the station transmits the frame with a probability p .
- With a probability $(1 - p)$, the channel waits for next time slot. If the next time slot is idle, it again transmits with a probability p and waits with a probability $(1 - p)$.
- The station repeats this process until either frame has been transmitted or another station has begun transmitting.
- If another station begins transmitting, the station waits for a random amount of time and restarts the algorithm.

Advantage of p-persistent CSMA

- It is the most efficient among 1-persistent CSMA and non-persistent CSMA.
- It reduces the number of collisions considerably as compared to 1-persistent CSMA. The channel utilization is much better than non-persistent CSMA.

Figure 4-4. Comparison of the channel utilization versus load for various random access protocols.



CSMA with Collision Detection

- Persistent and nonpersistent CSMA protocols are clearly an improvement over ALOHA because they ensure that no station begins to transmit when it senses the channel busy.
- Another improvement is for stations to abort their transmissions as soon as they detect a collision.
- In other words, if two stations sense the channel to be idle and begin transmitting simultaneously, they will both detect the collision almost immediately.

CSMA with Collision Detection

- The collision detection technology detects collisions by sensing transmissions from other stations.
- On detection of a collision, the station stops transmitting, sends a jam signal, and then waits for a random time interval before retransmission.
- Quickly terminating damaged frames saves time and bandwidth. This protocol, known as **CSMA/CD (CSMA with Collision Detection)** is widely used on LANs in the MAC sublayer.

Algorithms

The algorithm of CSMA/CD is:

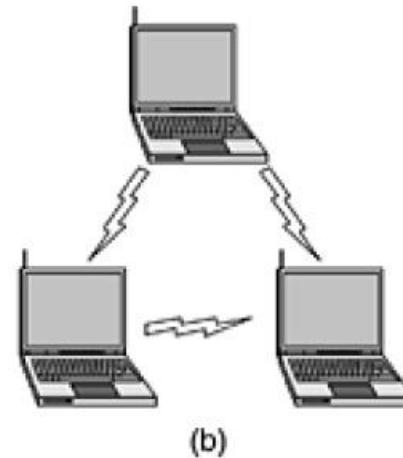
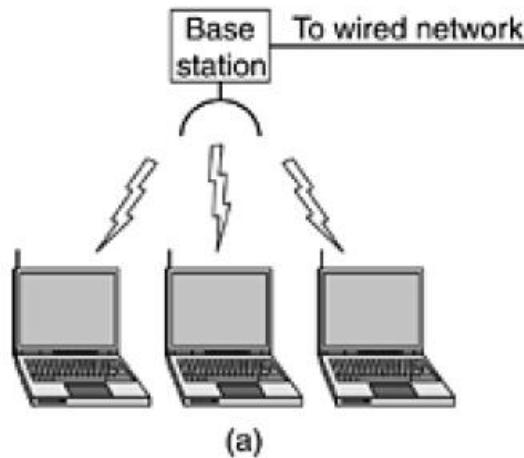
- When a frame is ready, the transmitting station checks whether the channel is idle or busy.
- If the channel is busy, the station waits until the channel becomes idle.
- If the channel is idle, the station starts transmitting and continually monitors the channel to detect collision.
- If a collision is detected, the station starts the collision resolution algorithm.
- The station resets the retransmission counters and completes frame transmission.

The algorithm of Collision Resolution is:

- The station continues transmission of the current frame for a specified time along with a jam signal, to ensure that all the other stations detect collision.
- The station increments the retransmission counter.
- If the maximum number of retransmission attempts is reached, then the station aborts transmission.
- Otherwise, the station waits for a backoff period which is generally a function of the number of collisions and restart main algorithm.

4.2.6 Wireless LAN Protocols

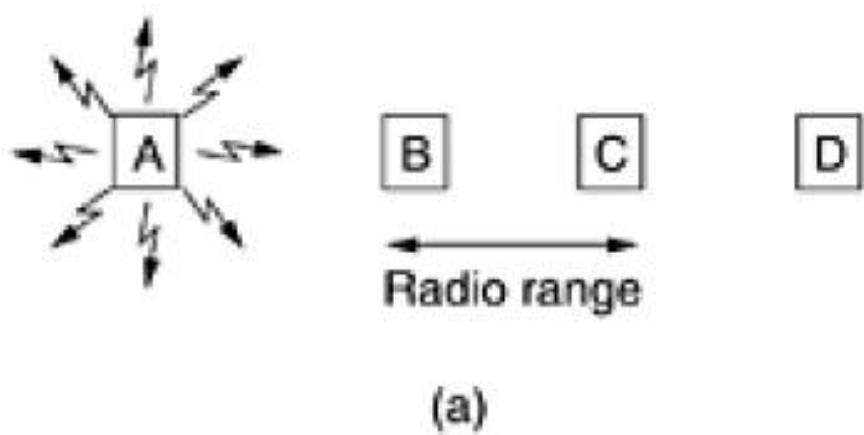
Figure 1-35. (a) Wireless networking with a base station. (b) Ad hoc networking.



- Wireless LANs have somewhat different properties than conventional LANs and require special MAC sublayer protocols.
- 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.
- If the transmission power of the base stations and notebooks is adjusted to have a range of 3 or 4 meters, then each room becomes a single cell and the entire building becomes a large cellular system.

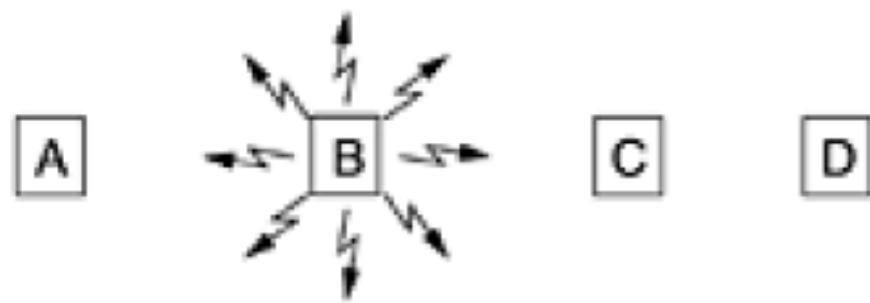
- **Assumption** : all radio transmitters have some fixed range. When a receiver is within range of two active transmitters, the resulting signal will generally be garbled and useless.
- It is important to realize that in some wireless LANs, not all stations are within range of one another, which leads to a variety of complications.
- Furthermore, for indoor wireless LANs, the presence of walls between stations can have a major impact on the effective range of each station.
- One approach in wireless LAN might be to try CSMA: just listen for other transmissions and only transmit if no one else is doing so.

Figure 4-11. A wireless LAN. (a) A transmitting.



- The radio range is such that station *A* and *B* are *within each other's range and can potentially interfere* with one another. *C* can also potentially interfere with both *B* and *D*, but not with *A*.
- **Hidden station problem:** when *A* is transmitting to *B*, If *C* senses the medium, it will not hear *A* because *A* is out of range, and thus falsely conclude 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 medium because the competitor is too far away is called the hidden station problem.*

(b) B transmitting.



(b)

Exposed station problem: Now let us consider the reverse situation: *B transmitting to A, If C* senses the medium, it will hear an ongoing transmission and falsely conclude that it may 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.*

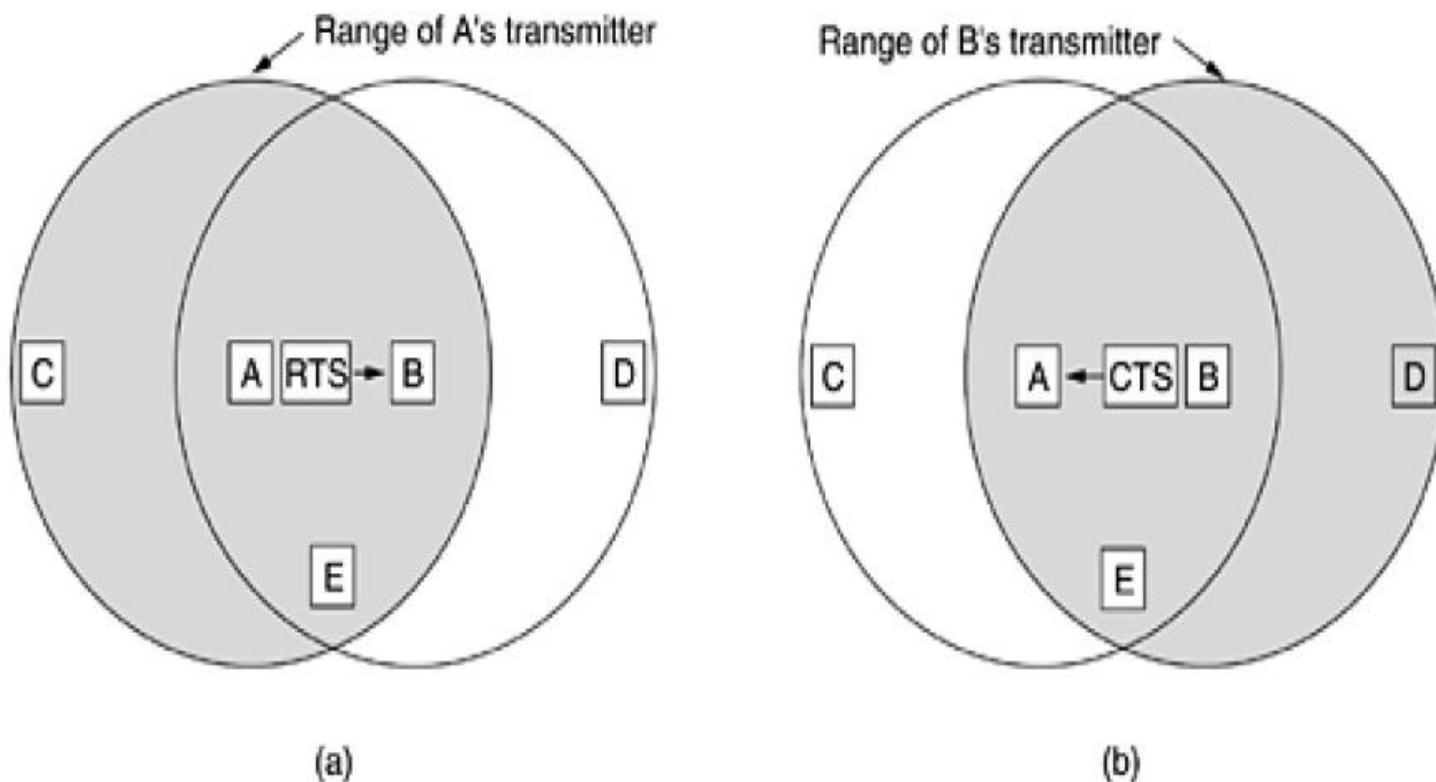
- The problem is that before starting a transmission, a station really wants to know whether there is activity around the receiver.
- CSMA tells it whether there is activity around the station sensing the carrier.
- With a wire, all signals propagate to all stations so only one transmission can take place at once anywhere in the system.
- In a system based on short range radio waves, multiple transmissions can occur simultaneously if they all have different destinations and these destinations are out of range of one another.

MACA and MACAW

- Protocol designed for wireless LANs is:
 - MACA (Multiple Access with Collision Avoidance) Protocol**
 - MACAW (MACA for Wireless) Protocol**

MACA protocol

Figure 4-12. The MACA protocol. (a) A sending an RTS to B. (b) B responding with a CTS to A.



- 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 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*.

- Now let us see how stations overhearing either of these frames react. 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.

- In Fig., *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.
- In contrast, *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.
- Station *E* hears both control messages and, like *D*, must be silent until the data frame is complete.

- Despite these precautions, 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.*

MACAW (MACA for Wireless)

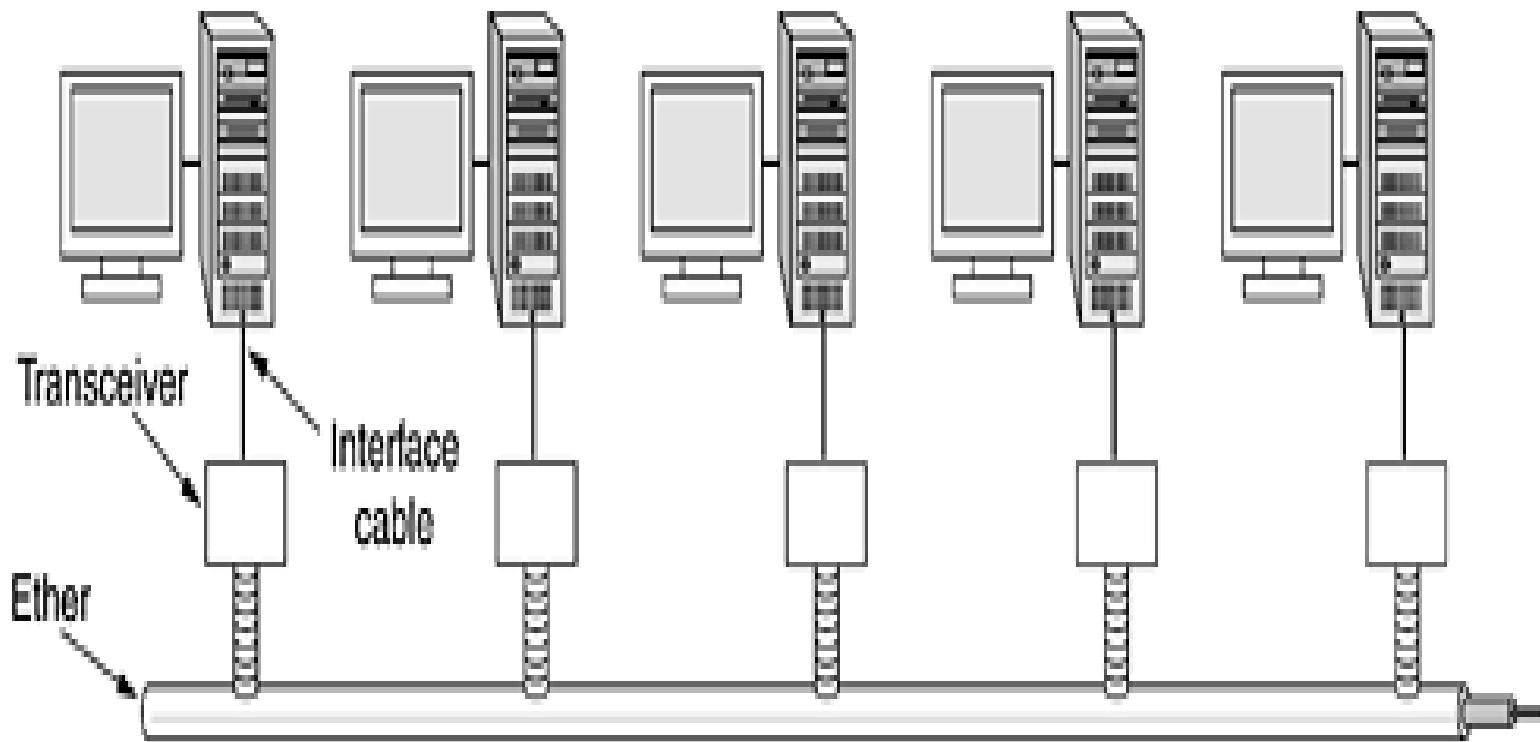
- Bharghavan et al. (1994) fine tuned MACA to improve its performance and renamed their new protocol **MACAW (MACA for Wireless)**. It includes:
 - An ACK frame after each successful data frame.
 - Backoff algorithm separately for each data stream (source-destination pair), rather than for each station. This change improves the fairness of the protocol.
 - A mechanism for stations to exchange information about congestion

MACAW (MACA for Wireless)

- Bharghavan et al. (1994) fine tuned MACA to improve its performance and renamed their new protocol **MACAW (MACA for Wireless)**.
- They noticed that without data link layer acknowledgements, lost frames were not retransmitted until the transport layer noticed their absence.
- They solved this problem by **introducing an ACK frame** after each successful data frame.
- They also observed that CSMA has some use, namely, to keep a station from transmitting an RTS at the same time another nearby station is also doing so to the same destination, so carrier sensing was added.
- In addition, they decided to **run the backoff algorithm separately for each data stream (source-destination pair)**, rather than for each station. This change improves the fairness of the protocol.
- Finally, they **added a mechanism for stations to exchange information about congestion** and a way to make the backoff algorithm react less violently to temporary problems, to improve system performance.

Ethernet

- Ethernet: In Ethernet the transmission medium is a thick coaxial cable (the ether).
- **Ethernet LAN standard: IEEE 802.3.** The committee also standardized a token bus (**802.4**) and a token ring (**802.5**).



Ethernet performance

- The performance of Ethernet under conditions of heavy and constant load, that is, *k stations always ready to transmit.*
- If each station transmits during a contention slot with probability p , *the probability A that some station acquires the channel in that slot is:*

$$A = kp(1 - p)^{k-1}$$

Ethernet performance

- If the mean frame takes P sec to transmit,
when many stations have frames to send,
- The longer the cable, the longer the contention interval, efficiency is less, So the Ethernet standard specifies a maximum cable length.
- Each slot has a duration 2τ ,

$$\text{Channel efficiency} = \frac{P}{P + 2\tau/A}$$

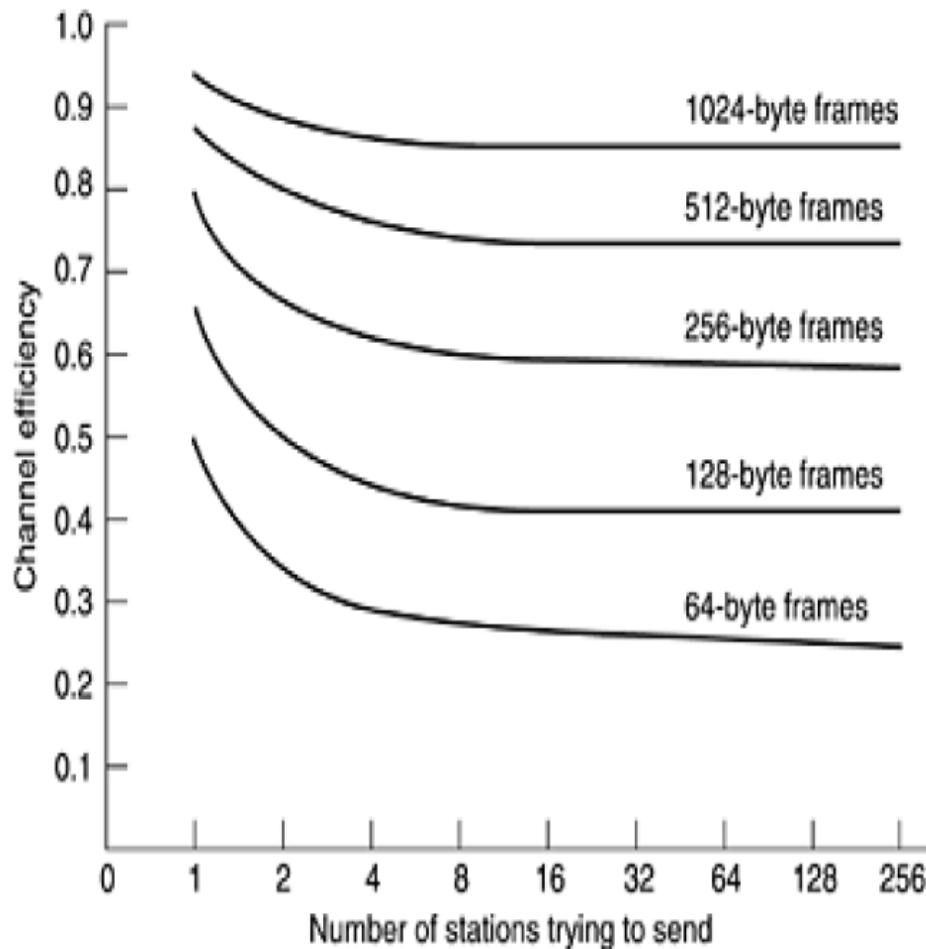
Ethernet performance

- In terms of the frame length, F , the network bandwidth, B , the cable length, L , and the speed of signal propagation, c , for the optimal case of e contention slots per frame. With $P = F/B$:

$$\text{Channel efficiency} = \frac{1}{1 + 2BLe/cF}$$

- When the second term in the denominator is large, network efficiency will be low.
- More specifically, increasing network bandwidth or distance (the BL product) reduces efficiency for a given frame size.
- If high bandwidth over long distances (fiber optic MANs, for example) is required, Ethernet implemented in this manner may not be the best system for these applications.

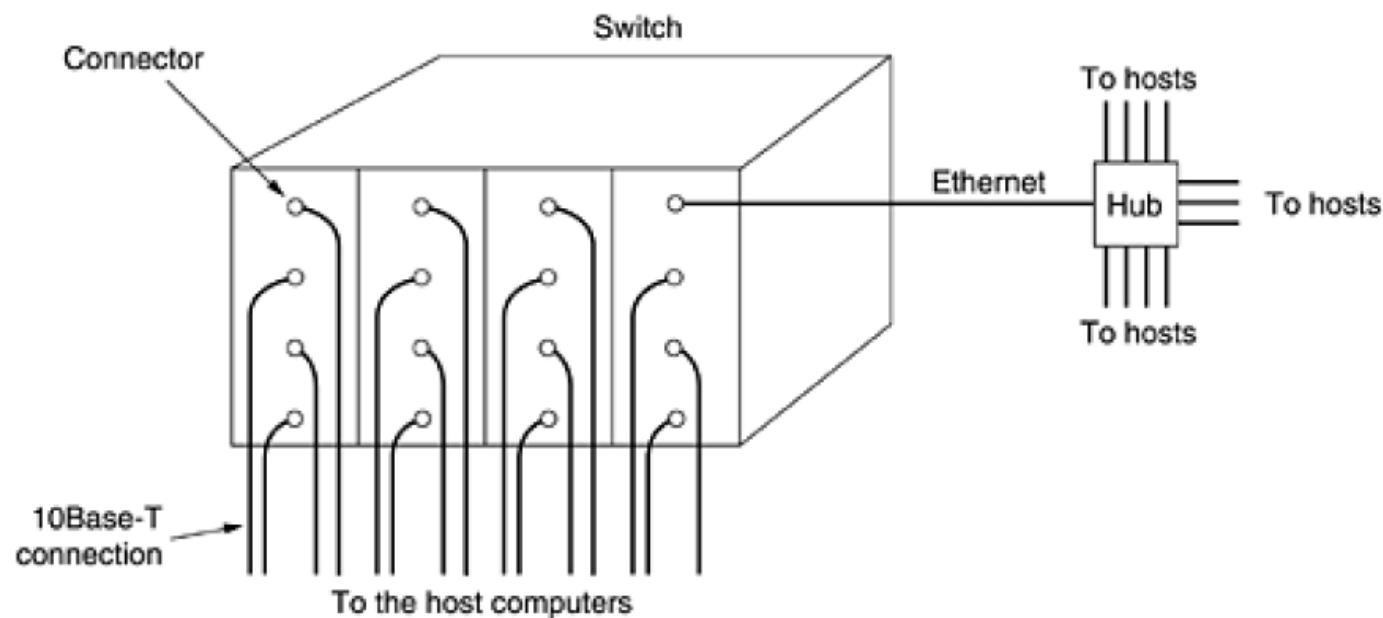
Figure 4-19. Efficiency of Ethernet at 10 Mbps with 512-bit slot times.



Switched Ethernet

- As more and more stations are added to an Ethernet, the traffic will go up.
- But with the growth of multimedia, even a 100-Mbps or 1-Gbps Ethernet can become saturated.
- An additional way to deal with increased load:
switched Ethernet

Figure 4-20. A simple example of switched Ethernet.



- 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 one to eight connectors.
- A backplane is an electrical connector that joins several electrical circuits together. The backplane connectors are parallel to each other in order to link each pin to its relative pin on each connector, forming a complete computer bus.
- Most often, each connector has a 10Base-T twisted pair connection to a single host computer.
- 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?
- 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

Fast Ethernet

- Fast Ethernet is one of the versions of the Ethernet standard.
- It support and provide 100 Mbps data transmission speeds on local area networks (LAN).
- It was launched in 1995 and was the fastest network connection of its time.
- Fast Ethernet is also known as 100 Base X or 100 Mbps Ethernet, and is defined by the IEEE 802.3u protocol.

- The basic idea behind fast Ethernet was simple: keep all the old frame formats, interfaces, and procedural rules, but just reduce the bit time from 100 nsec to 10 nsec.
- It was initially designed for copper-based twisted pair cable networks and included the 100 Base-TX, 100 Base-T4 and 100 Base-T2 standards.
- The length of the cable in copper-based fast Ethernet was restricted to 100 meters and supported different cable categories.

- The fiber-based fast Ethernet standards 100 Base-FX, 100 Base SX, 100 Base BX and 100 Base LX10 use one or more strands and modes of fiber optics to transmit data.
- The range of fast Ethernet for fiber mode is around 2000m.

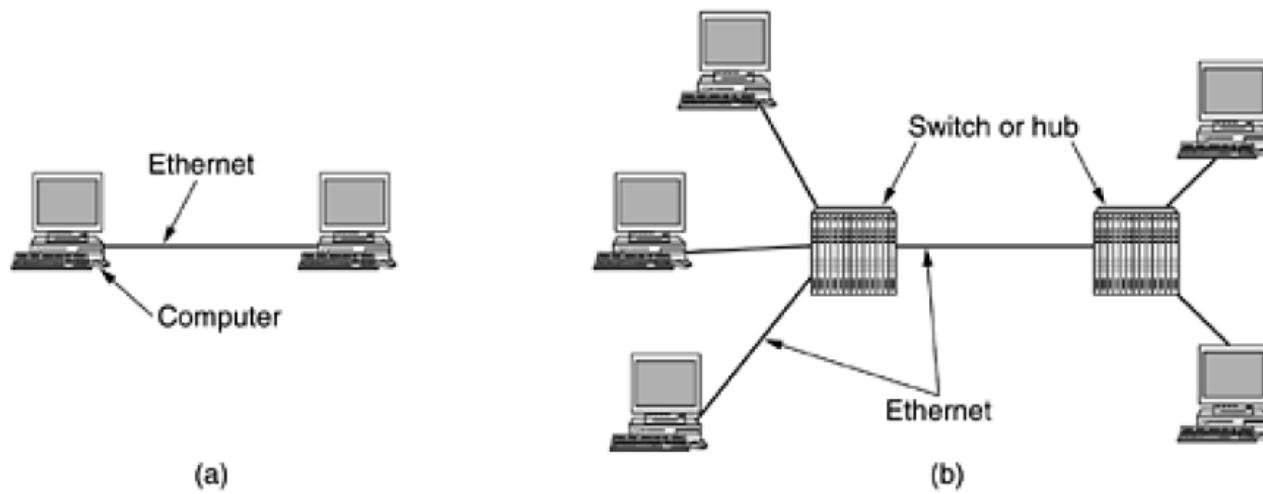
Figure 4-21. The original fast Ethernet cabling.

Name	Cable	Max. segment	Advantages
100Base-T4	Twisted pair	100 m	Uses category 3 UTP
100Base-TX	Twisted pair	100 m	Full duplex at 100 Mbps (Cat 5 UTP)
100Base-FX	Fiber optics	2000 m	Full duplex at 100 Mbps; long runs

Gigabit Ethernet

- Gigabit Ethernet (GbE), a transmission technology based on the Ethernet frame format and protocol used in local area networks (LANs).
- It support and provide 1 gigabit per second (1 Gbps) data transmission speeds.
- It was introduced in 1999 and Defined by IEEE 802.3z standard.
- It is currently being used as the backbone in many networks.
- Newer standards, such as 10 GbE, a networking standard that is 10 times faster than Gigabit Ethernet, are also emerging.

Figure 4-22. (a) A two-station Ethernet. (b) A multistation Ethernet.



- In the simplest gigabit Ethernet configuration, two computers are directly connected to each other.
- The more common case, however, is having a switch or a hub connected to multiple computers and possibly additional switches or hubs.
- In both configurations each individual Ethernet cable has exactly two devices on it, no more and no fewer.
- Gigabit Ethernet supports two different modes of operation: full-duplex mode and half-duplex mode.

- Gigabit Ethernet uses the same 802.3 framing structure as standard Ethernet.
- It supports 1 Gb per second (Gbps) speeds using Carrier Sense Multiple Access/Collision Detect (CSMA/CD).

Figure 4-23. Gigabit Ethernet cabling.

Name	Cable	Max. segment	Advantages
1000Base-SX	Fiber optics	550 m	Multimode fiber (50, 62.5 microns)
1000Base-LX	Fiber optics	5000 m	Single (10 μ) or multimode (50, 62.5 μ)
1000Base-CX	2 Pairs of STP	25 m	Shielded twisted pair
1000Base-T	4 Pairs of UTP	100 m	Standard category 5 UTP

MAC Sublayer Numericals

Problem 1

A group of N stations share a 56-kbps pure ALOHA channel. Each station outputs a 1000-bit frame on an average of once every 100 sec, even if the previous one has not yet been sent (e.g., the stations can buffer outgoing frames). What is the maximum value of N ?

Problem 1

A group of N stations share a 56-kbps pure ALOHA channel. Each station outputs a 1000-bit frame on an average of once every 100 sec, even if the previous one has not yet been sent (e.g., the stations can buffer outgoing frames). What is the maximum value of N ?

Solution:

There are N Stations Sharing 56kbps Pure ALOHA Channel

so with pure ALOHA Usable Bandwidth = $0.184 * 56\text{kbps} = 10.3\text{kbps}$

1 Station Outputs 1000 bits in every 100sec

so in 1sec One station will output at rate $1000/100 = 10\text{bits/sec}$

so For N stations in 1 sec Total Output Data is $10 * N$ bits this should be equal to the Channel Capacity in pure ALOHA

$$N * 10 = 10300$$

$N = 1030$ it is the maximum value of Number of Station Possible.

Problem 2

Ten thousand airline reservation stations are competing for the use of a single slotted ALOHA channel. The average station makes 18 requests/hour. A slot is 125 μ sec. What is the approximate total channel load?

Problem 2

Ten thousand airline reservation stations are competing for the use of a single slotted ALOHA channel. The average station makes 18 requests/hour. A slot is 125 μ sec. What is the approximate total channel load?

Solution:

The average station makes $\frac{18}{3600} = \frac{1}{200}$ requests/sec. The total channel

load is $10000 \times \frac{1}{200} = 50$ requests/sec. Using slot as the time unit, the total channel load is $50 \times (125 \times 10^{-6}) = \frac{1}{160}$ requests/slot.

Problem 3

A slotted aloha system has packets (both new and retransmissions) arriving at a rate of 50 per second. Packets take 40 ms to transmit.

- a) What is G (packets per slot)?**
- b) What is the probability of success of during a slot?**
- c) What is the average number of slots per successful transmission?**

Problem 3

A slotted aloha system has packets (both new and retransmissions) arriving at a rate of 50 per second. Packets take 40 ms to transmit.

- a) What is G (packets per slot)?
- b) What is the probability of success of during a slot?
- c) What is the average number of slots per successful transmission?

Solution:

a) $G = 50 * 0.04 = 2$ packets per slot.

b) $P_s = G e^{-G} = 0.27$

c) $1/P_s = 3.69$

Problem 4

Measurements of a slotted ALOHA channel with an infinite number of users show that 10 percent of the slots are idle.

- (a) What is the channel load, G ?**
- (b) What is the throughput?**
- (c) Is the channel underloaded or overloaded?**

Problem 4

Measurements of a slotted ALOHA channel with an infinite number of users show that 10 percent of the slots are idle.

- (a) What is the channel load, G ?
- (b) What is the throughput?
- (c) Is the channel underloaded or overloaded?

Solution:

a) $P(0) = e^{-G} = 0.1 \Rightarrow G = 2.3$

b) $T = G e^{-G} = 0.23$

c) Since $G > 1$, the system is overloaded.

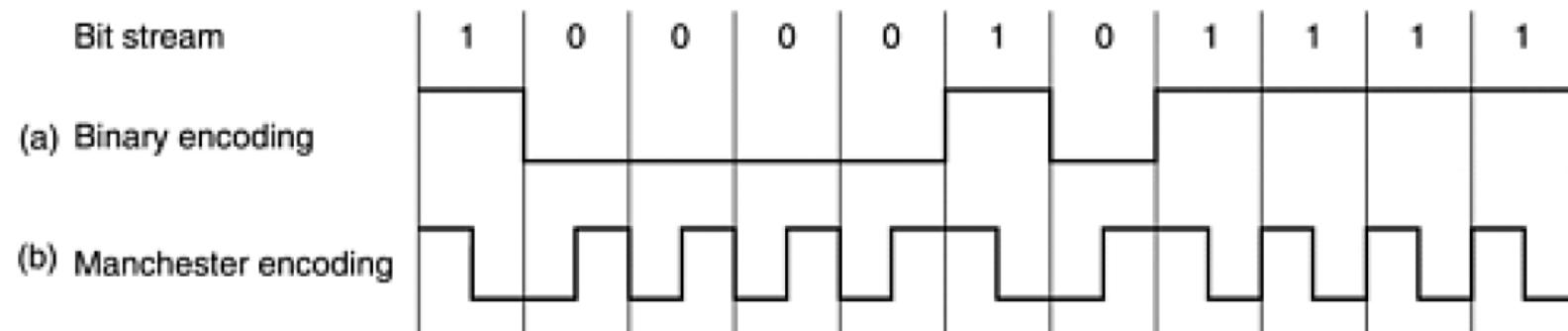
Problem 5

Sketch the Manchester encoding for the bit stream: 10000101111

Problem 5

Sketch the Manchester encoding for the bit stream: 10000101111

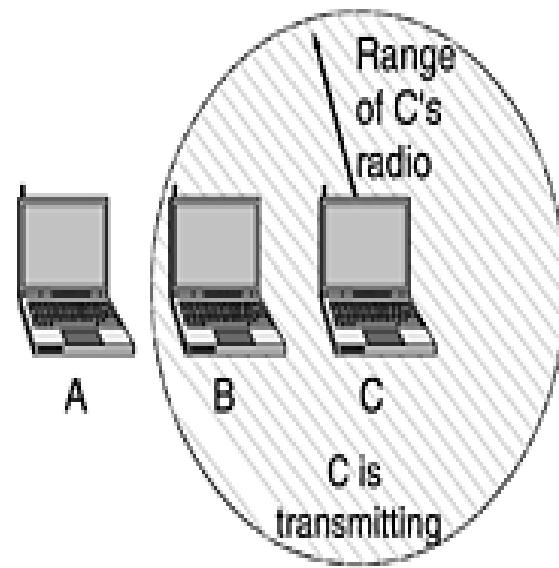
Solution:



The 802.11 MAC Sublayer Protocol

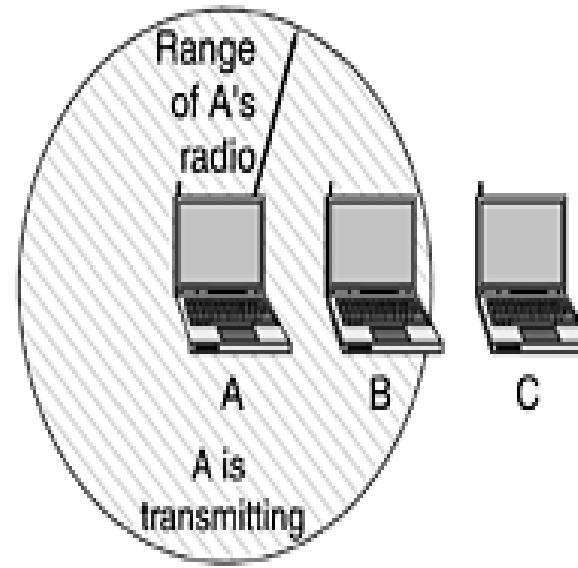
Figure 4-26. (a) The hidden station problem. (b) The exposed station problem.

A wants to send to B
but cannot hear that
B is busy



(a)

B wants to send to C
but mistakenly thinks
the transmission will fail



(b)

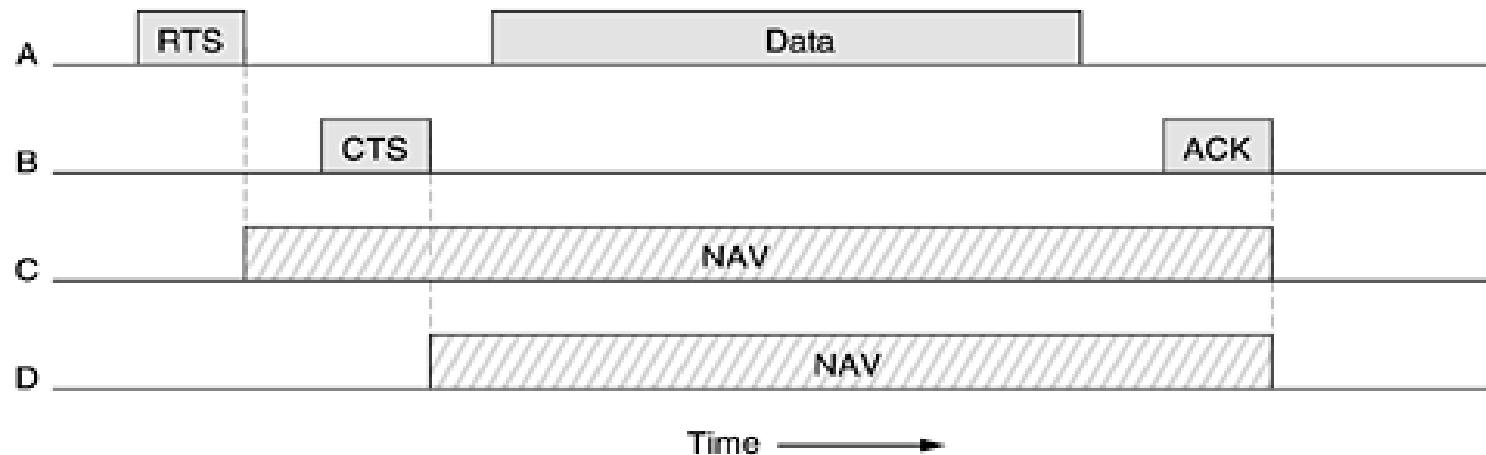
- Most radios are half duplex, meaning that they cannot transmit and listen for noise bursts at the same time on a single frequency.
- As a result of these problems, 802.11 does not use CSMA/CD, as Ethernet does.
- To deal with this problem, 802.11 supports two modes of operation:
 - DCF (Distributed Coordination Function)**, does not use any kind of central control
 - PCF (Point Coordination Function)**, uses the base station to control all activity in its cell.

- All implementations must support DCF but PCF is optional.
- When DCF is employed, 802.11 uses a protocol called **CSMA/CA** (**CSMA with Collision Avoidance**).
- In this protocol, both physical channel sensing and virtual channel sensing are used.
- Two methods of operation are supported by CSMA/CA.

- In the first method, when a station wants to transmit, it senses the channel. If it is idle, it just starts transmitting.
- If the channel is busy, the sender waits until it goes idle and then starts transmitting.
- If a collision occurs, the colliding stations wait a random time, using the Ethernet binary exponential backoff algorithm, and then try again later.
- The other mode of CSMA/CA operation uses **virtual channel sensing**.

- In this example, A wants to send to B. C is a station within range of A . D is a station within range of B but not within range of A.

Figure 4-27. The use of virtual channel sensing using CSMA/CA.

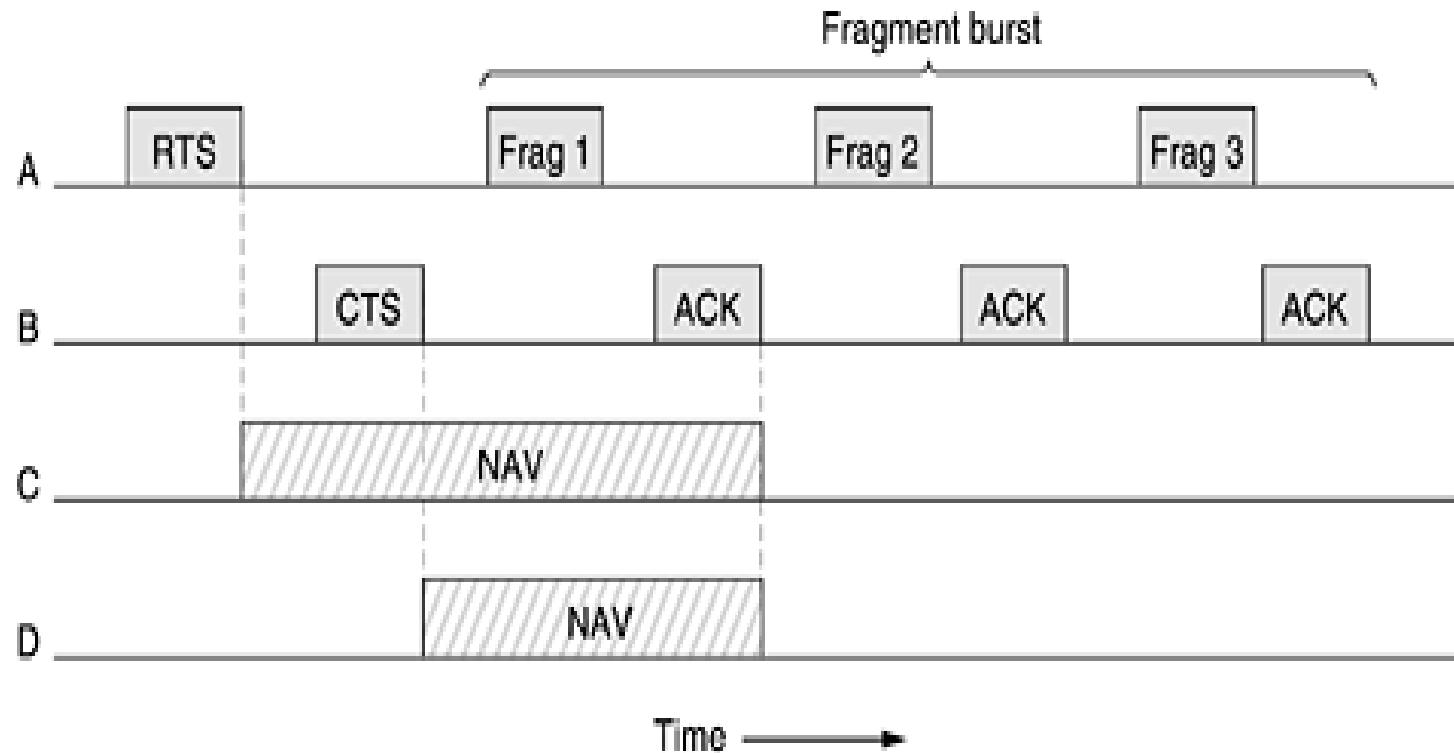


- The protocol starts when A decides it wants to send data to B . It begins by sending an RTS frame to B to request permission to send it a frame.
- When B receives this request, it may decide to grant permission, in which case it sends a CTS frame back.
- Upon receipt of the CTS, A now sends its frame and starts an ACK timer.
- Upon correct receipt of the data frame, B responds with an ACK frame, terminating the exchange.
- If A 's ACK timer expires before the ACK gets back to it, the whole protocol is run again.

- Now let us consider this exchange from the viewpoints of *C* and *D*. *C* is within range of *A*, so it may receive the RTS frame. If it does, it realizes that someone is going to send data soon, so for the good of all it desists from transmitting anything until the exchange is completed.
- From the information provided in the RTS request, it can estimate how long the sequence will take, including the final ACK, so it asserts a kind of virtual channel busy for itself, indicated by **NAV (Network Allocation Vector)**.
- *D* does not hear the RTS, but it does hear the CTS, so it also asserts the NAV signal for itself. Note that the NAV signals are not transmitted; they are just internal reminders to keep quiet for a certain period of time.

- If a frame is too long, it has very little chance of getting through undamaged and will probably have to be retransmitted.
- To deal with the problem of noisy channels, 802.11 allows frames to be fragmented into smaller pieces, each with its own checksum.
- Once the channel has been acquired using RTS and CTS, multiple fragments can be sent in a row. Sequence of fragments is called a **fragment burst**.

Figure 4-28. A fragment burst.

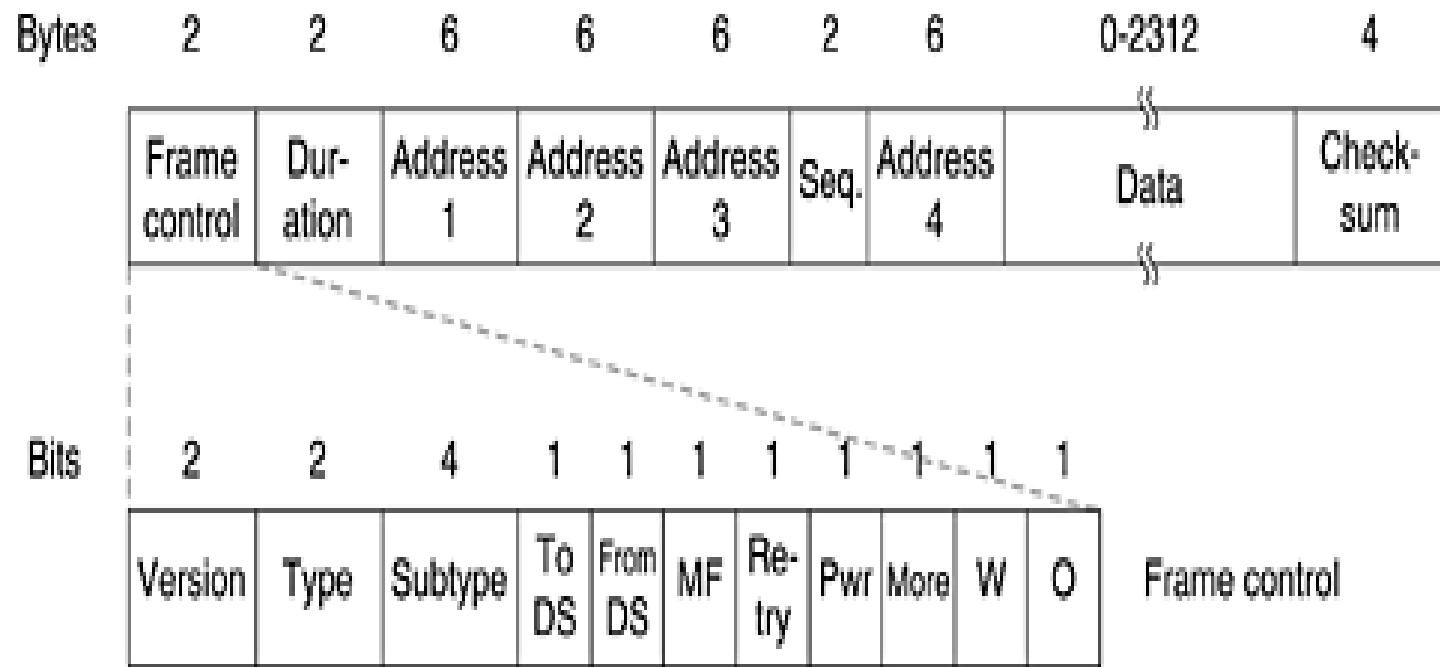


- Fragmentation increases the throughput by restricting retransmissions to the bad fragments rather than the entire frame.
- All of the above discussion applies to the 802.11 DCF(Distributed Coordination Function) mode. In this mode, there is no central control.
- The other allowed mode is PCF (Point Coordination Function), in which the base station polls the other stations, asking them if they have any frames to send.
- Since transmission order is completely controlled by the base station in PCF mode, no collisions ever occur.

The 802.11 Frame Structure

- The 802.11 standard defines **three different classes of frames** on the wire: **data, control, and management**.
- Each of these has a header with a variety of fields used within the MAC sublayer.

Figure 4-30. The 802.11 data frame.



- *Frame Control* field: It itself has 11 subfields.
- ***Protocol version:*** The first of these is the *Protocol version*, which allows two versions of the protocol to operate at the same time in the same cell.
- **Type:** data, control, or management
- **Subtype** fields (e.g., RTS or CTS).
- **To DS , From DS :** The *To DS* and *From DS* bits indicate the frame is going to or coming from the intercell distribution system.
- **MF:** The *MF* bit means that more fragments will follow.

- **Retry:** The *Retry* bit marks a retransmission of a frame sent earlier.
- **Pwr:** The *Power management* bit is used by the base station to put the receiver into sleep state or take it out of sleep state.
- **More:** The *More* bit indicates that the sender has additional frames for the receiver.
- **W:** The *W* bit specifies that the frame body has been encrypted using the **WEP (Wired Equivalent Privacy)** algorithm.
- **O:** Finally, the *O* bit tells the receiver that a sequence of frames with this bit on must be processed strictly in order.

- **Duration:** The second field of the data frame, the *Duration* field, tells how long the frame and its acknowledgement will occupy the channel. This field is also present in the control frames and is how other stations manage the NAV mechanism.
- **Address:** The frame header contains four addresses, all in standard IEEE 802 format.
- The source and destination are obviously needed, but what are the other two for?
- Frames may enter or leave a cell via a base station. The other two addresses are used for the **source and destination base stations** for intercell traffic.

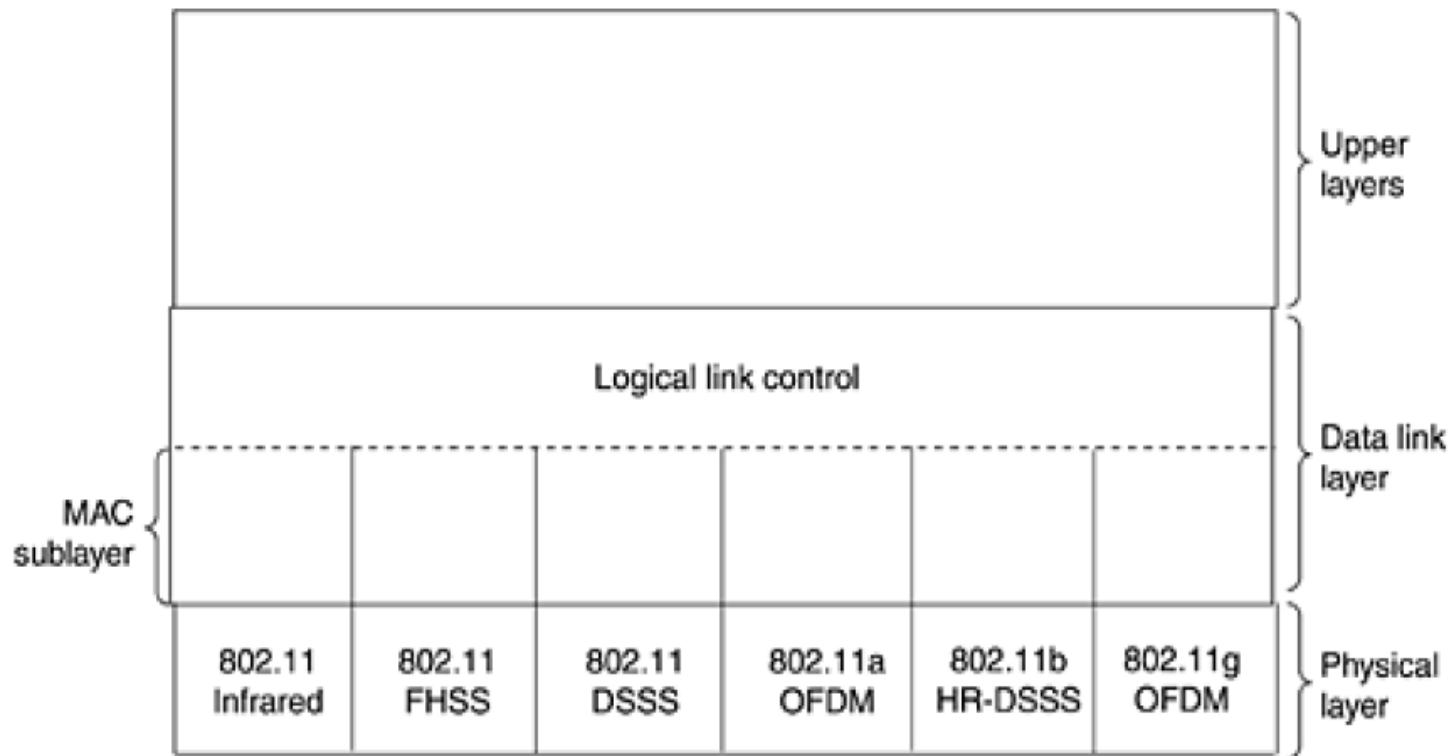
- ***Sequence***: The *Sequence* field allows fragments to be numbered. Of the 16 bits available, 12 identify the frame and 4 identify the fragment.
- ***Data field***: The *Data* field contains the payload, up to 2312 bytes, followed by *Checksum*.

- **Management frames:** Have a format similar to that of data frames, except without one of the base station addresses, because management frames are restricted to a single cell.
- **Control frames** are shorter still, having only one or two addresses, no *Data* field, and no *Sequence* field. The key information here is in the *Subtype* field, usually RTS, CTS, or ACK.

Wireless LANs

The 802.11 Protocol Stack

Figure 4-25. Part of the 802.11 protocol stack.



- The physical layer corresponds to the OSI physical layer fairly well, but the data link layer in all the 802 protocols is split into two or more sublayers.
 - the MAC (Medium Access Control) sublayer
 - Above it is the LLC (Logical Link Control) sublayer

- The 1997 IEEE 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** (Frequency Hopping Spread Spectrum) and **DSSS** (Direct Sequence Spread Spectrum).
- 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** (Orthogonal Frequency Division Multiplexing) and **HR-DSSS** (High Rate Direct Sequence Spread Spectrum)
- 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, in the technology used and speeds achievable.

Infrared

- The infrared option uses transmission at 0.85 or 0.95 microns.
- Two speeds are permitted: 1 Mbps and 2 Mbps. At 1 Mbps.
- Encoding scheme : Gray Code
- Infrared signals cannot penetrate walls, so cells in different rooms are well isolated from each other.
- due to the low bandwidth, this is not a popular option.

FHSS (Frequency Hopping Spread Spectrum)

- FHSS (Frequency Hopping Spread Spectrum) uses 79 channels, each 1-MHz wide, starting at the low end of the 2.4-GHz band.
- A pseudorandom number generator is used to produce the sequence of frequencies hopped to.
- The amount of time spent at each frequency, the **dwell time**, is an adjustable parameter, but must be less than 400 msec.

- FHSS' randomization provides more security.
- Over longer distances, multipath fading can be an issue, and FHSS offers good resistance to it.
- It is also relatively insensitive to radio interference, which makes it popular for building-to-building links.
- Its main disadvantage is its low bandwidth.

DSSS (Direct Sequence Spread Spectrum)

- DSSS (Direct Sequence Spread Spectrum), is also restricted to 1 or 2 Mbps.
- The scheme used has some similarities to the **CDMA** (Code Division Multiple Access) system.
- Each bit is transmitted as 11 chips, using what is called a **Barker sequence**.
- It uses phase shift modulation at 1 Mbaud, transmitting 1 bit per baud when operating at 1 Mbps and 2 bits per baud when operating at 2 Mbps.
- For years, the FCC (Federal Communications Commission) required all wireless communications equipment operating in the ISM(Industrial Scientific and Medical) bands (2 to 6 GHz) in the U.S. to use spread spectrum,
- But in May 2002, that rule was dropped as new technologies emerged.

OFDM (Orthogonal Frequency Division Multiplexing)

- The first of the high-speed wireless LANs, 802.11a, uses OFDM to deliver up to 54 Mbps in the wider 5-GHz ISM band.
- As the term FDM suggests, different frequencies are used—52 of them, 48 for data and 4 for synchronization.
- Since transmissions are present on multiple frequencies at the same time, this technique is considered a form of spread spectrum.

- Splitting the signal into many narrow bands has some key advantages over using a single wide band, including better immunity to narrowband interference and the possibility of using noncontiguous bands.
- A complex encoding system is used, based on phaseshift modulation for speeds up to 18 Mbps and on QAM above that.
- At 54 Mbps, 216 data bits are encoded into 288-bit symbols.
- The technique has a good spectrum efficiency in terms of bits/Hz and good immunity to multipath fading.

HR-DSSS (High Rate Direct Sequence Spread Spectrum)

- It is called **802.11b**
- Data rates supported by 802.11b are 1, 2, 5.5, and 11 Mbps.
- The two slow rates run at 1 Mbaud, with 1 and 2 bits per baud, respectively, using phase shift modulation.
- The two faster rates run at 1.375 Mbaud, with 4 and 8 bits per baud, respectively, using **Walsh/Hadamard codes**.
- The data rate may be dynamically adapted during operation to achieve the optimum speed possible under current conditions of load and noise.
- In practice, the operating speed of 802.11b is nearly always 11 Mbps.
- Although 802.11b is slower than 802.11a, its range is about 7 times greater, which is more important in many situations.

802.11g: OFDM

- An enhanced version of 802.11b, **802.11g**, was approved by IEEE in November 2001 .
- It uses the OFDM modulation method of 802.11a but operates in the narrow 2.4-GHz ISM band along with 802.11b.
- In theory it can operate at up to 54 MBps.
- IEEE 802.11 committee has produced three different highspeed wireless LANs: 802.11a, 802.11b, and 802.11g

Chapter 5. The Network Layer

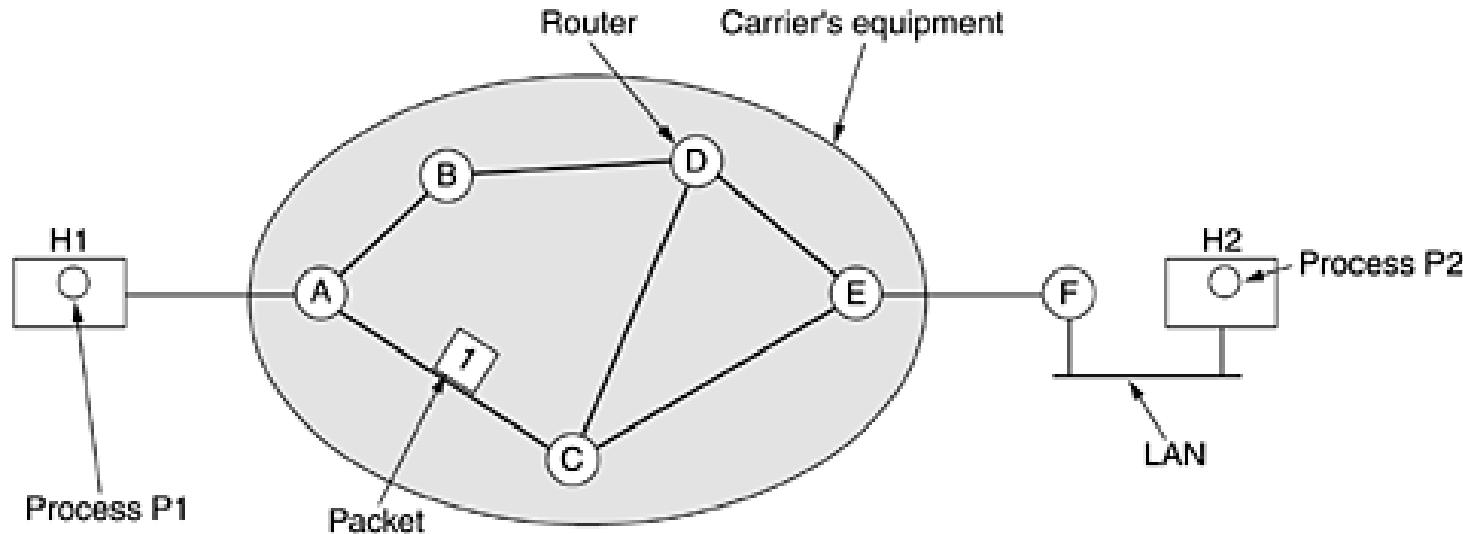
The Network Layer

- The network layer is responsible for packet forwarding including routing through intermediate routers.
- 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

Store-and-Forward Packet Switching

The environment of the network layer protocols



- 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 next router along the path until it reaches the destination host, where it is delivered.
- This mechanism is store-and-forward packet switching.

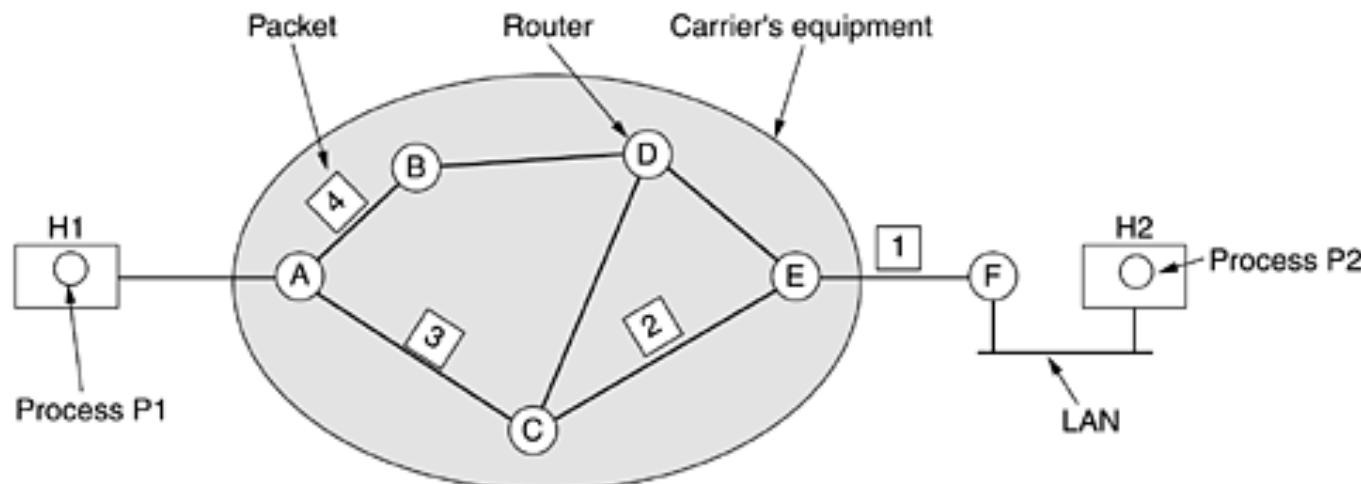
Services Provided to the Transport Layer

- 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.

- Given these goals, the designers of the network layer have a lot of freedom in writing detailed specifications of the services to be offered to the transport layer.
- This freedom often degenerates into a severe battle between two groups: whether the network layer should provide **connection-oriented** service or **connectionless** service.

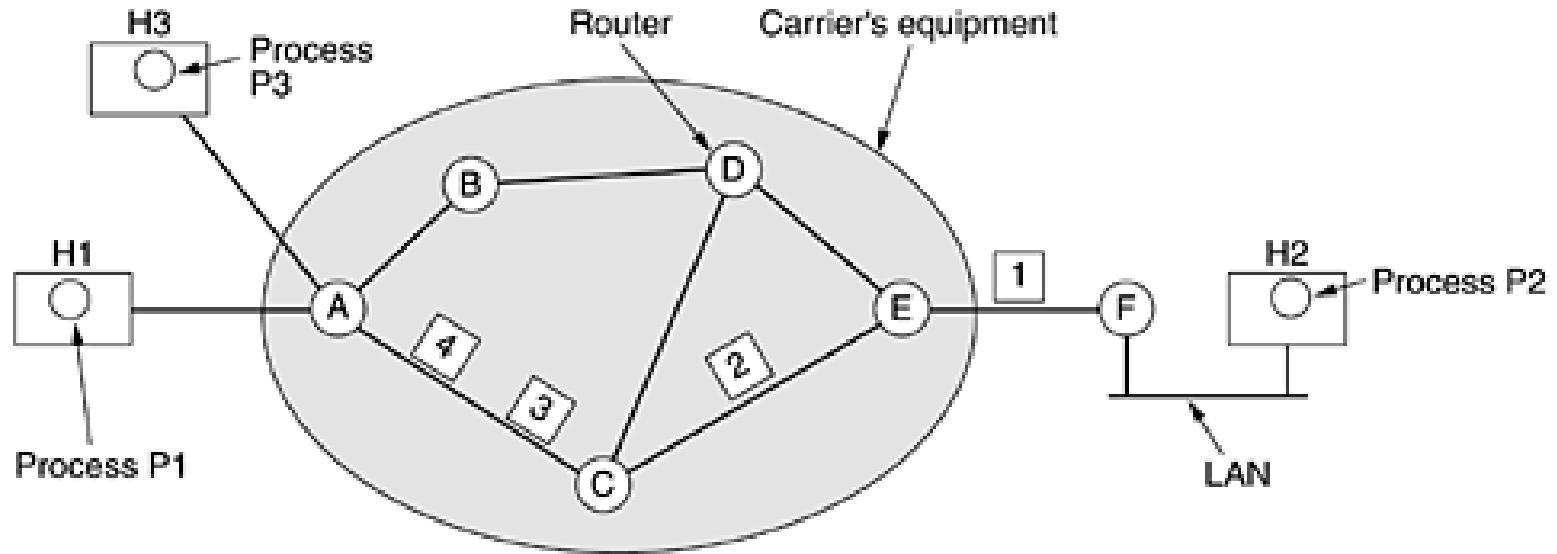
Implementation of Connectionless Service

Routing within a datagram subnet.



- In connectionless service, packets are injected into the subnet individually and routed independently of each other.
- No advance setup is needed. In this context, the packets are frequently called **datagrams** (in analogy with telegrams) and the subnet is called a **datagram subnet**.
- The algorithm that makes the routing decisions is called the **routing algorithm**.

Implementation of Connection-Oriented Service



- In connection-oriented service, a path from the source router to the destination router must be established before any data packets can be sent.
- All packets are routed through same path.
- This connection is called a **VC (virtual circuit)**, in analogy with the physical circuits set up by the telephone system, and the subnet is called a **virtual-circuit subnet**.