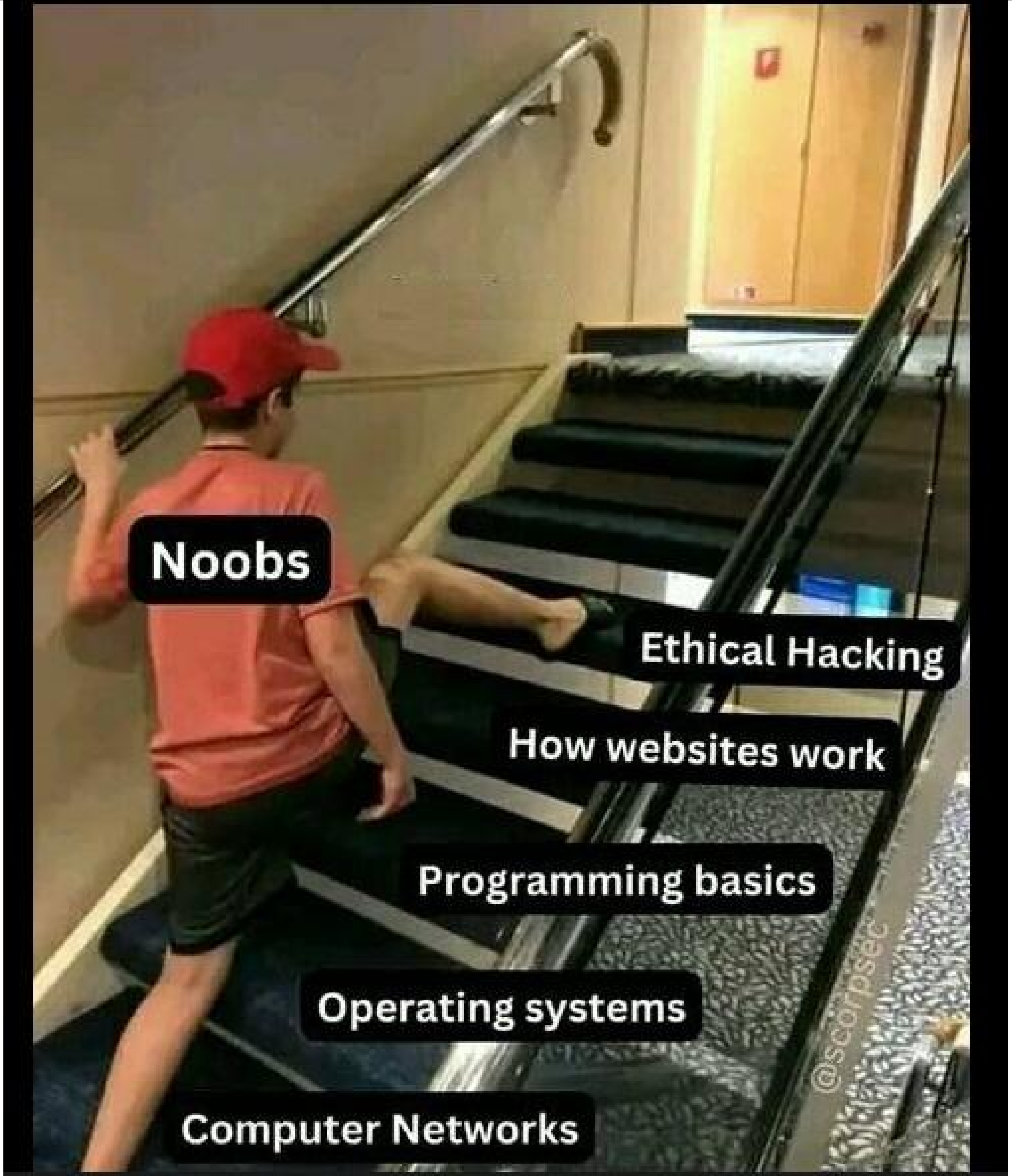
Chapter - 1

- Introduction to Computer Networks -
- Analog signals and Digital Signals -
- Data Transmission and Multiplexing -
- Data Encoding Techniques
- Packet Switching and Circuit S witching -
- Network Topologies -
- Reference Models: ISO/OSI Model and TCP/IP Model. -


Noobs
Ethical Hacking
How websites work
Programming basics
Operating systems
Computer Networks
@scorpsec

motivation

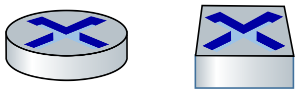
- The first thing that comes to our mind when someone talks about computer networks in the internet.
- But what internet actually is? and how it works?
- In this course, we will talk about the internal workings of internet.
- The actual definition of the internet can vary depending upon whom you are asking.

https://www.youtube.com/watch?v=Dxcc6ycZ73M&ab_channel=Code.org

The Internet: a "nuts and bolts" view

Billions of connected computing *devices*:

- *hosts = end systems*
- running *network apps* at Internet's "edge"



Packet switches: forward packets (chunks of data)

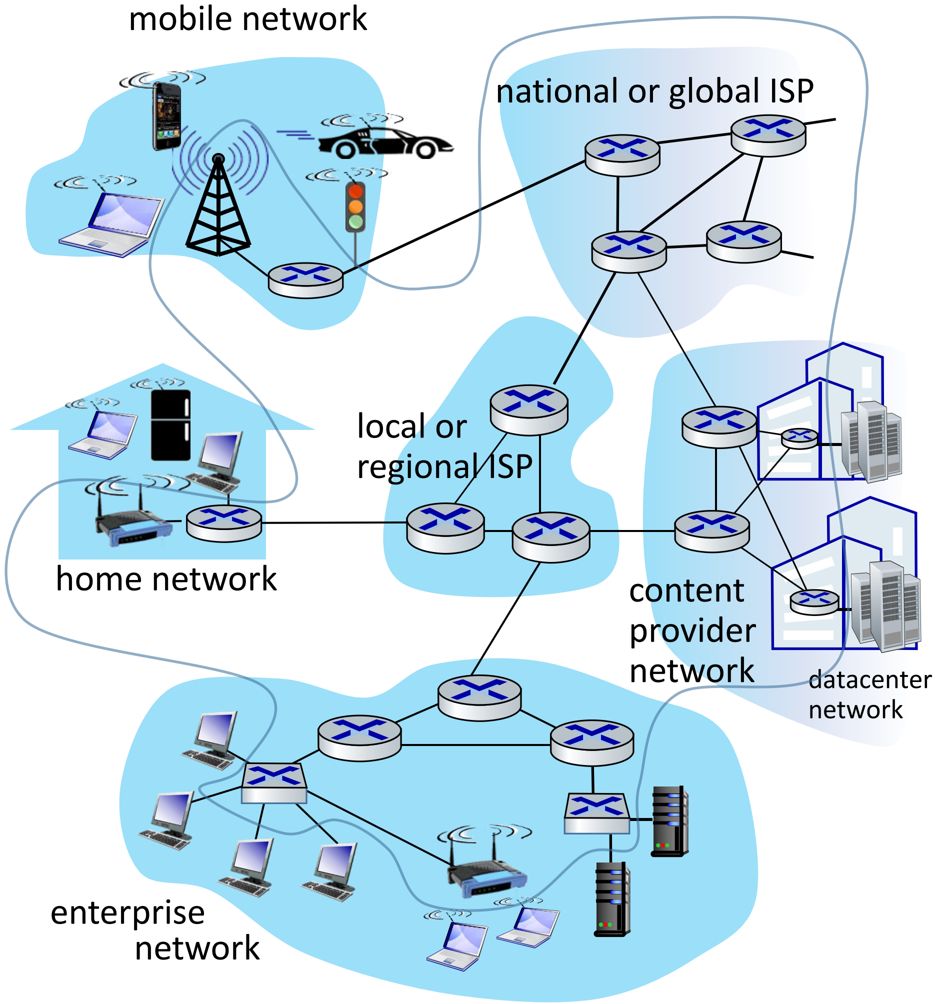
- *routers, switches*

Communication links

- fiber, copper, radio, satellite
- transmission rate: *bandwidth*

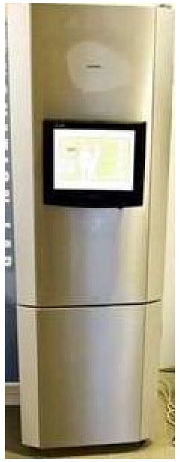
Networks

- collection of devices, routers, links: managed by an organization



“Fun” Internet-connected devices

Amazon Echo



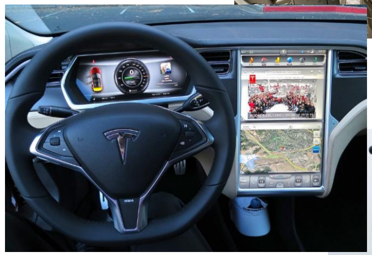
Internet refrigerator

IP picture frame

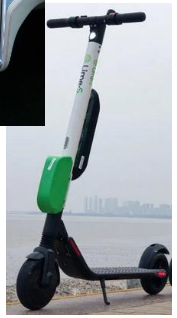
Pacemaker & Monitor

Tweet-a-watt: monitor energy use

bikes



cars



scooters

Web-enabled toaster + weather forecaster

Security Camera

Slingbox: remote control cable TV

AR devices

Fitbit

diapers

Internet phones

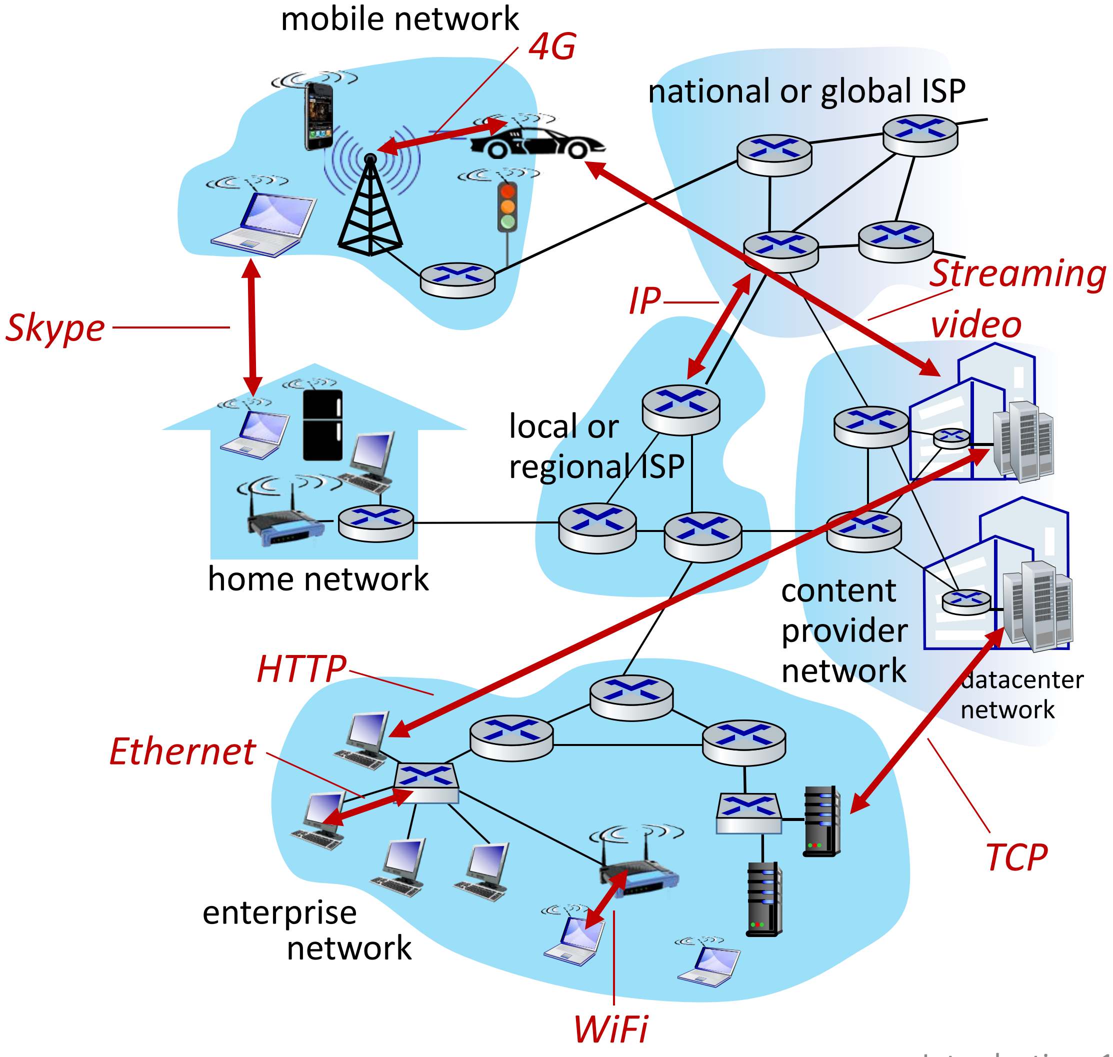
Gaming devices

sensorized, bed mattress

Others?

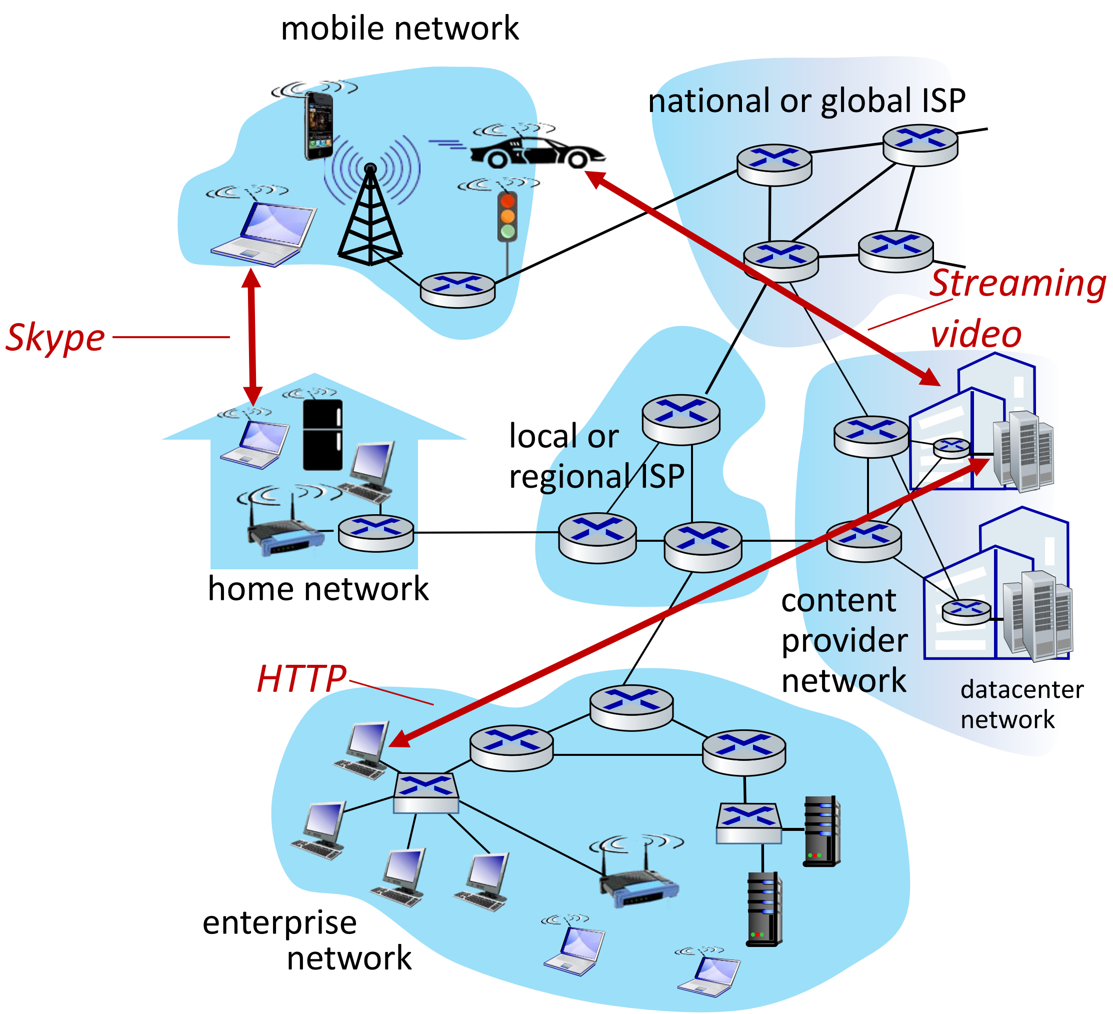
The Internet: a “nuts and bolts” view

- *Internet: “network of networks”*
  - Interconnected ISPs
- *protocols* are *everywhere*
  - control sending, receiving of messages
  - e.g., HTTP (Web), streaming video, Skype, TCP, IP, WiFi, 4/5G, Ethernet
- *Internet standards*
  - RFC: Request for Comments
  - IETF: Internet Engineering Task Force



The Internet: a "services" view

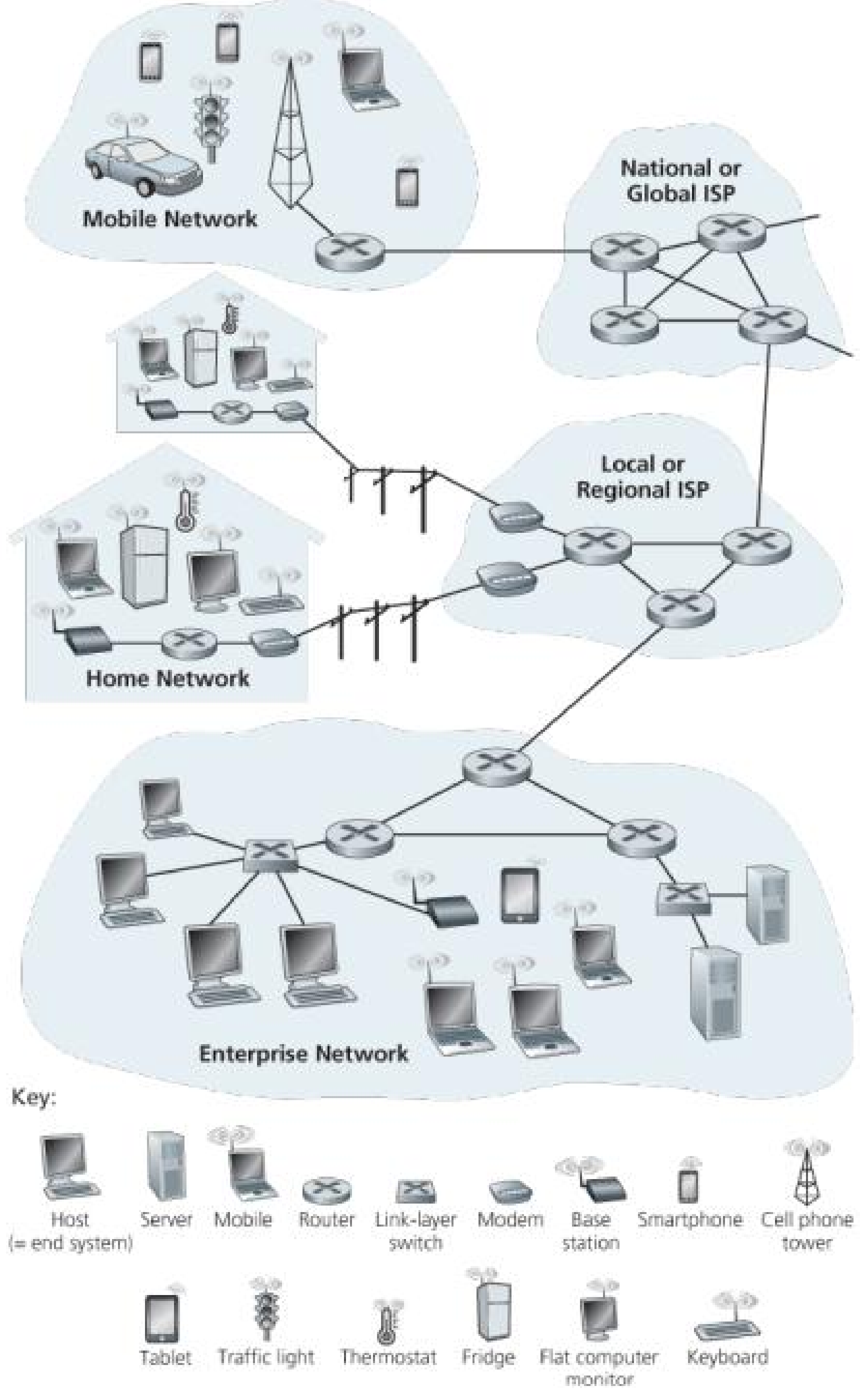
- *Infrastructure* that provides services to applications:
  - Web, streaming video, multimedia teleconferencing, email, games, e-commerce, social media, inter-connected appliances, …
- provides *programming interface* to distributed applications:
  - "hooks" allowing sending/receiving apps to "connect" to, use Internet transport service
  - provides service options, analogous to postal service



what is a computer network?

- A computer network is the **interconnection of a set of devices capable of communication**.
- These devices can be either a **host** (also called **end system** sometimes) such as a large computer, desktop, laptop, workstation, cellular phone, TV, etc.
- Or devices can be **connecting device** such as a router which connects the network to other network, a switch which connects devices together, a modem (modulator-demodulator) that changes the form of data, and so on.
- These devices in a network are connected using wired or wireless transmission media such as ethernet, WiFi, etc.

a computer network

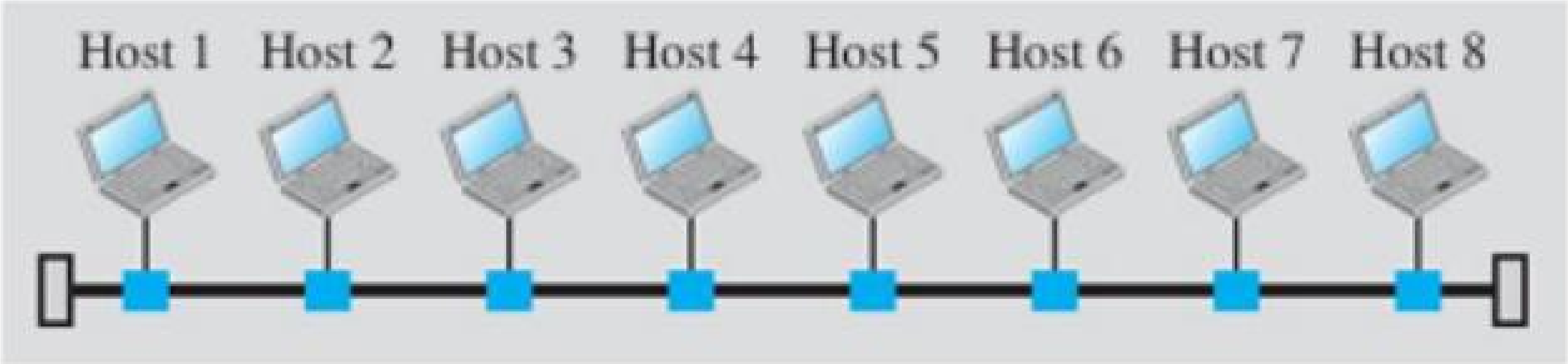


types of computer networks

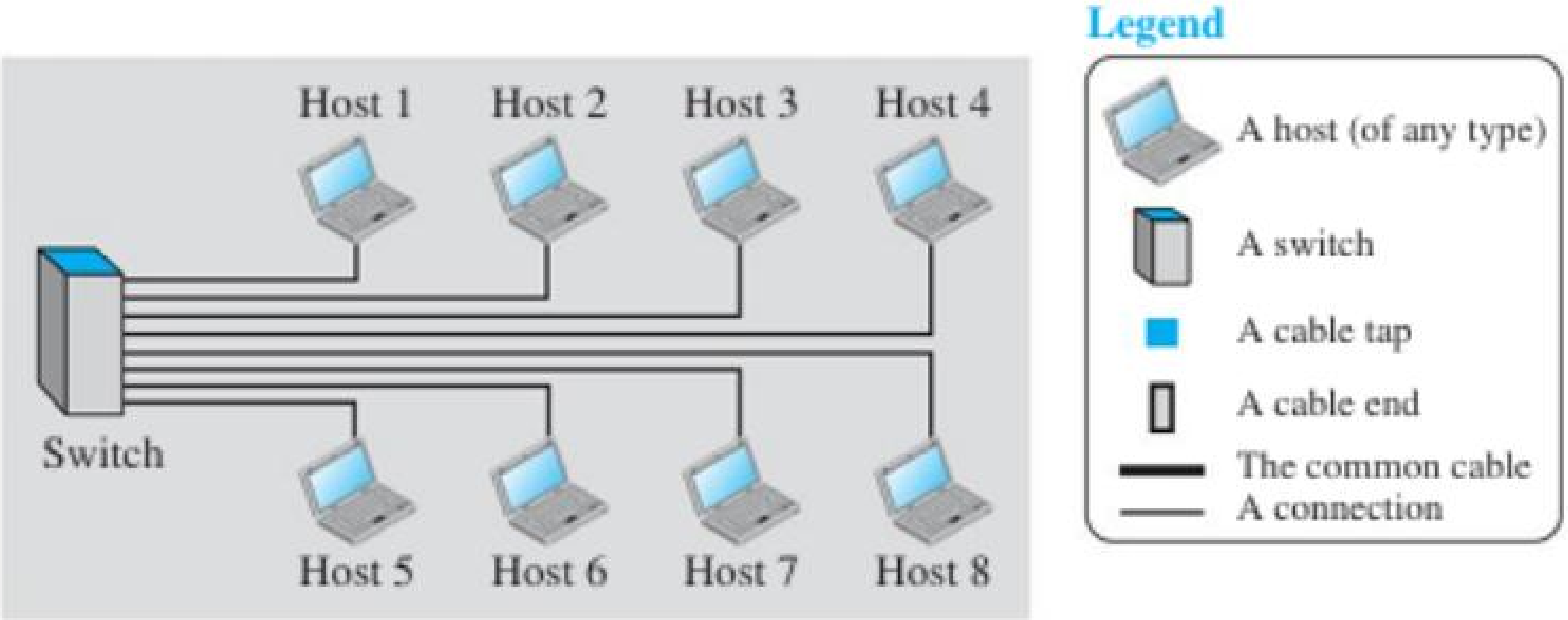
- Depending upon size and connection range, networks can be devided into several categories -
  - Personal Area Network (PAN)
  - Local Area Network (LAN)
  - Campus Area Network (CAN)
  - Metropolitan Area Network (MAN)
  - Wide Area Network (WAN)
    - Point-to-Point WAN
    - Switched WAN

Types of networks

- **Personal Area Network (PAN)**
  - A PAN is the smallest type of network, covering a very limited area, typically within the range of an individual person.
  - Range - Up to 10 meters
  - Devices - Smartphones, laptops, tablets, headphones, etc.
  - It uses Bluetooth and USB for connections.
- **Local Area Network (LAN)**
  - A LAN covers a small geographic area such as a single building, office, or home. It is used to connect computers and devices within close proximity.
  - Range - Up to a few kilometers.
  - Devices - Computers, servers, printers, etc.



a. LAN with a common cable (past)



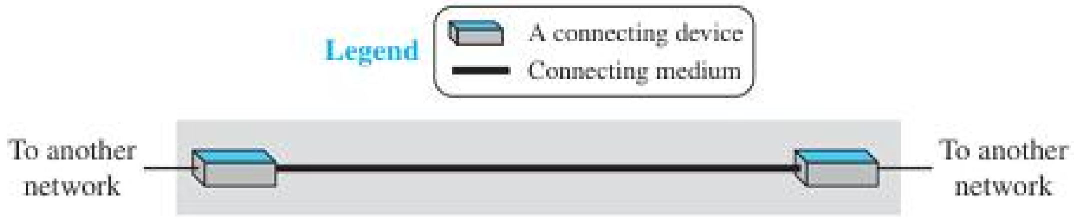
b. LAN with a switch (today)

types of networks

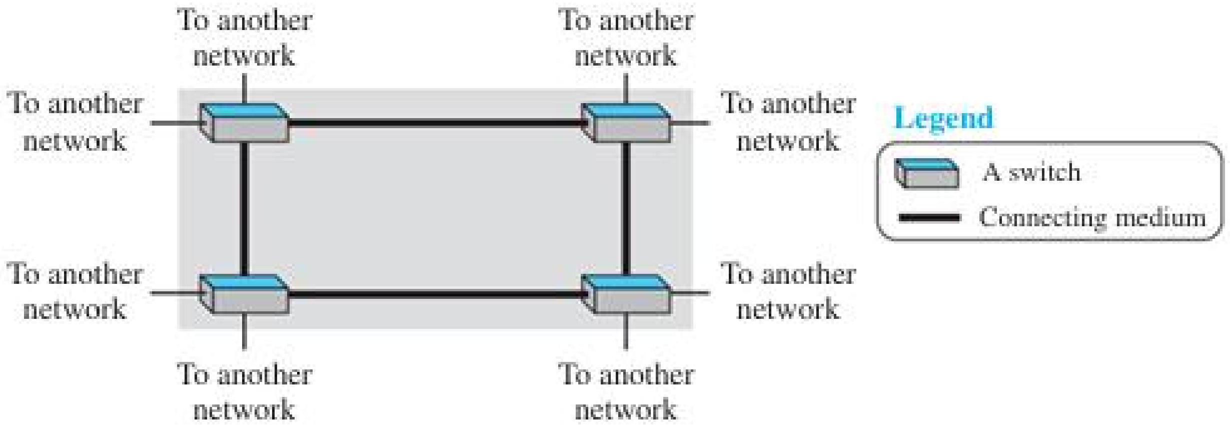
- **Metropolitan Area Network (MAN)**
  - A MAN covers a large geographic area than a LAN, typically spanning a city or a large campus. It connects multiple LANs within a specific region.
  - Range - Up to 50 kilometers.
  - Devices - Network infrastructure such as routers, switches, etc.
- **Wide Area Network (WAN)**
  - A WAN covers a broad area, often a country or continent. It connects multiple LANs and MANs.
  - Range - Thousands of kilometers.

wide area network

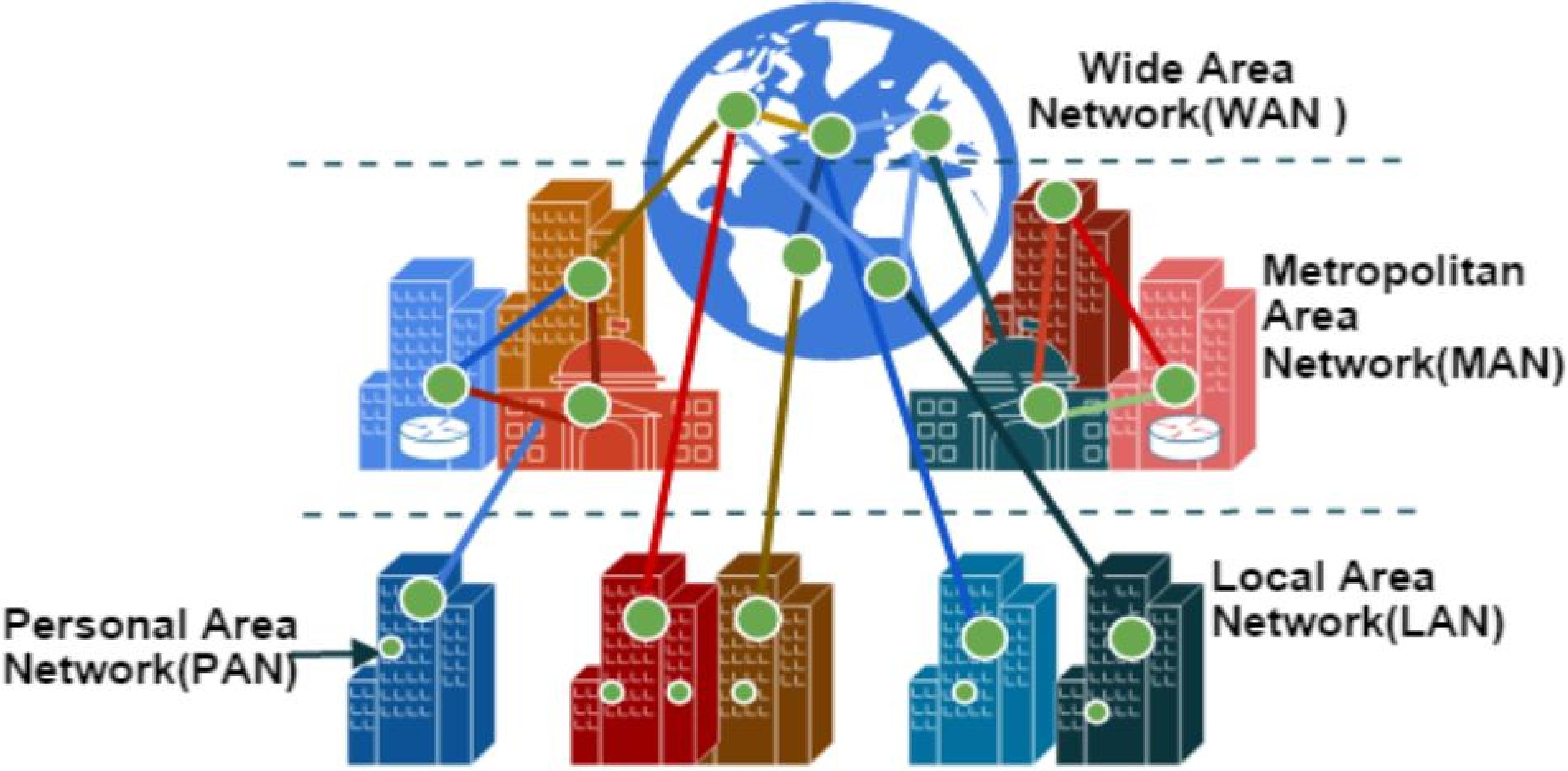
- **Point-to-Point WAN:** A point-to-point WAN is a network that connects two communicating devices through a transmission media (cable or air).



➢ **Switched WAN:**

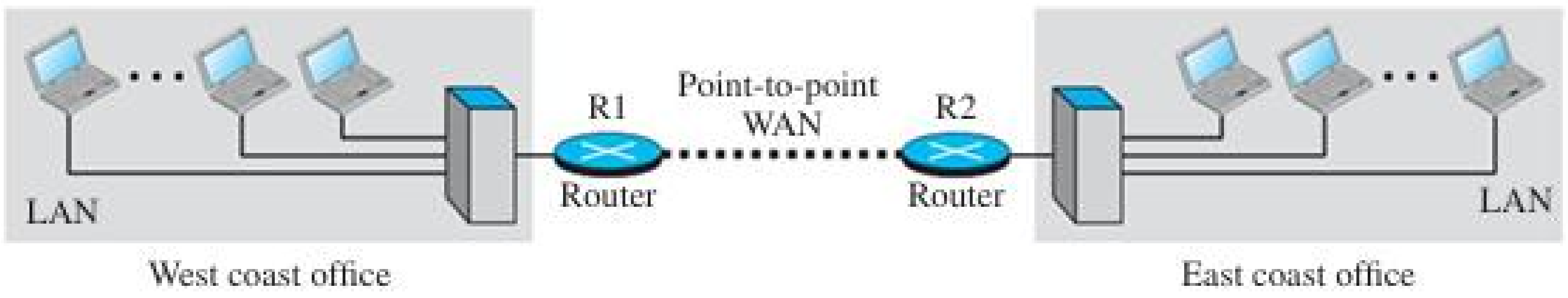


Types of Computer Networks

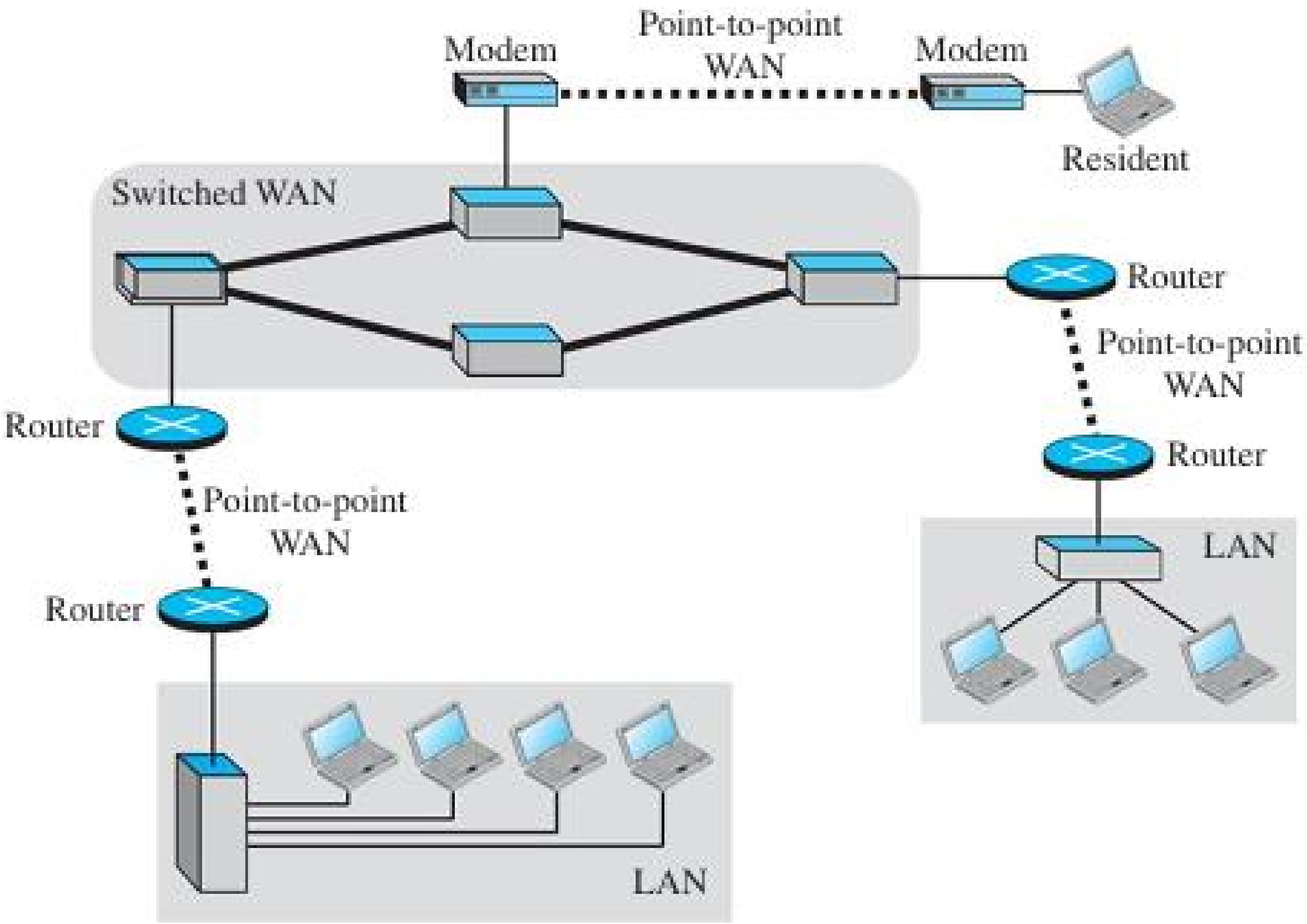


inter-network

- Today, it is very rare to see a LAN or a WAN in isolation; they are connected to one another. When two or more networks are connected, they make an *internetwork,* or **internet.**
- As an example, assume that an organization has two offices, one on the east coast and the other on the west coast. Each office has a LAN that allows all employees in the office to communicate with each other. To make the communication between employees at different offices possible, the management leases a point-to-point dedicated WAN from a service provider, such as a telephone company, and connects the two LANs. Now the company has an internetwork, or a private internet (with lowercase i).



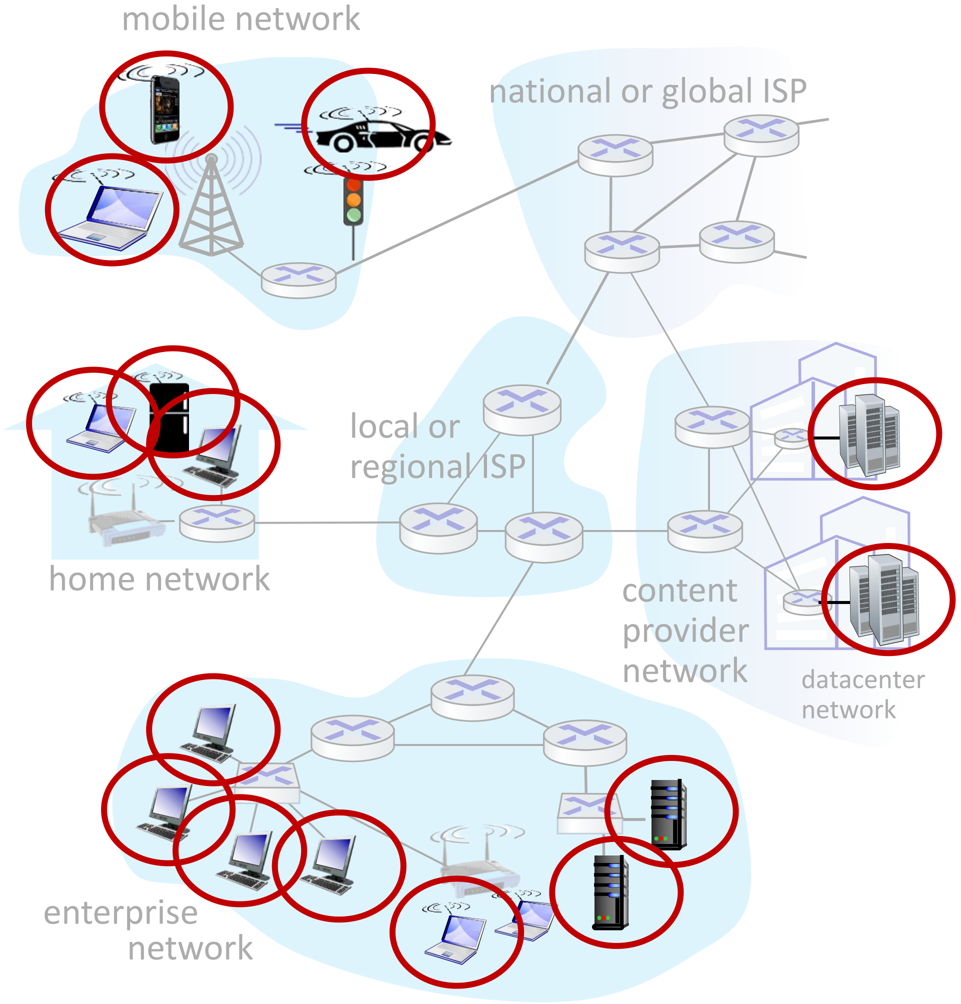
inter-network



A closer look at Internet structure

Network edge:

- hosts: clients and servers
- servers often in data centers



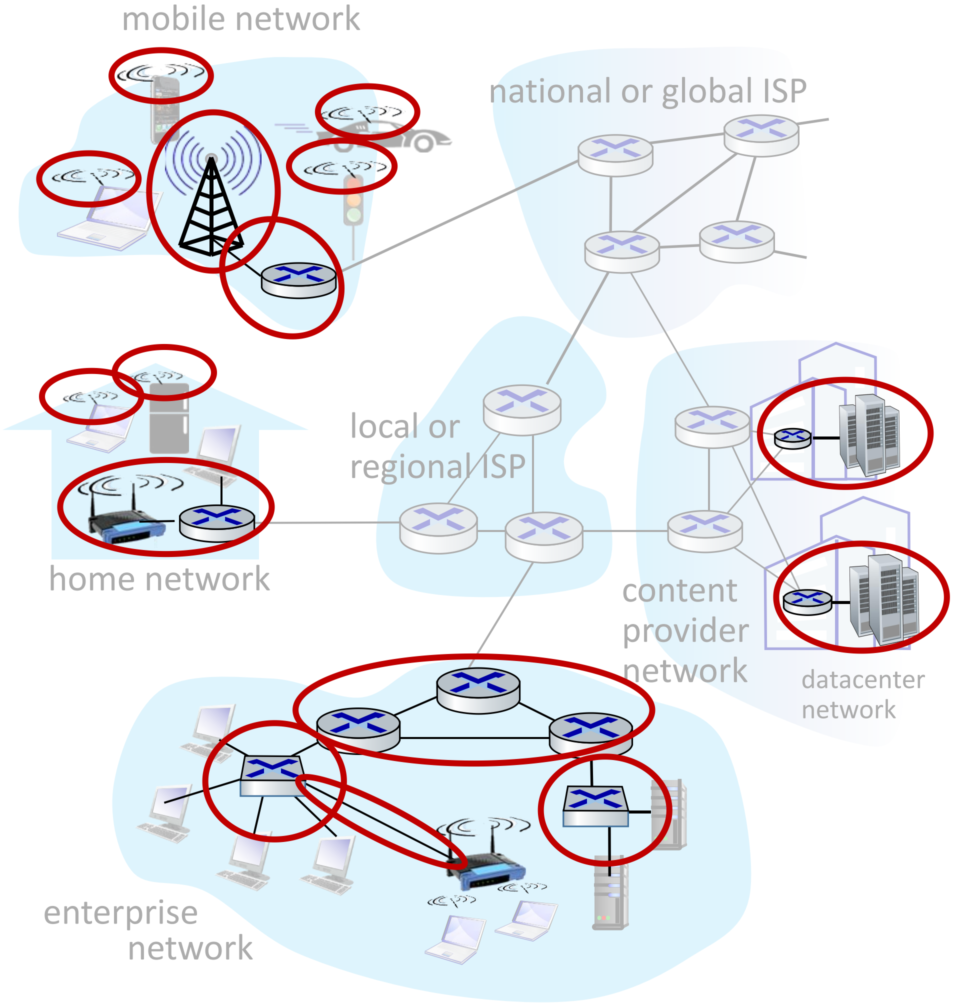
A closer look at Internet structure

Network edge:

- hosts: clients and servers
- servers often in data centers

Access networks, physical media:

- wired, wireless communication links



A closer look at Internet structure

Network edge:

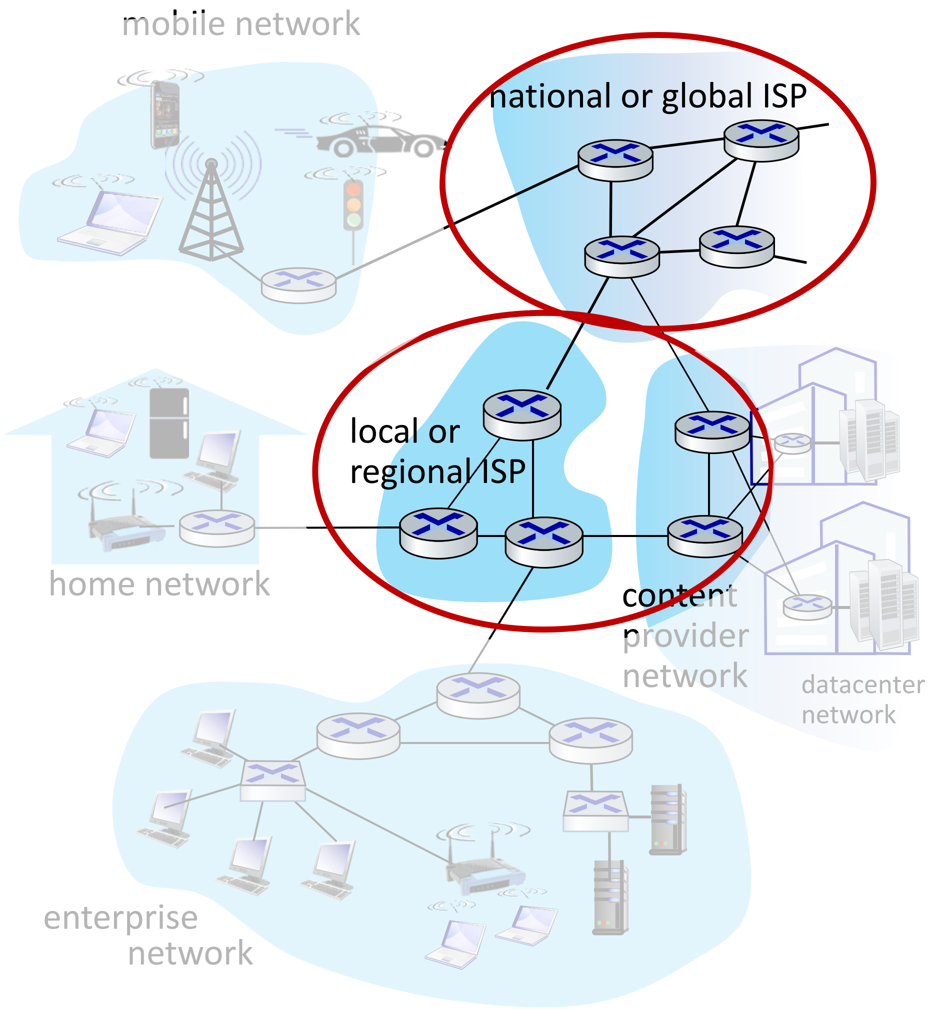
- hosts: clients and servers
- servers often in data centers

Access networks, physical media:

- wired, wireless communication links

Network core:

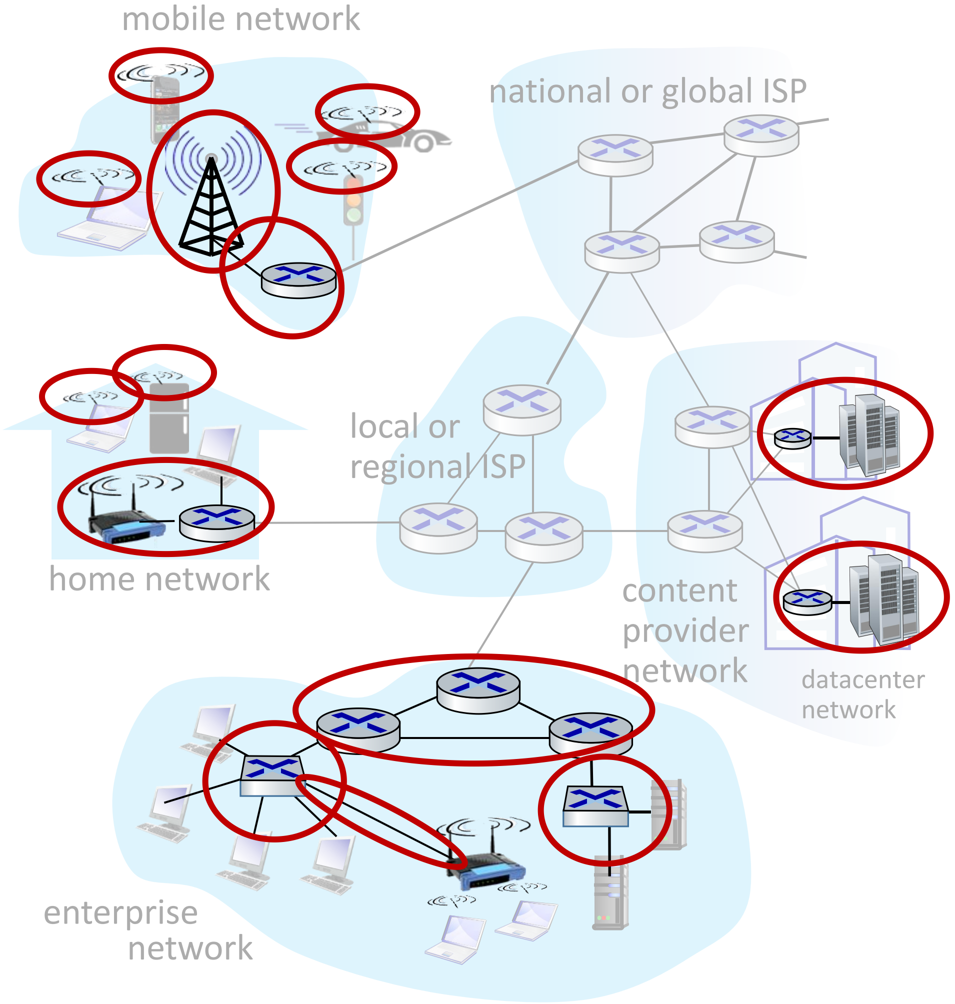
- interconnected routers
- network of networks



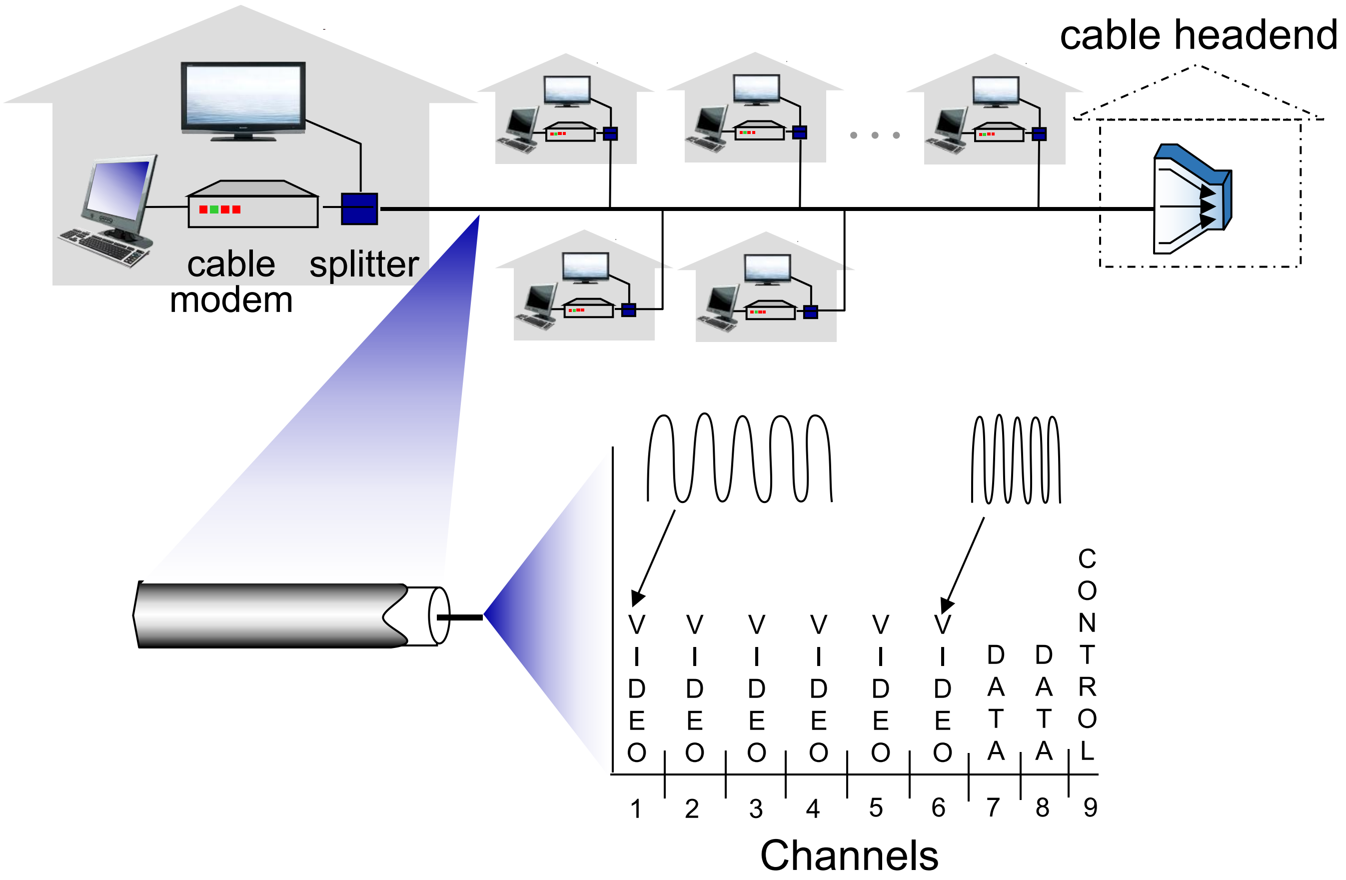
Access networks and physical media

Q: How to connect end systems to edge router?

- residential access nets
- institutional access networks (school, company)
- mobile access networks (WiFi, 4G/5G)

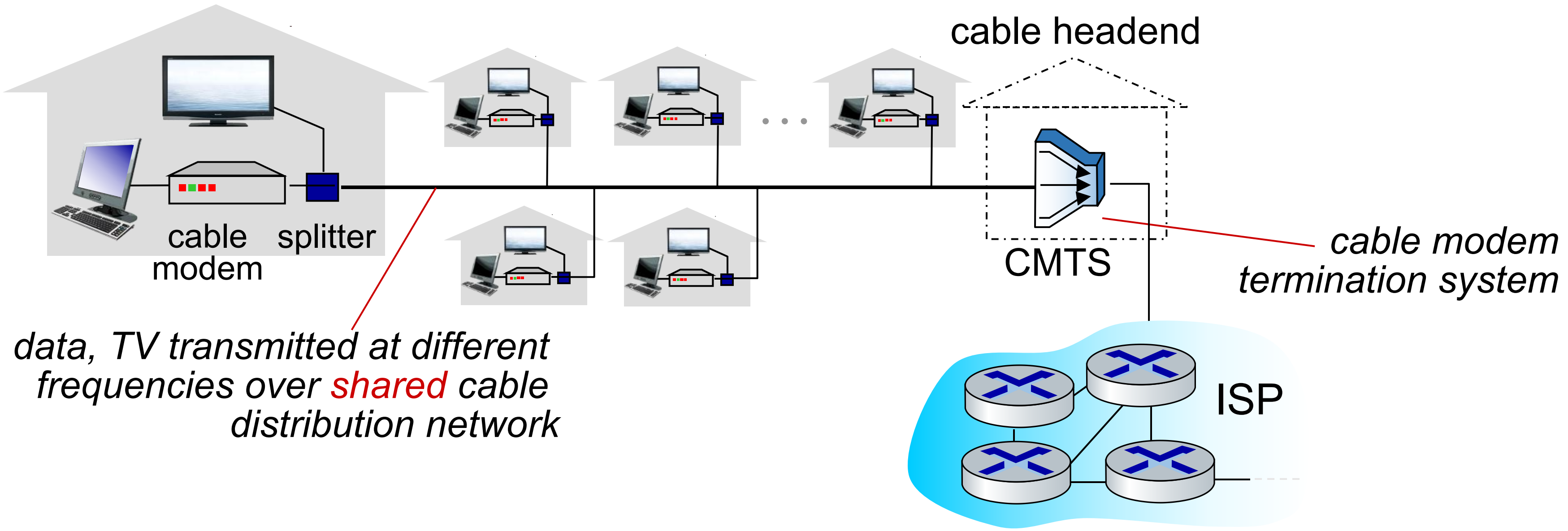


Access networks: cable-based access



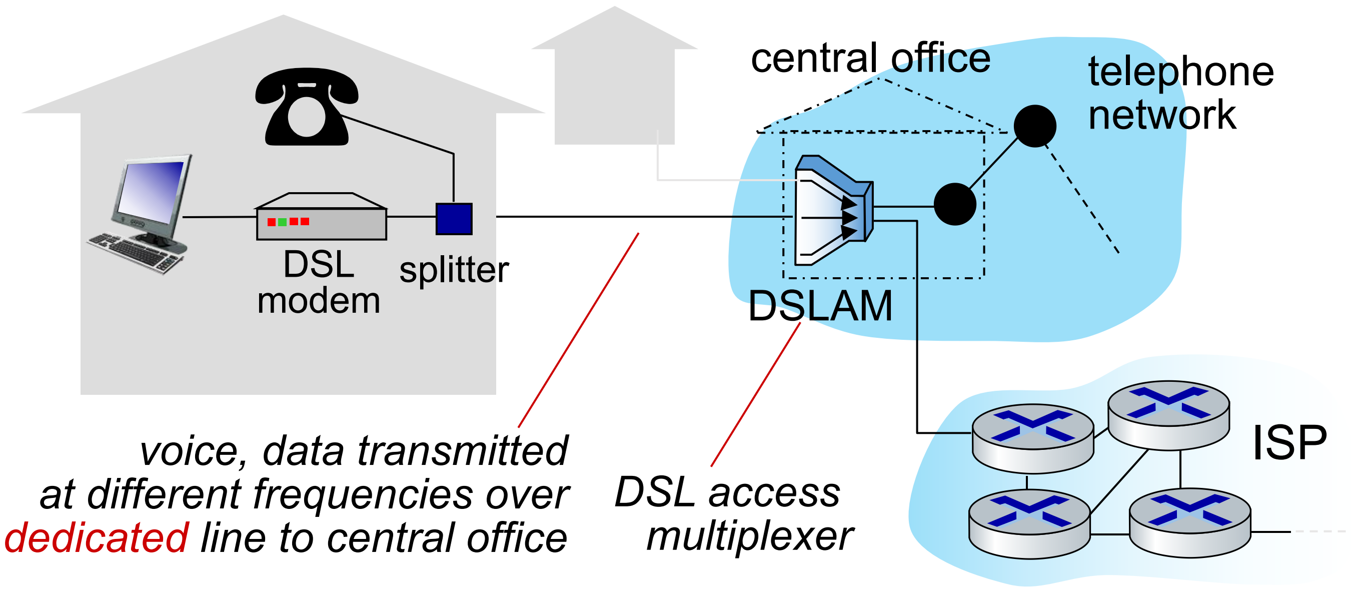
frequency division multiplexing (FDM): different channels transmitted in different frequency bands

Access networks: cable-based access



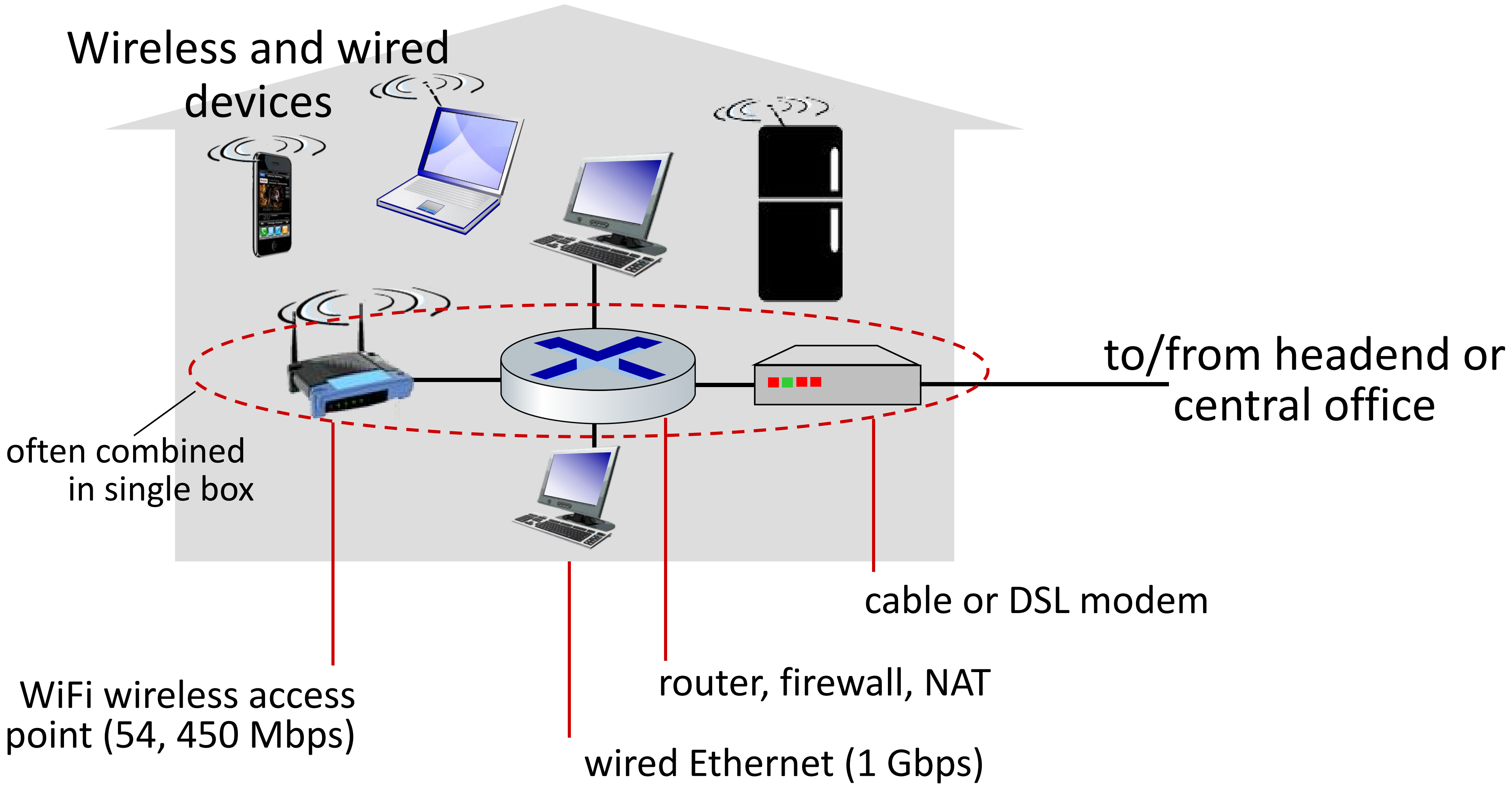
- HFC: hybrid fiber coax
  - asymmetric: up to 40 Mbps – 1.2 Gbps downstream transmission rate, 30-100 Mbps upstream transmission rate
- network of cable, fiber attaches homes to ISP router
  - homes *share access network* to cable headend

Access networks: digital subscriber line (DSL)



- use *existing* telephone line to central office DSLAM
  - data over DSL phone line goes to Internet
  - voice over DSL phone line goes to telephone net
- 24-52 Mbps dedicated downstream transmission rate
- 3.5-16 Mbps dedicated upstream transmission rate

Access networks: home networks



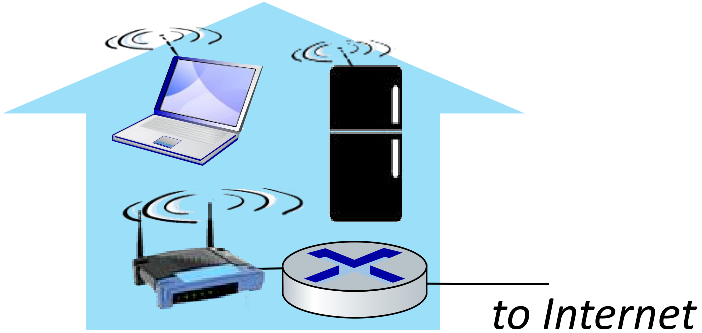
Wireless access networks

Shared *wireless* access network connects end system to router

- via base station aka "access point"

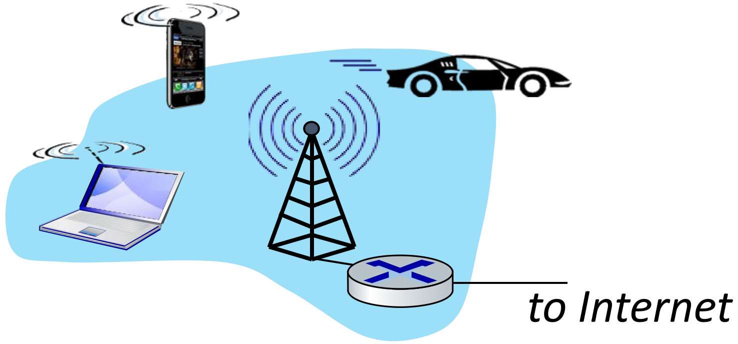
Wireless local area networks (WLANs)

- typically within or around building (~100 ft)
- 802.11b/g/n (WiFi): 11, 54, 450 Mbps transmission rate

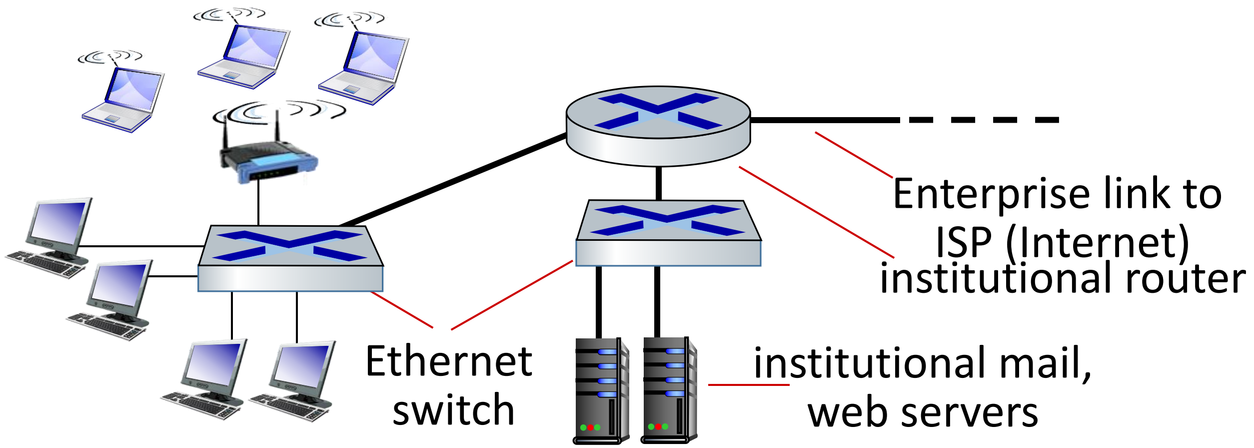


Wide-area cellular access networks

- provided by mobile, cellular network operator (10's km)
- 10's Mbps
- 4G/5G cellular networks



Access networks: enterprise networks

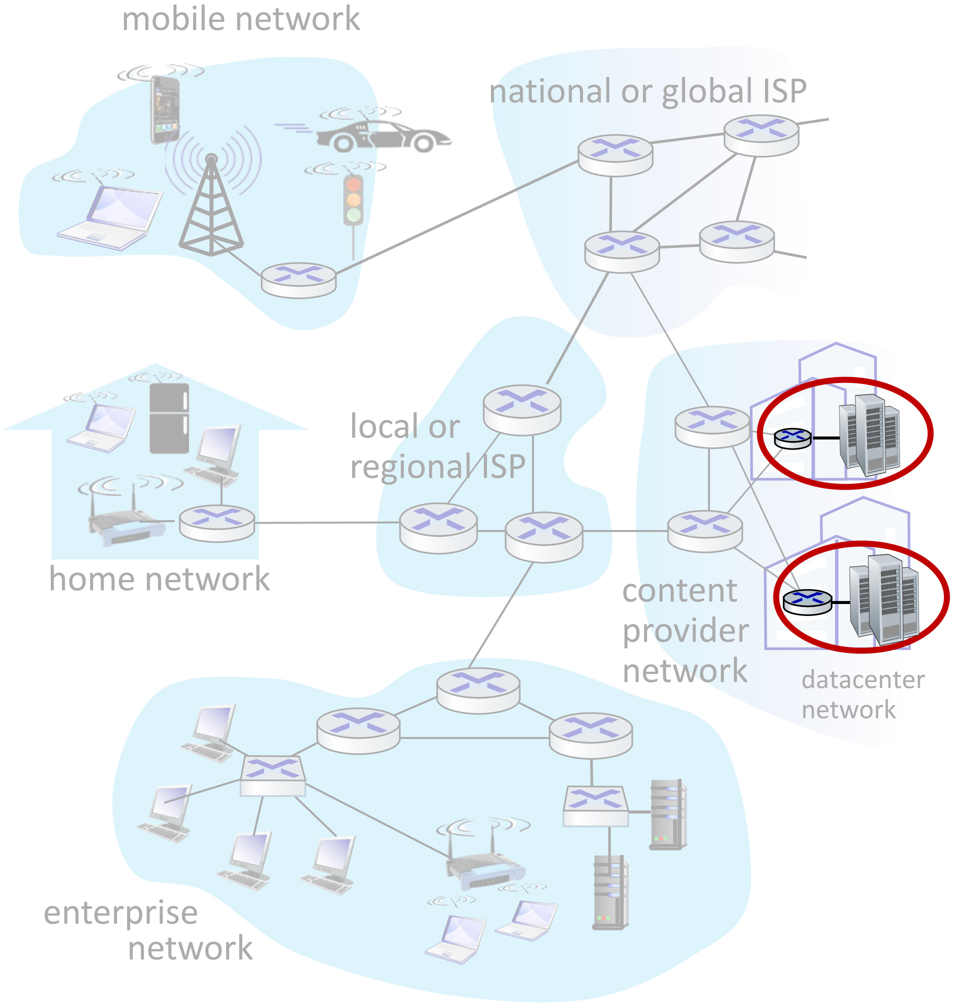


- companies, universities, etc.
- mix of wired, wireless link technologies, connecting a mix of switches and routers (we'll cover differences shortly)
  - Ethernet: wired access at 100Mbps, 1Gbps, 10Gbps
  - WiFi: wireless access points at 11, 54, 450 Mbps

Access networks: data center networks

- high-bandwidth links (10s to 100s Gbps) connect hundreds to thousands of servers together, and to Internet

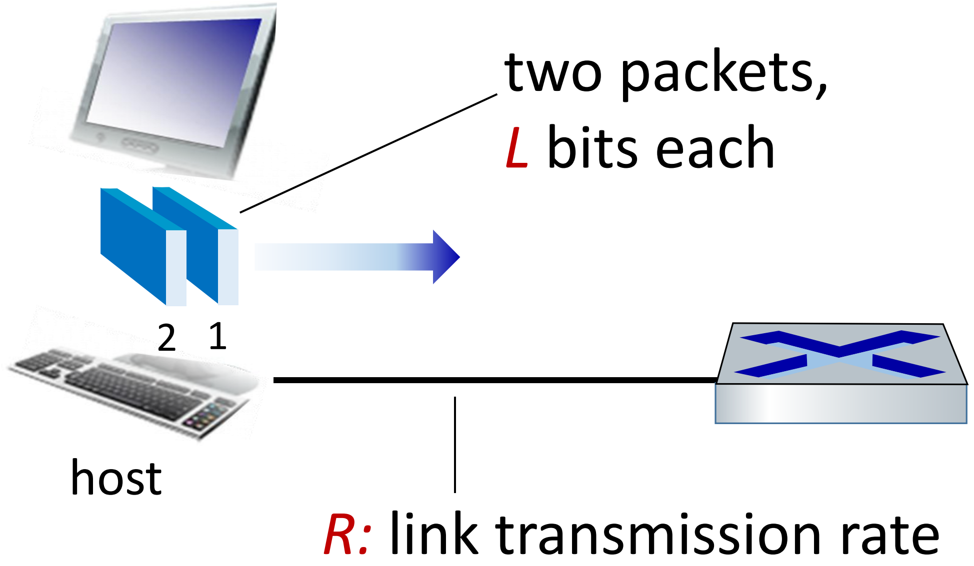
Courtesy: Massachusetts Green High Performance Computing Center (mghpcc.org)



Host: sends *packets* of data

host sending function:

- takes application message
- breaks into smaller chunks, known as *packets*, of length *L* bits
- transmits packet into access network at *transmission rate R*
  - link transmission rate, aka link *capacity, aka link bandwidth*



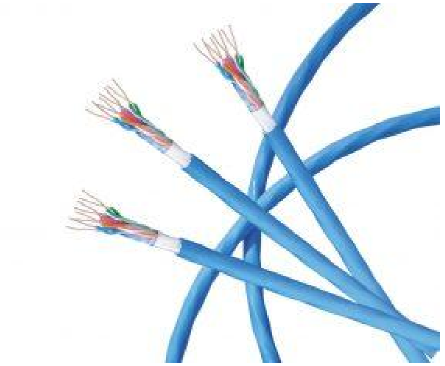
$$\text{packet transmission delay} = \text{time needed to transmit } L\text{-bit packet into link} = \frac{L \text{ (bits)}}{R \text{ (bits/sec)}}$$

Links: physical media

- bit: propagates between transmitter/receiver pairs
- physical link: what lies between transmitter & receiver
- guided media:
  - signals propagate in solid media: copper, fiber, coax
- unguided media:
  - signals propagate freely, e.g., radio

Twisted pair (TP)

- two insulated copper wires
  - Category 5: 100 Mbps, 1 Gbps Ethernet
  - Category 6: 10Gbps Ethernet



Links: physical media

Coaxial cable:

- two concentric copper conductors
- bidirectional
- broadband:
  - multiple frequency channels on cable
  - 100's Mbps per channel

Fiber optic cable:

- glass fiber carrying light pulses, each pulse a bit
- high-speed operation:
  - high-speed point-to-point transmission (10's-100's Gbps)
- low error rate:
  - repeaters spaced far apart
  - immune to electromagnetic noise

Links: physical media

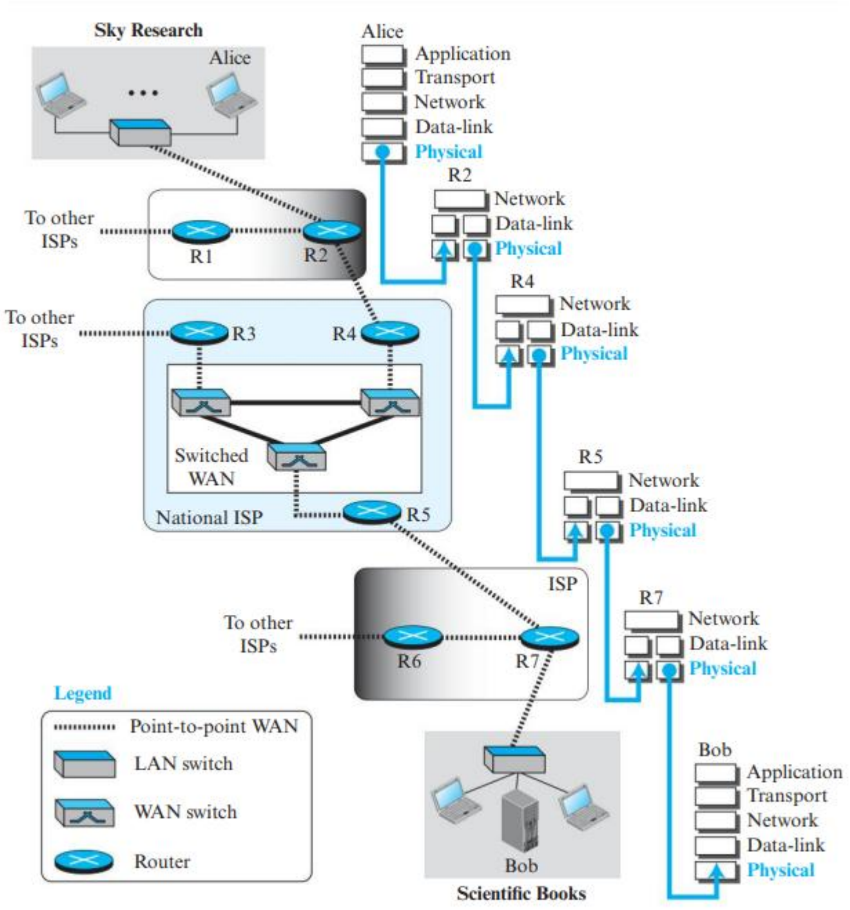
Wireless radio

- signal carried in various "bands" in electromagnetic spectrum
- no physical "wire"
- broadcast, "half-duplex" (sender to receiver)
- propagation environment effects:
  - reflection
  - obstruction by objects
  - Interference/noise

Radio link types:

- Wireless LAN (WiFi)
  - 10-100's Mbps; 10's of meters
- wide-area (e.g., 4G/5G cellular)
  - 10's Mbps (4G) over ~10 Km
- Bluetooth: cable replacement
  - short distances, limited rates
- terrestrial microwave
  - point-to-point; 45 Mbps channels
- satellite
  - up to < 100 Mbps (Starlink) downlink
  - 270 msec end-end delay (geostationary)

DATA AND SIGNALS

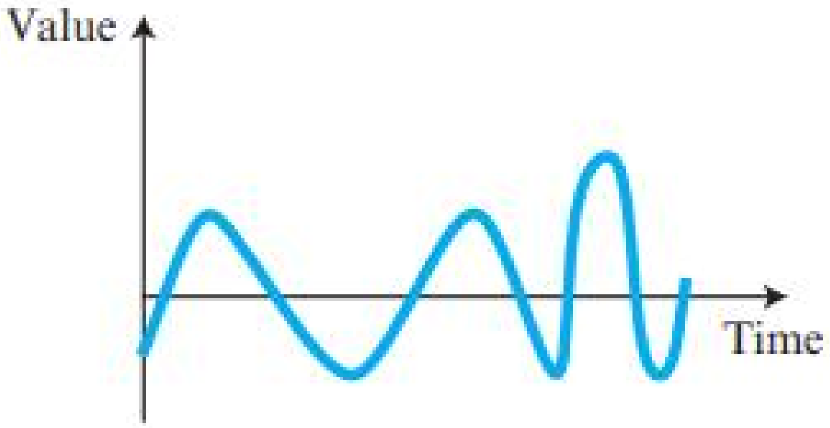


Analog and Digital

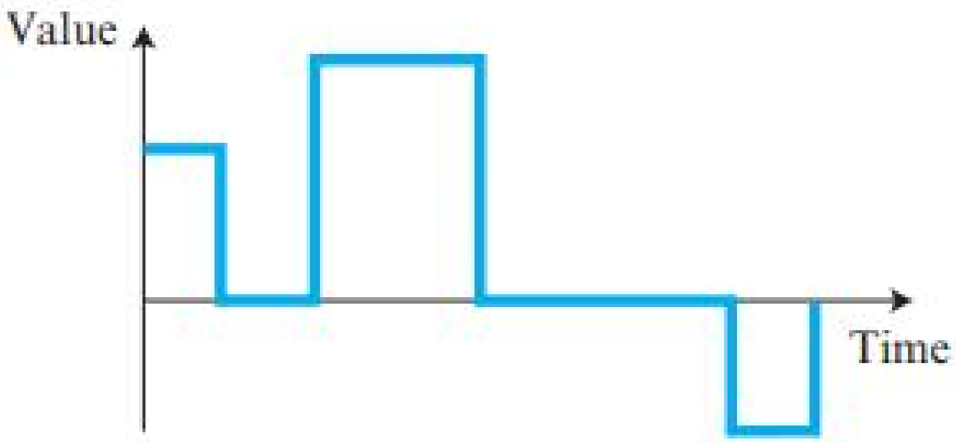
Analog data refers to information that is continuous. Analog data, such as the sounds made by a human voice, take on continuous values. When someone speaks, an analog wave is created in the air. This can becaptured by a microphone and converted to an analog signal or sampled and converted to a digital signal.

Digital data take on discrete values. For example, data are stored in computer memory in the form of 0s and 1s. They can be converted to a digital signal or modulated into an analog signal for transmission across a medium.

Like the data they represent, signals can be either analog or digital. An analog signal has infinitely many levels of intensity over a period of time. As the wave moves from value A to value B, it passes through and includes an infinite number of values along its path. A digital signal, on the other hand, can have only a limited number of defined values. Although each value can be any number, it is often as simple as 1 and 0. The simplest way to show signals is by plotting them on a pair of perpendicular axes. The vertical axis represents the value or strength of a signal. The horizontal axis represents time.



a. Analog signal

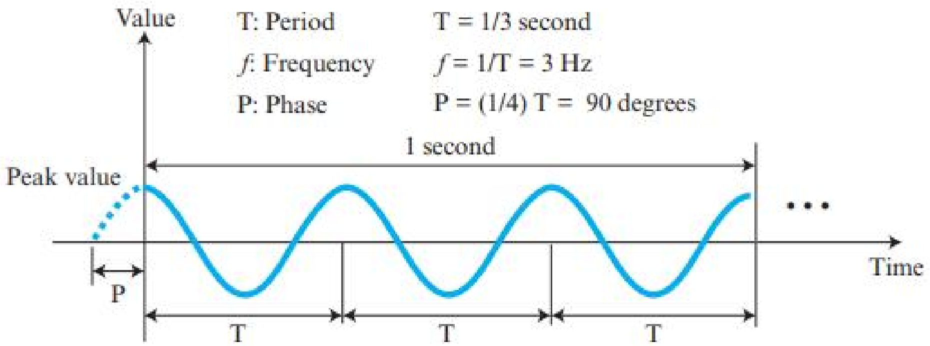


b. Digital signal

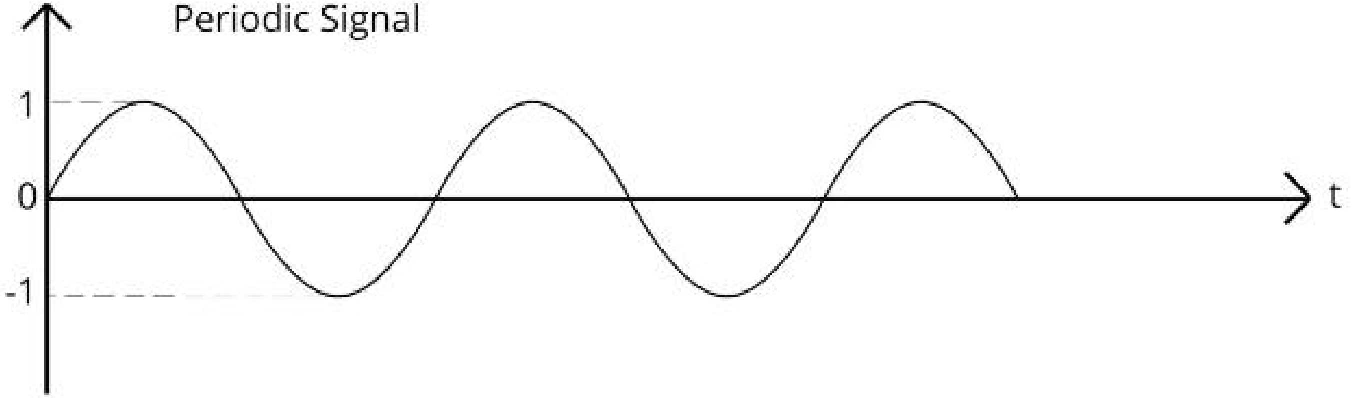
Analog Signals

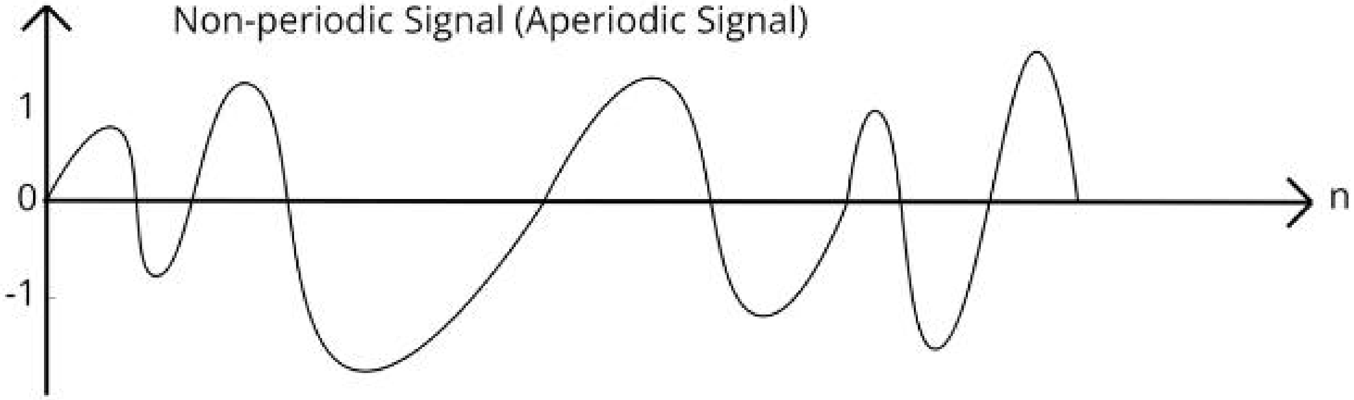
Let's first focus on analog signals, which can be either periodic or nonperiodic (aperiodic). A periodic analog signal completes a pattern within a measurable timeframe, called a **period**, and repeats this pattern over subsequent periods. Each full pattern is called a **cycle**. Conversely, a **nonperiodic analog signal does not exhibit a repeating pattern over time.**

In data communications, periodic analog signals are commonly used and can be classified as simple or composite. **A simple periodic analog signal, such as a sine wave, cannot be broken down into simpler signals**. A composite periodic analog signal comprises multiple sine waves. The sine wave is the most fundamental form of a periodic analog signal, characterized by a smooth, continuous oscillating curve. Each cycle of a sine wave consists of one arc above the time axis followed by one arc below it.

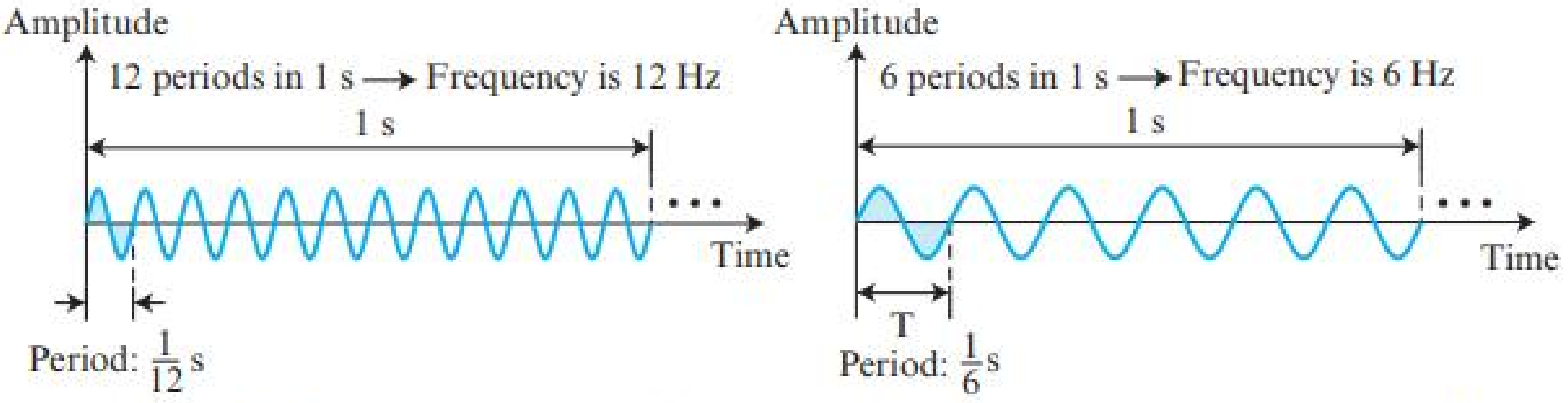


Periodic and Non-Periodic Signals





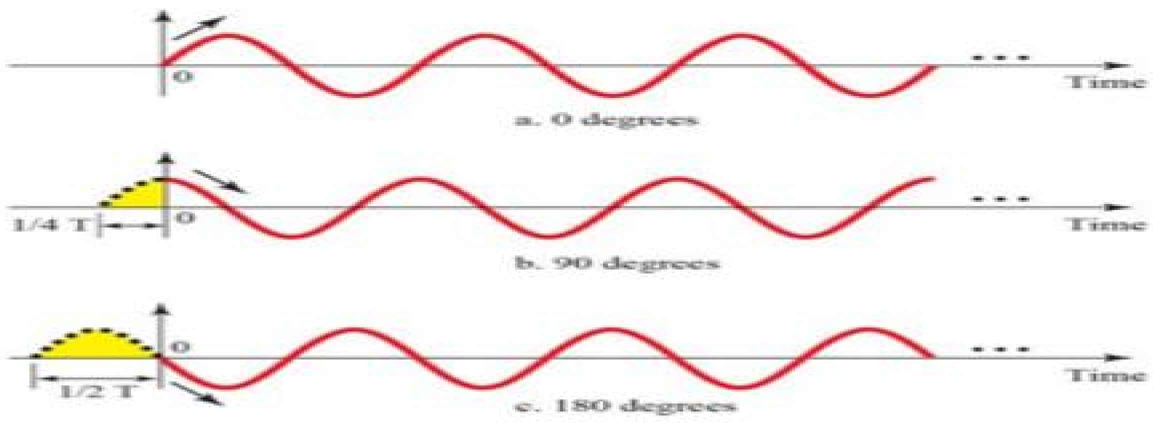
A sine wave can be described by three parameters: **peak amplitude, frequency(rate of change), and phase.** The peak amplitude is the a**bsolute value of the wave's highest intensity**, usually measured in volts for electric signals, and indicates the energy the wave carries. **Period refers to the time needed to complete one cycle**, measured in seconds, while **frequency, measured in Hertz (Hz), indicates the number of cycles per second.** Period and frequency are inversely related (f = 1/T).



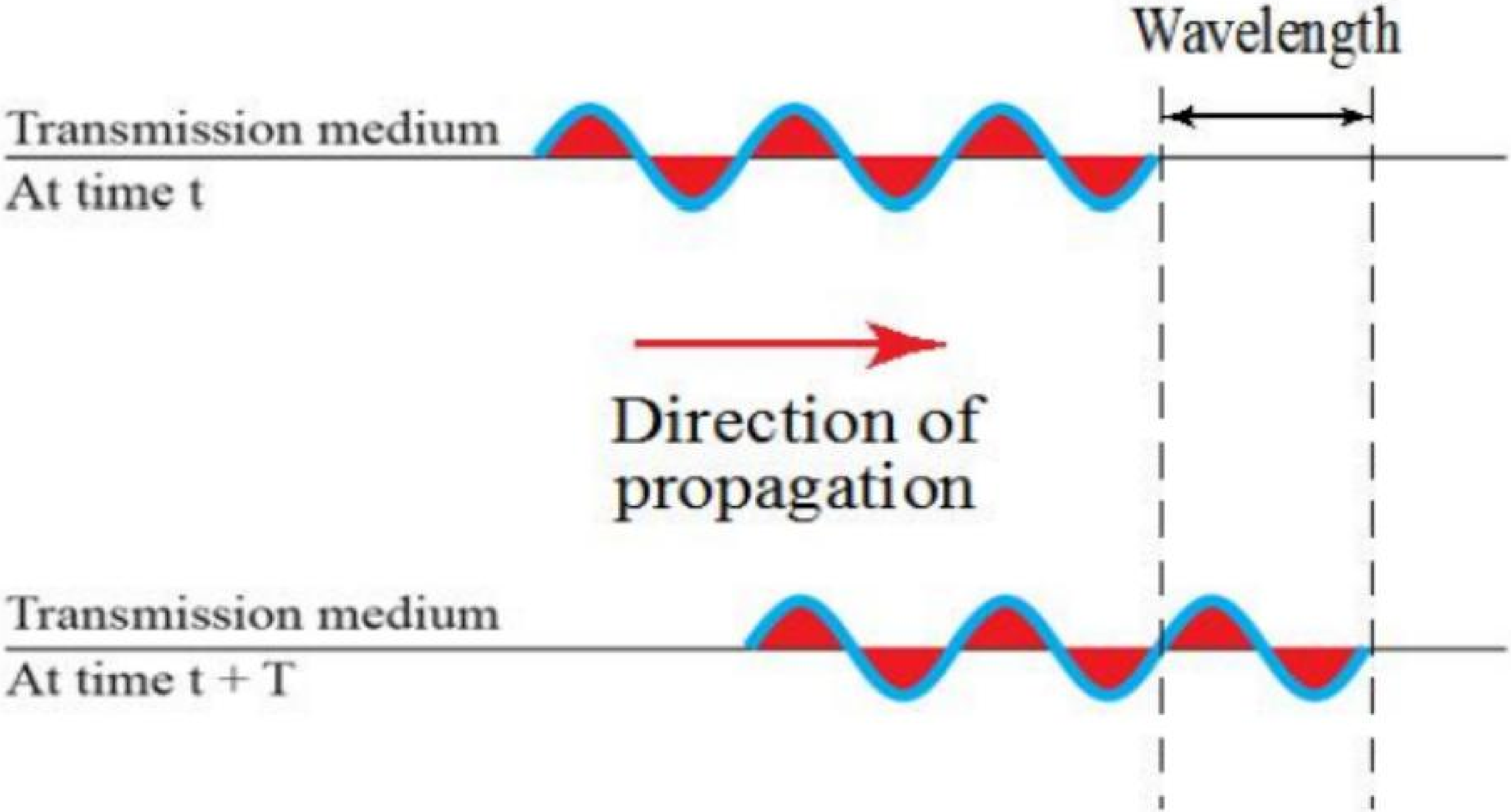
a. A signal with a frequency of 12 Hz

b. A signal with a frequency of 6 Hz

the position of a waveform relative to time 0, indicating the amount of shift along the time axis and the status of the first cycle. It is measured in degrees or radians (360° = 2π radians).



Wavelength is the distance a signal travels in one period, linking the period or frequency of a sine wave to the propagation speed of the medium. While frequency is independent of the medium, wavelength depends on both frequency and the medium. Wavelength is often used to describe light transmission in optical fibers. It can be calculated using the formula **λ = c / f, where λ is the wavelength, c is the propagation speed, and f is the frequency.**


Wavelength
Transmission medium
At time t
Direction of propagation
Transmission medium
At time t + T

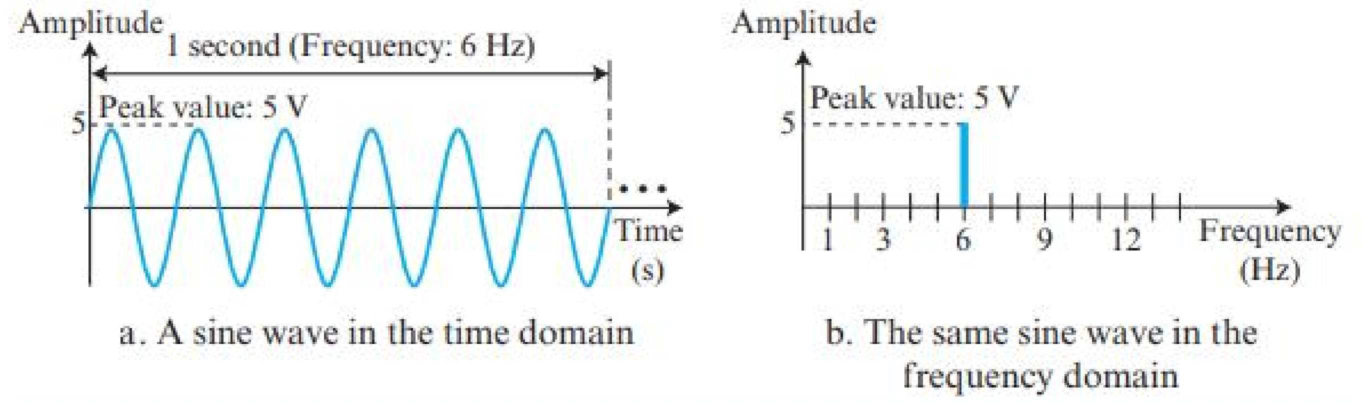
Signal Attributes

Property	Definition	Symbol	Measured in	Extra Information
Displacement	Distance a particle has moved from its equilibrium position		Meters(m)	
Amplitude	Maximum displacement		Meters(m)	
Wavelength	Length of one full wave(cycle)	λ	Meters(m)	
Frequency	Number of full cycles per second	f	Hz	
Time Period	Time to complete one full cycle	T	Seconds	F=1/T
Phase	Position of waveform relative to time 0	φ	Degree/Radian	

Time and Frequency Domains

A sine wave is comprehensively defined by its amplitude, frequency, and phase. We have been showing a sine wave by using what is called a time-domain plot, which shows changes in signal amplitude with respect to time. **To show the relationship between amplitude and frequency, we can use a frequency-domain plot.** Figure 7.5 shows a signal in both the time and frequency domains.

Figure 7.5 *The time-domain and frequency-domain plots of a sine wave*

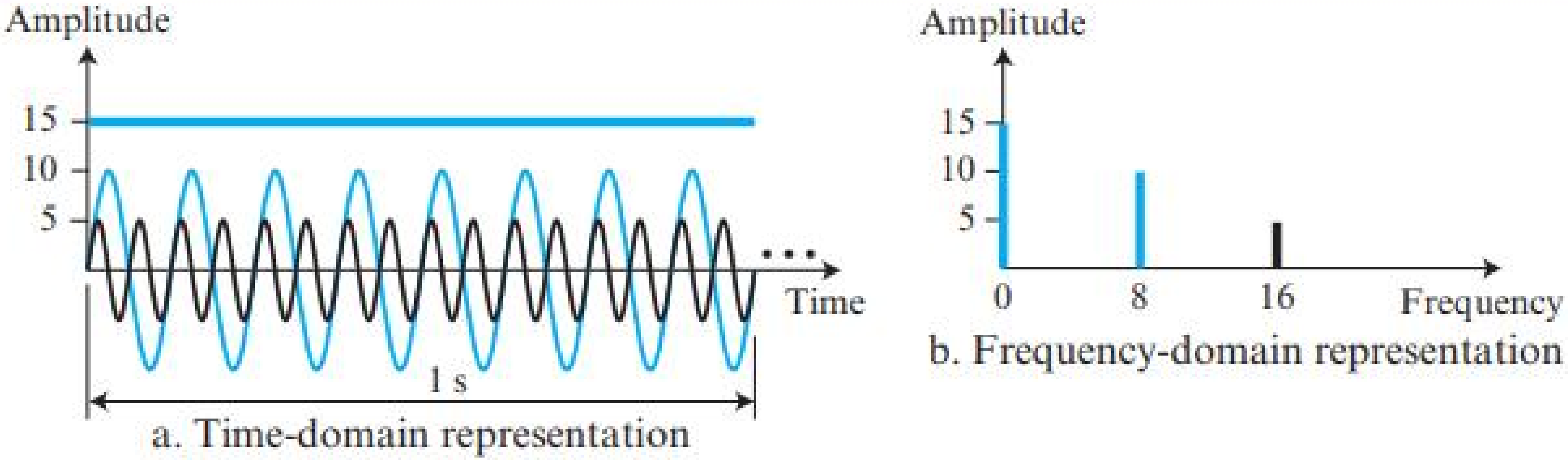


In the frequency domain, a sine wave is represented by one spike. The position of the spike shows the frequency; its height shows the peak amplitude.

Example 7.1

The frequency domain is more compact and useful when we are dealing with more than one sine wave. For example, Figure 7.6 shows three sine waves, each with different amplitude and frequency. All can be represented by three spikes in the frequency domain.

Figure 7.6 *The time domain and frequency domain of three sine waves*

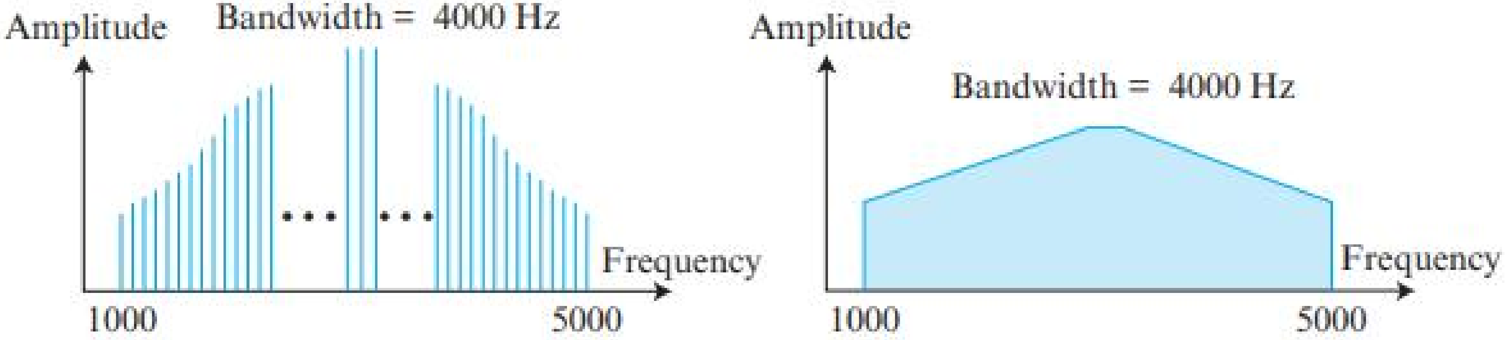


Composite Signals

a single sine wave can't convey meaningful information, such as a conversation over the phone. Instead, we use composite signals, made of many simple sine waves. In the early 1900s, mathematician Jean-Baptiste Fourier demonstrated that any composite signal is a combination of simple sine waves with different frequencies, amplitudes, and phases. Composite signals can be periodic or nonperiodic. **A periodic composite signal can be decomposed into simple sine waves with discrete frequencies, which are integral multiples of the fundamental frequency. A nonperiodic composite signal can be decomposed into an infinite number of simple sine waves with continuous, real-valued frequencies.**

Bandwidth

The bandwidth of a composite signal is the difference between the highest and the lowest frequencies contained in that signal.



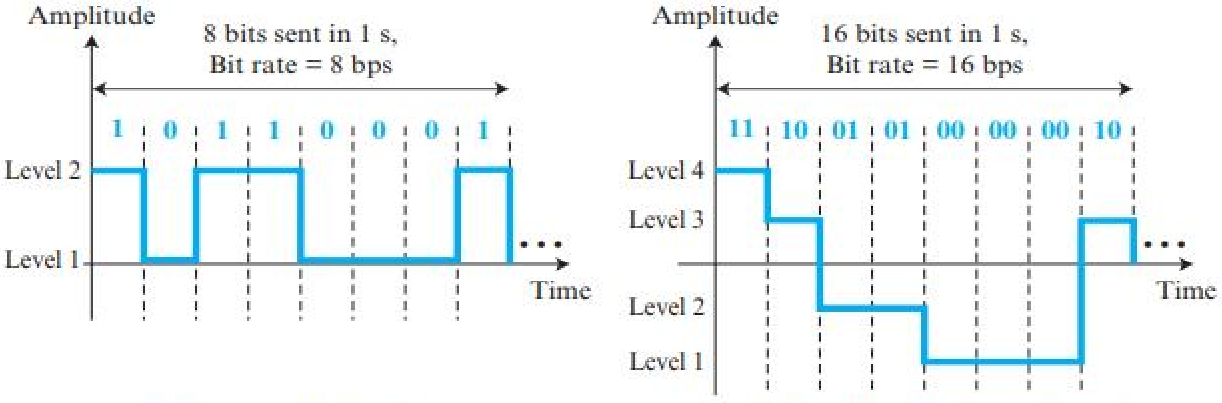
a. Bandwidth of a periodic signal

b. Bandwidth of a nonperiodic signal

Digital Signals

In addition to being represented by an analog signal, information can also be represented by a digital signal. For example, a 1 can be encoded as a positive voltage and a 0 as zero voltage. A digital signal can have more than two levels. In this case, we can send more than 1 bit for each level. Figure 7.8 shows two signals, one with two levels and the other with four. We send 1 bit per level in part a of the figure and 2 bits per level in part b of the figure. In general, if a signal has L levels, each level needs log2L bits.

Figure 7.8 *Two digital signals: one with two signal levels and the other with four signal levels*



a. A digital signal with two levels

b. A digital signal with four levels

Bit Rate

Most digital signals are nonperiodic, making period and frequency inappropriate characteristics. Instead, we use the term bit rate to describe digital signals, which is the number of bits sent in one second, expressed in bits per second (bps).

Bit Length

The bit length is the distance one bit occupies on the transmission medium.

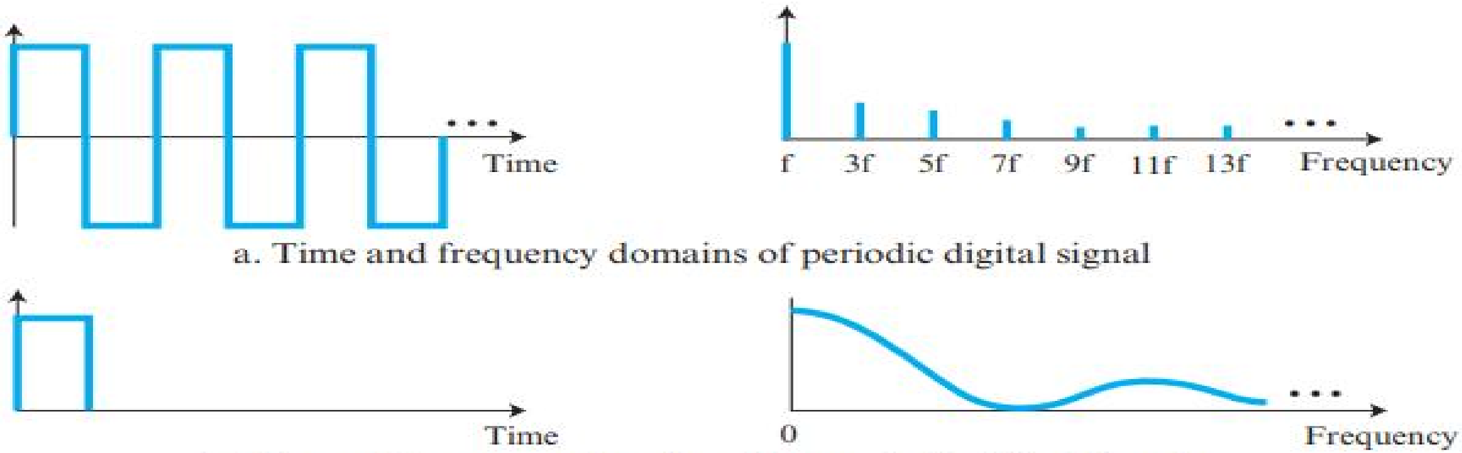
Bit length = propagation speed × bit duration

bit length = 1 / bit rate

Digital Signal as a Composite Analog Signal

a digital signal is a composite analog signal with infinite bandwidth. In the time domain, a digital signal consists of connected vertical and horizontal line segments. **A vertical line indicates an infinite frequency (sudden change), while a horizontal line indicates a frequency of zero (no change). This implies all frequencies between zero and infinity are present.**

Fourier analysis can decompose a digital signal. For a periodic digital signal (rare in data communications), the decomposed signal has an infinite bandwidth with discrete frequencies. For a nonperiodic digital signal, the decomposed signal also has infinite bandwidth, but with continuous frequencies.



a. Time and frequency domains of periodic digital signal

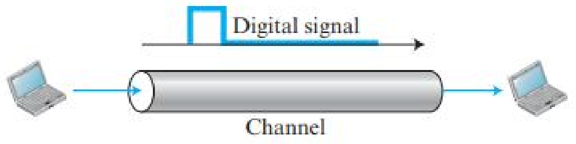
b. Time and frequency domains of nonperiodic digital signal

Transmission of Digital Signals

A digital signal, **whether periodic or nonperiodic, is a composite analog signal with frequencies ranging from zero to infinity.** For our discussion, we'll focus on nonperiodic digital signals common in data communications. The key question is how to transmit a digital signal from point A to point B. There are two approaches: baseband transmission or broadband transmission (using modulation).

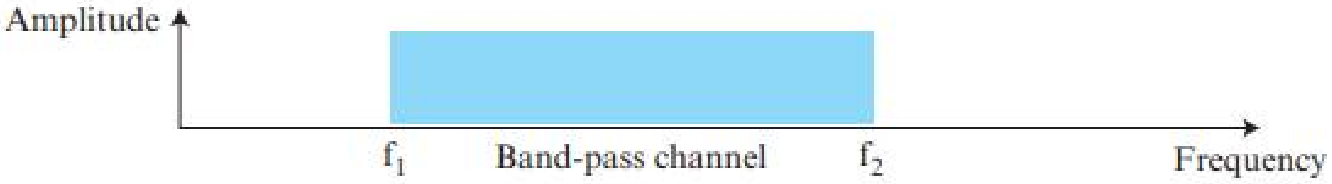
Baseband Transmission

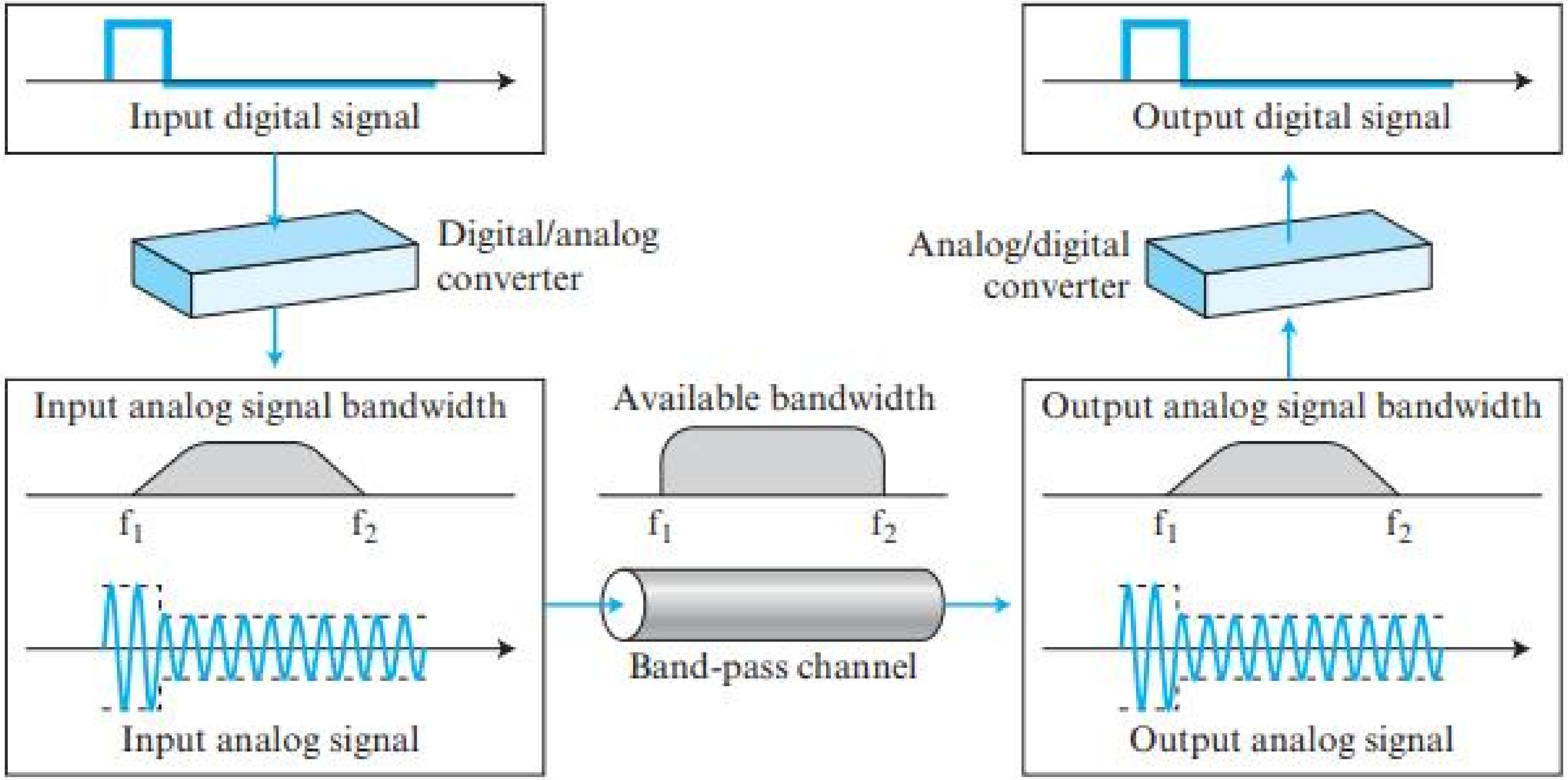
Baseband transmission involves **sending a digital signal over a channel without converting it to an analog signal**. This method requires a low-pass channel, which has a **bandwidth starting from zero**. Examples include a **dedicated cable** connecting two computers or a medium where only two stations communicate at a time. In both cases, the entire bandwidth constitutes a single channel suitable for baseband communication.



Broadband Transmission

Broadband transmission, or modulation, **involves converting a digital signal to an analog signal for transmission.** This method uses a band-pass channel, **which has a bandwidth that doesn't start from zero** and is more commonly available than a low-pass channel. **In modulation, a digital signal is converted to a composite analog signal using a carrier frequency.** The carrier's amplitude is varied to match the digital signal, resulting in a composite signal. At the receiver end, the analog signal is converted back to digital, replicating the original signal sent.





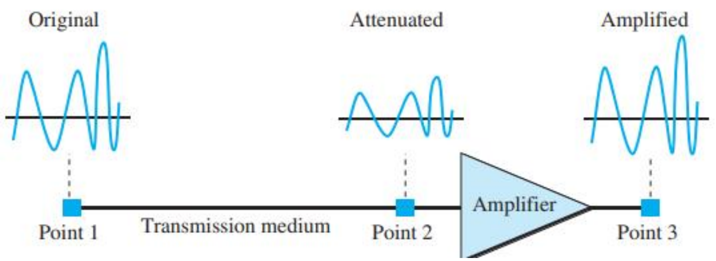
Input digital signal
Digital/analog
converter
Input analog signal bandwidth
f_1
f_2
Input analog signal
Available bandwidth
f_1
f_2
Band-pass channel
Output digital signal
Analog/digital
converter
Output analog signal bandwidth
f_1
f_2
Output analog signal

Signal Impairment

- Signals travel through transmission media, which are not perfect. The imperfection causes signal impairment.
- This means that the signal at the beginning of the medium is not the same as the signal at the end of the medium. What is sent is not what is received.
- Three causes of impairment are
  - ***attenuation***
  - ***distortion***
  - ***noise***.

Attenuation and Amplification

- Attenuation means a loss of energy. When a signal, simple or composite, travels through a medium, it loses some of its energy in overcoming the resistance of the medium. To compensate for this loss, we need amplification.

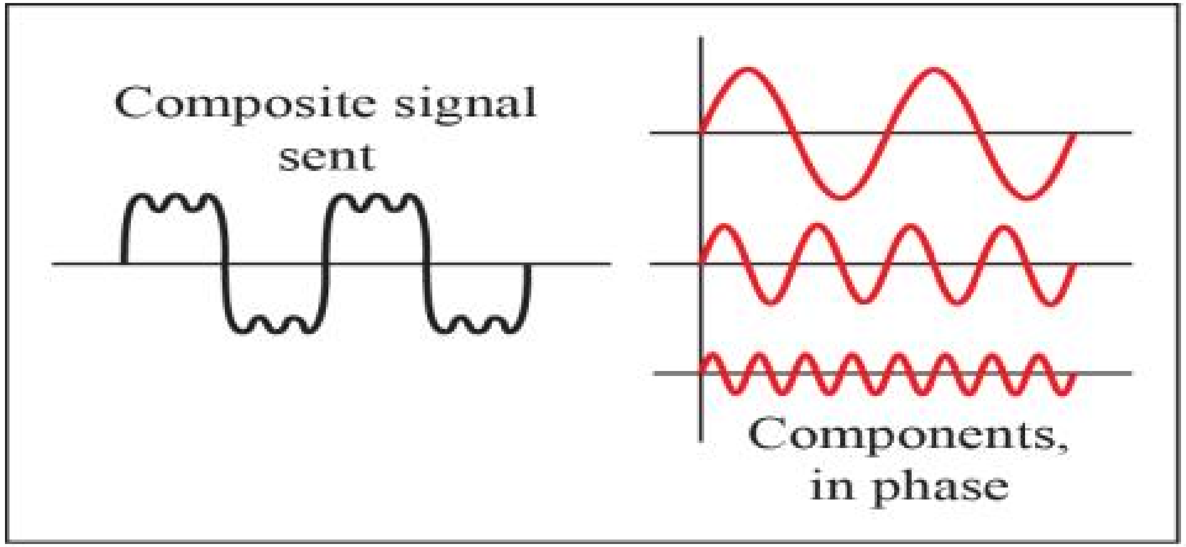


➢ The **decibel (dB)** measures the relative strength of two signals or one signal at two different points. Note that the decibel is negative if a signal is attenuated and positive if a signal is amplified. Variables P1 and P2 are the powers of a signal at points 1 and 2, respectively.

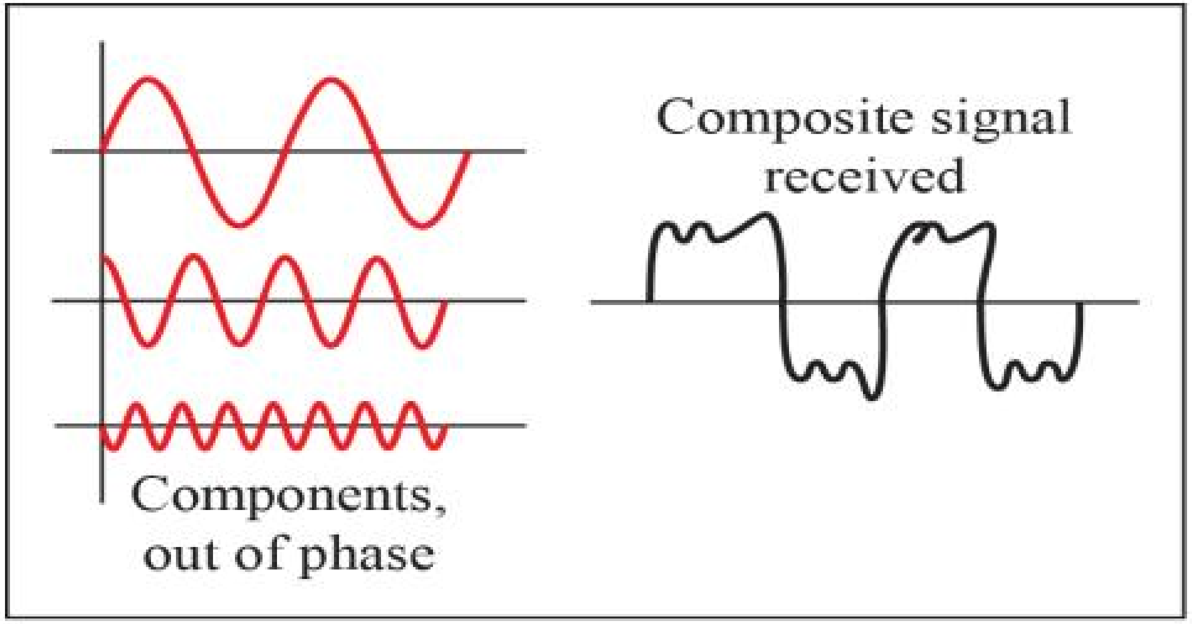
$$dB = 10 \log_{10} (P_2/P_1)$$

Distortion

- Distortion means that the signal changes its form or shape. Distortion can occur in a composite signal made up of different frequencies.
- Each signal component has its own propagation speed through a medium and, therefore, its own delay in arriving at the final destination.
- Differences in delay may create a difference in phase if the delay is not exactly the same as the period duration.



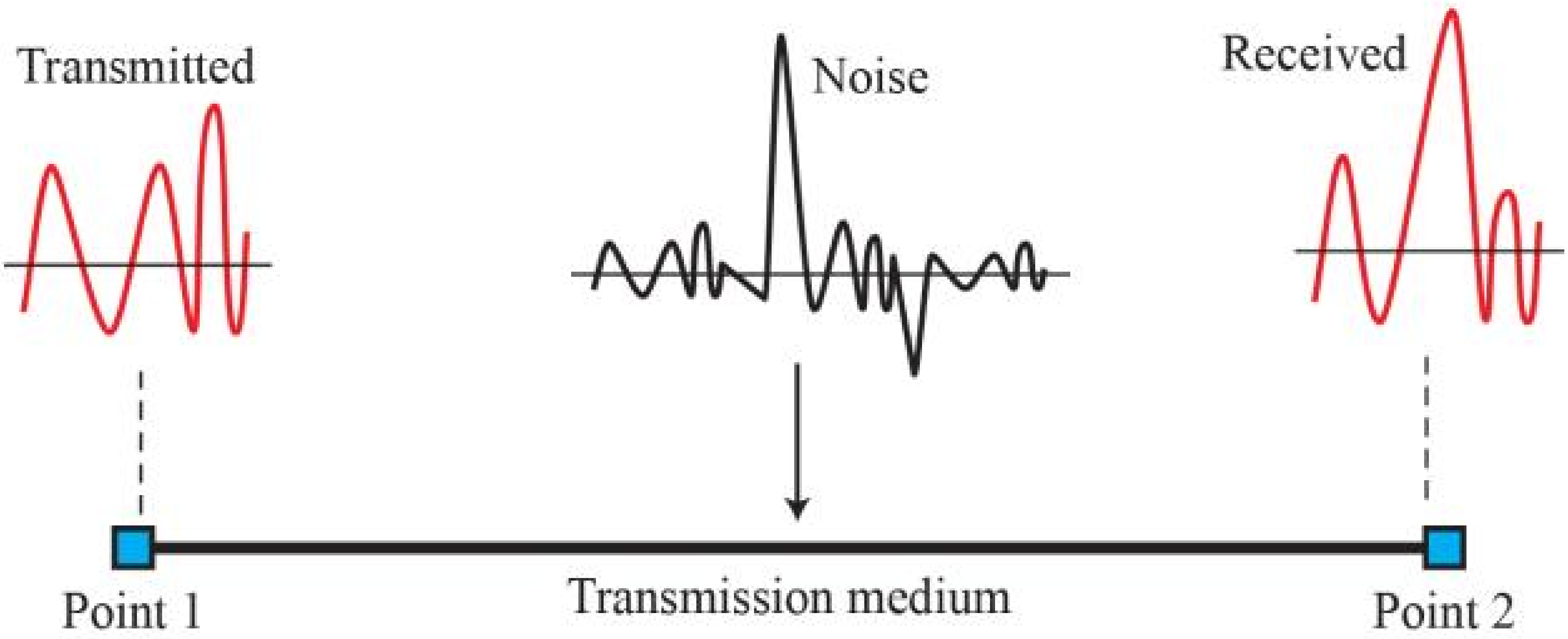
At the sender



At the receiver

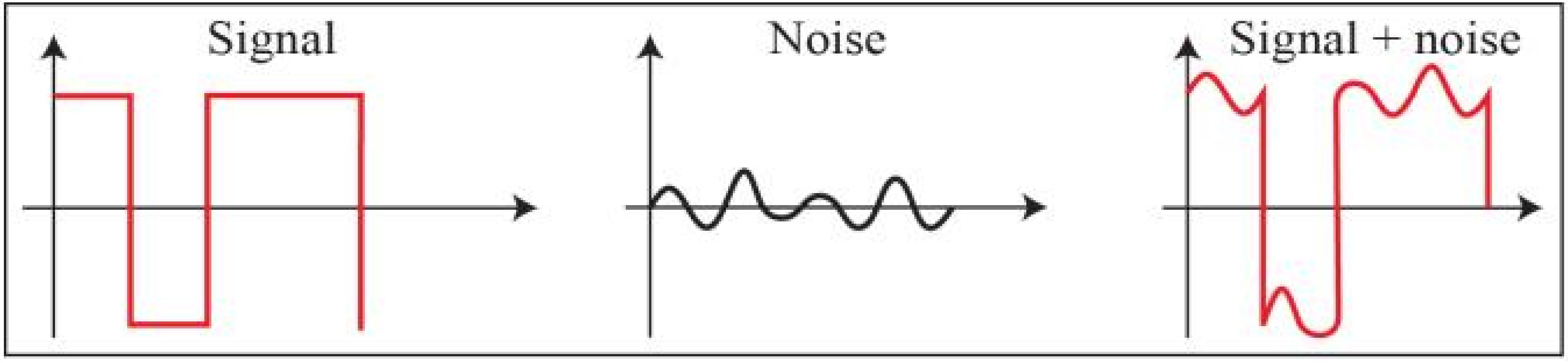
Noise

- Noise is another cause of impairment. Several types of noise, such as *thermal noise, induced noise, crosstalk, and impulse noise*, may corrupt the signal.
- ***Thermal noise*** is the random motion of electrons in a wire, which creates an extra signal not originally sent by the transmitter.
- ***Induced noise*** comes from sources such as motors and appliances. These devices act as a sending antenna, and the transmission medium acts as the receiving antenna.
- ***Crosstalk*** is the effect of one wire on the other. One wire acts as a sending antenna and the other as the receiving antenna.
- ***Impulse noise*** is a spike (a signal with high energy and very short duration) that comes from power lines, lightning, and so on.

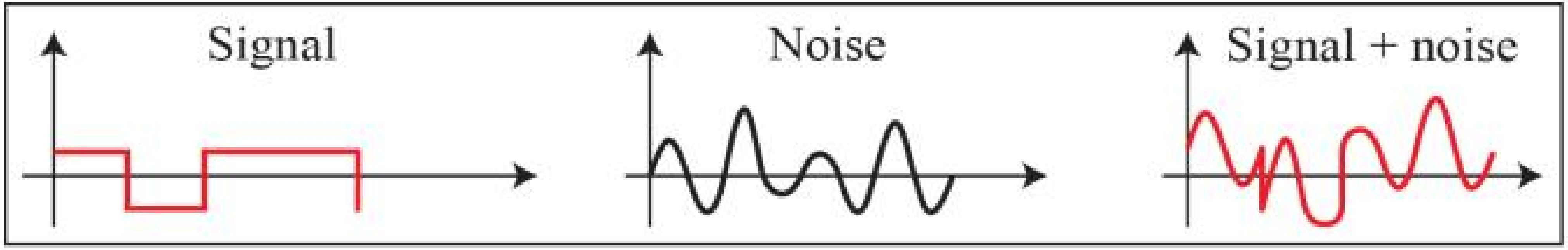

Transmitted
Noise
Received
Point 1
Transmission medium
Point 2

Noise

- To find the theoretical bit-rate limit, we need to know the ratio of the signal power to the noise power. The **signal-to-noise ratio** is defined as -
  - ***SNR = average signal power / average noise power***
- Because SNR is the ratio of two powers, it is often described in decibel units, SNR_{dB} as -
  - $\mathbf{SNR_{dB} = 10 \log_{10} SNR}$



a. High SNR



b. Low SNR

Data Rate Limits

- Data Rate Limit is how fast we can send data over a channel. It is measured in bits per second.
- Data rate depends on three factors:
  - The bandwidth available
  - The level of the signals we use
  - The quality of the channel (the level of noise)
- For a noiseless channel, the **Nyquist bit rate** formula defines the theoretical maximum bit rate
  - **BitRate = 2 x B x $\log_2$ L** where B is bandwidth of the channel and L is signal level
- For a noisy channel, the **Shannon capacity** formula to determine the theoretical highest bit rate -
- **C = B x $\log_2$ (1 + SNR)** defines the capacity of channel where ***SNR = average signal power / average noise power***

Performance Metrics

- Bandwidth
- Throughput
- Latency (Delay)
- **Jitter** - It is the variation in the time delay between received signals compared to the expected time intervals.

Digital Transmission

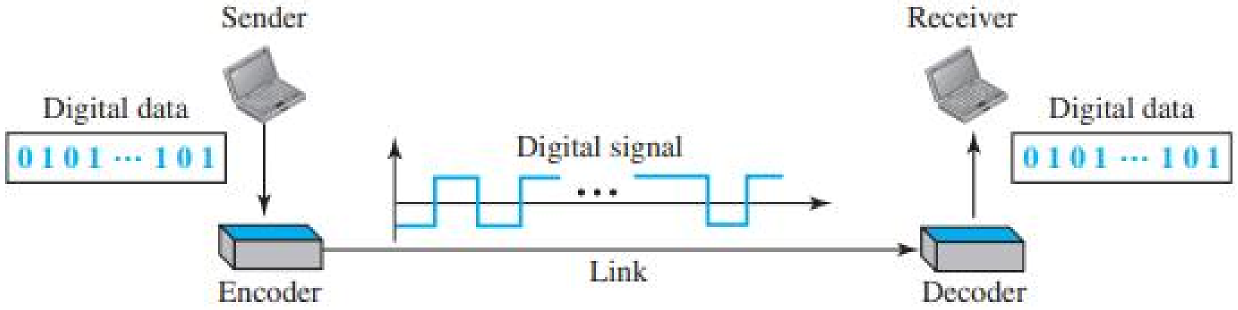
- A computer network is designed to send information from one point to another. This information needs to be converted to either a digital signal or an analog signal for transmission.
- In digital transmission, we transmit data using digital signals.
- If our data is in digital form (discrete data like text, images, etc.) then we need ***digital-to-digital encoding techniques.***
- If our data is in analog form (continuous data like temperature, etc.) then we need ***analog-to-digital encoding techniques.***

Digital-to-Digital Conversion

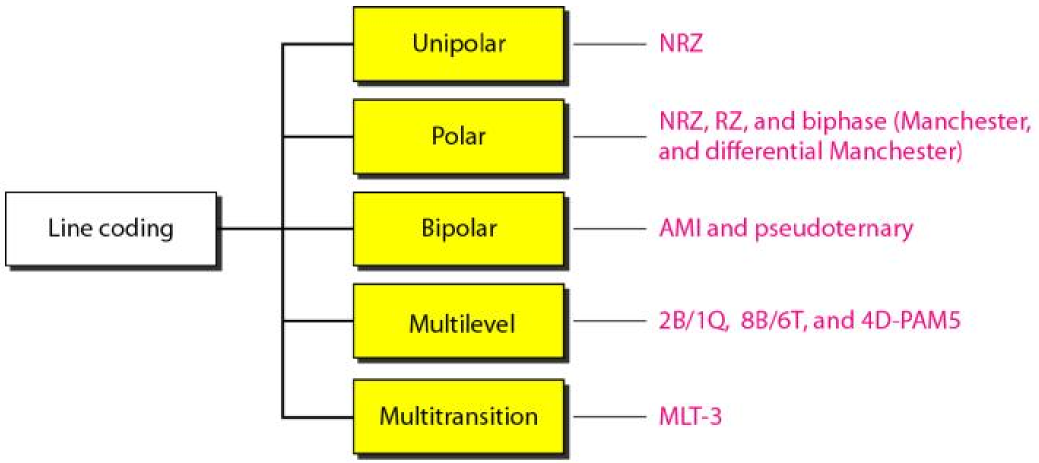
For digital transmission of digital data, digital-to-digital conversion is necessary, involving line coding, block coding, and scrambling. Line coding is always required, while block coding and scrambling are optional.

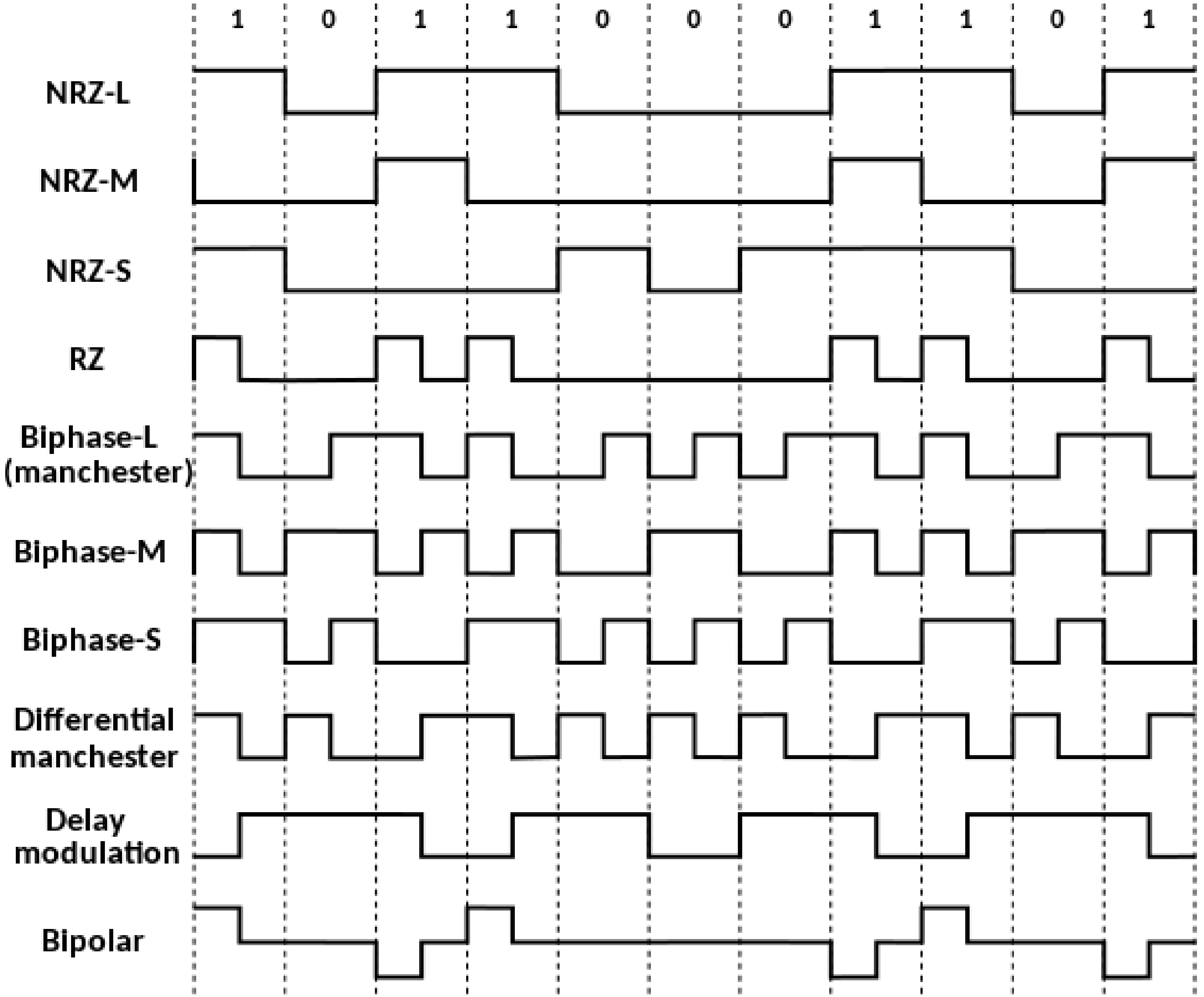
Line Coding

Line coding converts digital data into digital signals. Data stored as sequences of bits (e.g., text, numbers, images, audio, video) are encoded into a digital signal by the sender and decoded back into digital data by the receiver.



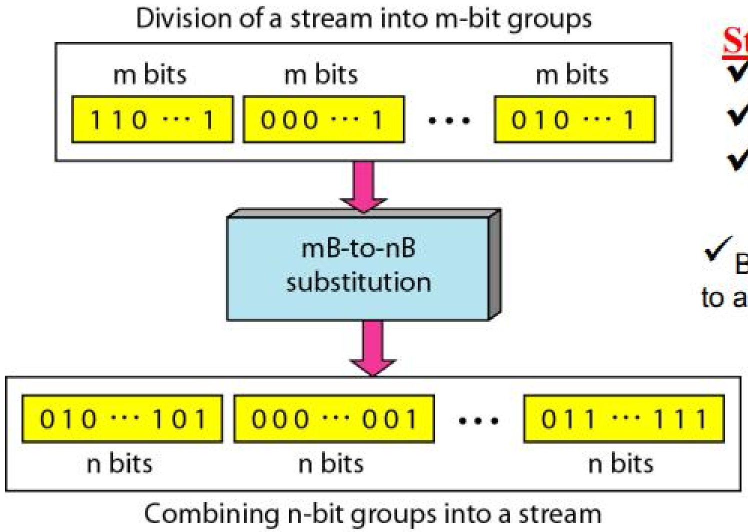
Line Coding Schemes





Block Coding

- Block coding changes a block of **m bits** into a block of **n bits**, where n is larger than m. That’s block coding is also known as ***mB/nB* coding**.
- Extra bits (redundancy bits) are added to ensure synchronization and to provide error detection.
- Block coding involves three steps - Division, Substitution, Combining

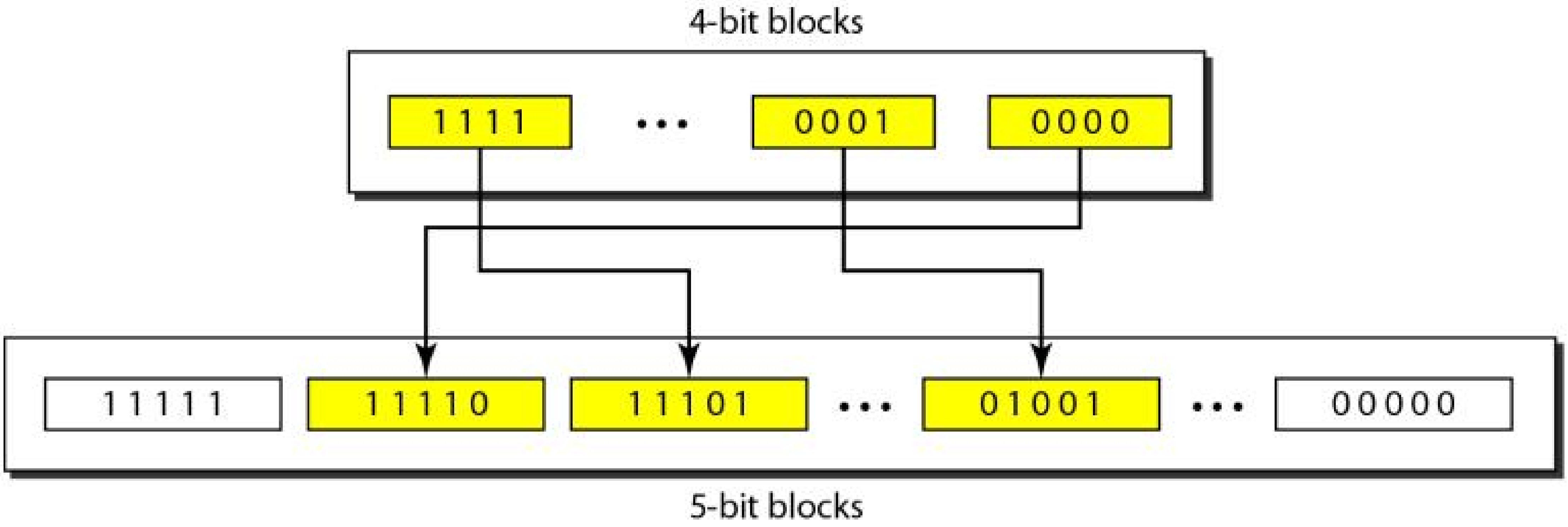


Steps of block coding:

✓ Division
✓ Substitution
✓ Combining

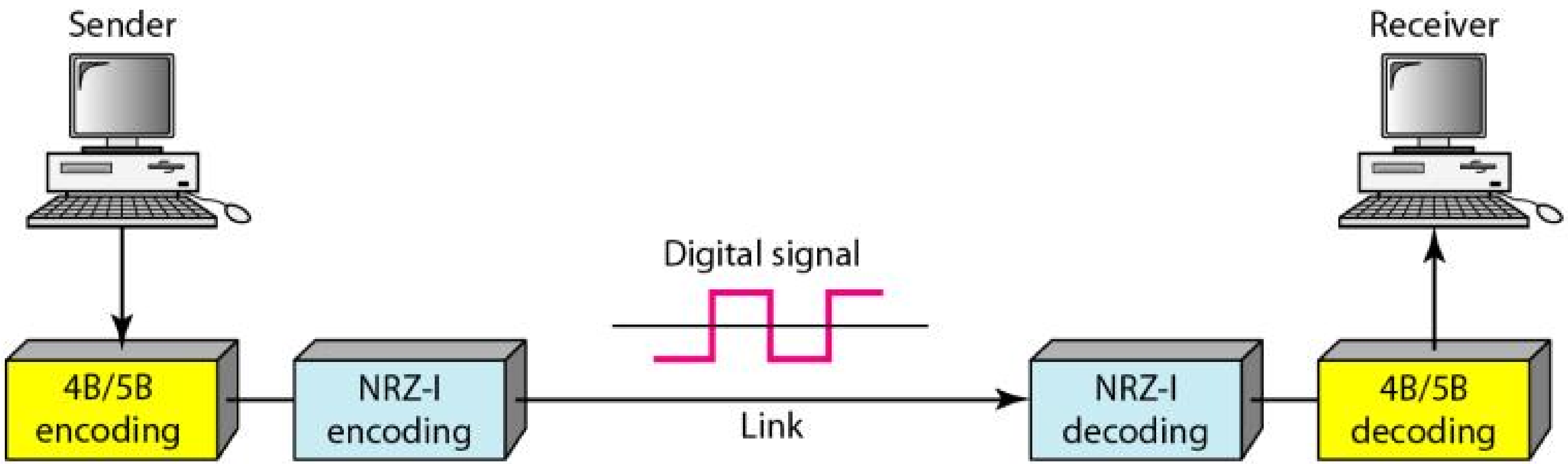
✓ Block coding is normally referred to as *m*B/*n*B coding.

Block Coding



Block coding with line coding

Figure 4.15 ***Using block coding 4B/5B with NRZ-I line coding scheme***



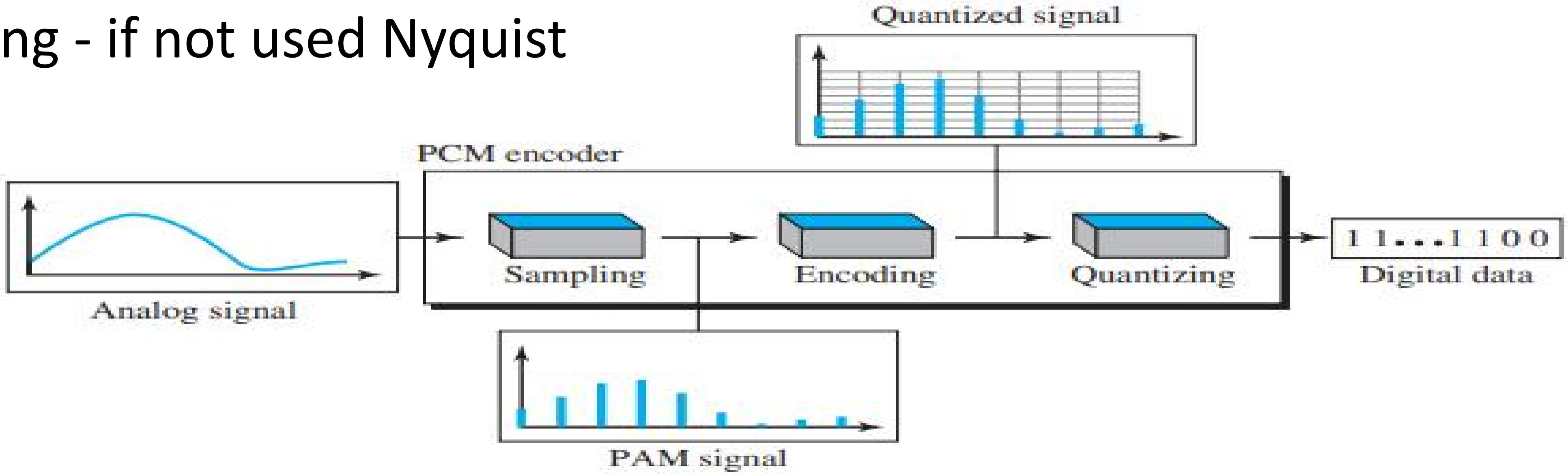
Analog-to-Digital Conversion

Analog signals from devices like microphones or cameras are often converted to digital data to reduce noise susceptibility. This section covers two digitization techniques: pulse code modulation and delta modulation. Once digitized, the data can be converted to a digital signal using methods from futher sections.

Pulse Code Modulation (PCM)

The most common technique used to change an analog signal to digital data (digitization) is called pulse code modulation (PCM). A PCM encoder has three processes, as shown in next page

- Nyquist - no. of sample required
- aliasing - if not used Nyquist



The three processes are:

1. The analog signal is sampled every T s.
2. The sampled signal is quantized, which means every sample is considered as a pulse.
3. The quantized values (pulses) are encoded as streams of bits.

Example

We want to digitize the human voice. What is the bit rate, assuming 8 bits per sample?

Solution - The human voice normally contains frequencies from 0 to 4000 Hz. So the sampling rate and bit rate are calculated as

$$\text{Sampling rate} = 4000 \times 2 = 8000 \text{ samples/s}$$
$$\text{Bit rate} = 8000 \times 8 = 64{,}000 \text{ bps} = 64 \text{ kbps}$$

PCM Bandwidth

It can be proved that the minimum bandwidth of the digital signal is

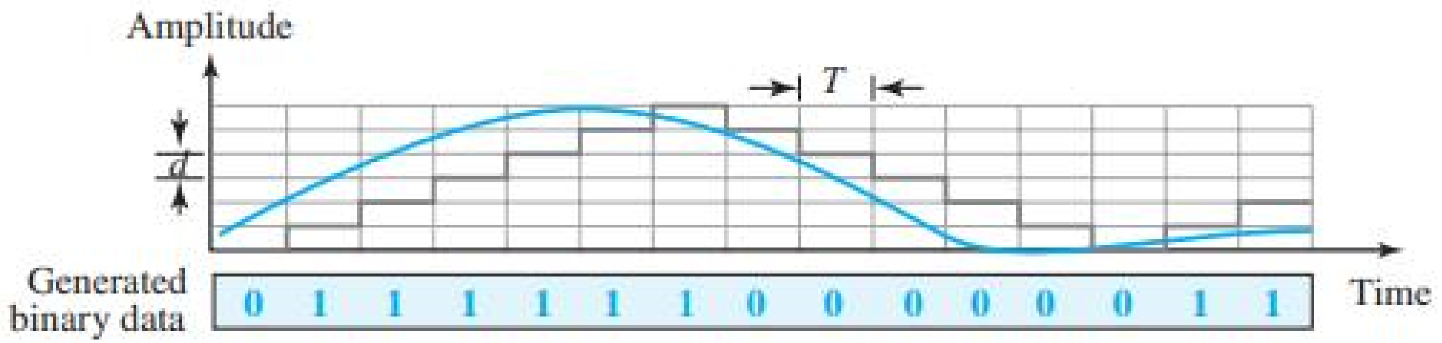
$$B_{min} = n_b \times B_{analog}$$

This means the minimum bandwidth of the digital signal is n_b times greater than the bandwidth of the analog signal. This is the price we pay for digitization.

Delta Modulation (DM)

PCM is a complex technique, while delta modulation (DM) is simpler. PCM measures the signal amplitude for each sample, whereas DM measures the change from the previous sample. Figure 2.10 illustrates this process, where bits are sent sequentially without code words.

Figure 2.10 *The process of delta modulation*



Analog Transmission

Digital transmission requires a low-pass channel (starting from 0), whereas analog transmission is necessary for a bandpass channel (not starting from 0). **Digital-to-analog conversion converts digital data to a bandpass analog signal, while analog-to-analog conversion changes a low-pass analog signal to a bandpass analog signal.** This section covers both conversion types.

Digital-to-Analog Conversion

Digital-to-analog conversion modifies one characteristic of an analog signal based on digital data. This involves altering amplitude, frequency, or phase, resulting in three modulation methods: amplitude shift keying (ASK), frequency shift keying (FSK), and phase shift keying (PSK). A more efficient method, quadrature amplitude modulation (QAM), combines changes in both amplitude and phase and is the most commonly used technique today.

Amplitude Shift Keying

In amplitude shift keying (ASK), the carrier signal's amplitude is varied to create signal elements, while frequency and phase remain constant.

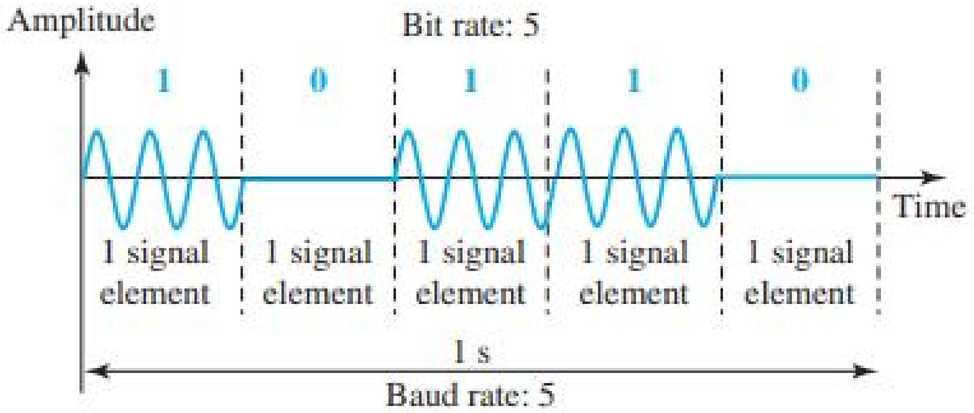
Binary ASK (BASK)

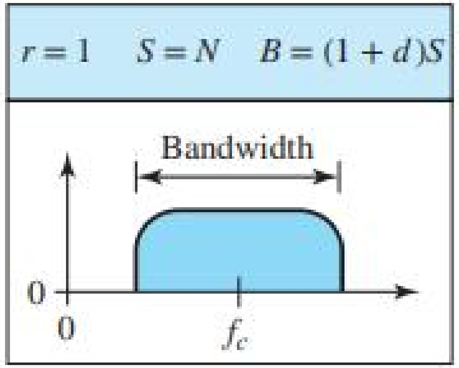
ASK typically uses two levels, known as binary amplitude shift keying (BASK) or on-off keying (OOK). One signal level's peak amplitude is 0, and the other matches the carrier frequency's amplitude. The modulation creates a nonperiodic composite signal with a continuous set of frequencies. The bandwidth, proportional to the signal rate (S), is influenced by a factor d (0 to 1), giving the bandwidth formula $B = (1 + d)S$. Thus, the required bandwidth ranges from S to $2S$. The carrier frequency f_c is at the center of this bandwidth, allowing the modulated signal to fit available bandpass channels, which is a key advantage of digital-to-analog conversion.

Baud rate :

Baud rate is the **rate at which the number of signal elements or changes to the signal occurs per second** when it passes through a transmission medium. The higher a baud rate is the faster the data is sent/received.

Baud rate = number of signal elements/total time (in seconds)





The Baud rate refers to the total number of signal units transmitted in one second. The Bit rate refers to the total Bits transmitted in one unit time.

Multilevel ASK

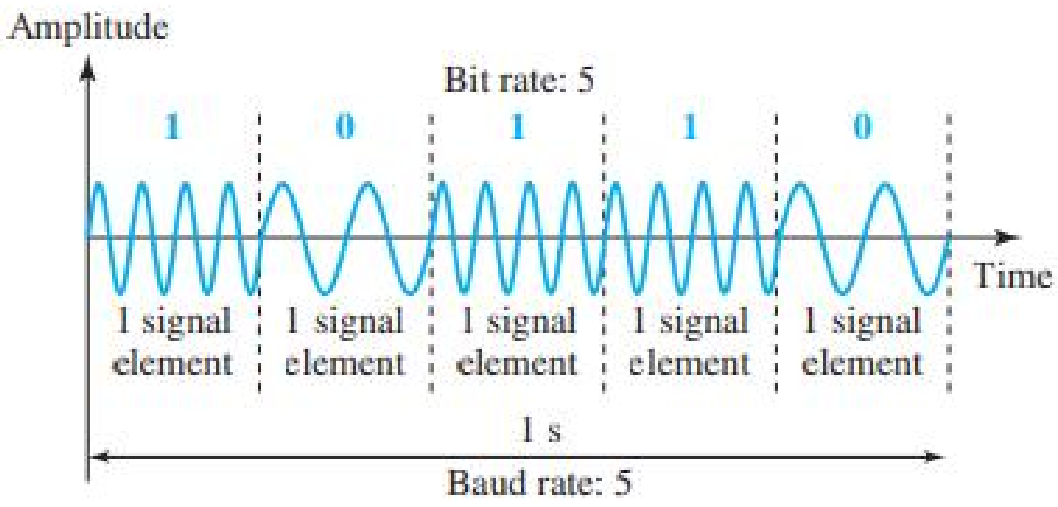
Multilevel ASK uses more than two amplitude levels, such as 4, 8, or 16, to modulate data using 2, 3, 4, or more bits at a time. This is commonly implemented with QAM rather than pure ASK.

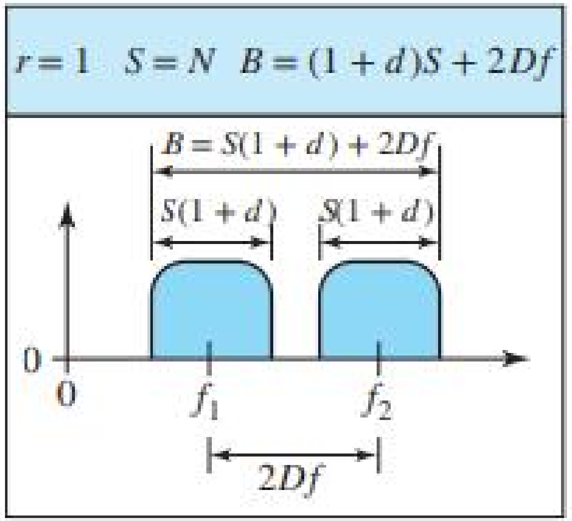
Frequency Shift Keying (FSK)

In frequency shift keying (FSK), the carrier signal's frequency is varied to represent data, while its amplitude and phase remain constant. The frequency stays the same for each signal element and changes only if the data element changes.

Binary FSK (BFSK)

Binary FSK (BFSK) uses two carrier frequencies: f_1 for data element 0 and f_2 for data element 1. Typically, these carrier frequencies are very high, and the difference between them is minimal. Figure 2.12 illustrates this concept for demonstration purposes.



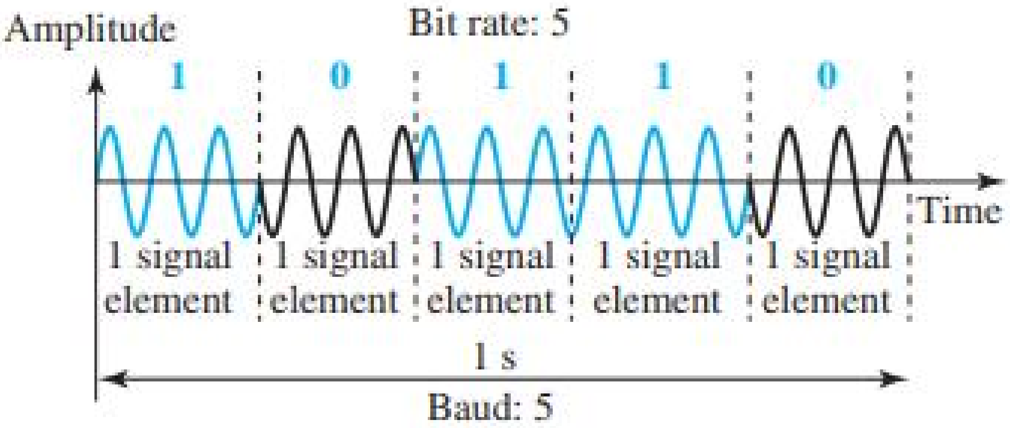


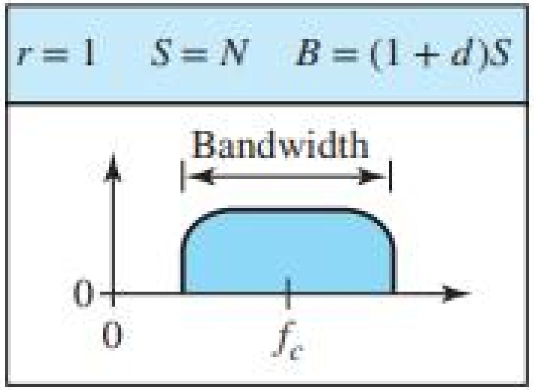
Phase Shift Keying

In phase shift keying (PSK), the carrier's phase is varied to represent different signal elements, while amplitude and frequency remain constant. PSK is more common than ASK or FSK, but QAM, which combines ASK and PSK, is the dominant digital-to-analog modulation method.

Binary PSK (BPSK)

Binary PSK (BPSK) is the simplest form of PSK, using two signal elements with phases of 0° and 180°. As shown in Figure 2.13, BPSK is as simple as binary ASK but has the advantage of being less susceptible to noise.



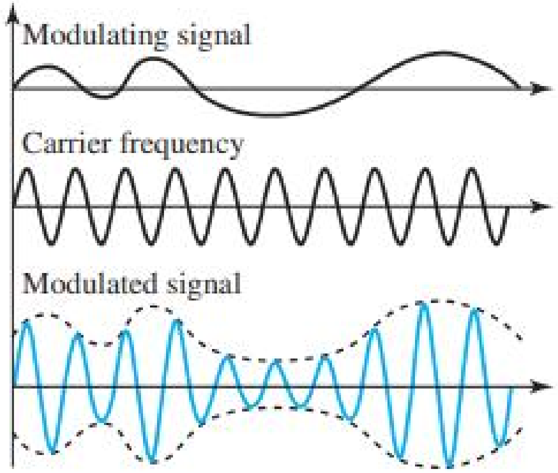


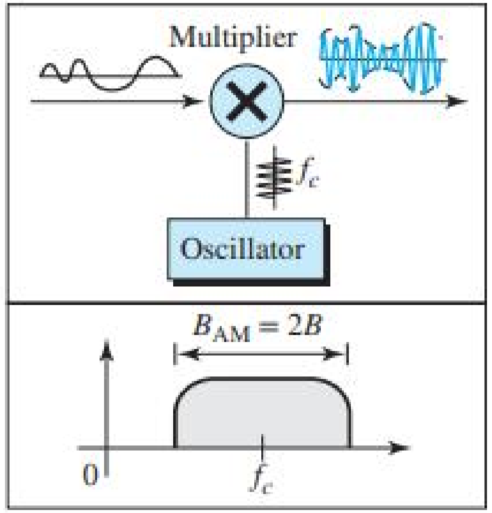
Analog-to-Analog Conversion

Analog-to-analog conversion, or analog modulation, represents analog information with an analog signal. It is necessary when using a bandpass medium or channel, such as radio, where each station is assigned a narrow bandwidth. Low-pass signals from different stations need to be shifted to different ranges for clear reception. This conversion can be done through amplitude modulation (AM), frequency modulation (FM), and phase modulation (PM), with FM and PM often grouped together.

Amplitude Modulation

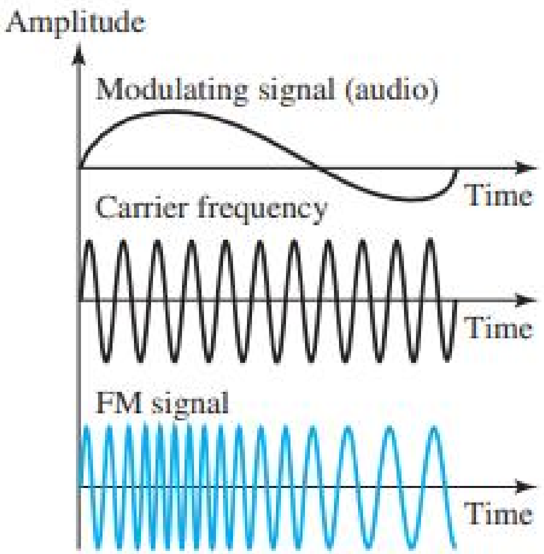
In AM transmission, the carrier signal's amplitude varies with the modulating signal's amplitude, while its frequency and phase remain constant. The modulating signal shapes the carrier's envelope. AM is typically implemented using a multiplier to adjust the carrier's amplitude. The bandwidth of an AM signal is twice that of the modulating signal and is centered on the carrier frequency. Since the upper and lower bands carry identical information, some implementations discard one-half of the signal to reduce bandwidth.

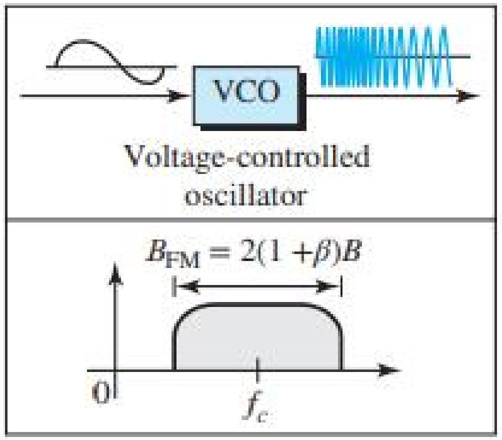




Frequency Modulation

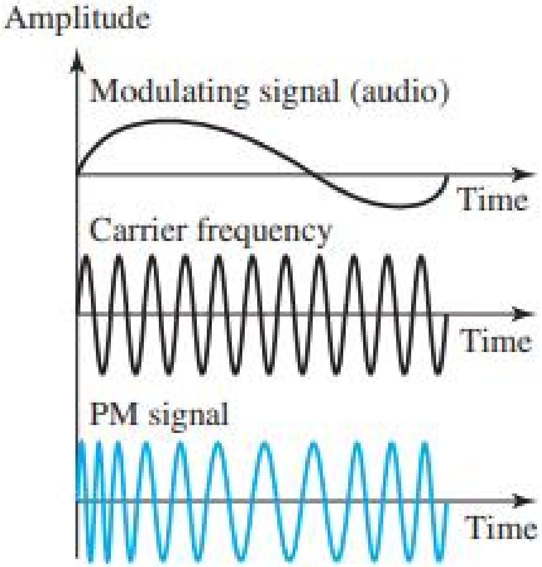
In FM transmission, the carrier signal's frequency varies with the modulating signal's amplitude, while its peak amplitude and phase remain constant. The frequency changes according to the input voltage from the modulating signal. FM is typically implemented using a voltage-controlled oscillator. The bandwidth of an FM signal is given by $B_{FM} = 2(1+\beta)B$, where β is a modulation factor, often around 4.

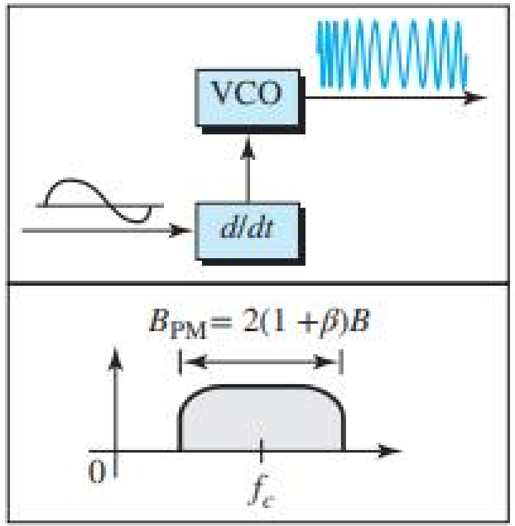




Phase Modulation

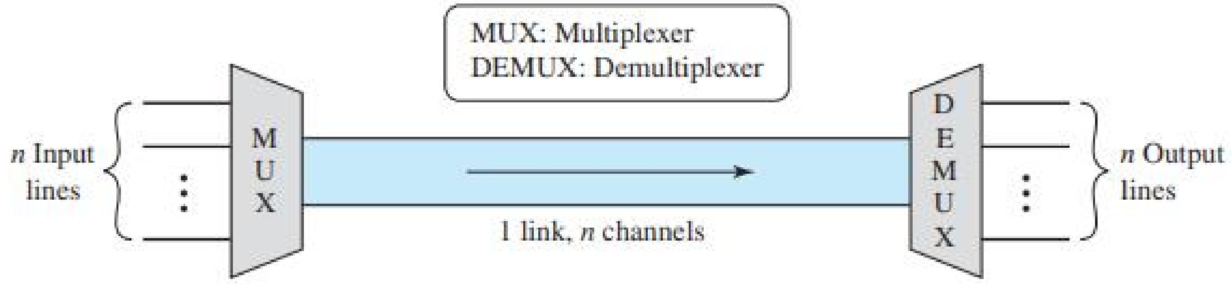
In PM transmission, the carrier signal's phase varies with the modulating signal's amplitude, while its peak amplitude and frequency remain constant. The phase changes as the modulating signal's amplitude changes. PM is similar to FM, but in FM, the frequency change is proportional to the modulating signal's amplitude, whereas in PM, it is proportional to the derivative of the amplitude. PM is typically implemented using a voltage-controlled oscillator and a derivative. The bandwidth of a PM signal is empirically several times that of the analog signal, with the modulation factor β being lower (around 1 for narrowband and 3 for wideband).





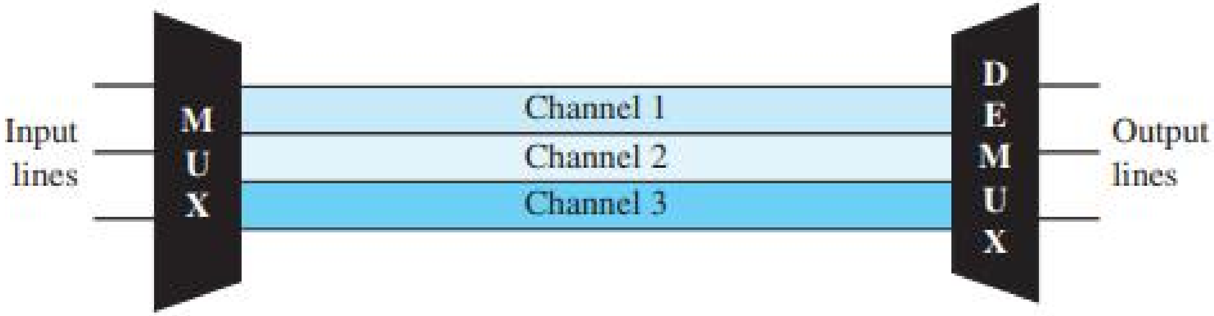
MULTIPLEXING

When the bandwidth of a medium linking two devices exceeds their bandwidth needs, the link can be shared using multiplexing techniques. Multiplexing allows multiple signals to be transmitted simultaneously across a single data link. This is essential as data and telecommunications traffic increases. Instead of adding individual links for each new channel, higher-bandwidth links can be installed to carry multiple signals. High-bandwidth media such as optical fiber and microwave technologies support this. In a multiplexed system, multiple lines share the bandwidth of one link, combining streams into a single one via a multiplexer and then separating them at the receiving end with a demultiplexer. The physical path is called a link, and each transmission path within it is a channel. The three basic multiplexing techniques are frequency-division multiplexing (FDM) and wavelength-division multiplexing (WDM) for analog signals, and time-division multiplexing (TDM) for digital signals.



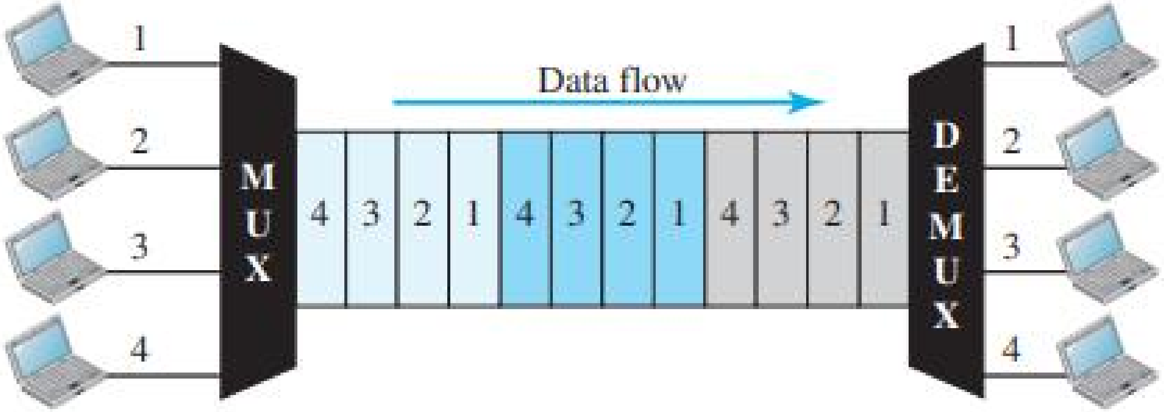
Frequency-Division Multiplexing

Frequency-division multiplexing (FDM) is an analog technique used when a link's bandwidth (in hertz) exceeds the combined bandwidths of the signals to be transmitted. Each signal modulates a different carrier frequency, and these modulated signals are combined into a single composite signal for transmission. Carrier frequencies are separated by sufficient bandwidth to avoid overlap, with guard bands used to prevent interference. Carrier frequencies must also avoid interfering with the original data frequencies. Although FDM is an analog technique, it can be used for digital signals by converting them to analog before multiplexing.



Time-Division Multiplexing

Time-division multiplexing (TDM) is a digital technique that allows multiple connections to share a high-bandwidth link by dividing time rather than bandwidth. Each connection uses a portion of the link's time sequentially. In TDM, data from different sources are combined into a single time-shared link, with each source's data occupying the link in turns. Unlike switching, TDM ensures fixed and unvarying delivery to specific destinations. Although primarily digital, TDM can also handle analog data by converting it to digital before multiplexing.

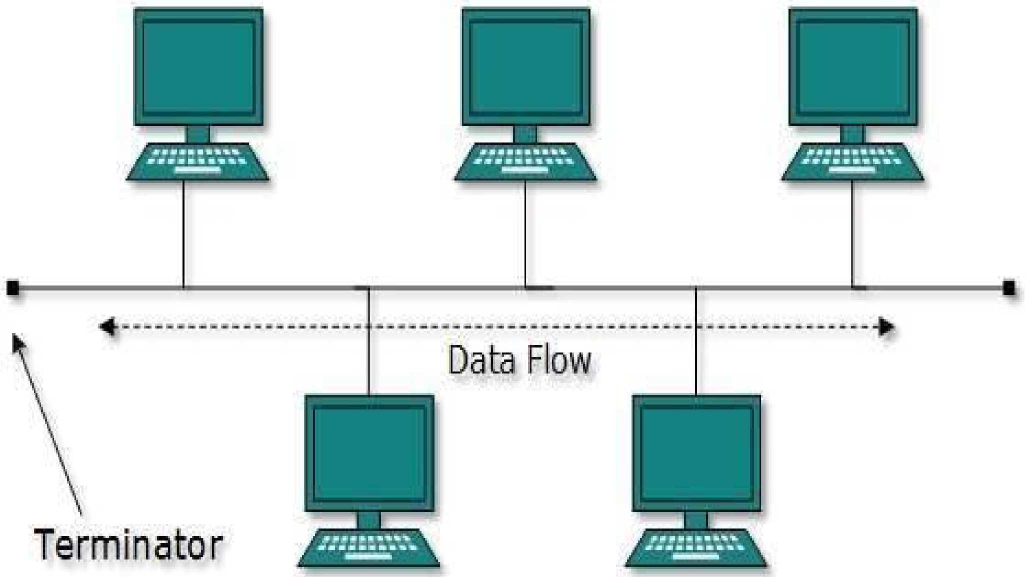


Network topologies

- Network topology refers to the arrangement of different elements like nodes, links, and devices in a computer network.
- It defines how these components are connected and interact with each other.
- Understanding various types of network topologies helps in designing efficient and robust networks.
- Common types include bus, star, ring, mesh, and tree topologies, each with its own advantages and disadvantages.

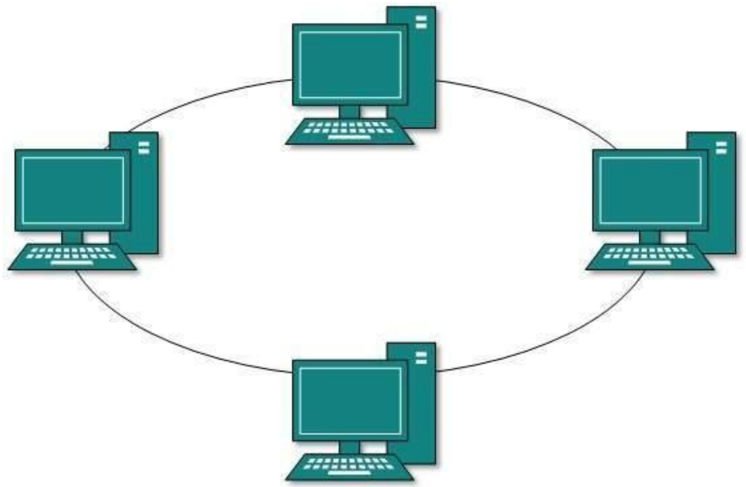
types of network topology

- Bus Topology
  - In case of Bus topology, all devices share single communication line or cable.
  - It is one of the simple forms of networking where a **failure of a device does not affect the other devices**. But **failure of the shared communication line can make all other devices stop** functioning.
  - Both ends of the shared channel have line terminator. The data is sent in **only one direction** and as soon as it reaches the extreme end, the terminator removes the data from the line.



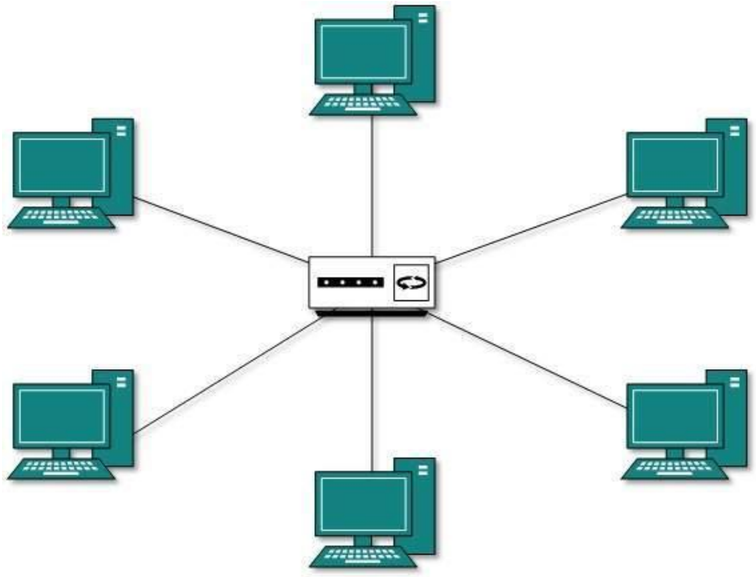
types of network topology

- Ring Topology -
  - In ring topology, each host machine connects to exactly two other machines, creating a circular network structure.
  - When one host tries to communicate or send message to a host which is not adjacent to it, the data travels through all intermediate hosts.
  - To connect one more host in the existing structure, the administrator may need only one more extra cable.
  - Failure of any host results in failure of the whole ring.Thus, every connection in the ring is a point of failure.



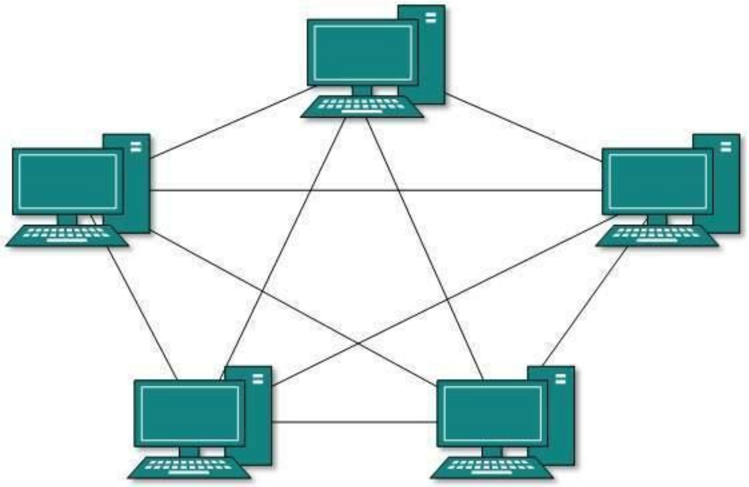
types of network topology

- Star Topology -
  - In Star Topology, all the devices are connected to a single hub through a cable.
  - This hub is the central node and all other nodes are connected to the central node.
  - If hub fails, connectivity of all hosts to all other hosts fails. Every communication between hosts, takes place through only the hub.
  - Star topology is not expensive as to connect one more host, only one cable is required and configuration is simple.



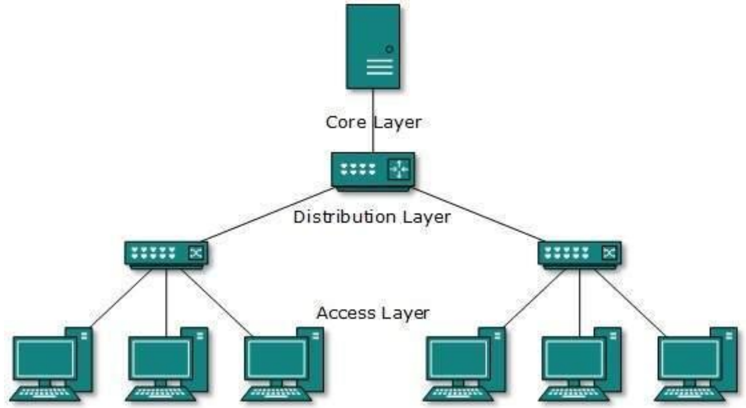
types of network topology

- Mesh Topology -
  - In this type of topology, a host is connected to one or multiple hosts.
  - This topology has hosts in point-to-point connection with every other host or may also have hosts which are in point-to-point connection to few hosts only.
  - Hosts in Mesh topology also work as relay for other hosts which do not have direct point-to-point links.
  - Full Mesh - All hosts have a point-to-point connection to every other host in the network. Thus for every new host n(n-1)/2 connections are required. It provides the most reliable network structure among all network topologies.
  - Partially Mesh - Not all hosts have point-to-point connection to every other host. Hosts connect to each other in some arbitrarily fashion. This topology exists where we need to provide reliability to some hosts out of all.



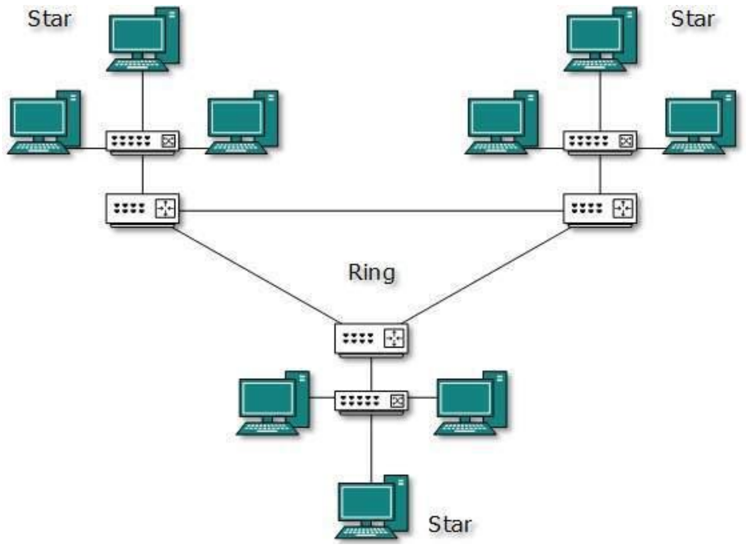
types of network topology

- Tree Topology or Hierarchical Topology
    - This topology divides the network in to multiple levels/layers of network.
    - Mainly in LANs, a network is bifurcated into three types of network devices. The lowermost is access-layer where computers are attached.
    - The middle layer is known as distribution layer, which works as mediator between upper layer and lower layer.
    - The highest layer is known as core layer, and is central point of the network, i.e. root of the tree from which all nodes fork.
    - Similar to the Bus topology, if the root goes down, then the entire network suffers even though it is not the single point of failure.
    - Every connection serves as point of failure, failing of which divides the network into unreachable segment.



types of network topology

- Hybrid Topology
  - A network structure whose design contains more than one topology is said to be hybrid topology. Hybrid topology inherits merits and demerits of all the incorporating topologies.
  - Most WANs are connected by means of Dual-Ring topology and networks connected to them are mostly Star topology networks.
  - Internet is the best example of largest Hybrid topology.



comparison of network topologies

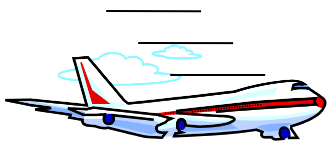
Topology	Advantages	Disadvantages
Bus	Simple, cost-effective, easy to extend	Difficult to troubleshoot, limited length and nodes, single point of failure (backbone)
Star	Easy to install/manage, failure isolation	Central point of failure (hub/switch), more cabling
Ring	High speed, no collisions	Failure of one device affects all, difficult to troubleshoot
Mesh	Highly reliable, redundant paths	Expensive, complex configuration
Tree	Scalable, easy management	More cabling, backbone failure affects entire network
Hybrid	Flexible, robust, scalable	omplex design/maintenance, costly

protocols & layered architecture of internet

protocols

- As we have discussed that internet is a complex network containing many hosts, routers, transmission links, etc. So how is this complex structure is managed or operates?
- The answer is Protocols. A protocol defines the rules that both sender and receiver and all intermediate devices need to follow to be able to communicate effectively.
- To reduce the complexity, the internet architecture is divided into protocol layers, where each layer serve a different purpose in the communication between two end-systems in the internet.

Example: organization of air travel



end-to-end transfer of person plus baggage

ticket (purchase)

baggage (check)

gates (load)

runway takeoff

airplane routing

ticket (complain)

baggage (claim)

gates (unload)

runway landing

airplane routing

airplane routing

How would you *define/discuss* the *system* of airline travel?

- a series of steps, involving many services

Example: organization of air travel

ticket (purchase)	*ticketing service*	ticket (complain)
baggage (check)	*baggage service*	baggage (claim)
gates (load)	*gate service*	gates (unload)
runway takeoff	*runway service*	runway landing
airplane routing	*routing service*	airplane routing

layers: each layer implements a service

- via its own internal-layer actions
- relying on services provided by layer below

Why layering?

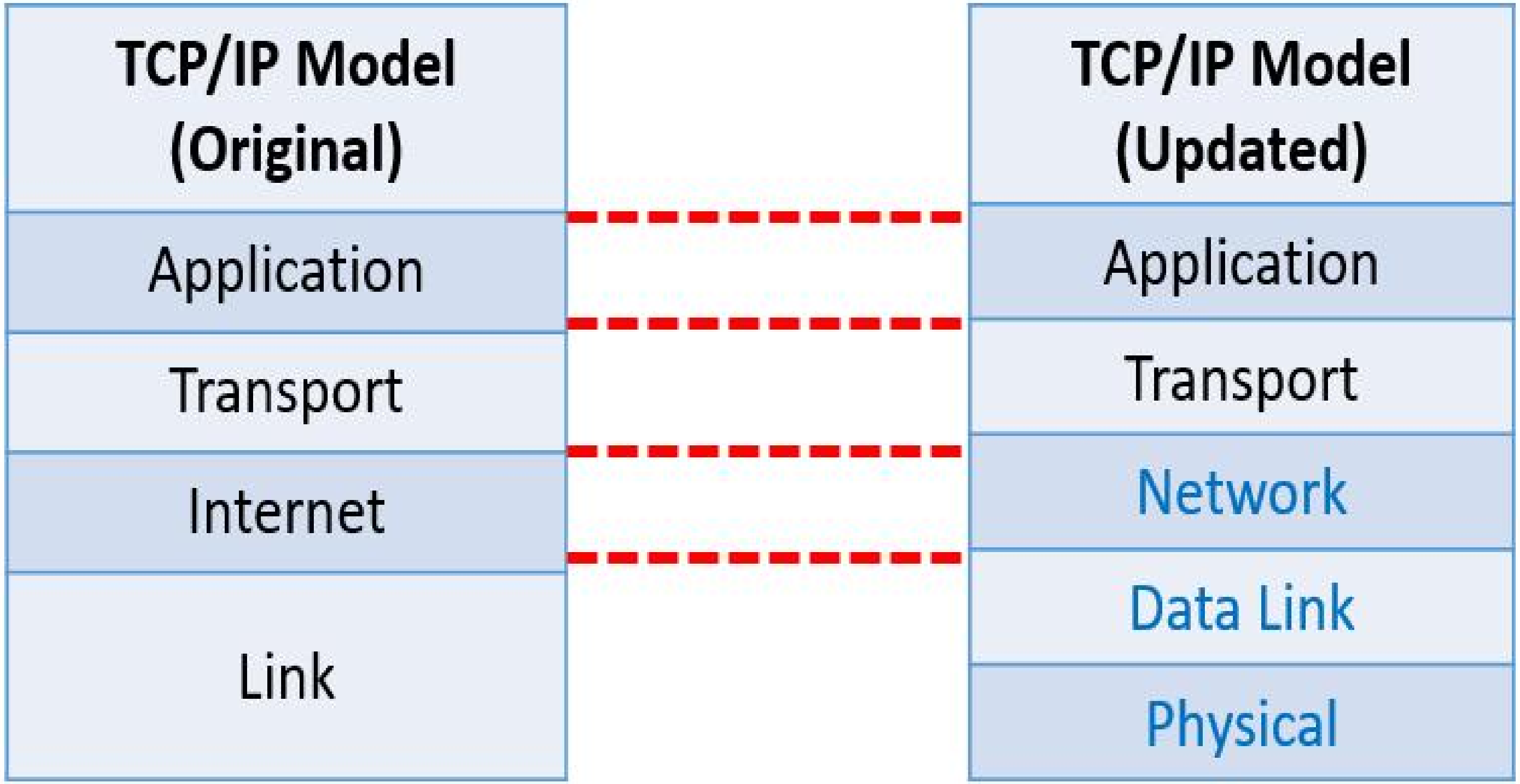
Approach to designing/discussing complex systems:

- explicit structure allows identification, relationship of system's pieces
  - layered *reference model* for discussion
- modularization eases maintenance, updating of system
  - change in layer's service *implementation*: transparent to rest of system
  - e.g., change in gate procedure doesn't affect rest of system

tcp/ip protocol suite

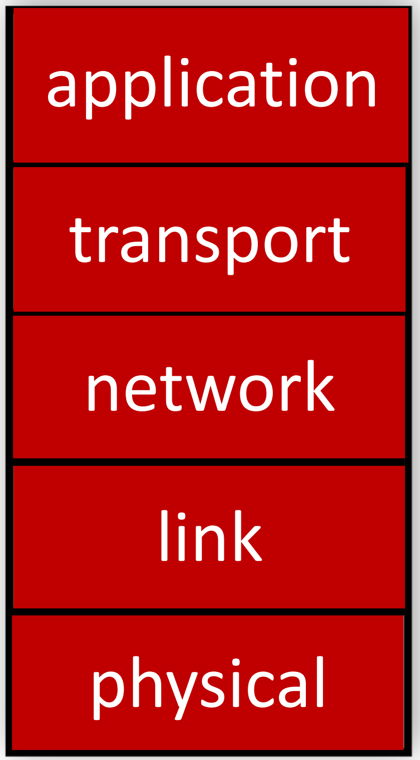
- TCP/IP (Transmission Control Protocol / Internet Protocol) is the protocol suite (a set of protocols organized in different layers) which is used in the internet today.
- TCP/IP is a hierarchical protocol made up of interactive modules, each of which provides a specific functionality.
- The term hierarchical means that each upper level protocol is supported by the services provided by one or more lower level protocols.
- The original TCP/IP protocol suite was defined as four software layers built upon the hardware. Today, however, TCP/IP is thought of as a five-layer model.

Layers in tcp/ip protocol suite

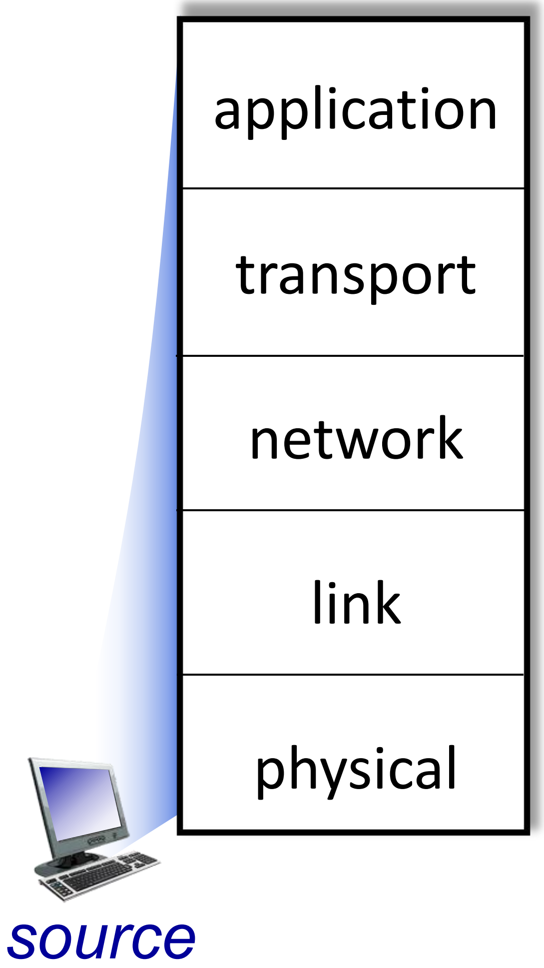


Layers in TCP/IP Stack

- *application:* supporting network applications
  - HTTP, IMAP, SMTP, DNS
- *transport:* process-process data transfer
  - TCP, UDP
- *network:* routing of datagrams from source to destination
  - IP, routing protocols
- *link:* data transfer between neighboring network elements
  - Ethernet, 802.11 (WiFi), PPP
- *physical:* bits “on the wire”



Services, Layering and Encapsulation



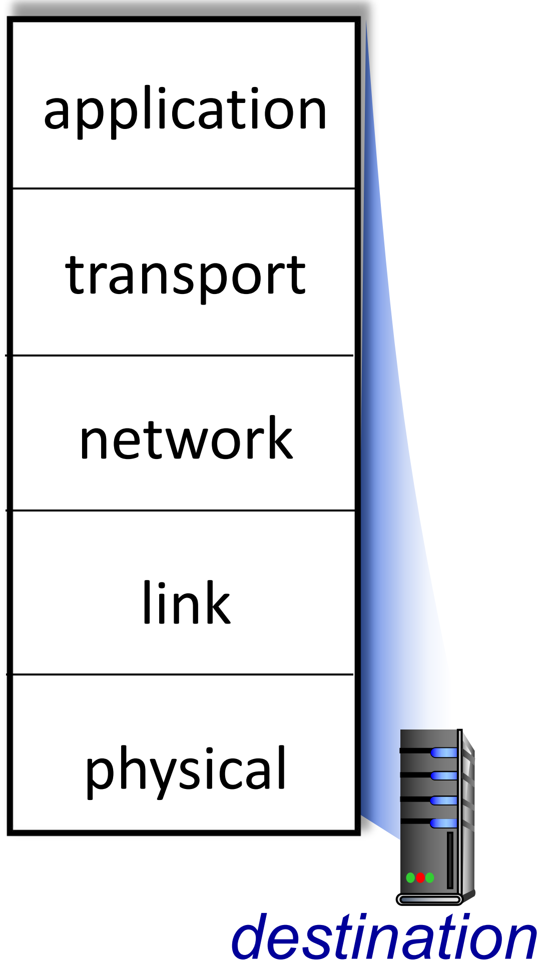
M

Application exchanges messages to implement some application service using *services* of transport layer

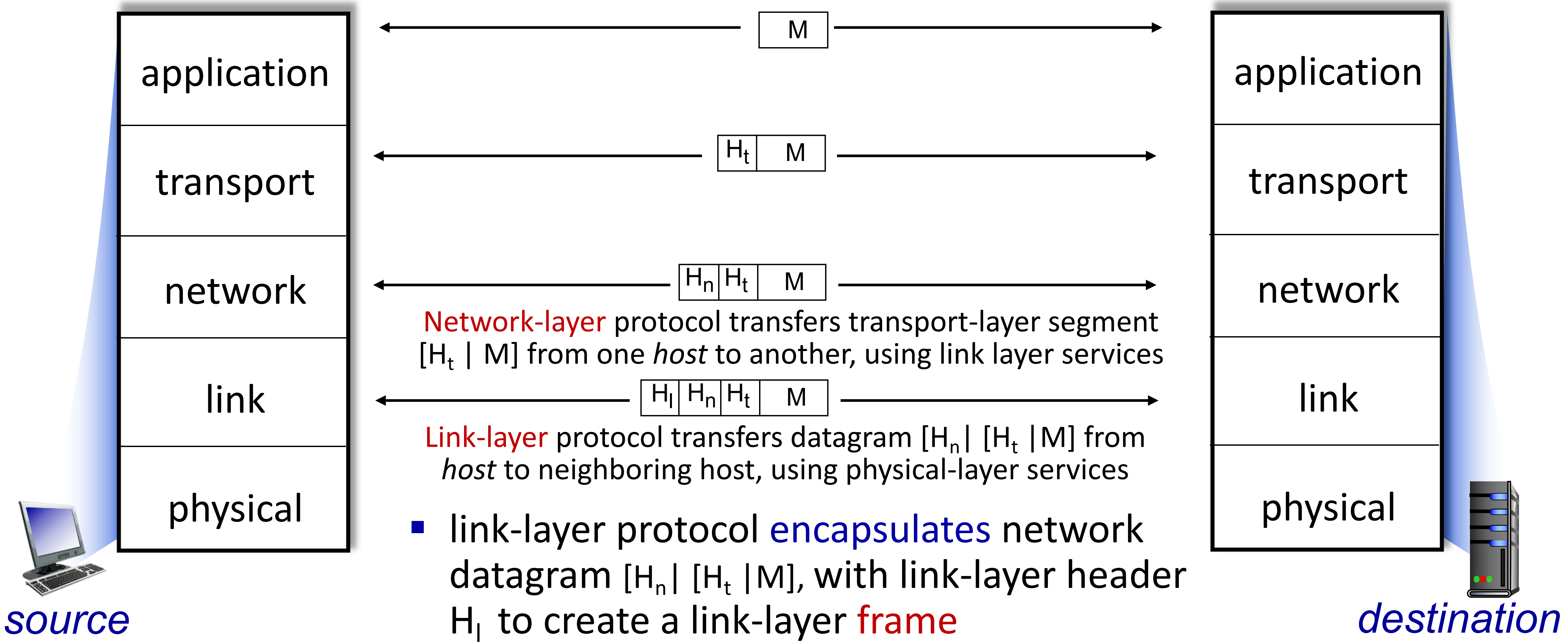
H_t M

Transport-layer protocol transfers M (e.g., reliably) from one *process* to another, using services of network layer

- transport-layer protocol encapsulates application-layer message, M, with *transport* layer-layer header H_t to create a transport-layer segment
  - H_t used by transport layer protocol to implement its service



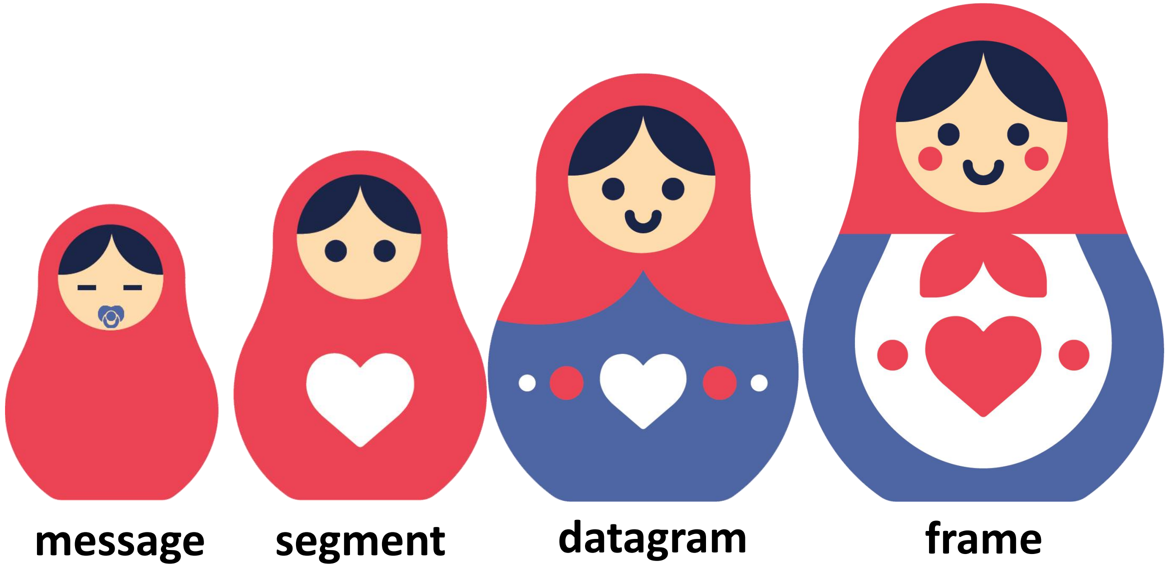
Services, Layering and Encapsulation



- link-layer protocol encapsulates network datagram [H_n| [H_t |M], with link-layer header H_l to create a link-layer frame

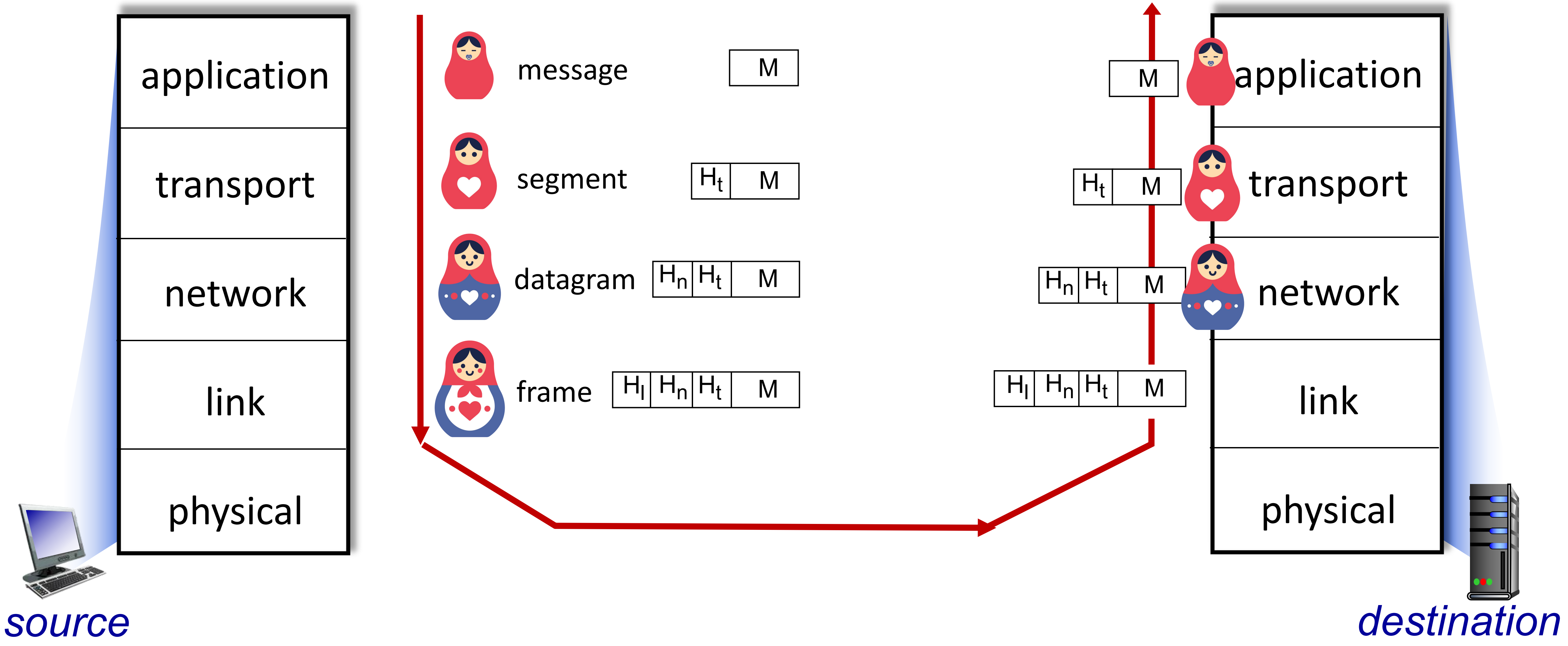
Encapsulation

Matryoshka dolls (stacking dolls)

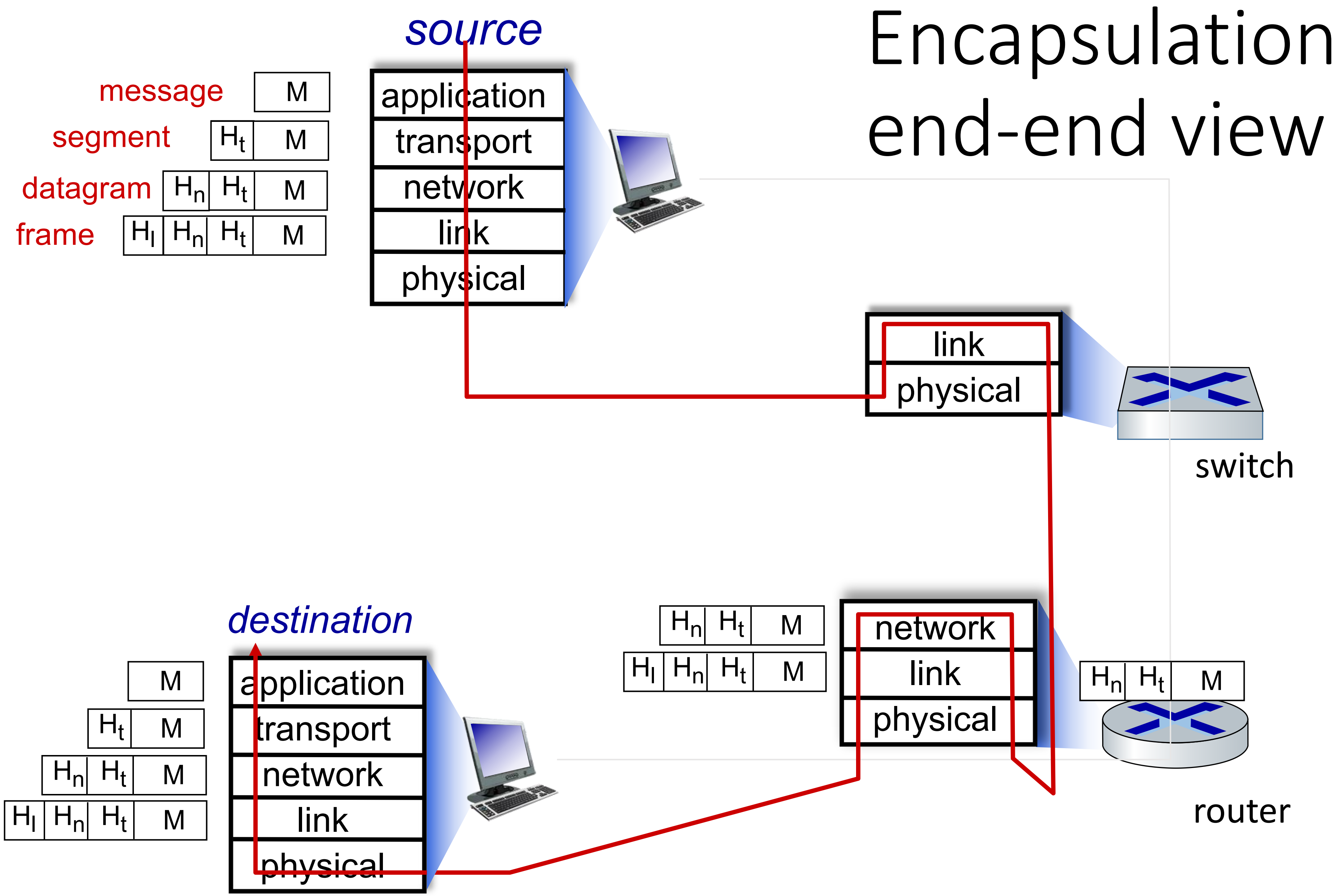


Credit: https://dribbble.com/shots/7182188-Babushka-Boi

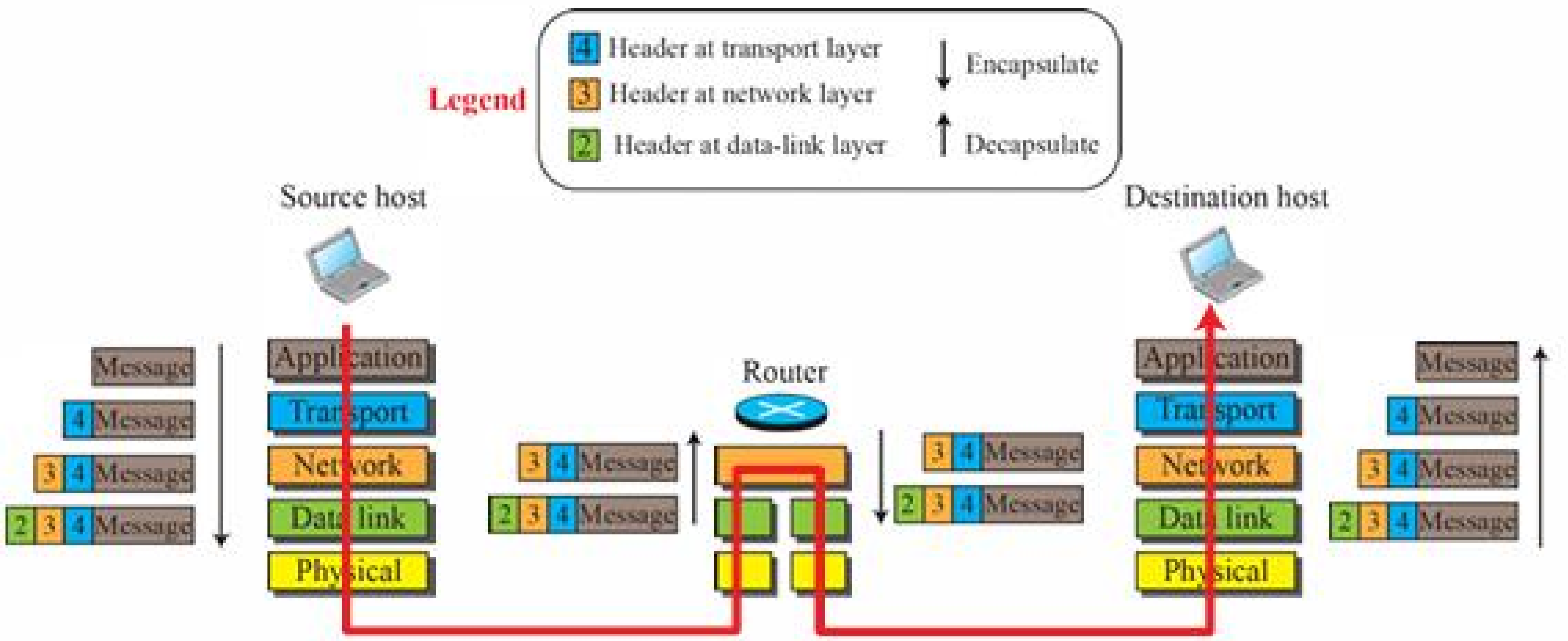
Services, Layering and Encapsulation



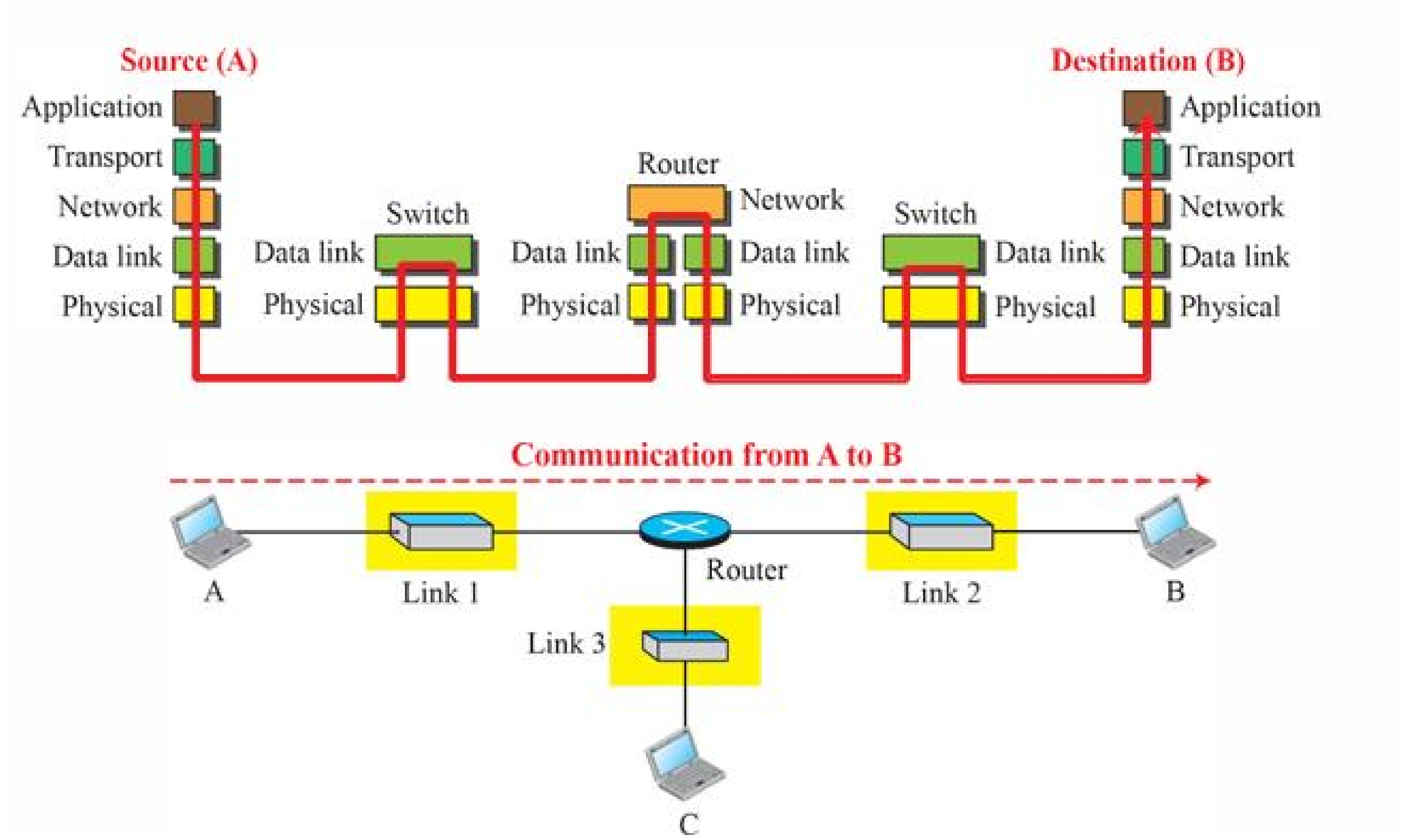
Encapsulation: an end-end view



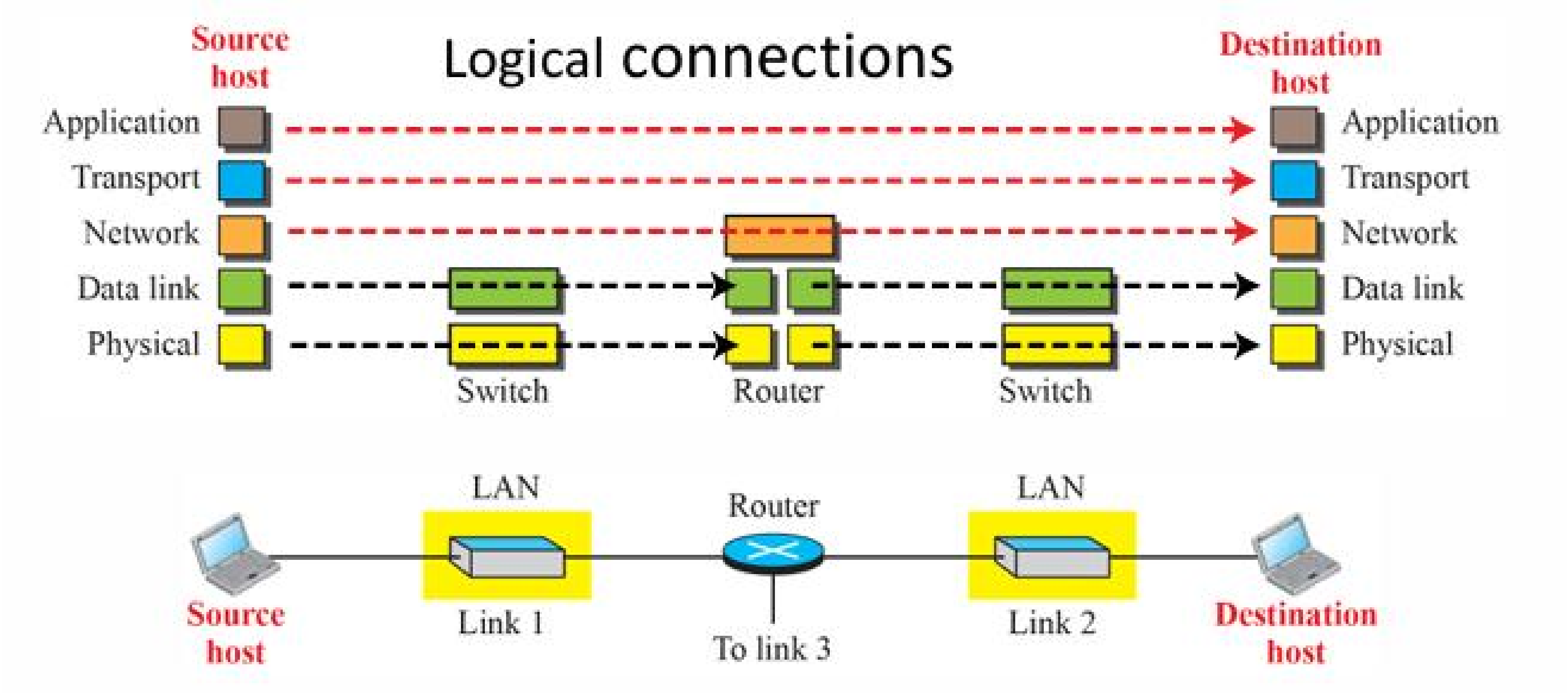
Encapsulation and Decapsulation



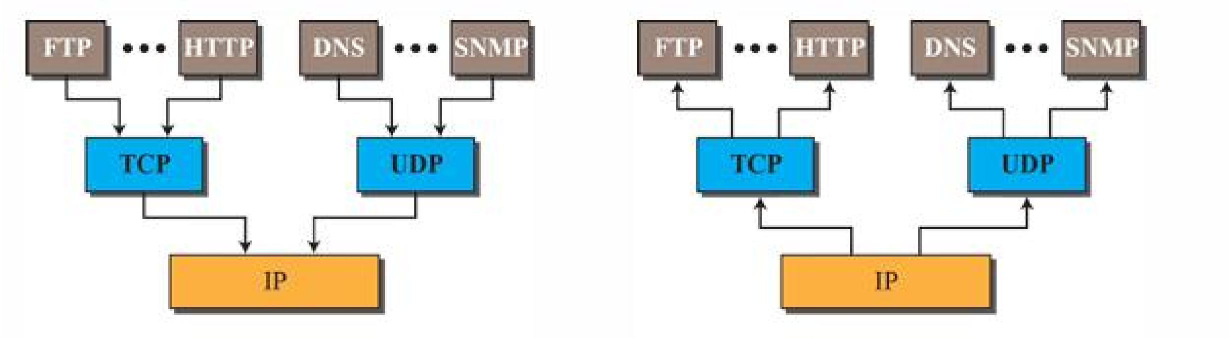
Packet Transmission in TCP/IP Stack



Logical Connections between layers in TCP/IP



Multiplexing and Demultiplexing



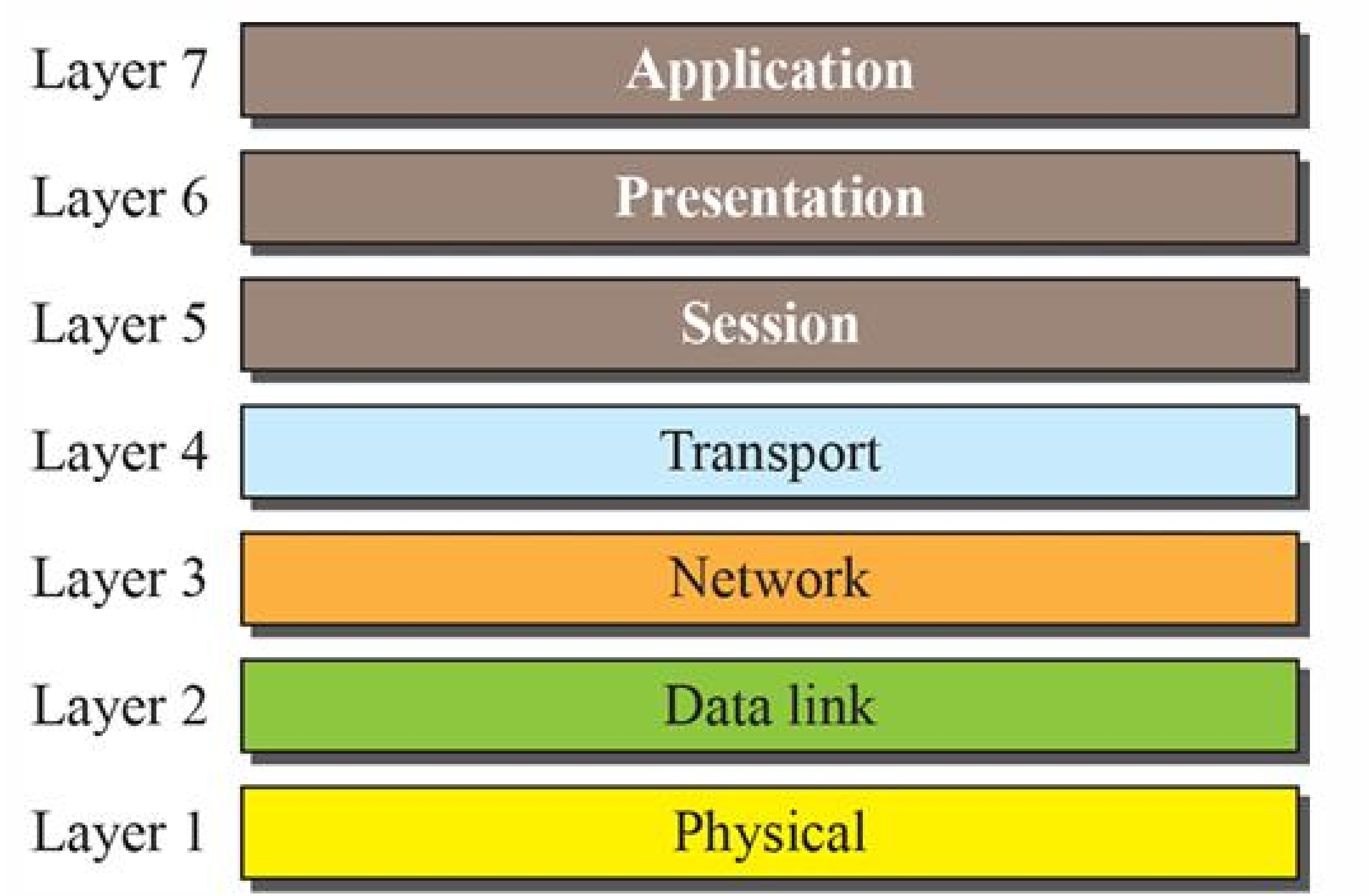
a. Multiplexing at source b. Demultiplexing at destination

- **Multiplexing -** At the source, a protocol at a layer can take data packets from several next-higher layer protocols one at a time and encapsulate them together. This means that data from different applications can be combined into a single packet.
- **Demultiplexing -** At the destination, the receiving protocol can identify and extract the encapsulated packets from the received data packet and deliver them to their respective next-higher layer protocols one at a time. This ensures that the data is correctly routed to the appropriate application or service.

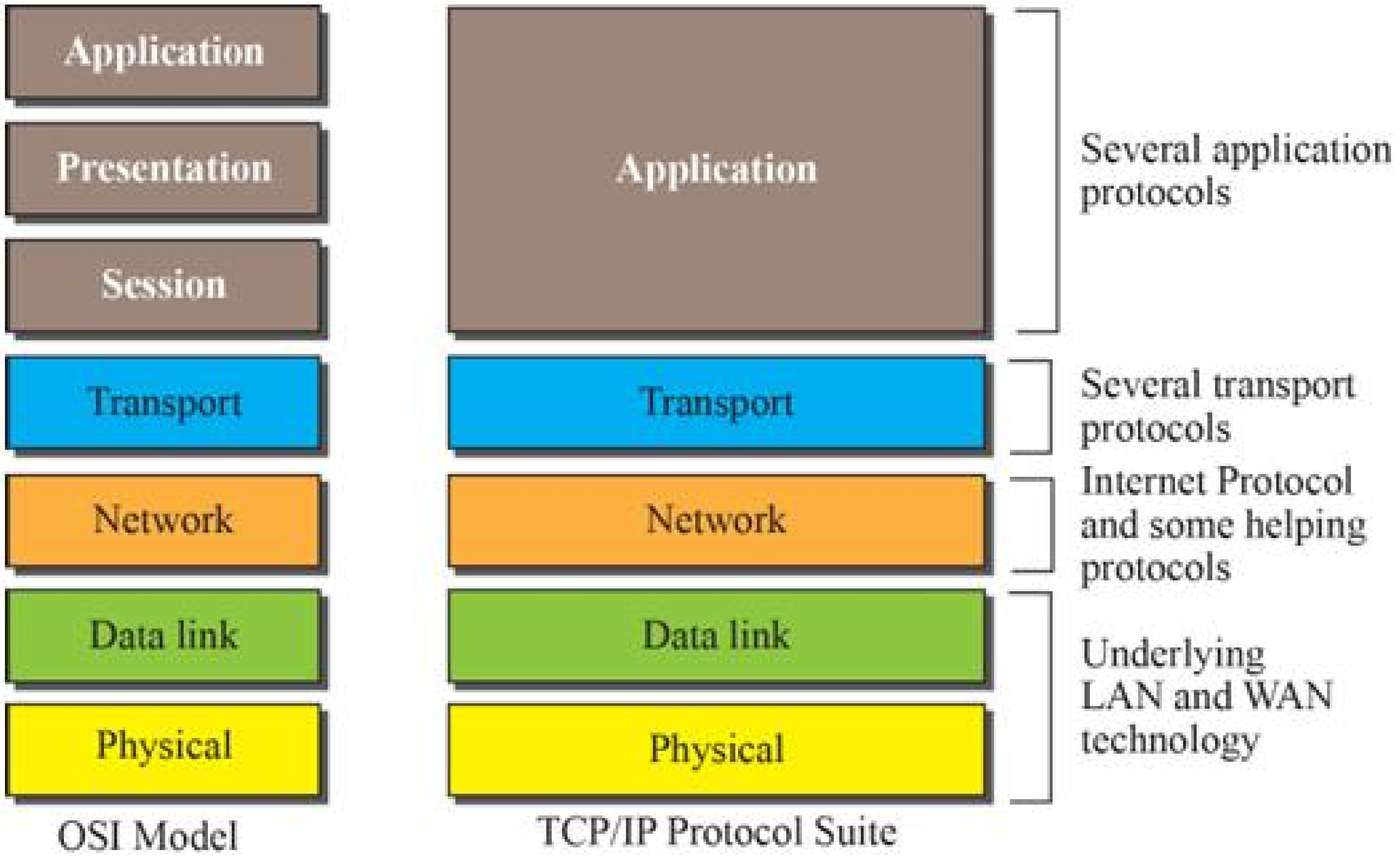
the osi model

- Established in 1947, the **International Organization for Standardization (ISO)** is a multinational body dedicated to worldwide agreement on international standards.
- An ISO standard that covers all aspects of network communications is the **Open Systems Interconnection (OSI) model**. It was first introduced in the late 1970s.
- An open system is like a universal translator that enables different systems to talk to each other, regardless of their differences. The OSI model is a blueprint that shows how to make this communication happen without changing the underlying hardware or software.
- The OSI model is a layered framework for the design of network systems that allows communication between all types of computer systems. It consists of seven separate but related layers, each of which defines a part of the process of moving information across a network

The OSI Model



OSI Vs TCP/IP



TCP/IP and OSI model

osi vs tcp/ip

- **Multiple Transport-layer Protocols**: TCP/IP has more than one transport-layer protocol, and some of them provide functionalities that were originally associated with the session layer in the OSI model.
- **Application Layer Variability:** The application layer in TCP/IP is not limited to a single piece of software. Instead, multiple applications can be developed at this layer. If a specific application requires functionalities similar to those in the session and presentation layers, those functionalities can be included in the application's development.
- The absense of session and presentation layers in TCP/IP was a deliberate decision, influenced by the presence of multiple transport layer protocols and the flexibility of the application layer to accommodate various application specific functionalities.

Lack of osi model’s success

- Timing and Cost: By the time the OSI model was completed, TCP/IP was already widely implemented and used. Switching to the OSI model would have required a significant amount of time, effort, and money to replace the existing TCP/IP infrastructure, which was not practical.
- Incomplete Definitions: Some layers in the OSI model, like the presentation and session layers, were not fully defined with concrete protocols and software. Although the services provided by these layers were listed, the actual implementation details were lacking, making it difficult for organizations to adopt them.
- Performance Issues: When organizations attempted to implement the OSI model in different applications, it didn't demonstrate significantly better performance than the established TCP/IP protocol suite. As a result, there was no compelling reason for the Internet authority to switch from TCP/IP to OSI.

switching

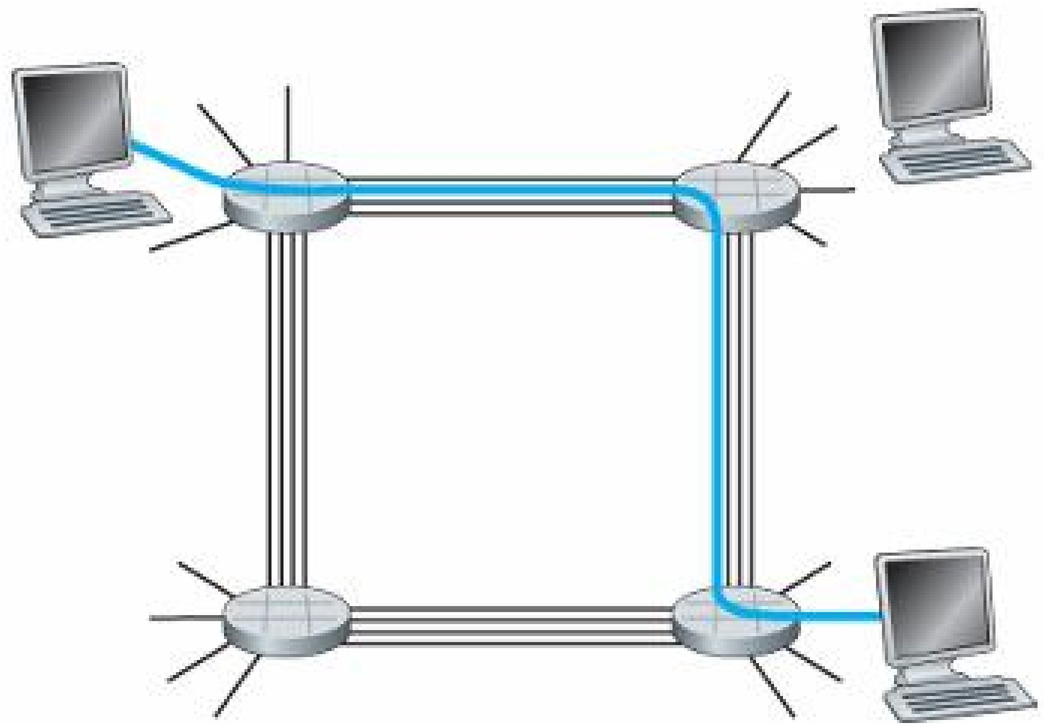
- We have discussed that internet is a network of networks in which two or more networks are connected to each other via switches and routers i.e. we can say that internet is a *swiched network.*
- Switching is the process of transferring data packets from one device to another in a network, or from one network to another via switches.
- There are two types of switching is popular in computer networks -
  - Circuit Switching
  - Packet Switching

Circuit switching

- In a circuit-switched network, a dedicated connection, called a **circuit**, is always available between two end systems; the switch can only make it active or inactive.
- In circuit-switched networks, the resources needed along a path (buffers, link transmission rate) to provide for communication between the end systems are reserved for the duration of the communication session between the end systems.
- Traditional telephone networks are examples of circuit-switched networks.

circuit switching

- In the example, a dedicated connection (in blue) is created between two end systems. This connection or circuit will be reserved for the entire duration of communication.
- Because each link has four circuits, for each link used by the end-to-end connection, the connection gets one fourth of the link's total transmission capacity for the duration of the connection. Thus, for example, if each link between adjacent switches has a transmission rate of 1 Mbps, then each end-to-end circuit-switch connection gets 250 kbps of dedicated transmission rate.
- Total time = setup time + TT + PD+tear down time

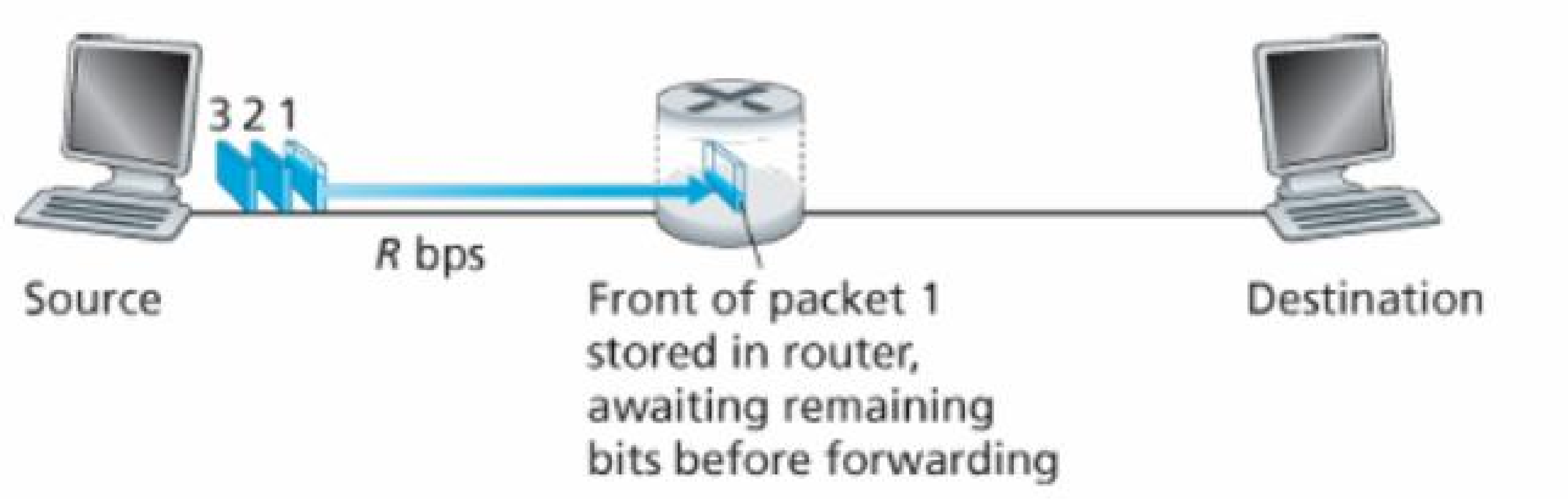


packet switching

- In a network application, end-systems exchanges messages with each other. These messages can be anything like a connection request or maybe some data transfer message like an email, file transfer, etc.
- To send a message from a source end-system to a destination end-system, the source breaks long messages into smaller chunks of data known as **packets**.
- Between source and destination, each packet travels through communication links and *packet switches* (routers, link-layer switches, etc.).
- Packets are transmitted over each communication link at full transmission rate of the link. So if a source end system or a packet switch is sending a packet of ***L bits*** over a link with transmission rate ***R bits / second*** then the time to transmit the packet is ***L / R seconds***.

packet switching: store-and-forward transmission

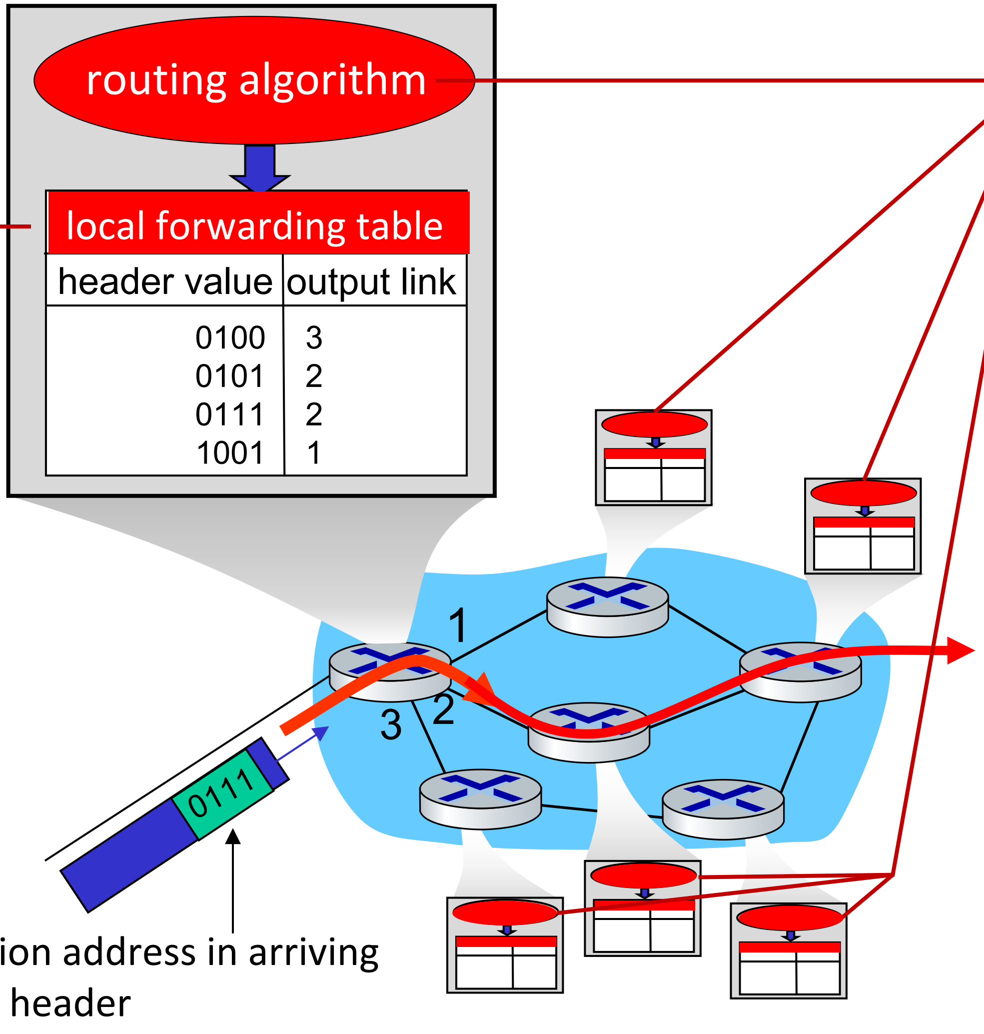
- Most packet switches use store-and-forward transmission at the inputs to the links. Store-and-forward transmission means that the packet switch must receive the entire packet before it can begin to transmit the first bit of the packet onto the outbound link.



Two key network-core functions

Forwarding:

- aka "switching"
- *local* action: move arriving packets from router's input link to appropriate router output link



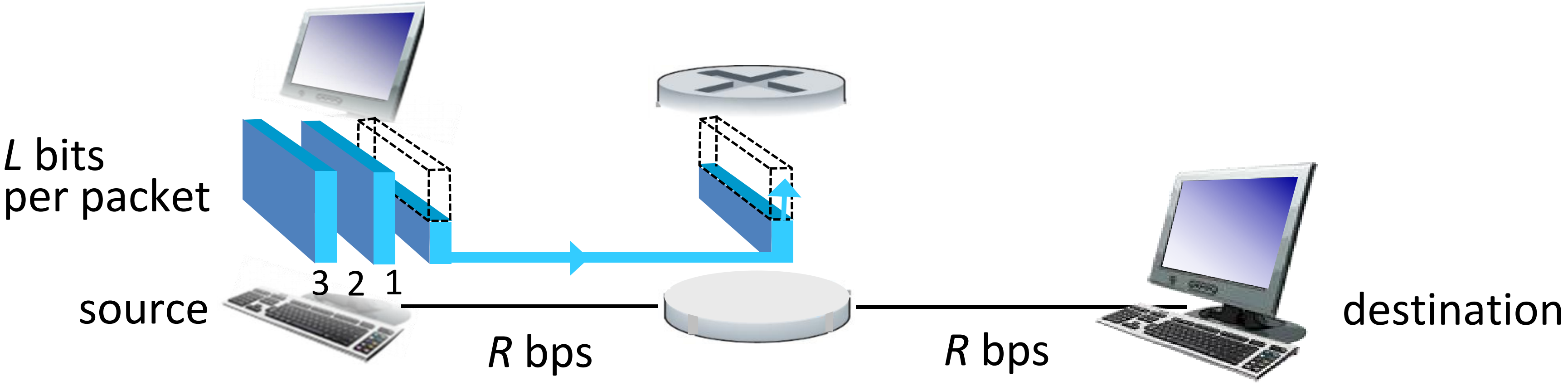
Routing:

- *global* action: determine source-destination paths taken by packets
- routing algorithms


45 h
3,053 miles
48 h
3,309 miles
San Jose
Northampton
United States
routing


forwarding
forwarding
45 h
3,053 miles
48 h
3,309 miles
San Jose
Northampton
Cleveland
Sacramento
United States

Packet-switching: store-and-forward

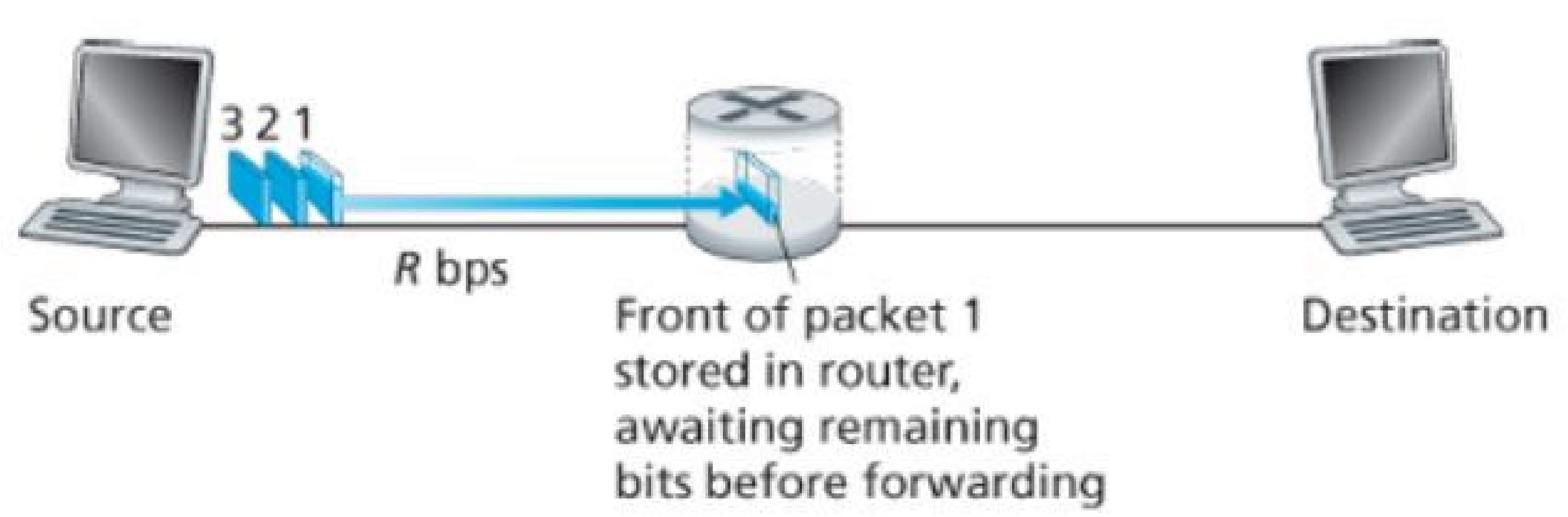


- packet transmission delay: takes *L*/*R* seconds to transmit (push out) *L*-bit packet into link at *R* bps
- *store and forward: entire* packet must arrive at router before it can be transmitted on next link

One-hop numerical example:

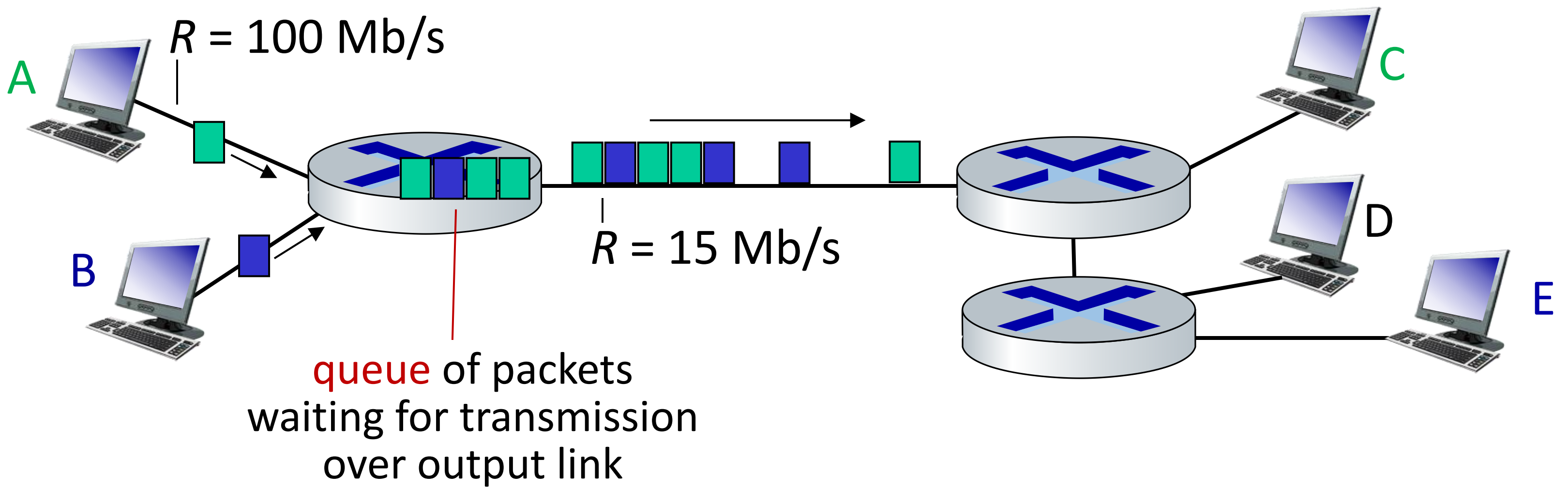
- *L* = 10 Kbits
- *R* = 100 Mbps
- one-hop transmission delay = 0.1 msec

exercise



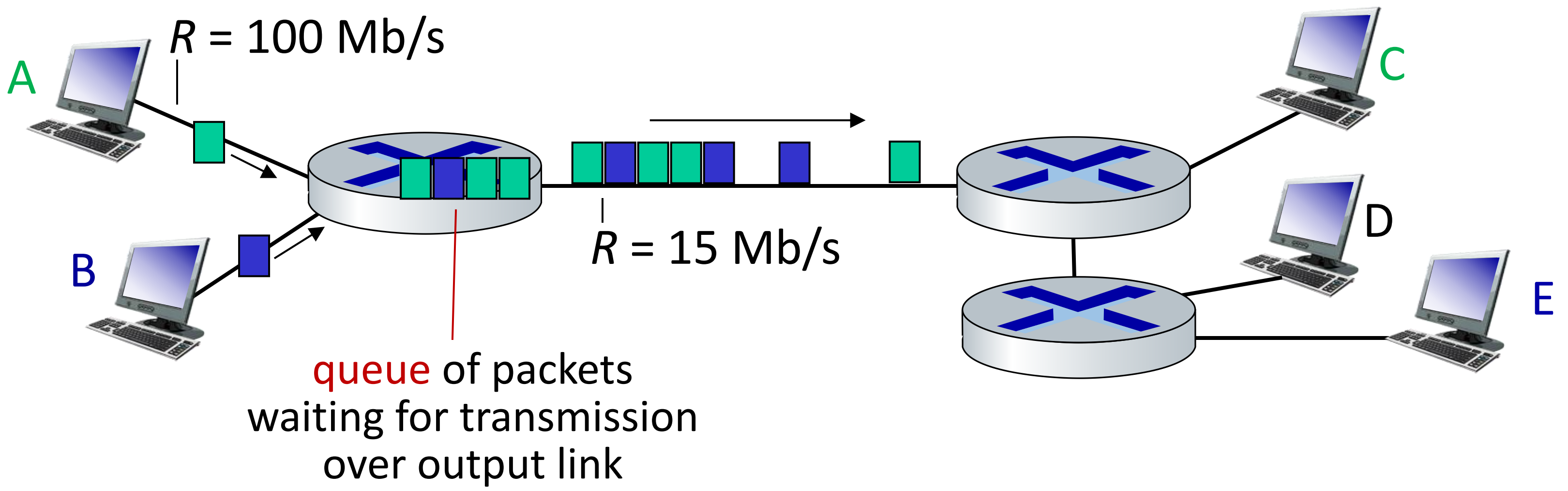
- In given example, how much time will it take for 1st packet to reach the destination? - (L/R)
- Total time taken in transmitting all three packets from source to destination? - n(Transmission time) + propagation delay

Packet-switching: queueing delays and packet loss



Queueing occurs when work arrives faster than it can be serviced:

Packet-switching: queueing delays and packet loss

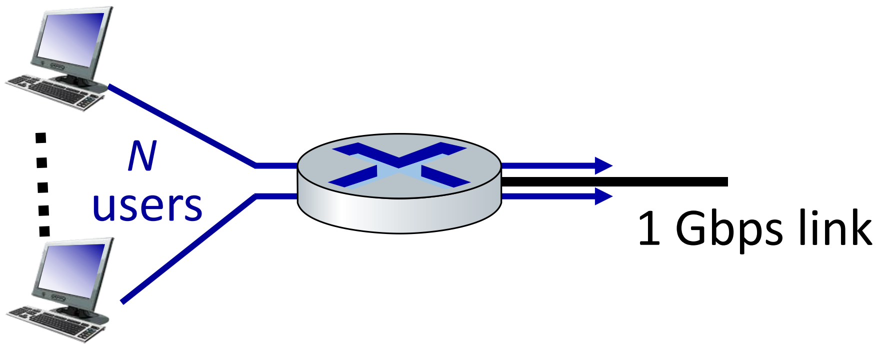


Packet queuing and loss: if arrival rate (in bps) to link exceeds transmission rate (bps) of link for some period of time:

- packets will queue, waiting to be transmitted on output link
- packets can be dropped (lost) if memory (buffer) in router fills up

Packet switching versus circuit switching

example:
- 1 Gb/s link
- each user:
  - 100 Mb/s when “active”
  - active 10% of time



Q: how many users can use this network under circuit-switching and packet switching?

- *circuit-switching:* 10 users
- *packet switching:* with 35 users, probability > 10 active at same time is less than .0004 *

Q: how did we get value 0.0004?

A: HW problem

Packet switching versus circuit switching

Is packet switching a “slam dunk winner”?

- great for “bursty” data – sometimes has data to send, but at other times not
  - resource sharing
  - simpler, no call setup
- excessive congestion possible: packet delay and loss due to buffer overflow
  - protocols needed for reliable data transfer, congestion control
- *Q:* How to provide circuit-like behavior with packet-switching?
  - “It’s complicated.” We’ll study various techniques that try to make packet switching as “circuit-like” as possible.

Q: human analogies of reserved resources (circuit switching) versus on-demand allocation (packet switching)?

Delay in Packet- switched Network

Delay in a packet-switched network refers to the time it takes for a packet of data to travel from the source to the destination. There are several components of delay:

- Transmission Delay
- Propagation Delay
- Processing Delay
- Queueing Delay
- Total Delay

- Transmission Delay: This is the time taken to transmit the entire packet onto the communication medium. It depends on the packet's size and the transmission rate of the link.
    - $Delay_{tr}$ = (Packet length) / (Transmission rate).
- Propagation Delay: This is the time it takes for a signal to travel from the source to the destination. It is determined by the physical distance between the two points and the speed of the signal through the transmission medium.
    - $Delay_{pg}$ = (Distance) / (Propagation speed).
- Processing Delay: When a packet arrives at a router or a network device, there is a small processing delay as the device inspects the packet headers and makes forwarding decisions.
    - $Delay_{pr}$ = Time required to process a packet in a router or a destination host

- Queueing Delay: If there are other packets ahead of the current packet waiting to be transmitted on the same link, the packet will experience a delay in the queue before it gets its turn for transmission.
  - $Delay_{qu}$ = The time a packet waits in input and output queues in a router
- Total Delay: assuming equal delays for the sender, routers, and receiver, the total delay (source-to-destination delay) a packet encounters can be calculated if we know the number of routers, n, in the whole path.

Total delay = (n + 1) ($Delay_{tr}$ + $Delay_{pg}$ + $Delay_{pr}$) + (n) ($Delay_{qu}$)

- Delay can have a significant impact on real-time applications such as voice and video, where low delay is essential to maintain smooth communication.

Throughput in Packet- switched Network

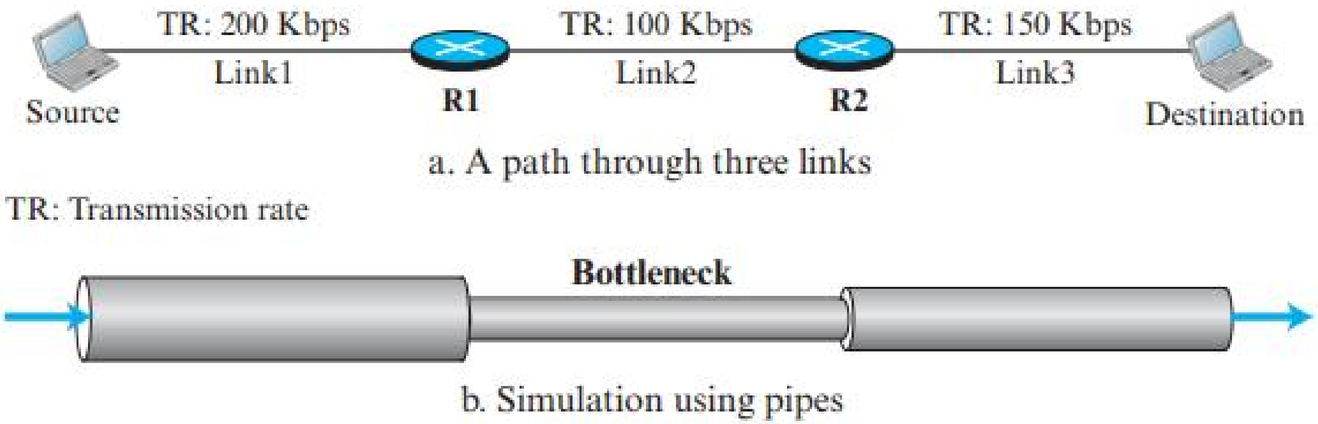
Throughput in a packet-switched network refers to the rate at which data can be transmitted over the network link. It is usually measured in bits per second (bps) or packets per second (pps). Throughput depends on various factors:

- Link Capacity
- Network Congestion
- Protocol Overheads
- Routing Efficiency

- Link Capacity: The maximum data rate that the physical link can support. It is determined by the link's transmission rate and bandwidth.
- Network Congestion: High levels of network congestion can reduce throughput as packets experience more queueing delay and compete for available bandwidth.
- Protocol Overheads: Packet-switched networks use various protocols to manage and control data transmission. These protocols add some overhead, which can slightly reduce the actual throughput.

- Routing Efficiency: The efficiency of the routing algorithms and protocols used in the network can impact how quickly packets reach their destination.
- Throughput is crucial for applications that require high data transfer rates, such as file downloads, streaming, and data-intensive processes.

Figure 4.10 *Throughput in a path with three links in a series*



In this figure, the data can flow at the rate of 200 Kbps in Link1. However, when the data arrives at router R1, it cannot pass at this rate. Data needs to be queued at the router and sent at 100 Kbps. When data arrives at router R2, it could be sent at the rate of 150 Kbps, but there is not enough data to be sent. In other words, the average rate of the data flow in Link3 is also 100 Kbps. We can conclude that the average data rate for this path is 100 Kbps, the minimum of the three different data rates. The figure also shows that we can simulate the behavior of each link with pipes of different sizes; the average throughput is determined by the bottleneck, the pipe with the smallest diameter.

In general, in a path with n links in series, we have

$$\text{Throughput} = \text{minimum } \{TR1, TR2, \ldots TRn\}.$$