CS1102
Lecture 6
Internet and WWW

Semester A, 2020-2021

Department of Computer Science

City University of Hong Kong

Communications

- Computer communications is the process of sharing data, programs, and information between two or more computers. Applications include:
  - E-mail
    - Provides a fast, efficient alternative to traditional mail by sending and receiving electronic documents
  - Texting
    - Provides very efficient direct text communication between individuals using short electronic messages
  - Videoconferencing
    - Provides a very-low-cost alternative to long-distance telephone calls using electronic voice and video delivery
  - Electronic commerce
    - Buying and selling goods electronically

Connectivity

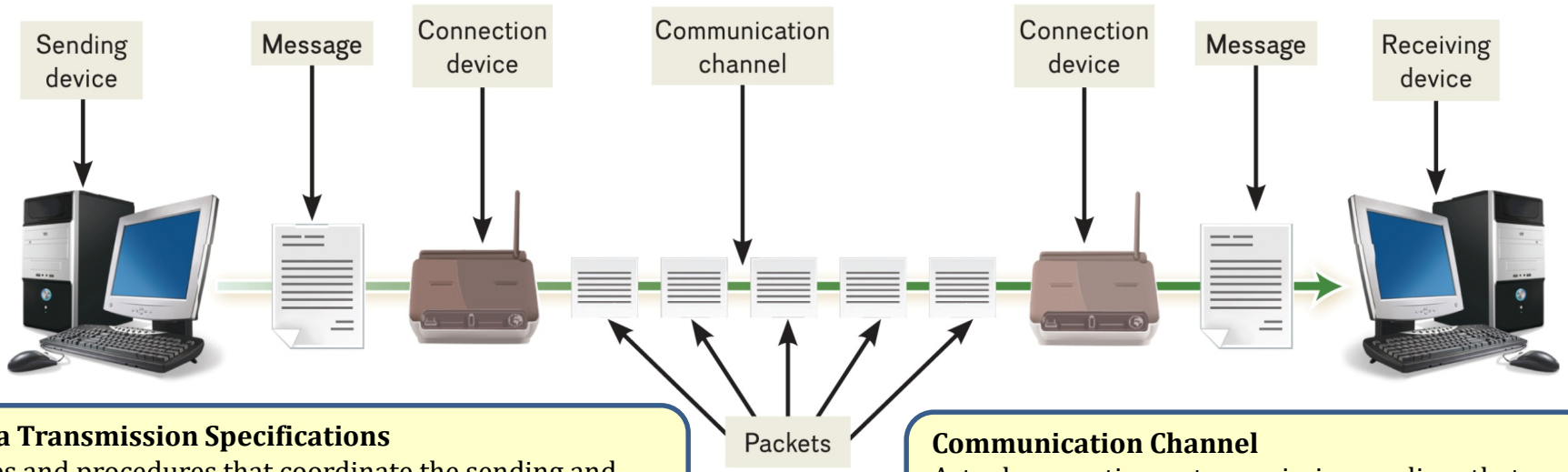
- Connectivity is a concept related to using networks to link people and resources
- You are linked to the world of larger computers and the Internet, including hundreds of thousands of web servers and their extensive information resources
- Being able to efficiently and effectively use computers becomes a matter of knowing not only about connectivity through networks to personal computers, but also about larger computer systems and their information resources

Communication Systems

- Communication systems are electronic systems that transmit data from one location to another

Sending and Receiving Devices
Often a computer or specialized communication device. They originate (send) as well as accept (receive) messages in the form of data, information, and/or instructions

Connection Devices
Act as an interface between the sending and receiving devices and the communication channel to convert outgoing messages into packets that can travel across the communication channel and to reverse the process for incoming messages



Data Transmission Specifications
Rules and procedures that coordinate the sending and receiving devices by precisely defining how the message will be sent across the communication channel

Communication Channel
Actual connecting or transmission medium that carries the message through physical wire or cable or wireless

Communication Channels

- Communication channels actually carry the data from one computer to another

1. Physical Connections

Ethernet Cable

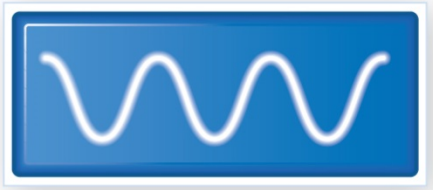
Coaxial Cable

Fiber-optic Cable

2. Wireless Connections
   - Bluetooth
   - Wi-Fi (Wireless Fidelity)
   - Microwave
   - WiMax (Worldwide Interoperability for Microwave Access)
   - Cellular
   - Satellite

Connection Devices

- Computers represent data with digital signals and continuous electronic waves are analog signals
- Modem, short for modulator-demodulator
  - Modulation: the process of converting from digital to analog
  - Demodulation: the process of converting from analog to digital
- Transfer rate is the speed with which modems transmit data, typically measured in Mbps, Megabits (million bits) per second or Gbps, Gigabits (billion bits) per second



Analog

Digital

Digital Subscriber Line

Generations of Cellular Networks

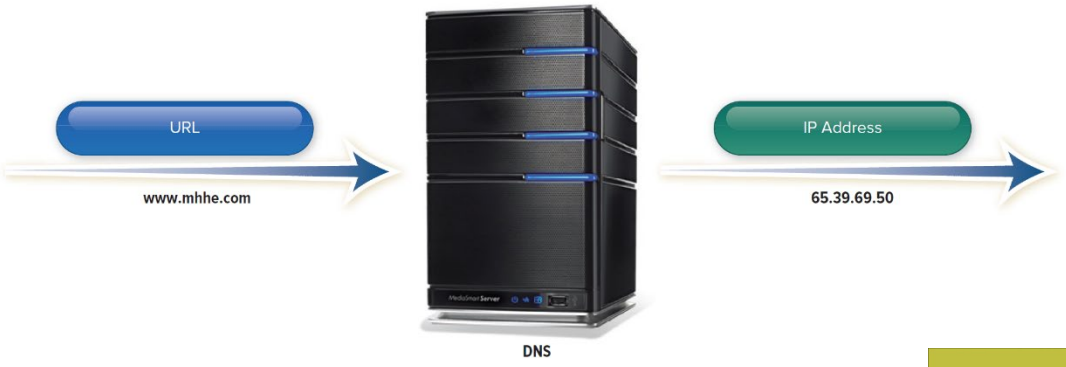
- First-generation mobile telecommunications (1G) started in 1980s using analog radio signals to provide analog voice transmission service
- Second-generation mobile telecommunications (2G) started in 1990s using digital radio signals
- Third-generation mobile telecommunications (3G) started in 2000s and provided services capable of effective connectivity to the Internet, marking the beginning of smartphones
- Fourth-generation mobile telecommunications (4G) use LTE (Long Term Evolution) connections to provide faster Internet access, allowing smartphones to easily stream videos and music
- Fifth-generation mobile telecommunications (5G) is being developed, with speeds that rival home Internet connections and is expected to be widely available by 2020

Data Transmission

- Bandwidth is a measurement of the width or capacity of the communication channel
- Protocols are rules for exchanging data between computers
  - HTTP: HyperText Transfer Protocol
    - Specifies the command and syntax for transmitting web pages and file
    - Has nothing to do with the data content & HTML
  - TCP/IP: Transmission Control Protocol/Internet Protocol
    - Identifies sending and receiving devices
    - Breaks information into parts called packets, for transmission across the Internet

Identification

- Every computer in the Internet has a unique numeric address called an IP address
- Similar to a postal service uses addresses to deliver mail, the Internet uses IP addresses to deliver email and to locate websites
- Because numeric addresses are difficult for people to remember and use, text-based addresses are adopted but they need to be converted to IP addresses by domain name server (DNS)



https://www.youtube.com/watch?v=2ZUxoi7YNgs

Packetization

- Before a message is sent, it is reformatted or broken down into small parts called packets
- Each packet is then sent separately over the Internet, possibly traveling different routes to one common destination
- At the receiving end, the packets are reassembled into the correct order

https://www.youtube.com/watch?v=PpsEaqJV_A0

Computer Network

- Computer network: a communication system that connects to two or more computers so that they can exchange information and share resources

Node
Any device that is connected to the network

Host
Any computer system connected to a network that provides access to its resources

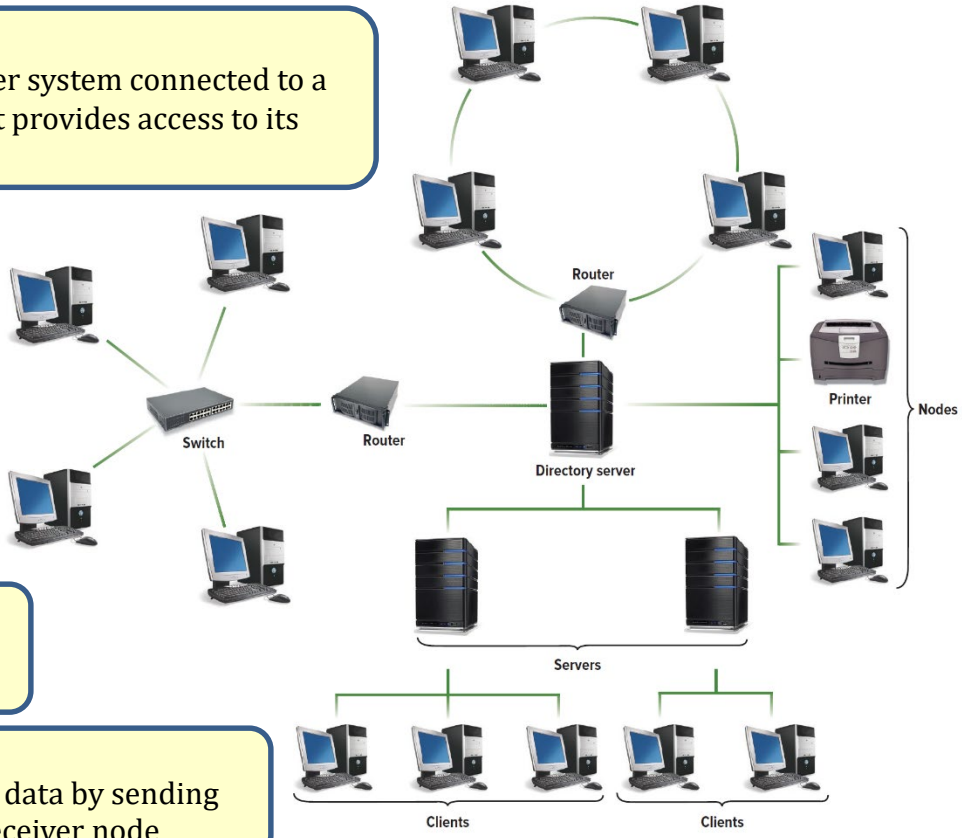
Client
A node that requests and uses resources available from other nodes

Server
A node that shares resources with other nodes

Directory Server
A specialized server that manages resources, such as user account, for an entire network

Router
A node that forwards or routes data packets from one network to their destination in another network

Switch
Central node that coordinates the flow of data by sending messages directly between sender and receiver node



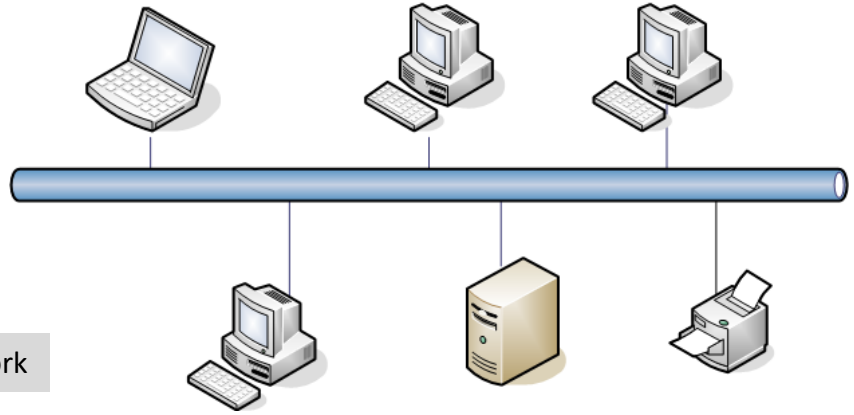
Network Types

Type	Description
LAN	Local area network; located within close proximity
Home	Local area network for home and apartment use; typically wireless
WLAN	Wireless local area network; all communication passes through access point
PAN	Personal area network; connects digital devices, such as PDAs
MAN	Metropolitan area network; typically spans cities with coverage up to 100 miles
WAN	Wide area network for countrywide or worldwide coverage

Network Architecture

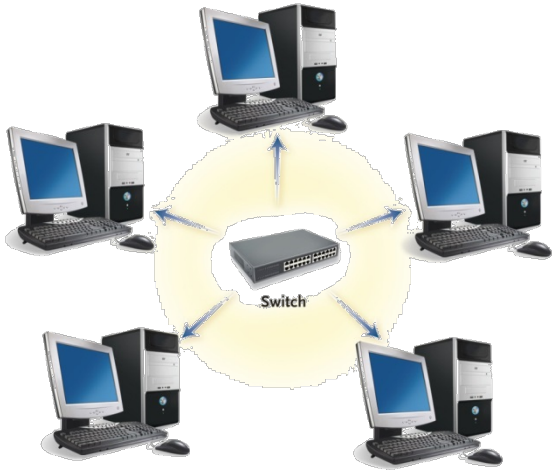
- Network architecture describes how a network is arranged and how resources are coordinated and shared
  - Network topology describes the physical arrangement of the network
  - Network strategies define how information and resources are shared

Network Topologies (1)



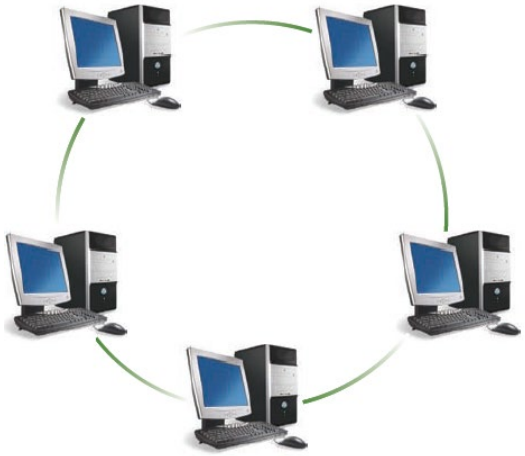
Bus Network

Each device is connected to a common cable called a bus or backbone, and all communications travel along the bus



Star Network

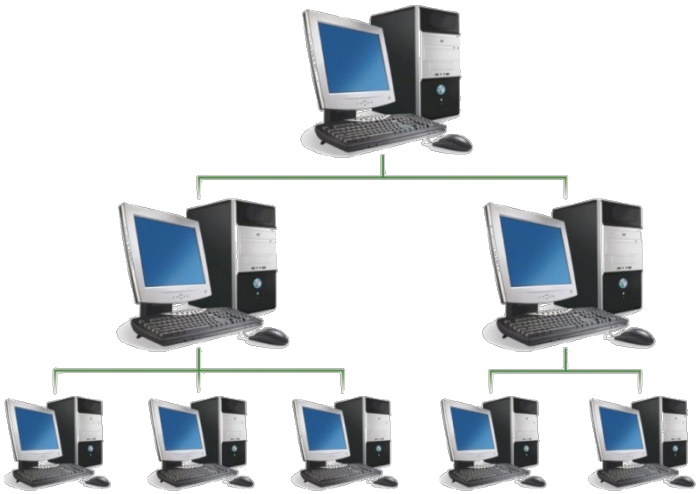
Each device is connected directly to a central network switch
Whenever a node sends a message, it is routed to the switch, which then passes the message along to the intended recipient



Ring Network

Each device is connected to two other devices
When a message is sent, it is passed around the ring until it reaches the intended destination

Network Topologies (2)



Tree Network

Each device is connected to a central node, either directly or through one or more other devices
The central node is connected to two or more subordinate nodes that in turn are connected to other subordinate nodes, and so forth, forming a treelike structure

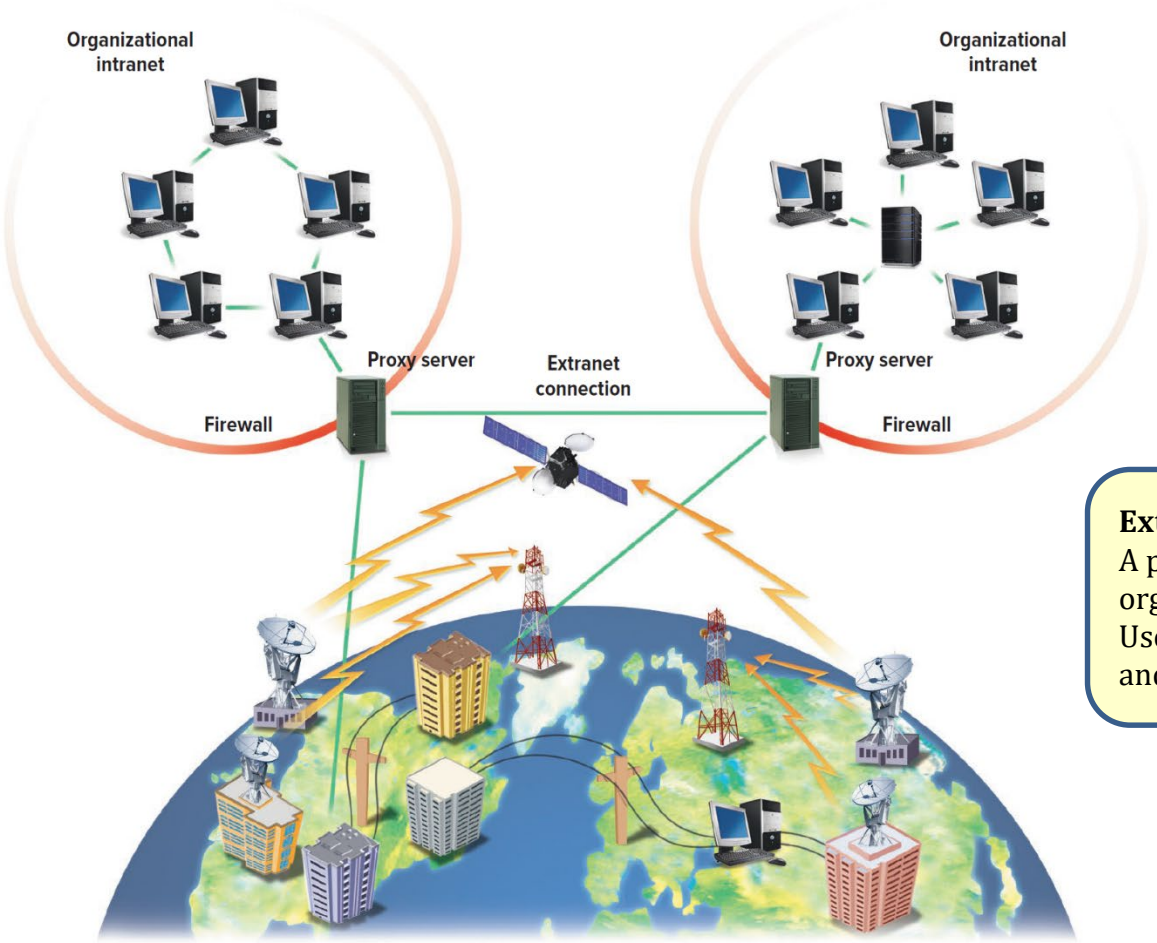
Mesh Network

Each node has more than one connection to other nodes
If a path between two nodes is somehow disrupted, data can be automatically rerouted around the failure using another path

Network Strategies

	Client/Server Network	Peer-to-Peer (P2P) network
Description	Uses central servers to coordinate and supply services to other nodes on the network Server nodes coordinate and supply specialized services, and clients request the services	Nodes have equal authority and can act as both clients and servers
Advantage	Ability to handle very large networks very efficiently Availability of powerful network management software to monitor and control network activities	Easy and inexpensive (often free) to set up and use
Disadvantage	Cost of installation and maintenance	Lack of security controls or other common management functions

Organizational Networks



Intranet
A private network within an organization that resembles the Internet

Extranet
A private network that connects more than one organization
Use Internet technologies to allow suppliers and others limited access to their networks

Network Security

- Firewall
  - Consists of hardware and software that control access to a company's intranet and other internal networks
  - Most use software or a special computer called a proxy server that decides whether it is safe to let a particular message or file pass into or out of the organization's network
- Intrusion Detection System (IDS)
  - Work with firewalls to protect an organization's network
  - Use sophisticated statistical techniques to analyze all incoming and outgoing network traffic
  - Recognize signs of a network attack and disable access before an intruder can do damage
- Virtual Private Network (VPN)
  - Create a secure private connection between a remote user and an organization's internal network
  - Special VPN protocols create the equivalent of a dedicated line between a user's home or laptop computer and a company server
  - The connection is heavily encrypted, and from the perspective of the user, it appears that the workstation is actually located on the corporate network

Internet

- Internet
  - A large network that connects together smaller networks all over the globe
  - Launched in 1969 in US
  - All text without graphics, animations, sound or video prior to the web
- The web, or World Wide Web (WWW)
  - Introduced in 1991
  - Provides a multimedia interface to resources available on the Internet
- Internet and the web are NOT the same thing
  - Internet is the physical network made up of wires, cables, satellites, and rules for exchanging information between computers connected to the network (being online)
  - The web is a multimedia interface to the resources available on the Internet

Web Generations

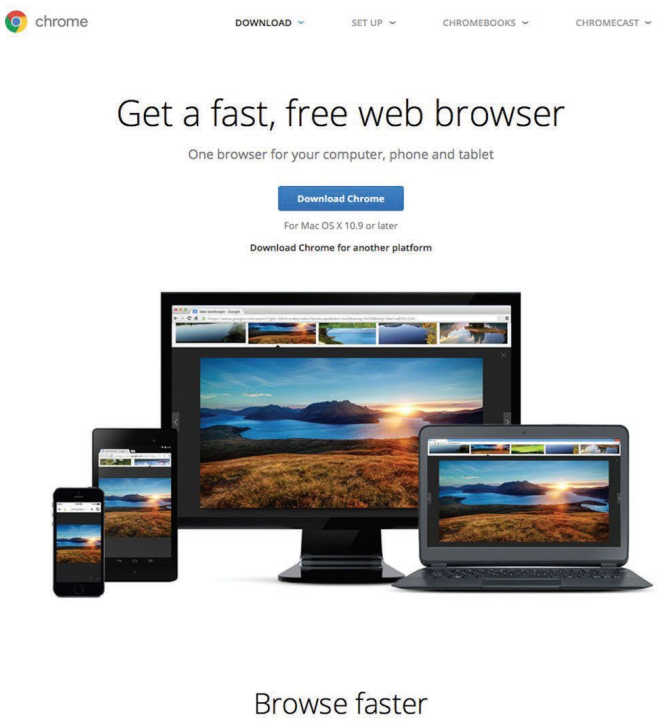
- Web 1.0
  - Focused on linking existing information
  - Search programs were created, e.g., Google
- Web 2.0
  - Supported more dynamic content creation and social interaction, e.g., Facebook
- Web 3.0
  - Identifies relationships between data, e.g., Siri and Google Assistant

Internet Service Provider

- Internet Service Providers (ISPs) are already connected to the Internet and they provide a path or connection for individuals to access the Internet
- The most widely used commercial ISPs use telephone lines, cable, and/or wireless connections

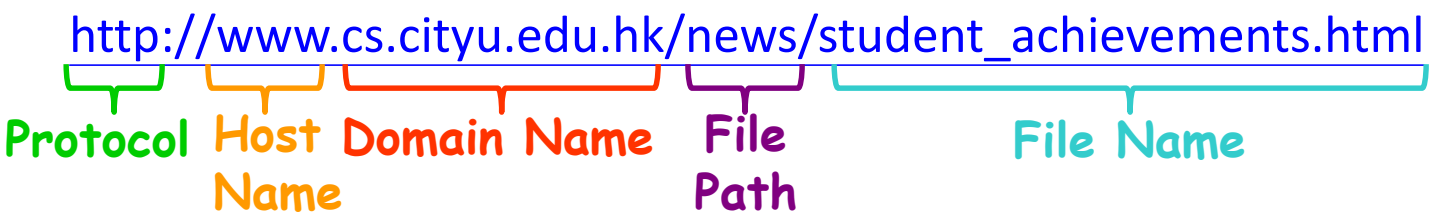
Browser

- Browsers are programs that provide access to web resources with an interface
- Popular web browsers include
  - Mozilla Firefox
  - Apple Safari
  - Microsoft Edge (replacing Internet Explorer)
  - Google Chrome



URL

- Uniform resource locator (URL) is the location or address of the resources
- Suppose you type in a URL on the browser



  - The browser breaks the URL into 4 parts
  - The browser asks a DNS server to help translate the host+domain name to IP address
  - The browser uses the IP address to set up a TCP connection to the destination server
  - Using HTTP protocol, the browser sends a request to the connected server asking for the HTML file
  - The server returns the corresponding HTML file to the browser
  - The browser reads the file, interprets the HTML tags and displays the page

Top-Level Domain

- Top-level domain is the last part of the domain name
- Also known as web suffix
- Typically identifies the type of organization

Top-Level Domain	Type
.com	Commercial
.edu	Educational (Mostly US)
.gov	Government
.mil	US military
.net	Network
.org	Organization
.hk	Hong Kong

HTML

- HyperText Markup Language (HTML) is a markup language for displaying webpages
- Webpages present information about the site along with references and hyperlinks that connect to other documents containing related information
  - text files
  - graphic images
  - audio and video clips

Interactive Website Technologies

- Cascading Style Sheets (CSS)
  - separate files referenced by or lines inserted into an HTML document that control the appearance of a webpage
- Javascript
  - A language within HTML documents or referred by them in separate file to trigger interactive features such as checking information entered in online forms
- AJAX
  - Advanced use of Javascript to create interactive websites that respond quickly
- Applets
  - Programs that can be downloaded quickly and run by most browsers and are used to display graphics, provide interactive games, etc.

HTTP

- Short form for HyperText Transfer Protocol
- Specifies the command and syntax for transmitting web pages and file
- Has nothing to do with the data content & HTML
  - e.g. HTTP can be used to transmit non-HTML data
- Allows browser to fetch web page as well as providing simple feedback information (e.g. form filling)
- Allows server to provide extra information, such as
  - Last updated date of web-page
  - Character set encoding (English, Chinese or Japanese)
  - Authorization
  - Cookies

Search Services

- A number of organizations operate websites that help users locate information they need
- Search services maintain huge databases relating to information (addresses, content descriptions or classifications, keywords, etc.) provided on the web and the Internet
- Special programs called **spiders** continually look for new information and update the search services' databases
- Search services provide special programs called search engines that can be used to locate specific information on the web

Search Engines

- Search engines are specialized programs that assist people in locating information on the web and the Internet

1. You enter a keyword or phrase reflecting the information you want
2. The search engine compares your entry against its database and returns a list of hits, or sites that contain the keywords
3. Search engines order the hits according to those sites that most likely contain the information requested and present the list to you in that order

- Different search engines may return different hits

Content Evaluation

- Anyone can publish content on the web
- Many sites such as Wikipedia allow anyone to post new material, sometimes anonymously and without critical evaluation
- Factors to consider in evaluating the accuracy of information found on the web:
  - Authority
    - Is the author an expert in the subject area?
    - Is the site an official site for the information presented or an individual's personal website?
  - Accuracy
    - Has the information been critically reviewed for correctness prior to posting on the web?
    - Does the website provide a method to report inaccurate information to the author?
  - Objectivity
    - Is the information factually reported, or does the author have a bias?
    - Does the author appear to have a personal agenda aimed at convincing or changing the reader's opinion?
  - Currency
    - Is the information up to date?
    - Does the site specify the date when the site was updated?
    - Are the site's links operational? If not, the site is most likely not being actively maintained

E-mail

- E-mail or electronic mail is the transmission of electronic messages over the Internet

  1. Client-based e-mail systems
     - Require a special program known as e-mail client to be run from your computer, e.g., Microsoft Outlook
  2. Web-based e-mail systems
     - Use browser to connect to e-mail service provider, e.g., Gmail
- Spam refers to unwanted and unsolicited e-mails which may cause distraction and nuisance, and are dangerous because computer viruses or destructive programs may be attached.

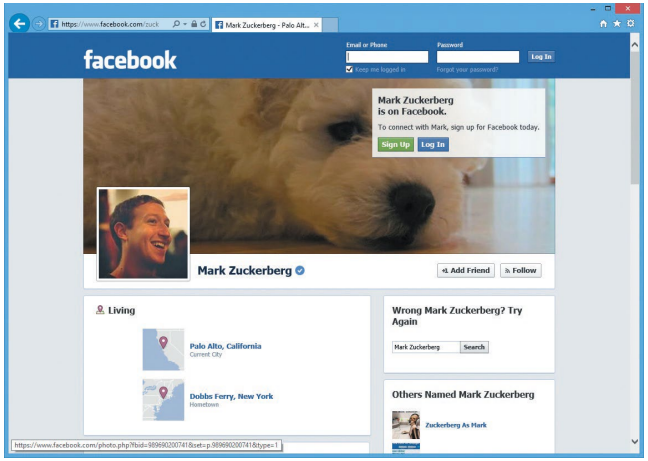
https://www.youtube.com/watch?v=1B0tiUSmlBI

Reducing Spam

- Keep a low profile
  - Many spammers collect e-mail addresses from personal web pages, social networking sites, and message boards. Be cautious when posting your address
- Use caution when giving out your address
  - Many companies collect and sell e-mail addresses to spammers. Be sure to read the privacy policy of a site before providing your address
- Do not ever respond to spam
  - Many spams trick users to validate active e-mail addresses. These addresses are worth more to spammers, who then sell the addresses to other spammers
- Use antispam and filter options
  - Most e-mail programs and web-based e-mail services have antispam and filter options that can be configured

Social Networking

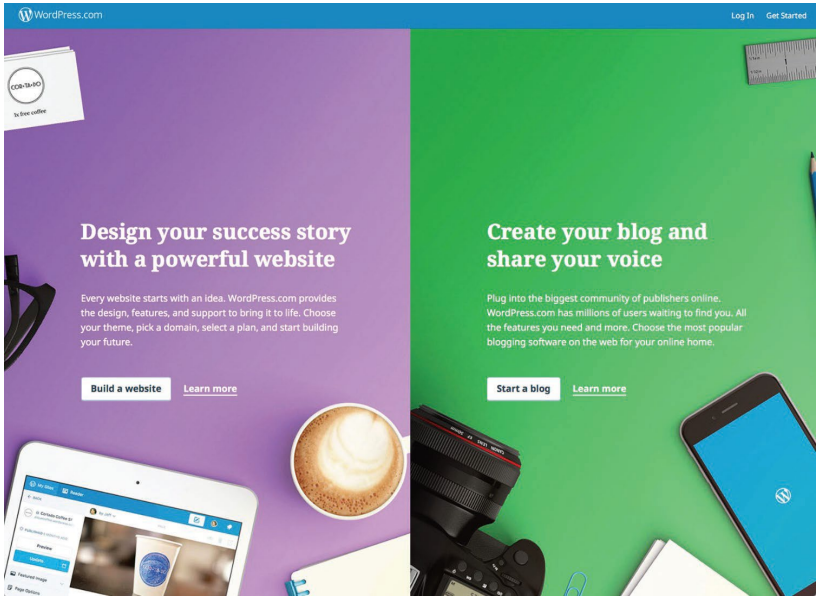
- Fastest-growing and most significant Web 2.0 applications
- Social networking sites
  - focus on connecting people and organizations that share a common interest or activity
  - typically provide a wide array of tools that facilitate meeting, communicating, and sharing



Organization	Site
Facebook	www.facebook.com
Google+	plus.google.com
LinkedIn	www.linkedin.com
Instagram	www.instagram.com
Pinterest	www.pinterest.com
Tumblr	www.tumblr.com
Vine	www.vine.com

Blogs

- Blogs
  - personal websites with postings that are time-stamped and arranged with the newest item first
  - Some blogs are online diaries with personal information, others focus on a hobby or theme
  - Most blogs are written by individual bloggers but there are group blogs with multiple contributors
  - Some businesses and newspapers have used blogging as a quick publishing method
  - Most widely used tool: Blogger and WordPress



Microblogs

- Microblog
  - Publishes short sentences that take only a few seconds to write
  - Designed to keep friends and other contacts up to date on your interests and activities
  - Most popular microblogging site: Twitter
    - Over 300 million active users send over 500 million Twitter messages, known as tweets, each day
    - Twitter is popular among celebrities and politicians to keep fans up to date on their moment-to-moment thoughts

Webcasts and Podcasts

- Webcasts
  - Use streaming technology in which audio and video files are continuously downloaded to your computer while you are listening to and/or viewing the file content
  - After a webcast is completed, there are no files remaining on your computer
  - Webcasts typically broadcast live events, e.g., YouTube Live 
- Podcasts
  - Do not use streaming technology
  - Before a podcast can be run, the media files have to be downloaded and saved to your computer
  - Once downloaded, the files can be run to listen to music or watch a movie as often as you would like
  - Podcasts are widely used to download music, tutorials, and educational training

Wiki

- A wiki is a website specially designed to allow visitors to use their browser to add, edit, or delete the site's content
- "Wiki" comes from the Hawaiian word for *fast*, which describes the simplicity of editing and publishing through wiki software
- Wikis support collaborative writing in which there is not a single expert author, but rather a community of interested people that builds knowledge over time
- The most famous example is Wikipedia, an online encyclopedia, written and edited by anyone who wants to contribute

Electronic Commerce (1)

- Electronic commerce, also known as e-commerce, is the buying and selling of goods over the Internet
- Advantages
  - Buyer
    - goods and services can be purchased at any time of day or night from any location that has an Internet connection
  - Seller
    - the costs associated with owning and operating a retail outlet can be eliminated
    - no in-store inventory and products are shipped directly from warehouses
- Disadvantages
  - Inability to provide immediate delivery
  - Inability to "try on" prospective purchases
  - Questions relating to the security of online payments

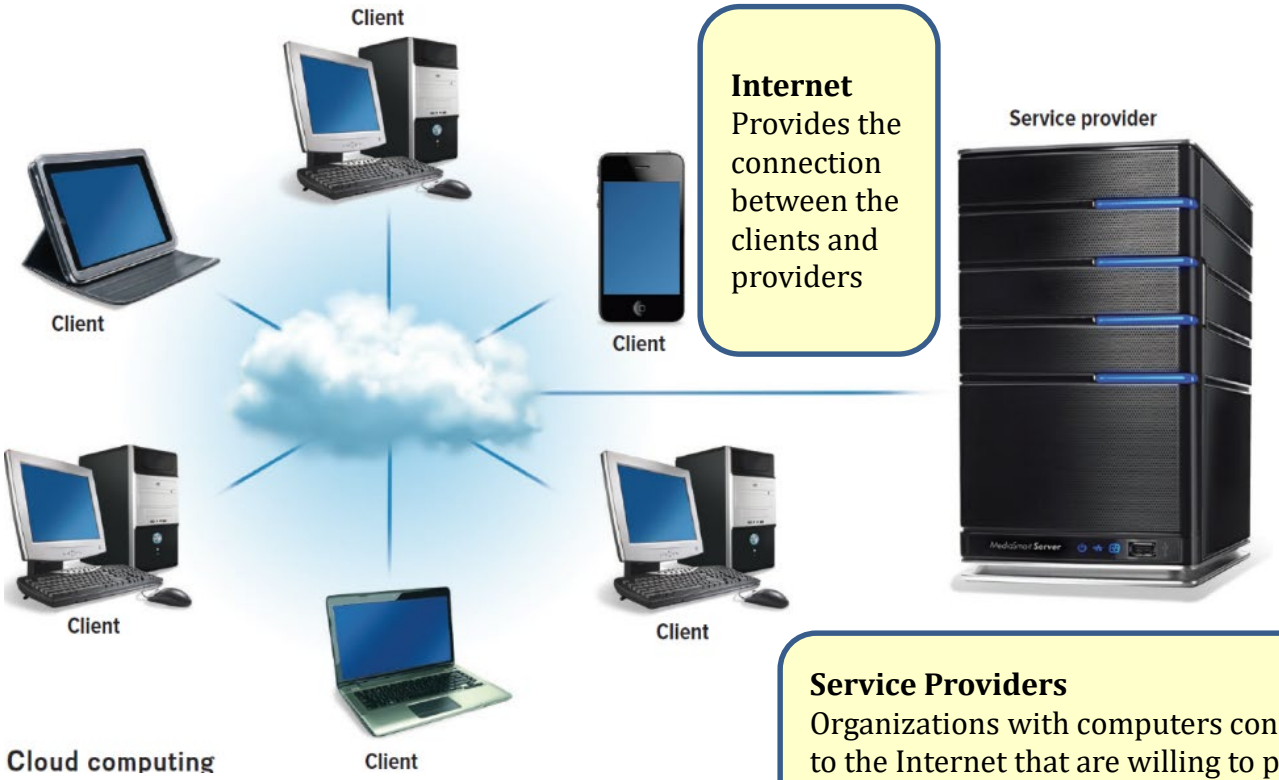
Electronic Commerce (2)

- Basic types of electronic commerce
  - Business-to-consumer (B2C)
    - Involves the sale of a product or service to the general public or end users
    - Applications: online banking, financial trading, shopping (e.g., Amazon)
  - Consumer-to-consumer (C2C)
    - Involves individuals selling to individuals
    - Application: web auctions
  - Business-to-business (B2B)
    - Involves the sale of a product or service from one business to another
    - Typically manufacturer-supplier relationship

Cloud Computing

- Cloud computing uses the Internet and the web to shift many of these computer activities from the user's computer to other computers on the Internet

Clients
Corporations and end users who want access to data, programs, and storage



Internet
Provides the connection between the clients and providers

Service Providers
Organizations with computers connected to the Internet that are willing to provide access to software, data, and storage

Internet of Things

- Internet of Things (IoT) allows everyday objects (smartphones, wearable devices etc.) embedded with electronic devices to send and receive data over the Internet
- For example: Apple's Health App (a Web 3.0 Application)

1. Access your Fitbit data
2. Combine it with other related health data
3. Analyze the data
4. Report back to you through your smartphone (e.g., heart rate, steps taken each day, estimate of calories burnt, etc.)

Lesson Summary

- A computer network is a communication system that connects to two or more computers so that they can exchange information and share resources
- Every computer in the Internet has a unique numeric address called an IP address
- Because numeric addresses are difficult for people to remember and use, text-based addresses are adopted but they need to be converted to IP addresses by domain name server (DNS)
- Uniform resource locator (URL) is the location or address of the resources
- HTML is a markup language for displaying webpages while HTTP specifies the command and syntax for transmitting web pages and file
- Network architecture describes how a network is arranged (network topologies) and how resources are coordinated and shared (network strategies)
- Internet and WWW are not the same thing
- Search services provide special programs called search engines that can be used to locate specific information on the web
- E-mail or electronic mail is the transmission of electronic messages over the Internet and it is important to reduce spams
- In addition to email, social networking, blogs, microblogs, webcasts, podcasts, Wiki are other types of Internet communication

Reading

- Computing Essentials 2019
  - Chapter 8, 2

Reference

[1] HTTP
- http://searchwindevelopment.techtarget.com/definition/HTTP

[2] What is TCP/IP and how does it make the Internet work?
- https://www.hostingadvice.com/blog/tcpip-make-internet-work/

[3] Introduction to Computer Network Topology
- https://camrojud.com/introduction-to-computer-network-topology/

[4] Client-server Model
- https://en.wikipedia.org/wiki/Client-server_model

[5] The World of Peer-to-Peer (P2P)
- https://en.wikibooks.org/wiki/The_World_of_Peer-to-Peer_(P2P)

[6] What is a Web Proxy Server?
- https://www.forcepoint.com/cyber-edu/web-proxy-server

[7] What is a Firewall?
- https://www.forcepoint.com/cyber-edu/firewall

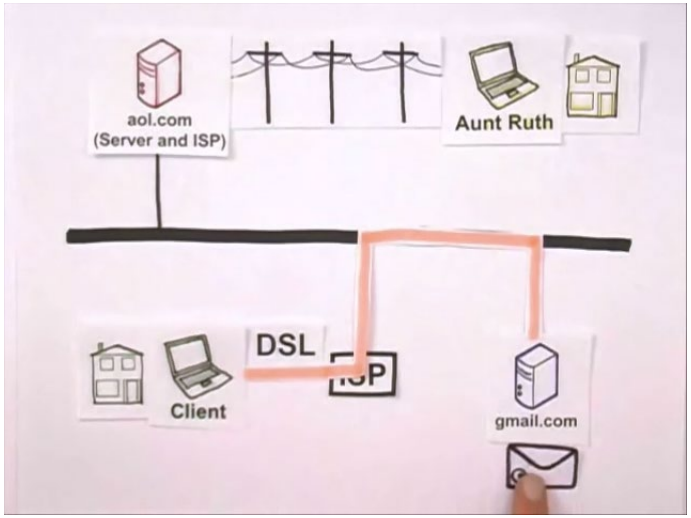
[8] W3C - A little history of WWW
- http://www.w3.org/History.html

[9] Web 2.0
- https://computer.howstuffworks.com/web-20.htm

[10] Web 3.0
- https://computer.howstuffworks.com/web-30.htm

[11] Internet of Things (IoT)
- https://computer.howstuffworks.com/internet-of-things.htm

Video



How the Internet Works in 5 Minutes
https://www.youtube.com/watch?v=7_LPdttKXPc

The client server model
https://www.youtube.com/watch?v=L5BlpPU_muY