The Internet

What is the internet, and how does it work?

The Internet

- What is the internet?
- Internet hardware and connection
- Internet infrastructure
- A brief history of the Internet

The Internet

- What is the internet?
- Internet hardware and connection
- Internet infrastructure
- A brief history of the Internet

What is the internet?

From an application perspective, the internet provides:

-Communication
  -School, work, family, friends, spam…
-Commerce
  -Amazon, ebay, travel, shopping, banking,…
-Entertainment
  -Music, Movies, games, news, blog, …
-Adventures
  -New applications, new techniques, millionaires!

The Internet

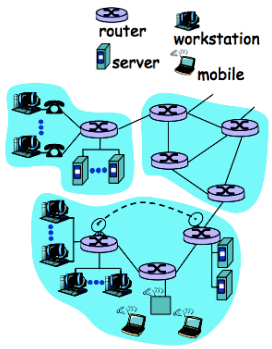
-A network of networks

-Millions of connected devices, or hosts

-Infrastructure: wire, fiber, radio, satellite,…

-Traffic controllers: routers,…

-Protocols: languages for sending and receiving messages: **TCP/IP, ftp, http,…**



The Internet is not…

The internet is not:

- **The World Wide Web (WWW).** The WWW is a way of accessing information over the internet
- **The physical media**: this is just one component of the internet
- **The applications**: again, this is just one component.

The Internet

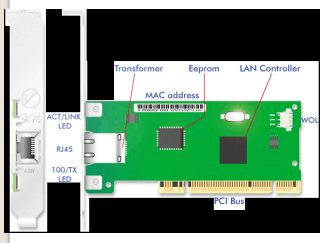
- What is the internet?
- Internet hardware and connection
- Internet infrastructure
- A brief history of the Internet

Hardware on the host: Network Interface Card: NIC

- Physically connects a computer to the network
- What is device driver?
  - NIC comes with different drivers for different types of operating systems.
  - A driver is the software that allows the operating system to communicate with the network interface card.

The MAC address: a unique signature

(Media Access Control address)



- When a NIC is manufactured, the card is given a unique hardware address.
- It never changes.

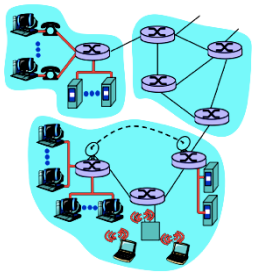
Internet Infrastructure

How to connect a host to the outside network (the "last mile" problem):

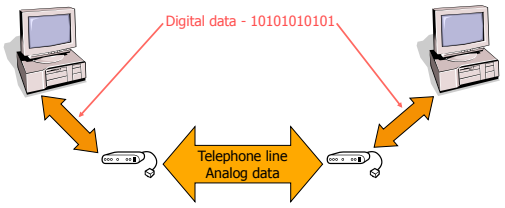
-home network
  -Phone line (modem, ADSL)
  -Cable

-Institution network
  -Ethernet,…

-Mobile access network
  -Wireless router,…



Home connection: The Modem



Dialup via modem:
-up to 56 Kbps
-Cannot use internet connection and phone at the same time

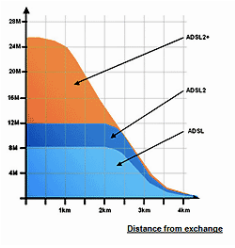
Home connection: ADSL

ADSL: Asymmetric Digital Subscriber Line

-up to 1 Mbs upstream (today typically < 256 Kbs)

-Up to 8 Mbs downstream (today typically < 1 Mbs)

-Distance is a factor (the further you are from the active router, the worst your connection is; ADSL is good in cities)

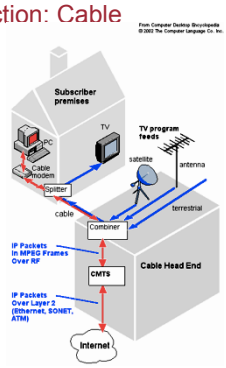


Home connection: Cable

- Computer attached to cable network through a cable modem.

-Up to 1Mbs upstream and 10 Mbs downstream

-Cable is shared by nearby houses: risk of congestion



Home connection: Powerline

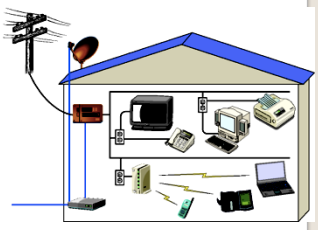
Powerline Communication (PLC):

-use the existing electricity network for voice and data transmission

-Superpose an analog signal to the AC current

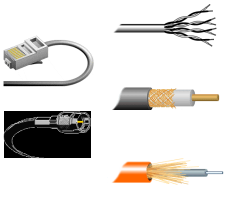
-Ubiquitous, cheap to setup

-High risk of interference



Home connection: the different cables

- Twisted pair
- RJ45 connector
- Coaxial cable
- BNC connector
- Fiber optic cable



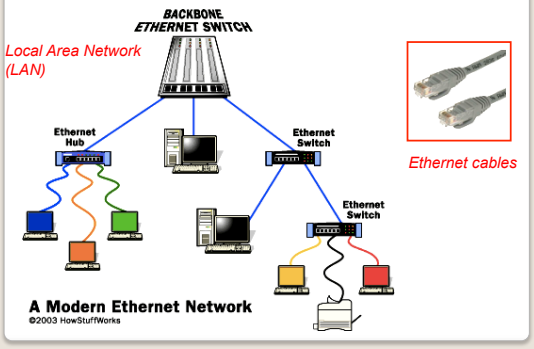
A cable for each length:

Optic > coaxial > Twisted pair

Institution Network

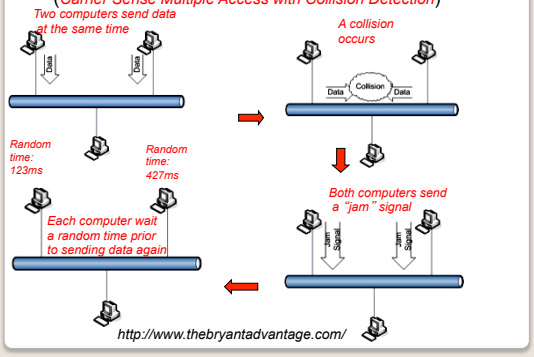
- Most institutions have their own network: Local Area Network (LAN). They are based on either:
  - Ethernet
    - It is the most popular architecture used to build networks.
    - It is the least expensive and easiest to setup
    - Popular in universities and research institutes
  - Token-ring architecture
    - It was developed by IBM in 1984.
    - They are popular in large organizations, such as banks and insurance companies.
  - Others: AppleTalk, …

Institution Network: the Ethernet



Ethernet: CSMA/CD

(*Carrier Sense Multiple Access with Collision Detection*)



Ethernet Hubs and Switches

-Both devices allow multiple computers to be connected to the internet through a single entry point

-Hubs are passive: they do not monitor traffic nor do they check for collision

-Switches are active: they are small computers with a CPU. They act as "cops" and control traffic.

Wireless Networks

Host is connected to the router by shared wireless access through a base station or "access point"

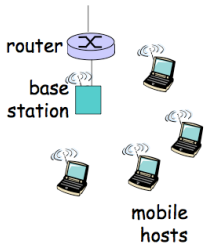
Wireless LANS:

-802.11b (WiFi): 11 Mbps, 2.4 GHz ISM band (ISM: Industrial, Scientific and Medical)

-802.11g: 54 Mbps, 2.4 GHz ISM band

-802.11a: 54 Mbps, 5 GHz ISM band

-802.11n: 100 Mbps



Wireless Transmission

Using radio wave:

-no physical wires
-Use electromagnetic wave as a carrier
-Suffer from environmental effects (obstructions)
-However…they are great to go around obstacles for physical cables (mountains, …)

Examples:

-terrestrial microwaves
-Wide area (cellular networks)
-Satellites

The Internet

- What is the internet?
- Internet hardware and connection
- Internet infrastructure
- A brief history of the Internet

Internet Infrastructure

To communicate over the Internet, the computers must:

- have a way to address one another.
- use a common language or a protocol to organize the exchange of messages.

Addressing:

- defines where to deliver the messages

Protocol:

- specifies exact format, order of messages sent and received among network entities, and actions taken on message transmission and receipt.

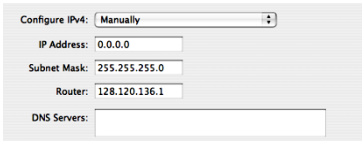
Internet Infrastructure
The IP address

- Each host in the Internet is assigned a specific and unique number for identification: it serves as a "postal address" on the network
- This number is called the IP address of the specific host.
- This number is divided into 4 parts to improving the readability.
- The range of each number is between 0 and 255.
  - E.g. 0.0.0.0
  - 255.255.255.255

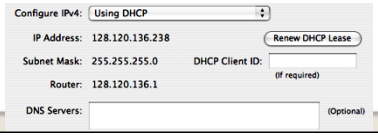
Internet Infrastructure
The IP address

Two ways to get an IP address:

- Manually



-Dynamically, through DHCP (Dynamic Host Configuration Protocol)



Internet Infrastructure
The IP address

Important elements about IP addresses:

-Prefer DHCP over manual assignment! If you choose an IP address that has already been assigned, this will create conflicts on the network!

-The subnet mask tells which computers are on the same subnet as your computer. For example, if your address is 128.120.136.238 and the network mask is 255.255.255.0, this means that any computer whose IP address starts with 128.120.136 is on the same network: information between your computer and this computer will travel directly. For any computer whose IP address does not start with 128.120.136, the information will first go through the router

-The DNS, or Domain Name Service, is a service that provides translation between a "human-readable" name for a computer and the IP address. For example, the IP address 128.120.33.39 corresponds to a computer named www.ucdavis.edu. The translation table is stored on a computer running DNS.

Internet Infrastructure
The IP address

Common top domain names in the US: .com, .mil, .edu, .org

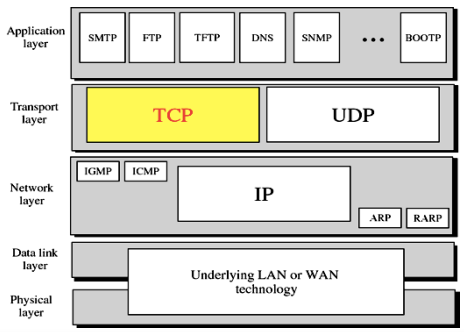
Outside of the US, the top-level domain identifies the country:
uk (England), fr (France), cn (China), …

Two computers can have the same high level name if they are not on the same domain

Internet Infrastructure

- Internet has a large collections of protocols organized in a layer model.
  - Application: enables the user, whether human or software, to access the network.
  - Transport: responsible for source-to-destination (end-to-end) data transfer.
  - Network: responsible for routing packets from source-to-destination across multiple networks.
  - Data link: responsible for data transfer between neighboring network elements.
  - Physical: coordinates the functions required to transmit a bit stream over a physical medium.

Internet Infrastructure



Internet Infrastructure

TCP/IP:

- A protocol is a collection of rules for formatting, ordering, and error-checking data sent across a network.
- In 1974, Vincent Cerf and Robert Kahn developed the **Transmission Control Protocol (TCP)** which was further split into the **Internet Protocol (IP)** and **TCP** in 1978.
- In 1982, DoD adopted TCP/IP as the standard protocol in the Internet.
- Because the significance of TCP/IP in the history of the Internet, Cerf and Kahn are considered to be the Fathers of the Internet.

Internet Infrastructure

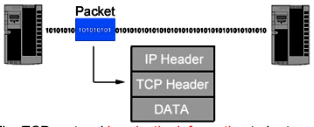
The main characteristics of TCP/IP:

- IP - is responsible for moving packet of data from node to node. IP forwards each packet based on a four byte destination address (the IP number).

- TCP - is responsible for verifying the correct delivery of data from client to server. Data can be lost in the intermediate network. TCP adds support to detect errors or lost data and to trigger retransmission until the data is correctly and completely received.

- Sockets - is a name given to the package of programs that provide access to TCP/IP on most systems.

Internet Infrastructure
The Transmission Control Protocol



- The TCP protocol breaks the information to be transmitted into multiple packets.
- Each packet is associated with a sequence number for identification.

137.189.90.184 Number 1 **Data here**
137.189.90.184 Number 2 **Data here**
137.189.90.184 Number 3 **Data here**

- Each packet is individually routed through the Internet, and arrive in random order.
- The data is reassembled on the destination computer in the correct order according to the sequence numbers.

Internet Infrastructure
The Transmission Control Protocol

The TCP/IP package

Bit offset	Bits 0 - 7	8–15	16–23	24–31
0	Source address			
32				
64				
96				
128	Destination address			
160				
192				
224				
256	TCP length			
288	Zeros		Next header	
320	Source port		Destination port	
352	Sequence number			
384	Acknowledgement number			
416	Data offset	Reserved	Flags	Window
448	Checksum		Urgent pointer	
480	Options (optional)			
480/512+	Data			

Computer sending the packet

Destination computer

Length of the packet

Checksum for integrity

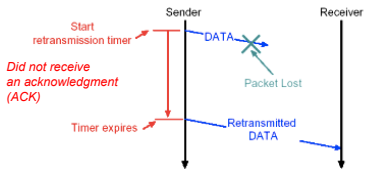
Internet Infrastructure
The Transmission Control Protocol

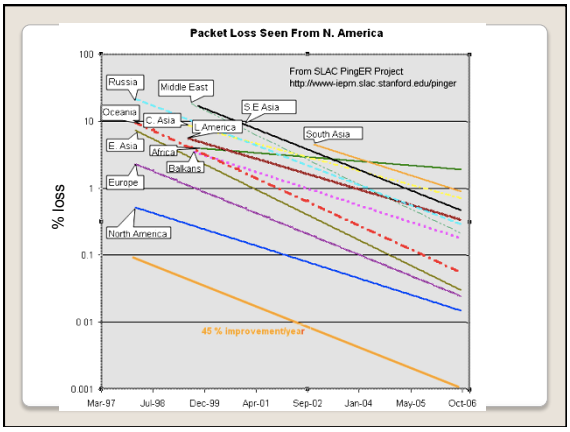
What happens when a packet is lost?

- A packet may be lost during the transmission across the Internet (host down, router down, link failure, …)
- When the destination host has been waiting for a particular packet for too long (timeout), it will request the source host to retransmit the packet.
- There is no need to retransmit all data packets. Instead, only the missing packet, which is identified by its sequence number, needs to be retransmitted.

Internet Infrastructure
The Transmission Control Protocol

How does the sender know it needs to retransmit:





Internet Infrastructure
The Transmission Control Protocol

The Transmission Control Protocol is designed to provide reliable data transfer. To ensure this is true, it needs to deal with two types of problems:
-Packet loss
-Packet corruption: the packet arrives, but it is different from what was sent by the sender.

The checksum procedure:

-On the sender host, TCP runs an algorithm on all bytes of the data and generates a single number, the "checksum". This number is then transmitted with the packet.

-On the destination host, the same algorithm is run on the data that arrive: if the corresponding checksum is equal to the checksum transmitted, the packet is considered safe; otherwise it is discarded.

Internet: Some Applications

The World Wide Web:
- HTML: Hypertext Markup Language
- HTTP: Hypertext Transfer Protocol

Mail
- SMTP: Simple Mail Transfer Protocol
- POP : Post Office Protocol

File transfer
- FTP: File Transfer Protocol

The Internet

- What is the internet?
- Internet hardware and connection
- Internet infrastructure
- A brief history of the Internet

Internet History

1961-1972: Early packet-switching principles

- 1961: Kleinrock - queueing theory shows effectiveness of packet-switching
  - PhD thesis
- 1964: Baran - packet-switching in military nets
- 1967: ARPAnet conceived by Advanced Research Projects Agency
- 1969: first ARPAnet node operational
- 1972:
  - ARPAnet demonstrated publicly
  - NCP (Network Control Protocol) first host-host protocol
  - first e-mail program
  - ARPAnet has 15 nodes

Internet History

1972-1980: Internetworking, new and proprietary nets

- 1970: ALOHAnet satellite network in Hawaii
- 1973: Metcalfe's PhD thesis proposes Ethernet
- 1974: Cerf and Kahn - architecture for interconnecting networks
- late70's: proprietary architectures: DECnet, SNA, XNA
- late 70's: switching fixed length packets (ATM precursor)
- 1979: ARPAnet has 200 nodes

Cerf and Kahn's internetworking principles:
- minimalism, autonomy - no internal changes required to interconnect networks
- best effort service model
- stateless routers
- decentralized control

define today's Internet architecture

Turing Award Winner 2005

Internet History

1980-1990: new protocols, a proliferation of networks

- 1983: deployment of TCP/IP
- 1982: SMTP e-mail protocol defined
- 1983: DNS defined for name-to-IP-address translation
- 1985: FTP protocol defined
- 1988: TCP congestion control
- new national networks: Csnet, BITnet, NSFnet, Minitel
- 100,000 hosts connected to confederation of networks

Internet History

1990, 2000's: commercialization, the Web, new apps

- Early 1990's: ARPAnet decommissioned
- 1991: NSF lifts restrictions on commercial use of NSFnet (decommissioned, 1995)
- early 1990s: Web
  - hypertext [Bush 1945, Nelson 1960's]
  - HTML, HTTP: Berners-Lee
  - 1994: Mosaic, later Netscape
  - late 1990's: commercialization of the Web

Late 1990's – 2000's:

- more killer apps: instant messaging, peer2peer file sharing (e.g., Naptser)
- network security to forefront
- est. 50 million host,
- backbone links running at Gbps

Internet Usage (2008)

INTERNET USERS AND POPULATION STATS FOR THE AMERICAS

THE AMERICAS	Population (2008 Est.)	% Pop. of World	Internet Users, Latest Data	% Population (Penetration)	User Growth (2000-2008)
All the Americas	918,822,841	13.7 %	413,183,671	45.0 %	227.5 %
Rest of the World	5,791,206,229	86.3 %	1,161,129,513	20.0 %	394.5 %
WORLD TOTAL	6,710,029,070	100.0 %	1,574,313,184	23.5 %	336.1 %

INTERNET USERS AND POPULATION STATS FOR THE AMERICAS

REGION	Population (2008 Est.)	% Pop. America	Internet Users, Latest Data	% Population (Penetration)	User Growth (2000-2008)
North America	337,572,949	36.7 %	246,822,936	73.1 %	128.3 %
South America	389,621,930	42.4 %	128,652,435	33.0 %	800.2 %
Central America	151,256,045	16.5 %	28,814,000	19.0 %	795.6 %
The Caribbean	40,371,917	4.4 %	8,894,300	22.0 %	1,489.9 %
TOTAL AMERICAS	918,822,841	100.0 %	413,183,671	45.0 %	227.5 %

http://www.internetworldstats.com/stats2.htm