CS120 The Information Era

LECTURE 3

TOPICS: Webservers, Routers, Packets, Protocols, and Web Infrastructure

How does the Internet Work?

- When you type a URL (Uniform Resource Locator) into a Web Browser and press Return, what do you think happens? That is, what steps do you think are required to obtain a web page on your computer?

The Basics: Getting a Web Page

- At a basic level, the procedure to obtain a webpage on your computer machine (the client) from a URL (http://www.pacificu.edu/finals.html)
  - 1) The web browser on the client breaks the URL into three parts:
    - protocol (http)
    - server name (pacificu.edu)
    - filename (finals.html)
  - 2) The client communicates with a domain name service (DNS) server to translate the web server URL into an IP address that it then uses to connect to this server using the http protocol.
  - cont->

The Basics: Getting a Web Page

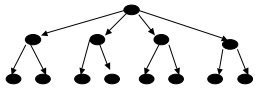
  - 3) The web browser on the client forms a connection to the web server at that IP address on port 80.
  - 4) Following the http protocol, the browser on the client sends a GET request to the server, asking for the file.
  - 5) The server sends the HTML text for the Web Page in packets to your client machine using the TCP/IP protocol and routers.
  - 6) The browser reads the HTML tags and formats the page on your screen.
- DON"T WORRY! We'll elaborate on and explain all of this terminology in today's lecture!

Client-Server model

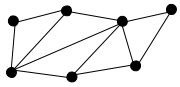
- All machines on the Internet are either Servers or Clients
- Servers: Machines that provide services to other machines
  - Web servers, e-mail servers, FTP servers, Application servers etc
- Clients: Machines used to connect to servers
  - All the machines in this room, the computer in your dorm room

The Internet: Basics

- What is the Internet?
- Do you think the internet is Hierarchical:



- or Heterarchical?



- What is an advantage of a heterarchical structure?

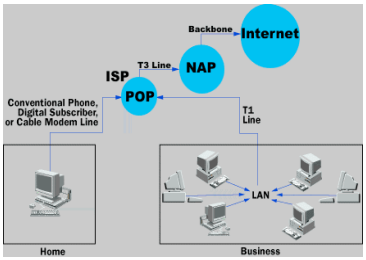
The Internet: Getting Connected

- Not one big network
  - more than 70,000 smaller networks
- Several large backbones, typically fiber optic and maintained by long distance companies
  - You try: http://global.mci.com/about/network/interactive/ then click on "**Interactive Internet Network Map**" link

The Internet: Getting Connected

- In Business or a University: Computer connected to a local area network (LAN) which then connects to internet using a high-speed line (T1 or faster) to a local ISP
- At Home: Computer connects to local ISP using a phone-line modem, DSL, or cable modem
- Both of these merge at a POP (or Point of Presence) for the ISP
- Many different ISP's then connect (typically via a T3 line) to a single NAP (or Network Access Point) on the backbone of the internet
- See picture next page->

Internet: Infrastructure



Internet: IP Addresses

- Each computer on the internet has a unique IP address of the form 123.456.78.910
- It can be fixed (i.e. static) or dynamic (i.e. different each time you connect to the internet)
- The client computers in LL21 and the dorms all have dynamic IP addresses obtained from the DHCP server on campus; servers have static IP addresses. Why do you think this is the case?
- What is the IP address of your computer in LL21? At home?
- Disadvantage is that it is hard to remember! Is there a better way?

Internet: Domain Name Service

- Prior to 1983, had to communicate with other computers on the internet using ONLY the IP address
- University of Wisconsin in 1983 developed DNS (or Domain Name Service) as a means of translating symbolic host domain names (i.e. www.pacificu.edu) into IP addresses automatically
- At Pacific, we have a several dedicated DNS Servers
- Verisign (www.verisign.com) is the primary company responsible for maintaining catalogue of Domain Names and have an IP address whois lookup service
  - Some smaller businesses or schools just use this service
- Try www.math.pacificu.edu and 64.59.233.252

Internet: Domain Names

- Top level domain names are the suffix.
- Examples:
  - com A commercial organization
  - edu Educational site in US
  - gov Government agency in US
  - mil Military site in US
  - net A network site
  - org A non profit organization
- Also country suffixes:
  - Au Australia de Germany
  - Ca Canada ie Ireland
  - etc

Transmitting Data: Routers

- Routers are the workhorse of the internet
  - Ensures that information doesn't go where it is not needed
  - Ensures that information makes it to the correct destination
  - Cisco Switch Router 12000: Moves 60 million packets a second
- Responsible for communications between networks
- Uses a configuration table to decide how to route information using
  - Which connections on the backbone lead to groups of IP addresses
  - Priorities for connections to be used
  - Rules for handling both routine and special cases of traffic
- Pacific has several routers (see pictures later)

Transmitting Data: TCP/IP

- Routers are capable of using the TCP/IP protocol (Transmission Control Protocol/Internet Protocol) among other protocols for transferring data
  - TCP divides a file into packets (typically about 1Kb or 1024 bytes) to be transmitted at the sending end
  - IP stamps messages with IP address and sends them
  - At receiving end, IP collects all packets
  - TCP reassembles the packets
  - If packet has an error, a message is sent back to the sender to resend the packet
- What information would a packet need to contain?

Transmitting Data: Packets

- A packet typically consists of the following information:
  - Header (96 bits): Sender's IP address, receiver's IP address, protocol, and packet number
  - Payload (or Body or Data) (896 bits): 896 bits
  - Trailer (32 bits): Data to show end of packet, checksum value
    - in CRC, or Cyclic Redundancy Check, the checksum value is the sum of all 1's in the payload expressed in hexadecimal
- Advanced Note: Packet also contains the port number and TTL (Time to Live): Starts at 255, decreased by one each router hop. At zero, packet expires

Transmitting Data: Dynamic Routing

- Routers will ping other computers to determine the condition of the network (including the load and any problem with the equipment)
  - www.internettrafficreport.com
- A sample ping to euler.math.pacificu.edu from 64.59.233.71
  - 64 bytes from 64.59.233.252: icmp_seq=0 ttl=64 time=0.611 ms
- A sample ping to www.pacificu.edu from 64.59.233.78
  - 64 bytes from 64.59.226.245: icmp_seq=0 ttl=63 time=0.753 ms
- Based on this information, it will send each packet off to its destination via the best available route using algorithms (one of the most common involves the use of Euler circuits in MATH 165)
- This means that the packets that form the original web page (or email, etc.) to be transmitted may take completely different routes to their destination!
- YOU TRY! From the command line, type ping followed by a website URL

Transmitting Data: Subnet Masks

- Routers also know when to keep information on the local network
- Subnet Masks: Look like IP address, and usually is 255.255.255.0
- A "255" in a particular spot tells the router that all packets with the sender and receiver having an IP address sharing this part of the address are on the same network and thus shouldn't be sent to another network
- Check System Preferences under Network
- Questions: Is your computer on the same network as the server www.pacificu.edu? Is your computer on the same network as the server euler.math.pacificu.edu?

Transmitting Data: Ports

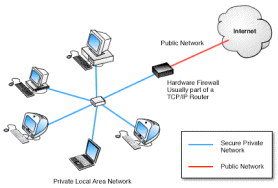
- Depending on the protocol (i.e. http, ftp, etc.), a webserver communicates via various ports
  - ftp: Port 21
  - telnet: Port 23
  - email: Port 25
  - time: Port 37
  - gopher: Port 70
  - http: Port 80
  - http://www.computerhope.com/jargon/p/port.htm for more
- Nothing forces a webserver to be on port 80; could set at port 918, for example, then could connect to old address http://xxx/yyy/com via the new HRL http://xxx.yyy.com:918
- Try www.pacificu.edu:80 and www.pacificu.edu:90

Tracing Data: Traceroutes

- Shows the routers used in transmitting packets
- In Windows, use command tracert followed by the website on the command line
- In Mac OS X/UNIX use traceroute followed by the website on the command line
  - You try the traceroute command
- There is also a website that will provide graphical traceroutes
  - http://itzacompany.com/tools/trace.cfm

Firewalls

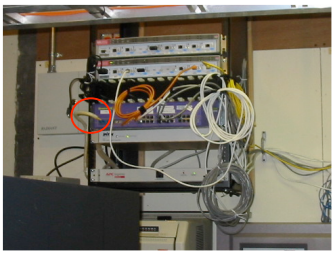
- Firewalls (among other things) control the ports and protocols that computers on the private local network can use to communicate with computers on the Public Network
- Restricts access to music filesharing, for example



Servers: Pacific (in UC)

Infrastructure: Pacific

- Fiber optic cable running from Pacific University to Comcast POP in Portland.



Infrastructure: Pacific Dorms

- CAT 5e/6 Cable running to each dorm room from Network switch (which connects to the main router system)

Router, Shaper, Firewall:Pacific

- TOP: Border Router (initial router for campus)
- MIDDLE: Packet Shaper (controls which packets have priority on our network)
- Bottom: Firewall (controls outside access)

Routers: Pacific

- Main router system
- All fiber optic cables, one generally for each building on campus