Introducing the Internet

1. Where did it come from?
2. Who run it?

The Internet: A Network of Networks

The Internet is an interconnected network of thousands of networks linking academic, research, government, and commercial institutions.
Internet Services

The Internet provides scientists, engineers, educators, students, business people, and others with a variety of services such as:

- **Electronic mail** (send/receive mail messages)
- **Remote login** (Telnet - access to other computers)
- **Transferring files** (FTP - accessing archives of data)
- **Newsgroups** (Usenet - on-line public discussions)
- **World Wide Web** (a collection of multimedia documents)

The Internet

- The Internet is a modern means of electronic communication.
- The Internet began as a way to permit rapid communication between the Defense Department and university research sites in the 1950s and 1960s.
- The basic Internet functions within the UNIX operating environment.
Early History

The Internet was the result of some visionary thinking by people in the early 1960s who saw great potential value in allowing computers to share information on research and development in scientific and military fields. J.C.R. Licklider of MIT, first proposed a global network of computers in 1962, and moved over to the Defense Advanced Research Projects Agency (DARPA) in late 1962 to head the work to develop it. Lawrence Roberts of MIT connected a Massachusetts computer with a California computer in 1965 over dial-up telephone lines. It showed the feasibility of wide area networking, but also showed that the telephone line's circuit switching was inadequate.

The Internet, then known as ARPANET, was brought online in 1969 under a contract let by the renamed Advanced Research Projects Agency (ARPA) which initially connected four major computers at universities in the southwestern US (UCLA, Stanford Research Institute, UCSB, and the University of Utah).

The Internet was designed in part to provide a communications network that would work even if some of the sites were destroyed by nuclear attack. If the most direct route was not available, routers would direct traffic around the network via alternate routes (packet switching).

The early Internet was used by computer experts, engineers, scientists, and librarians. Libraries began automating and networking their catalogs in the late 1960s independent from ARPA. In 1986, the National Science Foundation funded NSFNet as a cross country 56 Kbps backbone for the Internet. They maintained their sponsorship for nearly a decade, setting

Since the Internet was initially funded by the government, it was originally limited to research, education, and government uses. Commercial uses were prohibited unless they directly served the goals of research and education. This policy continued until the early 90's, when independent commercial networks began to grow. It then became possible to route traffic across the country from one commercial site to another without passing through the government funded NSFNet.
Internet Statistics

[Diagram: Internet Timeline]

[Diagram: Age split by Gender]
The Internet is a vast network of many different computers that are able to talk to each other in spite of the fact that they may be separated by large distances, may be very different computers built by different manufacturers, and may be running many different kinds of operating systems. This is similar to people in (say) Israel, Brazil, Japan, the United States, and Norway, each normally speaking a different language, being able to converse almost instantaneously with each other on very broad ranges of topics!

How is such efficient communication possible? Well, there are two big advantages that the computers on the Internet have over our hypothetical natives of different cultures and countries. The first is that the computers of the Internet are in reasonably constant contact with each other over a set of telephone lines, transoceanic cables, satellite links, and so on that allow rapid transfer of electronic information. The second is that the computers of the Internet have agreed to a common set of understandings that allows them to exchange vast amounts of information even though internally they may speak very different computer languages.
Internet Protocols

The language at the heart of the Internet is **TCP/IP**…

- Transmission Control Protocol/Internet Protocol
- Allows cross-network communication

A protocol is a set of rules that are used to transmit information on the Internet. There are several such *services*:

- **http**: hypertext transmission protocol
- **ftp**: file transmission protocol
- **gopher**: a means of exploring the non-graphic Internet for information.

**Internet Protocols**

**TCP** breaks messages into packets

- Each packet has all the information needed to travel from network to network
- Host systems called *Routers* determine how to route transmissions
Internet Addressing — The URL

• The URL is the *Universal Resource Locator* or Internet address for a particular location on the Internet.
• The URL was developed at CERN (European Organisation for Nuclear Research) which created the technological basis for the World Wide Web.

Internet Addressing — The URL

• IP address
  – The IP address consists of four sets of numbers punctuated by dots.
  – Example: 131.125.1.1 is the address for the Kean University web server also known as *turbo*.
  – The turbo server contains the Kean Website.
  – In this example the 131.125 part refers to Kean while the remaining numbers reference individual servers.
Internet Protocols

**IP** is the address for the packets

– Each Internet host computer has a unique *IP Address*

– Each address is comprised of four sets of numbers separated by periods, such as 123.23.168.22

Internet Addressing — The URL

- The URL is a unique descriptor that can identify any document, graphic, computer or other item anywhere on the Internet.
- URLs are like business cards for locations on the Internet.
- A browser generally has a place near the top of the window where which the current URL is placed.
Internet Addresses

131.125.66.4

To keep all of these machines straight, each machine on the Internet is assigned a unique address called an IP Address. IP stands for Internet protocol, and these addresses are 32-bit numbers, normally expressed as 4 "octets" in a "dotted decimal number." Every machine on the Internet has a unique IP address.

However, people generally can remember names better than numbers, so it is convenient to associate a name with such an IP address. In this case, the corresponding name is coe.kean.edu.

Internet Addresses

The translation between the numbers used by the network, and the name more commonly used by people is done by a computer called a nameserver. The purpose of a nameserver is to look up addresses, so its function may be likened to directory assistance on the telephone system.

- The host is named using DNS (domain name system), which translates IP addresses into a string of names.
IP Addresses

Every machine on the Internet has a unique identifying number, called an IP Address. A typical IP address looks like this:

216.27.61.137

To make it easier for us humans to remember, IP addresses are normally expressed in decimal format as a "dotted decimal number" like the one above. But computers communicate in binary form. Look at the same IP address in binary:

* 11011000.00011011.00111101.10001001

The four numbers in an IP address are called octets, because they each have eight positions when viewed in binary form. If you add all the positions together, you get 32, which is why IP addresses are considered 32-bit numbers. Since each of the eight positions can have two different states (1 or 0) the total number of possible combinations per octet is 256. So each octet can contain any value between 0 and 255. Combine the four octets and you get 4,294,967,296 unique values!

Out of the almost 4.3 billion possible combinations, certain values are restricted from use as typical IP addresses. For example, the IP address 0.0.0.0 is reserved for the default network and the address 255.255.255.255 is used for broadcasts.
Internet Addresses

Top level domains include:

- .edu - educational sites
- .com - commercial sites
- .gov - government sites
- .mil - military sites
- .net - network administration sites
- .org - nonprofit organizations

Reading a URL (1)

- **Parts of a URL**

  service: // hostname / directory-path

- **Service:** will be http or ftp.
- **//** - a UNIX separator indicating that the location is on a different computer
- **hostname** is the name of the host computer where the file you seek is to be found.
- **/** is the UNIX separator for directories (compared with DOS which is \\)
- **directory-path** tells the route to the file
Internet Addresses

An Internet address includes:
username@hostname.sub.dom

- **username** is the person’s “mailbox”
- **hostname** is the name of the host computer and is followed by one or more domains separated by periods:
  - host.subdomain.domain
  - host.domain
  - host.subdomain.subdomain.domain

Examples:

- **president@whitehouse.gov**
  - User President whose mail is stored on the host whitehouse in the government domain

- **hazel_filbert@lane.k12.or.us**
  - User hazel_filbert at the server for Lane County, Oregon, k-12 school district
Reading a URL (3)

- URL for EMSE 31222 Website
  http://www.kean.edu/~gkolodiy/3122/
- ~gkolodiy/3122/ – is the directory path.
- A directory is like a drawer in a file cabinet. It may contain other drawers, or it may contain files.
- ~gkolodiy – is the main directory assigned to Dr. Kolodiy
- 3122 is a directory within the main directory. It contains files for EMSE 3122.

Reading a URL (2)

- URL for EMSe 3122 Website
  http://www.kean.edu/~gkolodiy/3122/
- http:// - indicates hypertext transmission protocol
- www.kean.edu
  – www – indicates the server or location of the file
  – kean – is the domain name of the host computer
  – edu – is the zone that indicates an academic site
Internet Access Options

Direct (dedicated) Connection
- Computer has its own IP address and is attached to a LAN
- No need to dial up
- Files are stored on your computer
- Response time is quick

Internet Access Options

- Dialup Connections
  - limited connection using a modem
  - Full access dial up uses SLIP or PPP via modem
- Broadband Connections
- DSL service is newer, faster, and cheaper than ISDN
- Can share phone line with voice traffic
Internet Access Options

• Cable Modem Connection
  – allow Internet connections using shared TV cables
  – can exceed DSL speeds
  – Carry increased privacy and security risks

• Satellite Connections
  – provides connections using DirecTV satellite dishes

Internet Access Options

• Internet Service Providers (ISPs)
  – local ISPs provide connections through local telephone lines
  – national ISPs offer connections on a nationwide scale
The octets serve a purpose other than simply separating the numbers. They are used to create classes of IP addresses that can be assigned to a particular business, government or other entity based on size and need. The octets are split into two sections: Net and Host. The Net section always contains the first octet. It is used to identify the network that a computer belongs to. Host (sometimes referred to as Node) identifies the actual computer on the network. The Host section always contains the last octet. There are five IP classes plus certain special addresses:

* **Class A** - This class is for very large networks, such as a major international company might have. IP addresses with a first octet from 1 to 126 are part of this class.

* **Class B** - Class B is used for medium-sized networks. A good example is a large college campus. IP addresses with a first octet from 128 to 191 are part of this class.

A computer in a business or university has a Network Interface Card (NIC) that directly connects it to a Local Area Network (LAN) inside the business. The business then connects its LAN to an ISP using a high speed phone line like a T1 line. A T1 line can handle approximately 1.5 million bits per second, while a normal phone line using a modem can usually handle 30,000 to 50,000 bits per second.

ISPs then connect to larger ISPs, and the largest ISPs maintain fiber-optic "backbones" for an entire nation or region. Backbones around the world are connected through fiber-optic lines, undersea cables or satellite links (see this page for a nice backbone and connection diagram). In this way, every computer on the Internet is connected to every other computer on the Internet.
Web Addresses

A typical **URL** looks like this:

http://www.prenhall.com/beekman

- Dissecting Web Page address:

  - **Protocol for Web pages**
  - **Path to the host**
  - **Resource Page**
What is a URL?

A Uniform Resource Locator or URL address is a wedding of the information in the IP address for a machine and the information in its local file structure. Thus a URL address gives the location of a file, not with respect to a single computer, but with respect to the entire Internet!

Here is an example of a URL address:

http://csep10.phys.utk.edu/webcourse/browser/textfile.html

The first part gives (csep10.phys.utk.edu) the location of the server. The second part of the URL address specifies the Internet address of the machine housing the file and the location of the file at that internet address. In the first example cited.) The rest of the URL address then gives the location of the file in question on the machine specified by the preceding IP address.

When you type in a web name (URL such as http://www.kean.edu/gkolodiy/3122/Welcome.html) the browser breaks it into 3 parts

1. The protocol ("http")
2. The server name ("www.kean.edu")
3. Path to directory (gkolodiy/3122)
4. The file name ("Welcome.html")

* The browser communicates with a name server to translate the server name, "www.kean.edu", into an IP Address, which it uses to connect to that server machine. Every web page end with html or htm.
Inside the Web

Web site Jargon:

- **Web pages** are made up of text and images
- A **Web site** is a collection of web pages
- A **Home page** is the main entry to a Web site
- A **Web browser** like Netscape Communicator or Internet Explorer allows you to explore the Web by clicking links

Hyperlinks (links) are words or pictures that act as buttons, allowing you to go to another Web page.

Links are typically underlined or displayed in a different color.
Inside the Web

More Web site Jargon

- **Links** allow you to locate information without knowing its exact location (it may move from time to time)
- **Back** and **Forward** buttons let you retrace your steps
- **Bookmarks** and **Favorites** can be set up to mark your favorite Web locations

Country codes for some non-US websites

- .au - Australia
- .ca - Canada
- .de - Germany
- .jp - Japan
- .tv - Tuvalu
- .uk - United Kingdom
Publishing on the Web

HTML (Hypertext Markup Language)

- An HTML document includes codes that determine the format, layout, and structure of a Web document.

This text coded as HTML ...

```
<H1>Welcome to Computer Confluence</H1>
<b>Publishing on the Web</b>
```

Appears like this on the screen ...

Welcome to Computer Confluence!

Publishing on the Web
Publishing on the Web

Alternatives to HTML…

- Programs that convert document format features into HTML codes
  - Microsoft Word, FileMaker
- Web authoring programs
  - HomePage, GoLive, FrontPage

From Hypertext to Multimedia

Typical Web pages can contain:

- Tables
- Frames
- Forms
- Downloadable audio and video
- Streaming audio and video
- Real-time live audio or video
- 3-D environments
From Hypertext to Multimedia

Plug-Ins are software extensions that add new features. Examples include…
- QuickTime
- Shockwave/Flash
- RealPlayer
- Acrobat

Beyond HTML

Dynamic HTML: adds more programming power to HTML by allowing code to automatically modify itself under certain circumstances

JavaScript: a simple language for enhancing HTML Web pages
Beyond HTML

WML
- Wireless Markup Language helps create Web documents containing stock quotes, phone numbers, and other small nuggets of information

XML
- Will replace HTML plus provide additional features and extensions

VRML
- Virtual Reality Modeling Language creates 3-D virtual worlds

Beyond HTML

Java
- A full-featured, cross platform, object-oriented programming language

Java applets
- Small Java programs that can be automatically downloaded onto your client computer and can run on any platform