Internet Concepts

- Network, Protocol
- Client/server model
- TCP/IP
- Internet Addressing
- Development of the Global Internet

What is a Network?

- A group of two or more devices, that are able to communicate with one another:
  - The telephone network: interconnection of telephones and switches
  - Computer networks: interconnection of computers
  - Between two and millions of devices interconnected
The computers in a network are connected via hardware and software

Hardware
- Physically connects the computers together
- telephone lines, fiber-optic cables, routers and gateways, and the computers themselves

Software
- Enables communication and exchange of information, by following a set of rules, called *protocols*

Interoperability
- Different types of computers, using different operating systems, can be connected, communicate with each other, and share information
- Computers must agree to follow the same network protocols
Why use Computer Networks?

- Resource sharing
  - Share data, programs, and equipment
  - Across geographic boundaries
- Communications
  - Cooperation between dispersed groups
- Improved Reliability
  - Replication of files and resources allowing operation to continue despite hardware failures

Why use Computer Networks?

- Cost efficiency
  - Small computers have a better price/performance ratio than large ones
- Scalability
  - Add more computers to the network as required
Protocols

- A protocol is a standard set of rules that determines how computers communicate with each other across networks.
- A protocol describes:
  - the format that a message must take
  - the way in which computers must exchange a message
Protocols

- Usually defined for a particular activity
  - Sending messages across networks
  - Exchanging e-mail
  - Establishing remote connections
  - Transferring files
  - etc, etc...
- Enable different types of computers, running different operating systems to communicate

Mail Example

- E-mail protocols ensure that the e-mail message is correctly formatted and transmitted from originating computer to the destination computer
- Compare with post office mail
**Internet Protocols**

- Different protocols for different network services; some of the protocols used on the Internet:
  - Simple Mail Transfer Protocol (SMTP) to send and receive electronic mail
  - File Transfer Protocol (FTP) to transfer files between computers
  - Hypertext Transfer Protocol (HTTP) to transmit information on the World Wide Web
  - Network News Transfer Protocol (NNTP) to transmit discussion forum messages

**Layers and Stacks**

- To reduce their design complexity, most network protocols are organized as a series of layers or levels
- Each layer transparently provides services to the layer above it
- A group of such levels is called a protocol stack or protocol suite
Layers and Stacks

- Many different protocol stacks:
  - OSI  Open Systems Interconnection
  - TCP/IP  Internet Protocols
  - SS#7  Signalling System #7, used between telephone exchanges
  - SNA  System Network Architecture, IBM
  - DECnet Digital Network Architecture
- The number of layers, and the name, contents, and function of each layer differ for each protocol stack

Layering Analogy

- Japanese Philosopher
- Japanese/English Translator
- Post-office
- Irish-speaking Philosopher
- Irish/English Translator
- Post-office

Actual Transfer of Information
Logical Transfer of Information
Client/Server Computing

- A model for computer networking
- Efficient way to provide information and services to many users
- A network connection is only made when information needs to be accessed
  - Provides network efficiency
  - Decouples communicating parties

Client/Server Computing

- Client:
  - Requests services or information from a server computer
- Server:
  - Responds by sending the requested information back to the client computer
- Internet applications have used client/server model for decades
  - FTP, E-mail, Web
Client/Server Transaction

The Internet

- A network of networks
  - A large number of networks interconnected physically
  - Capable of communicating and sharing data
  - Able to act together as a single network (seamless)
- Based on TCP/IP
  - Transmission Control Protocol/Internet Protocol
  - The Internet is the network of networks which either use the TCP/IP protocol or can interact with TCP/IP networks via gateways (interpreters)
- Global
The Internet Protocol Stack

The Internet Protocol Architecture
TCP/IP

- TCP/IP is a family of protocols
  - The TCP/IP protocol suite (also includes UDP/IP)
- Many application protocols assume there is a way to communicate reliably between hosts
  - TCP/IP provides this service
- All TCP/IP protocol suite specifications are in the public domain
  - No license fees
  - Used extensively to create open system networking environments

Internet Protocol (IP)

- IP is a “Datagram” service
  - Provides a unreliable, connectionless datagram service
  - Datagrams == packets
- IP packets contain the source and destination addresses
  - “IP Spoofing” attack
- IP routes individual IP datagrams
  - Forwarded from router to router
- IP is never used directly by applications
  - Either UDP/IP or TCP/IP is used
Transmission Control Protocol (TCP)

- TCP provides reliable connection-oriented end-to-end communication
- TCP is a 'virtual circuit' protocol
  - Calls are setup, data exchanged and calls cleared
- Programs engage in stream based communication
  - The underlying TCP creates variable length packets from this
  - These are transmitted with sequence numbers, usually in a single IP datagram

Transmission Control Protocol (TCP)

- Stream-based communication being broken up and reassembled as messages (datagrams).
TCP/IP

Transmission Control Protocol (TCP)
- Breaks the higher level stream into IP datagrams
- Verifies that all packets arrive at their destination
  - Resends anything that gets lost
- TCP packets might take different routes
- Reassembles the data in correct order at receiver

TCP/IP

STEP 1
1 2 3
4 5 6

Data that makes up entire e-mail message

MESSAGE IS SENT

STEP 2

ROUTER

STEP 3
1 2 3
4 5 6

Data that makes up entire e-mail message

MESSAGE IS RECEIVED
User Datagram Protocol (UDP)

- Unreliable, meaning no guarantees that datagrams:
  - will arrive at their destination
  - will arrive in the order in which they were sent
  - will arrive intact
- Sits on top of IP but adds very little functionality
  - Essentially a front-end to IP

User Datagram Protocol (UDP)

- Less overhead than TCP → UDP may perform better than TCP when the network is not congested
  - In the case of congestion, UDP traffic anomalies become visible to application
- Hence, using UDP may require increased application functionality, e.g., to implement:
  - Data acknowledgements (positive, negative)
  - Data integrity checks
  - Datagram reordering
User Datagram Protocol (UDP)

- Suitable for some types of services, in particular those with real-time constraints
  - Media streaming
  - Some types of networked games

TCP and UDP Ports

- TCP and UDP use numeric port numbers to identify communication endpoints within a host.
- On the server side, port numbers are used to denote which service within the host machine you wish to talk to.
  - Each service (e.g., mail, web, dns, ...) runs on a separate port.
  - Only one server process can listen on a given port at the same time.
  - Multiple traffic flows can be received on the same server port.
TCP and UDP Ports

- On the client, port numbers are allocated automatically.
- Internet Assigned Numbers Authority (IANA) defines ports for common TCP applications:
  - daytime 13 (RFC867)
  - ftp 21
  - ssh 22
  - telnet 23
  - smtp 25

Other port numbers are available for private use.
- Both communicating machines must agree on what port numbers are used for what.
- http://www.iana.org/assignments/port-numbers

TCP and UDP ports

- More common port numbers:
  - http 80
  - http over ssl 443

- Other port numbers are available for private use.
Internet Addressing

- A way to locate computers, people, and resources on the Internet
- Hosts
  - 32 bit IP address e.g., 134.226.32.55
  - Domain name e.g., wilde.cs.tcd.ie
- Domain Name System (DNS) servers map domain names to IP addresses
  - wilde.cs.tcd.ie → 134.226.32.55

Internet Addressing

- People
  - E-mail addresses:
    - president@whitehouse.gov
    - pope@vatican.va
    - provost@tcd.ie
  - X.509 certificates
- Resources
  - URL’s identify specific sites and files available on the Web
    - http://www.random.org/mads/resume.html
Internet Domain Names

- A domain name identifies Internet hosts
- Top-level domain describes organization
  - COM - Commercial sites
  - EDU - U.S. Colleges
  - NET - Network providers
  - ORG - Non-profit, miscellaneous groups
  - GOV - U.S. government
  - MIL - U.S. military
  - INT - International organisations

Country Codes
- .IE - Ireland
- .UK, .FR, .NO, .SE, .NL, .DE, .BE, .JP, etc.

Structure for a host domain name
hostname.subdomain.second-level_domain.top-level_domain
eg. wilde.cs.tcd.ie

New top-level domains
- .info, .biz, .name, etc.
Internet Standards Bodies

- Internet Engineering Task Force (IETF)
  - Request for Comments (RFC’s)
  - Proposed Standard, Draft Standard, Internet RFC Standard
- Internet Society
  - Legal arm of the IETF
- World Wide Web Consortium (W3C)

The Size of the Internet

Internet Domain Survey Host Count

Source: Internet Software Consortium (www.isc.org)
Internet Significance

- September 2004 survey by Online Publishers Association (OPA):
  - The Internet has edged out TV as the medium of choice for 18 to 54-year-olds.
  - Some 45 percent of the 1,235 survey participants indicate the Internet is their top choice for media, followed by TV at 35 percent. Trailing much further behind are books, radio, newspapers, videos/DVDs, video/computer games, and magazines.
  - Source: ClickZ Stats, 21 September 2004

The Internet Population

- ~801.4M people online
  Source: www.glreach.com

- ~6,393M in the world
  Source: US Census Bureau

→

12.5% of the world’s population online

Source: 12.5% of the world’s population online

Table: Online Language Populations

<table>
<thead>
<tr>
<th>Language</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>38.5%</td>
</tr>
<tr>
<td>Chinese</td>
<td>13.7%</td>
</tr>
<tr>
<td>Spanish</td>
<td>9.0%</td>
</tr>
<tr>
<td>Japanese</td>
<td>8.4%</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Note: The table shows the distribution of online language populations with English being the most common, followed by Chinese, Spanish, and other languages.

Source: www.glreach.com
The Internet Population

~561M people online
Source: www.glreach.com

~6,200M in the world
Source: US Census Bureau

→

9% of the world's population online

Source: www.glreach.com

Autumn 2004 Trinity College, Dublin

369 million users (6% of world population)

505 million users (8% of world population)