How the Internet works

The single most important fact to understand about the Internet is that it can potentially link your computer to any other computer. Anyone with access to the Internet can exchange text, data files and programs with any other user. For all practical purposes almost everything that happens on the Internet is a variation of one of these activities. The Internet itself is the pipeline that carries the data between computers.

TCP/IP: the Universal Language of the Internet

The Internet works because every computer connected to it uses the same set of rules and procedures (protocols) to control timing and data format. The protocols used by the internet are called **Transmission Control Protocol/Internet Protocol** universally abbreviated as **TCP/IP**.

These protocols include the specifications that identify individual computers and that exchange data between computers. They also include rules for several categories of application programs so programs that run on different kinds of computers can talk to one another.

TCP/IP software looks different on different types of computers but it always presents the same appearance to the network. It does not matter if the system at the other end of a connection is a supercomputer, pocket size PC or anything in between – as long as it recognises TCP/IP protocols it can send an receive data through the internet.

Routing Traffic Across the Internet

Most computers are not connected directly to the Internet. Rather they are connected to smaller networks that connect to the Internet backbone through gateways. This is why the Internet is sometimes described as a *network of networks*. The core of the Internet is the set of backbone connections that tie the local and regional networks together and the routing scheme that controls the way each piece of data finds its destination.

The Internet includes many thousands of servers each with its own unique address. These servers in tandem with routers and bridges do the work of storing and transferring data across the network.

Because the Internet creates a potential connection between any 2 computers the data may be forced to take a long circuitous route to reach its destination. Suppose for example you request data from a server in another area,
1. Request must be broken into packets
2. Packets are routed through your local network and possibly though one or more subsequent networks to the Internet backbone.
3. After leaving the backbone the packets are routed through one or more networks until they reach the appropriate server and are reassembled into the complete request.

4. Once the destination server receives your request it begins sending you the requested data which winds its way back to you possibly over a different route.

Between the destination server and you PC the request and data may travel through several different servers each helping to forward the packets to their final destination.

Addressing Schemes – IP and DNS Addresses

Internet activity can be defined as computers communicating with each other using the common language of TCP/IP. Examples are
- Client system communicating with an Internet server.
- Internet server computer communicating with a client computer.
- 2 server computers communicating with each other.
- 2 client computers communicating via one or more servers.

The computer that originates a transaction must identify its intended destination with a unique address. Every computer on the Internet has a four part numeric address called the **Internet protocol address (IP address)** which contains routing information that identifies its location. Each of the four parts is a number between 0 and 255 so and IP address looks like 194.145.128.14

Computers have no problems working with long strings of numbers but we are not so skilled! Most computers on the Internet also have an address called a **domain name system (DNS)** address, an address that uses words rather than numbers.

Domains and Subdomains

DNS addresses have 2 parts
- Host name – name for a computer connected to the Internet
- **Domain** – generally identifies the type of institution that uses the address.
This type of domain name is often called a **top-level domain**.

Most common types of Internet domains,

Domain	Type of Organisation	Example
.com	Business (commercial)	ibm.com
.edu	Educational	centre.edu
.gov	Government	whitehouse.gov
.mil	Military	Navy.mil
.net	Gateway or host	Oceanfree.net
.org	Other organisation (typically non profit)	isoc.org

Some large institutions and corporations divide their domain addresses into smaller **subdomains**.
In 1996 the Internet Assigned Numbers Authority (IANA) and the Internet society began an organised movement to create an additional set of top-level Internet

domains. The action was necessary because many companies and private groups were finding it difficult to devise suitable domain names for their Internet sites. There was only so much room in the .com domain and some companies found that their mane or product name was already in use.

The group's goal was to expand the list of top-level domains to make it easier for organisations of all kinds to create an Internet domain for themselves. The group developed the Generic Top-Level Domain Memorandum of Understanding (TLD-MoU) which spells out proposals for the future management of Internet domains and proposes 7 new top-level domains for future use

Domain	Types of Organisation
.firm	Businesses or Firms (equivalent to .com)
.shop	Businesses offering purchases over the Internet
.web	Organisations involved in web related activities
.arts	Organisations promoting artistic or entertainment activities over the Internet
.rec	Organisations promoting recreational activities over the Internet
.info	Organisations providing information services over the Internet
.nom	Individual, family or personal nomenclature

Major Features of the Internet

The popularity of the Internet is due more to content than connectivity.

As a business tool it has many uses. Email is an efficient and inexpensive way to send and receive massages and documents around the world. The www is becoming an important advertising medium and channel for distribution. Databases and online information archives are often more up to date than any library. The Internet also has virtual communities made up of people who share interests.

Most individual users connect the computers modem to the phone line and set up an account with an **Internet Service Provides (ISP)** providing local and regional access to the Internet backbone. Many others connect through a school or business LAN.

The World Wide Web

The web was created in 1989 at the European Particle Physics Lab in Geneva as a method for incorporating footnotes figures and cross-references into online hypertext documents. A hypertext document is a specially encoded file that uses the **hypertext markup language (HTML)**. This language allows a documents author to embed **hypertext links** (called **hyperlinks** or **links**) into the document.

As you read a hypertext document (**web page**) on screen you can click on an encoded word or picture and immediately jump to another location. A collection of web pages is called a **web site**, and these are housed on a **web server**. Copying a page to the server is called posting the page (or publishing or uploading).

Popular web sites receive millions of **hits** (or **page views**) per day. Many web masters measure their sites success by the number of hits in a given timeframe. A **Webmaster** is the person or group responsible for designing and maintaining a website. The terms www and internet are used interchangeably however the www is just one part of the Internet.

Web browsers and HTML tags

Mosaic, a point and click web browser was developed at the University of Illinois in 1993. A **web browser** is a software application designed to find hypertext documents on the web and open them on the users computer. A web browser displays a web page as specified by the pages underlying HTML code. The code provides the browser with

- Fonts and font sizes
- Where and how to display graphics
- If and how to display sound, animation or other special content.
- Location of links and where to go if they are clicked.
- Whether special programming codes, which the browser needs to interpret, are used in the page.

HTML tags which are enclosed in angle bracket (<>) tell the browser how to display individual elements on the page. They are placed around the portions of the document that they affect. Most tags have a starting tag such as <H1> and an ending tag such as </H1>. A slash indicates a finishing tag.

HTTP and URLs

The internal structure of the World Wide Web is built on a set of rules called **hypertext transfer protocol (HTTP).** HTTP uses Internet addresses in a special format called a **uniform resource locator (URL)** that look like, type://address/path. Type specifies type of server on which the URL is located. Address is the address of the server. Path location within the file structure of the server.

Home Pages

- Personalised start page
  On your browser you can choose a web page that opens immediately when you launch the browser.
- Web site home page
  A web sites primary page is also called a home page. This is the first page that you see when you type the sites basic URL.

Helper Applications and multimedia content

Large files such as audio and video require special applications in order to be played in real time across the Web. These applications are called **helper applications** or **plugins**.

Plugins are used to support several types of content including **streaming audio** and **streaming video**. One of the most commonly used plugin applications is macromedias Shockwave, enabling web designers to create high quality animation or video with sound that plays directly within the browser window.

Finding content with a search engine

Specialised Web sites called **search engines** use powerful data searching techniques to discover the type of content available on the Web. By using a search engine and specifying your topic of interest you can find the right site of information.

Popular Search Engines
- Alta vista – http://www.altavista.com/
- Excite – http://www.excite.com/
- Hotbot – http://www.hotbot.com/
- Webcrawler – http://www.webcrawler.com/
- Yahoo! – http://www.yahoo.com/
- Google – http://www.google.com/
- Metacrawler – http://www.metacrawler.com/

Electronic Mail

Popular Internet email programs include
- Eudora
- Microsoft Outlook
- Netscape Messenger

If you have an account with an ISP then you can establish an email address. This unique address allows other users to send messages to you and allows you to send messages to others. A user can set up an account by specifying a unique **user name.** When you send a message you must include a persons username in the address eg. nora@maths.tcd.ie

When you send email the message is stored on a server until the recipient retrieves it. This type of server is called a **mail server** and many use the **post office protocol** and are called **POP servers.**

Listserv systems

One type of mailing list that uses email is an automated list server or **listserv**. Users on the list can post their own messages so the result is an ongoing discussion.

News

The Internet supports a form of public bulletin board called **news**. Many of the most widely distributed **newsgroups** are part of a system called **Usenet**. Users post **articles** about the groups topic and as others respond they create a thread of linked

articles. A newsreader program obtains articles from the **news server** using the **network news transfer protocol** (**NNTP**). To see articles posted on a specific topic you subscribe to the newsgroup addressing that topic.

Major Usenet domains
- Comp – computer related
- Sci – science and technology (not computers)
- Soc – social issues and politics
- News – topics related to Usenet
- Rec – arts hobbies and recreational activities
- Misc – all other topics.

Telnet – Remote Access to distant computers

This is the Internet tool for using one computer to access a second computer. You can send commands that run programs and open text or data files. Connecting to a Telnet host is easy, enter the address and the telnet program establishes a connection.

FTP

File Transfer Protocol is the Internet tool used to copy files from one computer to another. When a user has accounts on more than one computer FTP can be used to transfer data or programs between them.

Internet Relay Chat (IRC)

Internet Relay Chat, or just chat is a popular way for Internet users to communicate in real time with other users. Chat does not require a waiting period between the time you send a message and the time the other party receives it. IRC is often referred to as the CB radio of the Internet because it enables few or many people to join a discussion.

Accessing the Internet

- Direct Connection
  Programs run on the local computer which uses TCP/IP protocols to exchange data with another computer through the Internet. An isolated computer connects to the Internet through a serial data communications port using SLIP (serial line interface protocol) or PPP (point to point protocol).
- Remote Terminal Connection
  Exchanges data and commands in ASCII format with a host computer that uses UNIX or similar OS. TCP/IP application programs and protocols all run on the host. This is known as a shell account as the command set in UNIX is called a shell.
- Gateway Connection
  Even if a LAN does not use TCP/IP commands and protocols it may provide some Internet services. Such networks use gateways that convert commands and data to TCP/IP format.

- Connecting Through a LAN
  If a LAN uses TCP/IP protocols for communication within the network it is simple to connect to the Internet through a router, another computer that stores and forwards data to other computers on the internet.
- Connecting Through a Modem
  If there is no LAN on site a stand alone computer can connect to the internet through a serial data communications port and a modem using either a shell account and a terminal emulation or a direct connection with a SLIP (serial line interface protocol) or PPP (point to point protocol) account.
- High Speed Data Links
  Using fibre optics, microwave and other technologies it is entirely practical to establish an internet connection that is at least 10 times faster than a modem connection.
  - ISDN service
  - xDSL services
  - Cable modem service

Working on the Internet

Businesses and Firewalls

A firewall is set up to control access to a network by people using the Internet. Firewalls act as barriers to unauthorised entry into a network that is connected to the Internet, allowing outsiders access to public access areas but preventing them from exploring proprietary areas of the network.

A firewall system can be hardware, software or both. It works by inspecting requests and data that travel between the private network and the Internet. If the request or data does not pass the firewalls security inspection it is stopped from travelling any further.

Intranets and Extranets

An intranet is a LAN or WAN that uses TCP/IP protocols but belongs exclusively to a corporation, school or organisation. It is accessible only to the organisation's workers. If it is connected to the Internet then it is secured by a firewall to prevent unauthorised access.

An extranet is an intranet that can be accessed by outside users over the Internet. To gain entrance an external user typically must log on to the network by providing a valid user ID and password.

Intranets and Extranets are popular for several reasons including
- Use standard TCP/IP protocols → simpler and less expensive to install and configure.
- Enable users to work in standard web browsers providing a consistent interface.
- Function readily with firewalls and other standard security technologies.

Issues for Business Users and Telecommuters.

- Ownership

  Any piece of text or graphic retrieved from the Internet may be covered by trademark or copyright law making it illegal to use it without the owners consent.
- Libel

  If email messages are sent through an employers network then the employer may become involved if the sender is accused of libel.
- Appropriate use

  When using a business network to access the Internet users must be careful to use network resources appropriately.

Commerce on the www

Using an e-commerce site is like browsing through an online catalogue. When you are ready to make your purchases you can pay in several ways,

- One time credit card purchases

  Provide your personal and credit card information each time you make a purchase.
- Set up an online account

  If you think you will make more purchases with the online vendor you can set up an account at the web site. The vendor stores your personal and credit card information on a secure server and a cookie is placed on your computer disk. Later when you access your account using a user ID and password the site uses the information in the cookie to access your account.
- Use electronic Cash

  Also called digital cash. Takes the form of a redeemable electronic certificate which can be purchased from a bank that provides electronic cash services. Not all e-commerce web sites accept digital cash yet.
- Electronic wallet

  Program on your computer that store credit card information, a digital certificate that verifies your identity and shipping details. Not accepted by all e-commerce sites.

Security

Reputable e-commerce sites use sophisticated measures to ensure that customer information cannot fall into the wrong hands. One way is to use **secure web pages**.

One way to provides secure websites is to encode pages using **secure sockets layer (SSL)** technology which encrypts the data. Another way is to use **secure HTTP (S-HTTP)**. SSL can be used to encode any amount of data, S-HTTP is used to encode individual pieces of data.