



INTRODUCTION: HISTORY OF THE INTERNET

The internet is made up of thousands of small national and regional networks, which link together to form a global network of computers that can communicate and exchange information. The internet's origins lie in **ARPANET**, commissioned in 1969 by the US Defense Department to provide a secure communications channel for US military research which would be resilient to nuclear attack. ARPANET proved so successful that in 1983 the military use was split off and the remaining service opened up to other researchers. At that point, ARPANET connected 60 universities in the USA, one in Norway and two in the UK. By the following year, over 1000 host computers were connected.

The base network in the UK was **JANET** (the Joint Academic Network, www.janet.ac.uk). Its origins lie in several small scientific networks which were developed in the UK from the late 1960s onwards: for example, the National Physical Laboratory's network established in 1968. JANET was inaugurated on 1 April 1984. In 1991, JANET was directly linked to the internet. JANET is now funded by the **Joint Information Systems Committee (JISC)**, www.jisc.ac.uk, of the Higher Education Funding Councils and managed by the UK Education and Research Networking

Association (UKERNA). Now many private networks in the UK have been connected to the internet.

BASIC NETWORKING TECHNOLOGIES

Networks

The two basic types of network are **local area networks** (LANs) and **wide area networks** (WANs). A LAN links computers that are physically close to each other and usually 'hard wired' together via cables. Typically, LANs are used in single organizations at single sites. WANs, however, can cover large distances, within and beyond national boundaries; they are generally connected through telecommunications links, which may use a mixture of advanced technologies such as fibre-optic cables and satellites. It is the wide area networks that are of most concern here.

Computers handle digital data in the form of discrete bits and bytes. Networks typically transmit data in one of two forms, either **analog** (a continuous signal) or **digital**. Most networks, like the internet, use dedicated digital telecommunications lines. Organizations which make heavy use of an external network may install leased telecommunications lines to provide permanent, digital links from their in-house multi-user machines or LANs to external network services.

Protocols

Networks require a common framework of routines and rules to allow computers to communicate with each other. These are called **protocols**. Protocols are technically complex but, put simply, they specify, for example:

- how data are to be encoded for transmission
- the physical transmission media that are allowable

- the conventions for addressing items of data so that they can be delivered to the correct network destination
- the applications (the types of tasks) that are to be supported on the network.

The internet uses the **TCP/IP** protocol, (Transmission Control Protocol/Internet Protocol). Local networks typically use the **Ethernet** protocol. The main task of a protocol is break down data to be transmitted into **packets** (small chunks with an identifier, a destination and a sender) so that those packets can be **routed** (directed across intermediate connections) and travel from the sending machine to the destination machine.

Servers

Any use of a network involves at least two computers: the one the user is on, and another one that is being accessed for some purpose via the network. The user's computer may be a personal computer directly connected to the network, or a **gateway server** to which the user connects from their own machine which acts as a terminal. The internet generally operates on a **client/server** model. One computer, the client, on behalf of the user, requests services of another computer, the server. At any one time a single server computer can be dealing with any number of client computers. Thus server computers tend to be more powerful than client computers, whose job very often entails merely the display of data passed along from a server.

Bandwidth

The speed and power of the server, and the **bandwidth** (how much data can be moved per second) of the network connection, determine how quickly these services are fulfilled. Dedicated network lines, especially if

they are made of **optical fibre**, can handle far more data than an ordinary telephone line can, even using the latest, fastest modems. The limit on data transfer, the bandwidth of a telecommunications line, can still be reached even on the fastest of lines when lots of users try to move vast amounts of data.

Digital networks

There is a trend away from the use of single-function telecommunications lines moving data in analog form (the traditional telecommunications infrastructure) to one in which lines handle data in digital form and can transmit voice, computer data, video, etc. in a single common format. Apart from all traffic using one line, digital lines are easier to upgrade to higher capacity than analog ones. BT recently announced that its core network is to be based on TCP/IP, to exploit this flexibility. This trend is beginning to affect even the standard telephone service. VOIP (voice over IP) offers a voice service over an IP-based network. This service is expected to slowly replace the existing analog domestic and business phone system. VOIP phones plus onto a broadband connection and function exactly as a 'normal' phone, but rental costs and call charges are lower. They are available now from a range of suppliers, including British Telecom (www.btbroadbandvoice.com) and Sipgate (www.sipgate.co.uk/).



Services - VOIP, <http://directory.google.com/Top/Business/Telecommunications/Services/VoIP/>, lists services and/or software which use VOIP.

CONNECTING TO THE INTERNET

Your organization may already provide you with internet access (as at academic sites, for example). Alternatively you can pay a commercial internet service provider (or ISP), some of which offer low-cost

connections aimed at individual users, while others specialize in linking corporate networks, with a range of options (and costs) in between.

Essentially an individual user needs a telephone socket connected to a

For a listing of ISPs in the UK and the connection services and plans they offer see *ISP Review*, www.ispreview.co.uk/.



computer with a **modem** (a device which enables computer signals to traverse a telephone line). The modem, under the control of **communications software**, calls a number which, once a user identification and a password have been given, opens a connection into the internet. Just about any sort of computer sold by a computer dealer for business or home office use can be used to access the internet. This is known as **dial-up** access.

Dial-up connections function only while the line is live and modems have reached a ceiling in terms of speed. Increasingly users are opting for permanently 'on' **broadband** connections to the internet, providing the home user with the same level of access as someone working for an organization. These could be provided either by an **asymmetric digital subscriber line (ADSL)**, which works over ordinary telephone lines, or **cable modems**, which work on the optical fibre networks of local cable companies. In ADSL data flow faster towards the user (**downstream**), than from the user to the network. This is a good thing since most internet usage is pulling content off the network. Most broadband services are **contended**, shared between a certain number of users. At the time of writing a range of faster speed (greater than 512k per second downstream) broadband connections is appearing. Some are based on a **symmetric digital subscriber line (SDSL)** in which data flow upstream and downstream at the same rate.

Broadband Help, www.broadband-help.com/home.asp, is an authoritative guide to broadband services and suppliers.



Not all locations in the UK can access broadband: some are too remote from a telephone exchange that offers ADSL. **Integrated services digital network** (ISDN) connections will be available and these are always on, but their speed is not much faster than modem dial-up. **Satellite broadband** is currently the only option for broadband in these locations: satellite broadband, however, suffers from **signal lag** in between the ground and the routing satellite in orbit. This significantly hampers highly interactive services (see Chapter 7).

Home networking

Home networking is growing because of always-on broadband domestic connections. This involves sharing the single broadband connection among a number of computers (e.g. the parental work laptop and the children's recreational desktop). Such sharing requires a **router**, a device hitherto seen only in workplace networks, and means of connecting devices. Wires can be used but more convenient (and less unsightly in the home) is **wi-fi** (wireless-fidelity) which is an implementation of the networking Ethernet protocol which works without wires. A wi-fi capable router broadcasts and receives data and each connected device needs a wi-fi network card (add-on which gives a new facility). Increasingly, devices such as **internet radio** receivers and **digital music players** are being linked into home networks. The day of the networked fridge is not far away!



DIY Home Networking Guides and Tutorials, www.homenethelp.com/, is good for information on this fast developing area.

Internet on the move

Mobile internet connections have appeared: they do not require a wired connection and operate in public places. Wi-fi connection points (known as

hotspots) have been offered in busy public places (e.g. train stations, hotels, airports, etc.), some provided by wi-fi service providers and some by chains of shops (e.g. Starbucks). For a subscription a laptop or a PDA with a wi-fi card can be connected within range of a wi-fi server. **Roaming** enables a customer of one wi-fi ISP to use the connection points of another. However, wi-fi connections are short range (around 50 metres) and not yet ubiquitous.

Intel's Hotspot Finder, <http://intel.jiwire.com/>, is the best global guide to hotspots and has links to wi-fi service providers. FreeNetworks.org, www.freenetworks.org, lists free wi-fi hotspots throughout the world, provided by people willing to share their broadband connections.



Mobile phone companies are offering internet connections wherever a mobile phone signal exists. Many mobile phones come with **general packet radio service (GPRS)** which is always on and transmits digital data at around the same speed as a modem. A new generation, **3G (3rd generation)**, is also always on but connects at near broadband speeds. These connections are meant to deliver internet facilities (like e-mail and web) to a mobile phone but small screen size and tiny keyboards make mobile phones (and related devices like the Blackberry) fiddly to use. It is possible to connect some mobile phones via wires or wirelessly, using the limited range networking protocol **Bluetooth**, to laptops or PDAs, which have bigger screens and better keyboards. Mobile internet connections are, at the time of writing, expensive because they are charged on data transmitted, but it is expected that charges will fall, as they have in the fixed-line internet connection market. Of course, if you lack the appropriate technology (mobile phone and/or laptop) then your choice of internet connection on the move will be restricted to **cybercafés**.

For a searchable list of cybercafés see Cybercafes.com, www.cybercafe.com/.





USING THE INTERNET ON THE MOVE

Mobile internet services are currently very much like fixed line dial-up services were ten years ago, i.e. slow, unreliable and expensive. GPRS connections are not ubiquitous. Areas outside cities or major transport links have few GPRS connections. 3G is being rolled out only in big cities. 3G connections promise high speed but do not seem to deliver. Content provider plans to sell streamed music and video over such links might well be premature. Some services provide automatic connection detection (between GPRS, 3G and wi-fi), leaving the user to choose the fastest (or cheapest) available. Charging by content sent and received is fair but very disruptive of internet use patterns. (Why bother to turn off graphics in web pages or eschew downloading e-mail in case someone has sent a large attachment?)

uk.telecom.mobile, <http://groups.google.co.uk/group/uk.telecom.mobile?hl=en>, and 3G Forum, www.3g.co.uk, are good sources of advice and opinions.

SITE ADDRESSING - HOST AND DOMAIN NAMES

People and resources on a network have to have a **site** that is defined in terms of an address of a computer on the network. Thus people are located by the particular computer they use, and information resources by the computer on which they are stored. On the internet, computer names are made up of two parts, a **host name** and a **domain name**. A domain name is similar to the STD or area code in a telephone number. It tells you where the computer is, and what organization owns it. A host name is an identifying name for a computer within a domain, just like a telephone number identifies an individual in a particular area. Here are some typical examples:

- www.bbc.co.uk
- www.ukoln.ac.uk
- www.homeoffice.gov.uk.

Domain names are normally in three parts and read right to left. The right-hand part identifies the country of the domain. Most countries have a two-letter country code – uk for the United Kingdom, de for Germany, etc. If there is no country code, then the United States is implied although there is a ‘us’ code which is occasionally seen. To the left of the country code is a code showing organization type. In the UK domain these are:

ac	academic
co	private company
gov	government
org	non profit-making organizations.

A number of new domain names were introduced in 2001/2. They include .biz, .info, .name and .pro as well as .aero, .coop and .museum.

In domains for other countries, organization type codes may vary. For example, domains in the United States and Australia use ‘edu’ for an academic organization, while in most European countries ‘ac’ is used. To understand the organization type in a domain name, a little judgement is called for.

The last element in a domain name is an abbreviated name for an organization. Thus Leeds Metropolitan University is abbreviated to ‘lmu’ and the University of Bath to ‘bath’. Some organizations are recognizable; others are not. The host name is the final element. It can be something mundane (like ‘hp3’ for the third Hewlett Packard minicomputer) or it can be something more memorable (like ‘sloth’).

There are exceptions to these conventions such as:

- portico.bl.uk
- pipex.net.

In both the above, there is no organization type code.

Host and domain names together identify a computer on the internet. An alternative form of computer name that can be used is the **IP address**. This is a string of four numbers separated by full stops, e.g. 138.38.32.45 for UKOLN. These numbers are used in the actual addressing done on the internet and an automatic translation is made (invisibly to the user) of host and domain names into numbers, by a system known as the **Domain Name Scheme (DNS)**. The DNS scheme is managed by a non-profit organization called **ICANN**.

Domain names can be valuable commodities. A company is usually very concerned to get a domain name that either matches its company name or is appropriate for its service/product. ICANN oversees a number of



Nominet, www.nominet.org.uk/, is the ultimate registrar for .uk domains.

accredited domain-name **registrars** around the world who are responsible for assigning and maintaining domain names.

The internet may be difficult to use and understand primarily because the resources it contains have no physical presence for a user, other than on a computer screen. Host and domain names not only locate people and resources but can also help the user to infer something more about new resources as they are discovered. Thus a UK television schedule provided by www.bbc.co.uk (the BBC) ought to have more authority than one provided on sloth.cs.du.edu, a computer (host name 'sloth') located at the Department of Computer Studies at Denver University in the United States (a hypothetical example). This is only a guiding principle, not a universal rule.

Domain names are commercial properties. **Domain squatting** is the strategy of buying domain names expected to be wanted by big companies or rich individuals later, to whom they can be resold at profit. **Typo squatting** is the practice of buying domains identical in name to very

popular domains (e.g. www.google.com) except for a minor typo. Mistakes which cause people to arrive at these domains still bring in substantial traffic. Typo squatting can be used in conjunction with phishing (see Chapter 5) to create almost convincing fakes of other domains (e.g. www.nationalwestminsterbankuk.co.uk instead of www.natwest.co.uk, the real one!).

WHOIS

How do you know who owns a domain? **WHOIS** is an internet search facility of the ICANN domain registries. It can be searched either via a client search package or a website which offers this facility, such as Arin WHOIS Database Search, www.arin.net/whois/. **IPBlock** is a related search function which turns an IP address into a domain name.

Traceroute

Domains can also be used in diagnosing internet-related problems. An internet connection fault (e.g. a website does not load) can either be the fault of your internet service provider in which case you can contact it with the problem, or it might be a fault somewhere else on the internet. The most useful facility for diagnosing where internet problems lie is **traceroute**. It can be used via either a client search package or a website which offers this facility. Traceroute shows you each machine or router on the internet that traffic from your machine to another machine (e.g. a website) traverses. Using this display you can see where the problem occurs – in your ISP, at an intermediate stage or at the final site you are trying to access. The traceroute display shows graphically where a break occurs.



For a global listing of browser-based traceroute tools, which can route to a destination site and also back to your own address (so that together the paths show how packets flow between your machine and the website you are trying to reach) see Traceroute.org, www.traceroute.org/. The London Internet Exchange, the main switching point for internet traffic into/out of the UK, provides a public traceroute facility at Network Tools, www.linx.net/www_public/our_network/network_tools/. A client package that performs all the above functions is Sam Spade, www.samspade.org/ssw/.

SECURE INTERNET

Your internet connection can be spied on at the packet level by **packet sniffers** (software which records packets and their contents) anywhere between your machine and the packet's destination. Packets reveal your machine's IP address and from this someone can find out who owns that machine, where that machine is located and who provides its internet connection. Looking into the contents of packets will reveal what services you are using, and possibly critical details like your passwords, your credit card number, etc. If the IP address of your machine is discovered then **port scanner** software on a hostile machine can probe your machine's internet connection looking for a way into your machine to plant trojans (see Chapter 5). **War driving** is cruising using special software to detect private unsecured wi-fi connections and illegally using them for internet access.

What can be done about these security weaknesses? The internet was never designed with security in mind so solutions are not easy to use nor all-encompassing.



Wireless LAN Security, www.wardrive.net/, gives advice on how to lock-down wi-fi networks.

Port scanners can be defeated by a **firewall** which watches and blocks unauthorized external accesses to a machine (and watches outgoing connections). Firewalls can be software-based or hardware-based (most routers have one).

Zone Alarm is an example of a firewall for Windows systems. It has a freeware version downloadable from Zonelabs, www.zonelabs.com/. A firewall must be configured to stop or allow certain outgoing or incoming connections although they do have default settings.



Passwords and credit card details can be hidden by encryption (scrambling characters). This can be done by **secure sockets layer (SSL)** which is used by internet shops: its activation is shown by the padlock symbol in the lower bar of your browser. **IP spoofing** software can hide your IP address as can public or for-pay anonymizer websites.

WebTunnel from Primedius, <http://www.primedius.com/>, is client software which hides your real internet address, while Proxying and Filtering - Hosted Proxy Services, http://directory.google.com/Top/Computers/Internet/Proxying_and_Filtering/Hosted_Proxy_Services/?il=1, is a directory of online anonymizer services.



Finally **virtual private network (VPN)** software can encrypt all your packets, if they are being sent to a chosen destination that can host its own matching VPN software, e.g. you are working at home and connecting up to your office remotely.

Virtual Private Networks, http://directory.google.com/Top/Computers/Security/Virtual_Private_Networks?tc=1/, lists VPN facilities.



