The history and growth of Internet and WWW

Technical Report · January 2000

CITATIONS
0

READS
2,835

1 author:

Vassilys Fourkas

43 PUBLICATIONS **16** CITATIONS

SEE PROFILE

1. The history and growth of the Internet and the World Wide Web

The idea of remote connection to a machine can be indirectly traced to George Stibitz, of Bell Laboratories, who in the 1940's annual conference of the American Mathematical Society demonstrated a remote connection over a distance of 230 miles (Kitchin, 1998:28). Twenty years later and during the 1960s, an exciting concept of a decentralised, packet-switching network was kicked around the Defence Department of Advanced Research Projects Agency (DARPA), America's foremost Cold War military corporation. Leonard Kleinrock at MIT published the first paper on packet switching theory in July 1961 and one year later J.C.R. Licklider of MIT, and first head of the computer research programme at DAPRA, wrote a number of memos about his idea of a 'Galactic Network', thus an idea for globally interconnected set of computers through which everyone could quickly access data from any site (Leiner, *et al*:2000). However, Donald Davies and Roger Scantlebury at the National Physical Laboratory in UK set up the first test network of these principles in 1968 (Sterling, 1993: 2). Shortly afterward the US Pentagon's decided to fund a larger, more ambitious project called ARPANET to prevent any attempt by the Soviets to destruct or takeover the American telecommunications system in case of a nuclear war.

The outcome was a complex computer-mediated communication network - an experimental network designed to support military research about how to build networks that could withstand partial outages, like bomb attacks and still function. According to Castells (1996: 6) it was the electronic equivalent of the Maoist tactics of dispersal of guerrilla forces around a vast territory to counter an enemy's might with versatility and knowledge of terrain. Due to Kleinrock's early development of packet switching networking, his Network Measurement Center at UCLA was selected to be the first node on the ARPANET, in September 1969. In 1973, the first international connection to the ARPANET was made to University College of London via the Norwegian network of NORSAR.

However, the first event in the life of the Internet as we know it today came in 1974, when Vint Cerf and Bob Khan defined the Transmission Control Protocol (TCP) and Internet Protocol (IP) by which information could be put into a 'packet' and addressed so that computers on the network would pass it along, in the right direction, until it arrived at its destination (Leiner, et al: 2000). Eventually, the Internet was officially launched in 1983 when ARPANET, having 562 nodes around the world, split into the public Internet (494 nodes) and the military Milnet (68 nodes) that became integrated with the USA's Defence Data Network created the previous year. In 1984, Paul Mockapetris of USC/ISI invented the Domain Name System (DNS) that converts a numerical IP address (e.g. 193.134.196.35) in an assigned domain name (e.g. man.ac.uk), making it easier for people to use the network. Thus, by 1985, Internet was already well established as a technology supporting a broad community of academics and developers, and was beginning to be used by other communities for daily computer communications. The use of TCP/IP standards for computer networking became global and lot of institutions, foundations, and private firms have become Internet Service Providers, providing connectivity to thousands of computers on a large number of networks.

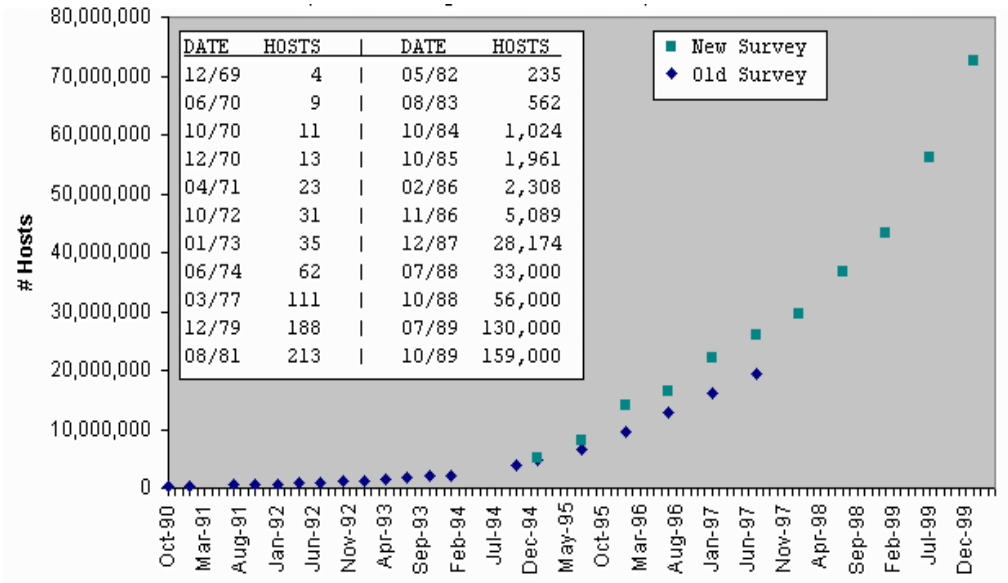
In early 1990s the number of registered domain names around the world was less than 200,000 but within the last four years it was rapidly proliferated, and by the end of July 2000 more than 17.8m domains registered worldwide (DomainStats, 2000). Most of them belong to commercial sector, as the number of domains ending in .com was 9,482,427 (55%). Accordingly, there is a true strong competition in domain name registration, which stimulates the growth of Internet as a global electronic market. For instance, in 1996 domain name 'tv.com' sold to CNET for US$15,000; in 1997 domain name 'business.com' sold for US$150,000 and two years later was re-sold for US$7.5million; in 1998 Compaq paid US$3.3million for 'altavista.com' (Zakon, 2000). Therefore, a domain name registration is

anymore considered as an internal part of marketing plans, and an increasingly important strategic asset for many companies, as well as for public bodies that are seriously thinking to exploit Internet as a means for developing their services.

On the other hand, according to surveys contacted by the Internet Software Consortium[1], and as Figure I.1 shows, the number of connected to Internet computers known as 'Internet hosts' in 1990s is similarly proliferated: from 10,000 in 1987, to more than 1,000,000 in 1992, and by December 1999 the number of hosts exceeds over 70,000,000. It must be noticed that, although the definition of host includes everything from desktop computers to powerful web servers, it does provide a good indicator of the growth and minimum size of the Internet and has been effectively used by many researchers (i.e. Kellner, 1999; UNDP, 1999; Zook, 2000; Zakon, 2000).

Figure I.1. The growth of Internet hosts

(Source: Zakon, 2000: http://info.soc.org/guest/zakon/Internet/History/HIT.html)



DATE	HOSTS	DATE	HOSTS
12/69	4	05/82	235
06/70	9	08/83	562
10/70	11	10/84	1,024
12/70	13	10/85	1,961
04/71	23	02/86	2,308
10/72	31	11/86	5,089
01/73	35	12/87	28,174
06/74	62	07/88	33,000
03/77	111	10/88	56,000
12/79	188	07/89	130,000
08/81	213	10/89	159,000

The latest results ISC's measurements (based on data collected by Network Wizards) show that the Internet currently 109,574,429 "advertised" connected computers (hosts) in 230 countries and territories. The figure of 109 million hosts represents a significant new benchmark for the number of Internet hosts. The current annual growth rate now stands at 51% - within the 46-67 % rates seen over the past two years. If the same growth rate is sustained, the Internet would cross the 1 billion host mark in mid 2005. The largest domain is .com, jumping 7.8 million hosts since January to a new high of 36.3 million hosts (see Appendix I: a). That represents a current annualised growth rate of 32%. As a percentage of the entire Internet, the .com hostcount stayed about the same at 33.2% of all Internet hosts[2].

Sterling (1993) and Georgiadou (1995), argue that the USA's National Science Foundation through its Office for Advanced Scientific Computing benefaction was very important because, with its networking effort in the late 1980s ($200 million from 1986 to 1995), Internet became accessible not only by computer scientists and the government as it was before, but by the wider academic and scientific community. In addition, and during the 1980s, networking projects were also developed in other western countries, as well in Japan and Australia. In 1982 the EUNET was developed in Germany to serve European research and business networking and in 1984 the JANET, an academic and research network was launched

[1] www.isc.org

[2] The general data plus some charts and information about the source data are available at the ISC site www.isc.org/ds. Post-processed data is available from Matrix.Net, Inc. (formerly MIDS) of Austin TX www.matrix.net

in UK. Both became soon compatible and thus connected to the Internet, and such kind of developments certainly gave further boost to its growth. However the widely recognised fact that lent wings to Internet's diffusion was the invention of the World Wide Web in 1989.

Behind the reality of the World Wide Web or *Web* are the ideas of Vannevar Bush and Ted Nelson. Vannevar Bush - President Franklin Roosevelt's science advisor - is the first known person who in 1945 envisaged a 'hyper-textual' - virtual - learning environment. He proposed a scholar's workstation called 'memex'; a hypothetical mechanical device that predated computers and that could display text and pictures at a press of a button. Twenty years later Ted Nelson was a student at Harvard University when he imagined a future in which one could get any type of information, images, sound, text, films all by selecting the appropriate icon within an on-screen document – the *hypertext* (Georgiadou, 1995: 35). All traditional text, whether in printed or digital form is sequential, meaning that there is a single linear sequence defining the order of the text. On the other hand hypertext is non-sequential; there is no single order that determines the sequence in which the text is to be read because words in a document could be 'linked' to other files initially only on his computer. (Nielsen, 1995: 1).

In 1989 Berners-Lee[3], British graduated from the Queen's College, Oxford in 1976, proposed the actual WWW hypermedia project at the European Laboratory for Particle Physics[4], in Geneva. Building on ideas that were current in software design at the time, Berners-Lee fashioned a kind of *hypertext notebook*. Moreover, the Lee's vision was to develop a network scheme with no central manager, no central database, and no scaling problems, something like the open-ended and infinite Internet. To Berners-Lee words, "the dream behind the Web is of a common information space in which we communicate by sharing information. Its universality is essential: the fact that a hypertext link can point to anything, be it personal, local or global, be it draft or highly polished" (Berners-Lee, 1999).

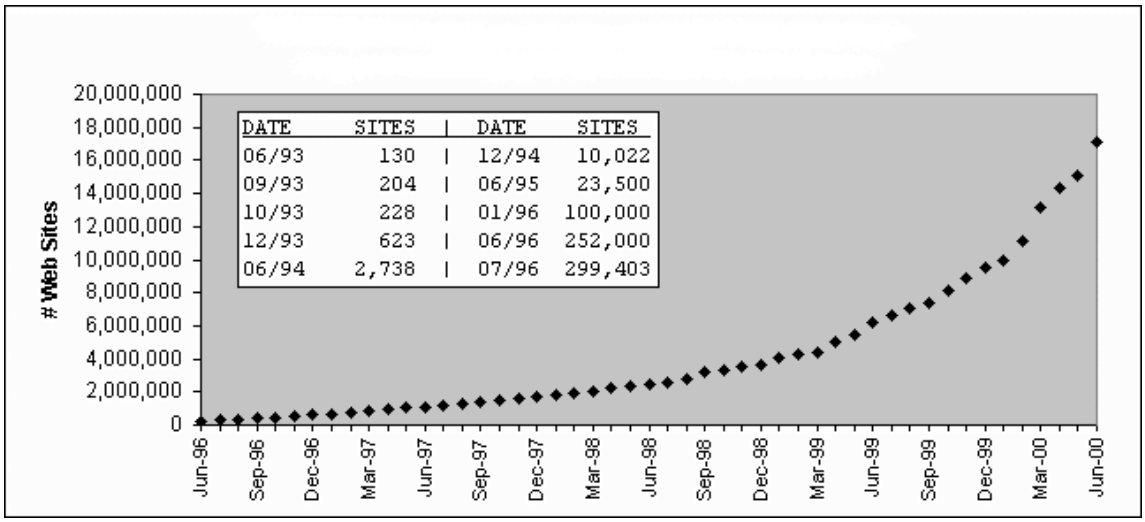
Hence, he developed a relatively easy-to-learn coding system – HTML (HyperText Mark-up Language) - that has come to be the lingua franca of the Web. He designed an addressing scheme that gave each Web page a unique location – the Universal Resource Locator (URL) - hacking also a set of rules – the HTTP (HyperText Transfer Protocol) - that permitted these Web pages to be linked together on computers across the Internet. Thus, the URL (e.g. http://www.man.ac.uk) is the Web version of an Internet domain name (IP address). In 1991, the Web debuted on the Internet, but it did not become widely accessible to those outside scientific and academic community until the invention of the *Mosaic* graphical interface in 1993 and the launch of the *Netscape* Internet browser a year later. From that moment on, Internet and WWW grew as one (see Winder, 1995:34). But what is the difference between the Net and the Web? "On the Net, you find computers -- on the Web, you find documents, sounds, videos... information. On the Net, the connections are cables between computers; on the Web, connections are hypertext links. The Web exists because of programs, which communicate between computers on the Net. The Web could not be without the Net" (Berners-Lee, 1999).

Accordingly, the overall importance of the Web lies on the fact that it allows the exploration of Internet in a simple and interactive way, by a simple act of clicking the mouse on an active point/ link (text or picture), contrary to the old days of navigating, when users used to follow a much more complicated path with a series of commands. "The Web made the Net useful because people are really interested in information and don't really want to have know about computers and cables" (Berners-Lee, 1999). The Web, has no central point but most people find their way around utilising search engines (like Yahoo at www.yahoo.com and AltaVista at www.altavista.com), which are massive web directories. Another option is the sites that interrogate several search engines at once, such as Ask Jeeves (www.askjeeves.com) or Metacrawler (www.metacrawler.com). Apart from, through Internet, the Web offers a range of

[3] www.w3.org/People/Berners-Lee
[4] www.cern.ch

asynchronous (like e-mail) and synchronous (like teleconferencing) communication channels that are based on a global system of electronic mail (e-mail) and domain name addresses. In addition, at the core of the idea of the Web was collaboration. Berners-Lee wanted Web users to be involved in a two-way process, reading web pages but also adding to and amending them, creating links, and, of course, creating new pages. The simplicity of HTML computer language and the increasing development of easy-to-use Web design software (e.g. Microsoft FrontPage) permit to everyone interested to create his/her own web site, which include links, graphics, and multimedia components. Indeed, the Web performs a tremendous growth and according to latest estimates the number of existing web sites in July 2000, is over 17,000,000.

FigureI.2. The growth of the World Wide Web sites
(Source: Zakon, 2000: http://info.soc.org/guest/zakon/Internet/History/HIT.html)



DATE	SITES	DATE	SITES
06/93	130	12/94	10,022
09/93	204	06/95	23,500
10/93	228	01/96	100,000
12/93	623	06/96	252,000
06/94	2,738	07/96	299,403

The design of a web site is the design of a dynamic and interactive user interface that requires very different skills from the design of traditional one-way media like brochures and even television commercials. Thus, 'web design' has been developed to ensure that a user visiting a web site is able to navigate around easily, including going backwards and forwards between images, spaces, sounds and videos, as well as being able to retrieve, display, save and print the information they are interested in. Basic design considerations, such as use of interlaced gif images, continuity and identity between pages, and so on, should be seen as helpful hints to facilitate the rapid delivery of information in a structured way, but not so as to constrain the creation of innovative page layouts and the use of attractive graphics (Green, 1997). Content usability raises a whole new range of issues, and Nielsen (2000), points out the need for a new interface paradigm that is likely to be a human-centred and content-oriented design rather than the function-oriented windowing systems in common use today. But human-oriented design requires a very different approach to design than is commonly done. It requires paying attention to the needs of end-users and using proper spatial metaphors to give a spatial, and thus a tangible, meaning to users when navigate into the Web. It requires a multidisciplinary approach, where so scientists work together with engineers, architects, designers and computer scientists to create a more friendly and easy-to-navigate, Web (Benedict, 1993; Bohrer, 1997; Nielsen Norman Group, 2000). On the other hand, congestion or 'on-line queuing' is one of the main problems that Internet and the Web face and that are why people often refer to 'World Wide Wait'. It is widely believed that queuing by congestion is how the Internet and the Web is run right now, and that will not change until usage sensitive pricing is introduced, since demand, driven by flat rate pricing, is insatiable (Odlyzko, 2000: 4).

2. Domain names and URLs

The communication over the Internet and the Web is based on a global system of e-mail and domain name addresses. While Internet hosts recognise each other by their Internet Protocol (IP) numbers, fortunately each IP number has a corresponding name – the domain name. Domain names, therefore, are the basis for network communication between machines, and e-mail addresses are the basis for network communication between people. Both are usually ending in one of following global top-level domains:

- .edu – Education
- .com – Commercial firms
- .net- Computer networks (usually non-profit)
- .mil – Military (US)
- .org – Non-commercial organisations
- .gov – Governmental agencies

In most of European countries the domain names are slightly different as the above codes are usually accompanied by the national code at the end (i.e. *.uk*, for Great Britain and *.jp* for Japan). The domains names in the United Kingdom are coded as follows:

- ac.uk – Higher Education
- sch.uk – Schools
- co.uk – Commercial
- .org.uk – non-commercial organisations like charities, political parties, community groups etc
- .gov.uk – Government departments (except Defence) and organisations, local governments
- mod.uk – Ministry of Defence
- nhs.uk – National Health Service organisations

According to Network Wizards[5], the top ten host names are: *www*, *host*, *mail*, *router*, *ns*, *ftp*, *gw*, *user* and *server*. The ten top-level domain names are *.com* (commercial), *.net* (networks), *.edu* (educational), *.mil* (military), *.jp* (Japan), *.us* (United States), *.uk* (United Kingdom), *.de* (Germany) *.ca* (Canada) and *.au* (Australia). To make it clearer, we must illustrate how exactly an Internet address is built up hierarchically, examining for instance the fs1.ar.man.ac.uk Internet address. The leftmost code (*.fs1*) indicates the name of the host machine or local server; the next one (*.ar*) indicates the department in which this machine is located; the *.man* indicates the organisation or institution, in this case University of Manchester; the *.ac* stands for Academic Community and the rightmost code (*.uk*) stands for the United Kingdom.

The basis of Web's ability to link documents anywhere on the network is the Uniform Resource Locator (URL), which is a standard way of specifying networked information, similarly to the Internet Domain Name functionality. Let us take the following URL as an example:

`http://www.leicester.gov.uk/reports/economics99.html`

The first element in the URL, in this case *http*, specifies the access method to be used in reading the resource. This, in effect, tells to WWW browser how to communicate with the server that has the information. Most often WWW browsers communicate with other WWW servers, in with case, the language (or protocol) used is *Hypertext Transfer Protocol*, or *http*. Anyway, this protocol information is followed by a colon and after the colon comes the description of where to find the resource. The // code indicates that what follows is the domain

[5] www.nw.com

name of an Internet host, in our case www.leicester.gov.uk that according to how we previously analysed the domain name hierarchy, stands for British governmental agency – obviously for the Leicester City Council. Finally following the host name is the path of the resource – the file or Web page – *economics99.html* is in the directory *reports*.

3. Networking-Cyberspatial technologies

In the networking world, people commonly make a distinction between the Internet, an intranet and an extranet. The *Internet* is a worldwide *network of networks* whose owners agree to implement common standards and protocols and which provides services such as electronic mail, file transfer, access to information (most commonly using the standards and access software which make up the World Wide Web) and other facilities. An *intranet* is a single network, owned by a single organisation, which uses the same standards and applications software as the Internet but which is separated from the rest of the Internet, usually by a 'firewall' – a combination of hardware and software that keeps the internal network private. An *extranet* is created when two or more organisations agree to allow each other (some of) the privileges of being on each other's intranet, usually be careful reconfiguration of each organisation's firewall to allow members of the other organisation to be treated as if they were 'insiders'. An extranet, then, is something more than an Internet exchange point, where networks exchange incoming and outgoing Internet traffic, because it requires an extra degree of trust between the extranet partners. How much trust is needed depends on the kind of access which is made available to the other organisation: this might range from an agreement to allow one organisation's traffic to traverse the other organisation's network on the way to a third party (which has security implications relating to areas such as junk e-mail and hacking) to an agreement to give the other organisation access to parts of a private, internal database.

Actually the complex structure of the Internet that is considered as the most powerful characteristic of it can simultaneously be as its weak and fragile point. It is estimated that in usual conditions 3% of the Internet's nodes (access points, web sites, Internet portals) are out of order because of network maintenance, or because of an unforeseen technical problem. Nevertheless, the millions of Internet users do not notice significant problems in their navigation because the traffic of data is continuing through other paths nodes and networks. Internet resembles air transport network. There are many small/ regional airports which are connected to some bigger/ national, and which by turn are connected to the really big international airports. In case that, for any reason, there is a sudden problem in the operation of the JFK airport at New York the result would be a unprecedented disorder that would influence considerable part of the international air traffic. Correspondingly, in case that a central Internet' node that is responsible for transferring big amount of data would be out of function for an unexpected problem (sabotage by hackers/ activists or cyber-terrorists) the Internet would be affected as a whole.

Through the over two decades of Internet activity, we have seen a steady evolution of *organizational structures* designed to support and facilitate an ever-increasing community working collaboratively on Internet issues. In the late 1970's, recognizing that the growth of the Internet was accompanied by an increased need for coordination mechanisms, Vint Cerf, then manager of the Internet Program at DARPA, formed several coordination bodies: An International Cooperation Board (ICB), chaired by Peter Kirstein of UCL, to coordinate activities with some cooperating European countries centred on Packet Satellite research; an Internet Research Group which was an inclusive group providing an environment for general exchange of information; and an Internet Configuration Control Board (ICCB) to assist Cerf in managing the burgeoning Internet activity. In 1983, the ICCB was disbanded and in its place a structure of Task Forces was formed, each focused on a particular area of the technology (e.g. routers, end-to-end protocols, etc.). The Internet Activities Board (IAB) was formed from the chairs of the Task Forces. The growth in the commercial sector brought with it increased

concern regarding the standards process itself. This coupled with a recognized need for community support of the Internet eventually led to the formation of the Internet Society in 1991, under the auspices of Kahn's Corporation for National Research Initiatives (CNRI) and the leadership of Cerf, then with CNRI. Besides, in 1992, the Internet Activities Board was re-organized and re-named the Internet Architecture Board operating under the auspices of the Internet Society. A new coordination organization was formed, the World Wide Web Consortium (W3C). Initially led from MIT's Laboratory for Computer Science by Tim Berners-Lee (the inventor of the WWW) and Al Vezza, W3C has taken on the responsibility for evolving the various protocols and standards associated with the Web.

Regarding important features/technologies in terms of modelling the built environment, we witness the fast development and widespread use of *Geographic Information Systems* (GIS) and other geographical and architectural software such as the ARC-INFO/VIEW, AutoCAD and ArchiCAD. Eventually, the integration between AutoCAD models, maps created with GIS, videos and photographs, and VRML/QuickTime programming is the latest advancement in representing interactive 3D spaces over the Web[6]. VRML/QuickTime or QTVR is a software with the ability to place VRML objects within a photographic rendering scene. The use of rendered photograph eliminates the need, when using VRML, to extensively texture map to achieve a high degree of realism. Very important is also the Shared Virtual Reality (SVR) system, which differs from Virtual reality in that the experience of virtual places is no longer individual, but rather shared across the Internet with other users simultaneously. Thus, for planners, architects and public decision-makers SVR offers a new way of communication their ideas and taking into account ideas from the general public (Brusasco, Caneparo and Zorgno, 1998). The technological integration of these technologies over the Web, has been summarised in terms such as ‘spatial technologies’ (Couclelis, 1994) and ‘**cyberspatial technologies**’ (Kitchin, 1998), defining a subset of technologies that provides both spatial and cyberspatial information, and may contribute to the modification of spatial relations by affecting people’s ways of interaction and mobility.

[6] More information about, and good examples of, such kind of 3-D virtual spaces can be found at the web site of the Centre for Advanced Spatial Analysis (CASA- UCL, dir: Prof. M. Batty) www.casa.ucl.ac.uk.