Internet in a Nutshell

J Indira Subbaram*

ABSTRACT

The article provides necessary information about Internet, its origin, growth and becoming an important information resource all over the world. The hardware and software requirements, agencies providing Internet connection in India, their charges and the commands for connecting to Internet are given. The common terms used with Internet are explained for a beginner and even some Indian service providers on Internet discussed.

1. INTRODUCTION

Internet, is a popular buzzword, more so in the present decade. If you are not on Net, you are history. We are in the age of electronic communication where having an e-mail address on the business card is becoming a necessity like telephone and fax numbers. And for corporate entities, the letterhead is not complete without a WWW address like: http://www.mycompany.com/

On the Internet, one can communicate with around 50 million users worldwide, a population which is doubling each year. Researchers from all corners of the earth are finding that their work thrives in a networked environment. Immediate access to the work of colleagues and a virtual library of millions of volumes and thousands of papers afford them the ability to incorporate a huge amount of knowledge hitherto unthinkable. People from all walks of life in different countries use the Internet for exchanging messages, carry on real time chatting, play games, sell goods and

services, advertise products, download software, search for useful and usable information and do many more things. Work Groups can now conduct interactive conferences with each other, regardless of physical location—the possibilities are endless.

1.1 What is Internet?

The Internet is a network of networks spread worldwide. It is a conglomeration of smaller networks and other connected machines spanning the entire globe. Each country has at least one backbone network that operates at a very high speed and carries bulk of the traffic. Other networks connect to that backbone.

1.2 Background

The Internet was born in 1969, out of an effort to connect together a US Defence Department network called ARPANET (Advanced Research Projects Agency Network), linking four communication hosts IMPS (Interface Message Processors), in such a fashion which would ensure that even if one host fails or gets bombed out, communication would route around the unaffected area and stay alive. In 1974, ARPANET allowed international bodies to link to this network by releasing TCP/IP. ARPANET finally retired in

* BHEL Corporate R&D
Vikasasnagar
Hyderabad - 500 093
e-mail: bhelrnd@shakti.ncst.ernet.in

1990, supplanted by National Science Foundation (NSF). In the same year, NSF created the Advanced Network Series (ANS). This backbone became the NSFNET with more than 5 million hosts, covering 50 percent of the Internet outside USA.

No one person or group runs the Internet. Voluntary organisations like ISOC (Internet Society), IAB (Internet Architecture Board) and IETF (Internet Engineering Task Force) hold online discussions to create standards for the Internet connectivity.

Similarly, no one pays for it but at the same time it is not free. Some service providers on the Internet are non-government funded agencies, like UUNET, PSI, etc. These *commercial operators have interconnected their* backbones to form CIX (Commercial Internet Exchange).

1.3 How it Works?

The key to the Internet treasure is the transmission protocol for networks. The TCP/IP is actually a set of protocols developed to allow cooperating computers to share resources across a network. Today TCP/IP is the most successful mechanism for networking of computers worldwide. TCP (Transmission Control Protocol) is responsible for breaking up the message into datagrams, re-assembling them at the other end, re-sending anything that gets lost and putting things back in the right order. IP (Internet Protocol) is responsible for routing (sending) individual datagrams. Every IP datagram contains the source address and destination address so that each datagram can be delivered independently. It is TCP module that contains the logic necessary to provide a reliable and virtual circuit. By employing TCP/IP, one can contact computers of different platforms. It is immaterial whether the computers and networks connected work on DOS or UNIX platform and where it is located.

2. TOOLS AND UTILITIES

Everybody talks about what a great source of information the Internet is, but no one ever mentions how hard it can be to find what one is looking for.

The Internet contains a wealth of information—if only one could find it. There are people who often talk about vast amount of information they gathered from the Internet. Yet there are people who comment that Internet is used only for e-mail services, groping in the dark and wasting time, and in the process one is frustrated and disappointed. All one needs is patience and understanding to develop the skill and technique, and once one has acquired the skill, gold can be mined on the Internet. Its limitation depends on the user's ability to explore its resources. The succeeding sections describe various services available on the Net.

2.1 Remote Login (Telnet)

Telnet is the main Internet protocol for login connection with a remote machine. Telnet is a utility which will allow you to log on to another system and gives the user an opportunity to be on one computer system and work on another which may be thousands of miles away.

Many Universities like Colorado Alliance of Research Libraries (CARL), Boston University, Stanford University and public libraries, like Library of Congress (LC), have made their library catalogues available for the Internet user. CARL offers a number of library databases, including searches of government periodicals, book reviews, indices for current articles and access to other library databases.

One of the interesting databases is Science and Technology Information System (STIS). For searching the database, telnet to stis.nst.gov and login as 'public'. The STIS is maintained by NSF and provides access to many NSF applications. The full text publication can be searched online and copied.

2.2 Moving Files (Ftp)

File Transfer Protocol (Ftp) is the primary method of transferring files over the Internet. Any useful software or data found on any of the sites can be downloaded by using simple commands. It normally requires user ID. There is a good way around this restriction—the 'anonymous' ftp service. It essentially lets anyone in the world have access to a system which dedicates some portion of the disk or entire disk or even one computer by

maintaining archive of source code. There are two ways to get connected to a system either by using its host name or its Internet number. Using the hostname is preferred.

E.g., ftp.uu.net

After establishing the connection to 'ftp.uu.net' you should respond with 'anonymous', for the prompt USERID. When the system prompts you for a password, respond with your e-mail address.

2.3 Finding Files (Archie)

The biggest problem is to find what exists on Internet. Some students of McGill University created a database system called Archie. There are servers all over the world and one can ask Archie to find a file via Telnet, or by e-mail.

telnet archie.au

and login archie

or send e-mail archie@archie.rutgers.edu

Archie helps you locate a file if you know its name from any of the thousands of anonymous ftp sites.

2.4 Electronic Mail (E-mail)

E-mail facilitates exchange of messages across computers. E-mail combines the advantages of writing with the immediacy of the telephone. E-mail makes the memo cycle immediate—no more time is wasted in dictating, typing, printing, copying and distributing.

One can use Internet e-mail, besides sending personal messages to people overseas, to request software, files and information from a number of sources. Full range of Internet services can be harnessed by using e-mail facility in the following way:

Ftp by using e-mail to ftpmail@sunsite. unc.edu

Message : 'open ftp.microsoft.com get/ index.txt'

ARCHIE by using e-mail to archie@archie.unl.edn

Many organisations who have only e-mail use this facility for file transfers for getting softwares.

2.5 Mailing Lists (Listserv)

A mailing list is maintained on a Listserv—a special computer hosting such facility. There are thousands of topics on which mailing lists exist. Subscribers to a particular list can 'post' messages to all other subscribers by simply sending one message to the Listserv address. The message will automatically be forwarded to all the members. Areas of interest are wide and varied. On BITNET there is an automated system for maintaining discussion lists called Listserv.

For example, you can subscribe to: Library and Information Science Forum of IISC, Bangalore, India by sending e-mail to:

listserv@ncsi.iisc.ernet.in

with the message: subscribe LIS-FORUM <your name>

for cancellation send the message : Unsubscribe <your name>

If you subscribe to many mailing lists at a time, your mail box can be the biggest mess and it is worse when your system is down for a couple of days, or if you go on vacation.

2.6 Network News (Usenet)

Usenet is a set of machines that exchange articles stacked with one or more universally recognised tasks called Newsgroups. It is like a meeting place where people gather to meet their friends and keep abreast of the latest trends in their field of interest or express whatever is on their mind. The transmission of Usenet News is entirely cooperative. Feeds are generally provided out of goodwill and the news is distributed everywhere. The newsgroups are organised in several hierarchies.

Comp: Topics of interest to both computer professionals and hobbyists.

Misc: Subject groups include fitness, job-hunting, law, and investment.

Sci: Discussions relating to research in sciences.

News: Groups concerned with the news, network, group maintenance and software.

K12: Educational topics.

2.7 Bulletin Board Systems (BBS)

There are thousands of bulletin boards on various subjects. Normally they are of local interest to some geographical area. Subscribers reach bulletin boards via telephone lines. Nowadays several BBSs are reachable via Internet. Hence our interest.

BBS is similar to Newsgroups. They contain programmes, documents, graphics, sound or video clips—if it can be put in a computer file it can be attached to a file. Most of what is found on a BBS can be downloaded. Members send (upload) information files to BBS also. Exchange of 'personal mail' is also another feature of BBS.

Examples of BBS on the Internet would include: FAO, NTIS, DDA, DOE, etc. To illustrate how useful BBSes are, let us see one example:

'The North American Reliability Council (NERC) is a non-profit corporation owned by regional councils. NERC has a home page on the Web. Its publications developed by electric utility experts have become standards for the industry. Their publications available on BBS can be downloaded free of charge. Their publications on GADS (Generating Availability Report), boiler tube failure trends, predicting unit availability report, etc., would never have been available without the Internet facility even if we are prepared to pay. NERC provides large volumes of power plant outage data along with software for analysing the data to get meaningful reports. This software has provision to add our own internal data also and generate useful reports ourselves using standardised formats for input of data. This service is extremely useful'.

2.8 Finding Someone (Whois)

There is no single unified Internet directory. The main 'whois' database is maintained by the Network Information Centre (NIC). The 'whois' command will let you search a database of every registered domain. It is primarily used by postmasters of Web sites.

2.9 Finger

The finger command gives information about each user who is currently logged in with reference to a specific machine on the Internet. If you want specific information about a person, finger command has the capability to give more information about a user on Internet host. This helps to find complete e-mail address or the telephone number of the person.

e.g., finger sanjay@ernet.in

To see who is currently logged in at NCST you can use

%finger@ncst.ernet.in

2.10 Tunnelling through the Net (Gopher)

Gopher represents the first generation of attempts to replace the 'where' with the 'what' in cyberspace. It is a way of accessing what you want without having to know where it is. This is like browsing the remote library card catalogue and automatically getting the wanted material. When we want something we like, we can read and access it through the gopher without knowing domain names, address, changing paradigms etc.

Although each gopher site uses its own organisation for materials in its domain, Gopher menus are like a directory listing the information sources. In case the user does not find everything desired in one gopher 'library', the gopher allows the user to pass electronically to the next gopher 'library' and check there.

2.11 World Wide Web (WWW)

In the recent past, the World Wide Web or WWW has completely reshaped the Internet. It is allowing the users to experience sights and sounds in an innovative style of navigation by introducing graphical user interfaces to facilitate access to the Internet. The tools for locating relevant information mentioned earlier like gopher, telnet, ftp, etc., are being merged and are becoming accessible through a single common interface, with the overwhelming popularity of World Wide Web.

There are three main concepts that characterise the WWW:

1. URL (Uniform Resource Locator) specifies the protocol and locations for information retrieval.

2. The Hyper Text Transfer Protocol (HTTP) specifies how hypertext documents are transferred through the Internet on the WWW.

3. The Hypertext Markup Language (HTML) contains rules for document publishing on the WWW. HTML contains the text and structure of the document, pictures and links to other Internet sites and locations.

A 'Home Page' is the first page or the introduction page of a set of related information put up on the Net. It is easy to set up a home page using several tools available for this purpose. The home page acts as powerful advertisement for your product/service. The number of home pages on the WWW was around 25 million in 1995 and it is likely to be 200 million by 1999!.

One can browse the Web in character mode also. Although, much of the charm of the graphic capabilities of the WWW are lost in this case. Cursor can be used to browse up and down and press Enter on any file (a link, indicated by underlined portion or prefixed by a number) of interest. Information currently accessible through telnet, gopher. WAIS, ftp, usenet, archie, etc., can be served via Web with wide variety of information databases associated with these existing Internet resources without bothering to know its URL number.

While gopher allows a line of text information to serve as a link to another gopher site or a document, the WWW allows a page to be displayed with multiple fonts, graphics and links to movies and sounds. These links provide access to other pages of information, display movies or play sounds that are referenced on the page. To see the graphical contents of a web page you need browser software like Netscape, Mosaic, Cello or Internet Explorer.

3. COMMERCIAL DATABASE VENDORS

A large number of online database vendors nowadays permit connectivity via Internet. Telecommunication charges are cheapest via Internet. Another advantage is that the search output can be sent directly to the user's e-mail account at cheaper rates. However, the connect time and document charges also have to be reckoned with. All the commercial services require prior arrangement for becoming a valid user by obtaining a USERID and a password from the vendor concerned. Some examples are given here.

3.1 Knight-Ridder Info Service

KR Information Service Provider is the largest commercial database vendor. Most of the databases are full text, including many major newspapers, patents, etc. Accessing via Internet cuts the communication charges to $4.20 per hour. It can be accessed by the command:

telnet dialog.com

or http:/www.krinfo.com/

3.2 BRS

Bibliographic Research Service (BRS) acquired by CD Plus Technologies offers access primarily to databases of medical and scientific collections. Access via Internet charges are $5 an hour, versus $13 via X.25 networks.

3.3 Data-star

Data-star is online information service with a heavy emphasis on European trade and business data. It is now owned by Knight-Ridder. It can be accessed by the command:

telnet rserve.rs.ch

Access charges via Internet are $3 an hour, as against $11 via X.25 networks.

3.4 STN

STN is a database available online to the scientists, particularly chemists, and engineers. It is a joint effort of several scientific organisations including Chemical Abstract Service of American Chemical Society. It can be accessed by the command:

telnet stanc.ca.org

Access via Internet is $5 an hour, whereas through X.25 network it is $9 to $10 an hour.

4. WHAT IS NEEDED?

In order to get connected to Internet, one needs a PC, a phone line and a modem on the hardware side. One also need a subscription for Internet connection from an Internet service provider. Some special software is also required. We shall discuss these now.

4.1 Hardware Needed

While any PC will do to get hooked on to Internet, in practice, a good high speed 486 machine with 4 to 8 MB RAM, 16 color VGA display and a modem of 14.4 Kbps with error correction and compression are necessary. To print pictures after downloading, a good laser printer is also needed.

4.2 Software

As far as software is concerned, besides TCP/IP stacks, programs for terminal connectivity like Procomm, browsers like Netscape, Lynx, etc. are also needed. Recently introduced Windows 95 Operating System comes as an Internet OS as it has inbuilt software for all Internet needs.

The operating systems provide complete information highway access solution and navigation is possible through the Internet with easy-to-use graphical interfaces. It gives an automatic dial-up connection to Internet services or access via any SLIP enabled provider. It includes web explorer, gopher, telnet, ftp, newsreader, e-mail and TCP/IP dial capabilities. It will enable the user to travel in the information highway fast lane, with high speed modem support.

4.2.1 Netscape

Netscape is a software browser that helps users to navigate the World Wide Web. Installing and configuring is easy compared with other Web browsers. Netscape's capability includes display of images with the text on pages as you download them, whereas other Web browsers require downloading of all the images on a page before you start reading it and if you see a link you want to follow, you can click on it and connect to the next page. With Netscape you just select, copy and paste. The evaluation version is free and can be downloaded from Netscape's web site. Nescape enjoys an 80 percent share of the Web browser market.

4.3 Internet Access Providers

In India, access to Internet for libraries is possible through ERNET, VSNL and NICNET (RENNIC).

4.3.1 ERNET

The Educational and Research Network (ERNET) implemented by Department of Electronics (DOE) with UNDP funding initially, has over 400 organisations connected within India and neighbouring countries. These institutions are mainly academic and research organisations, non-government organisations and to a limited extent private and international organisations. ERNET is supported by several backbone sites, covering DOE (New Delhi), IIT (Madras), IISc (Bangalore), IMTECH (Chandigarh), VECC (Calcutta), IUCCA (Pune), NCST (Bombay) and Central University of Hyderabad, which enable organisations located at different geographical locations to access various services of Internet. One can opt to any one of the nearest backbone sites to get the connectivity. There are three categories of connectivity which are primarily based on the range of tariff from Rs.5,000 to Rs.5 lakhs per annum.

4.3.2 VSNL

The Videsh Sanchar Nigam Ltd. (VSNL) offers full range of utilities in three options for accessing Gateway Internet Access Service (GIAS).

(a) Shell Account - using dial-up access to GIAS Host.

(b) TCP/IP Account - VSNL provides direct TCP/IP connection on dial-up or leased basis to GIAS system and access through online, all services of the Internet.

(c) Packet Switched Network Access - VSNL provides access via GPSS using I-Net, RABMN and HVNet.

GIAS tariff ranges from Rs.500 (only for students), Rs.5000 (with quota of 250 hours) to

Rs.25,00,000 depending upon the option of services.

4.3.3 NICNET (RENNIC)

The National Informatics Centre Network (NICNET), was started about ten years ago, supported by the Planning Commission, and connects 700 earth stations, with links to all district headquarters and a number of public sector undertakings. NICNET has become an important network facilitating decision making in the government.

NICNET offers electronic mail facilities only to research and educational networks (RENNIC) with a nominal annual subscription of Rs.2700 plus Rs.1000 for registration. NICNET also offers services for the public (tariff is not yet advertised). The highlights of their services are their Management Information Services for Rural Development Schemes, Medical Literature and Analysis and Retrieval Systems (MEDLARS) and National Library of Medicine (NLM). Information for patient care is being provided within an hour of receiving the request, it is claimed.

4.3.4 Softlink Services

The Software Technology Park (STP) is another scheme of DOE, dealing with export of computer software using data communication links or in the form of physical media including export of professional services. Speeds upto 64 kbps and 128 kbps are offered. This is suitable for organisations having high bandwidth for implementing various networking protocols.

Their subscription rate is Rs. 5 lakhs (including 1 lakh towards deposit), irrespective of number of hours of usage and volume of data. This scheme does not fall in the line of library scope of activity.

The subscribers of any of the above service providers may opt for full range of Internet utilities via I-Net, RABMN and HVnet, either by using dial-up line or leased line.

4.3.5 Private Internet E-mail Providers

Besides the above governmental agencies, there are a number of private e-mail service providers in the country who help reach Internet e-mail addressees. BI Infotech, DART, Datapro, UUNET, CMC, Sprint/RPG and ICNET are some of such companies. Libraries can get low cost Internet e-mail connection for fees ranging from about Rs. 3000 per year to about 15000 per year.

5. RECENT TRENDS

There were many technical developments, new sites and services on the Internet during 1994-96. Some areas are worth special mention:

5.1 Development of the WWW

This hypermedia system, originally desired to provide access to scientific documents is being regarded as almost synonymous with Internet itself. Its graphical interface, browsing style of access, support for multimedia have influenced the Internet life style.

5.2 Increased Interest in 'Subject Access' of Internet Resources

The terms of new systems in addition to the existing archie, gopher, veronica, various other browser systems have appeared, complementing the browsing and navigating approach.

Let us look at an example:

Electric Power Research Institute (EPRI) home page gives details such as :

EPRI Products
EPRI Publications
EPRI Projects
EPRI Technical Expertise Directory

If you click at 'EPRI Technical Expertise Directory', you will find 224 experts. The details of the experts can be selected by a set of keywords.

If a research project is on 'Polycrystalline Silicon Solar Cells', the information relating to the expert in the field is given under.

e.g., Frank Goodman
e-mail fgoodman@eprinet.epri.com
His areas of expertise -
Crystalline - PV cells
Photovoltaic testing and applications
PV cells, crystalline - material
Silicon Photovoltaic Cell Technology

From the data of his expertise, anyone in a similar field of interest can establish contact via e-mail for exchange of developments in the subject.

5.3 Commercial Interest on the Net

Indications pointing towards business-like environment in the Internet are predominating, incorporating a greater degree of controls and standardisation, security facilitating commercial transactions, home shopping and home banking.

5.4 Web Search Tools

Perhaps the greatest limitation of current distributed hypermedia systems is that there is no single way to find out what information has changed, what information is new or even what information is out there on the Internet. Unless the user has extensive knowledge of the Internet, it is hard to locate resources of interest on WWW.

A good number of search tools to search the soaring millions of Web documents are now coming up. There are about 15 tools now. Webcrawler, Metacreawler, All-in-one, Netseek, Excite, Lycos and Yahoo are some of the famous ones. These search tools facilitate searching web documents using keywords, phrases, boolean operators, exact matches, concepts, free natural language query, etc. Excite, for example, searches 16 million records in less than a minute. Learning to use each of these tools effectively is becoming an expert's job now. Hence the invasion of new breed of 'Internet Librarian' or 'Electronic Librarian'!.

6. CONCLUSION

The Internet now provides access to literally thousands of information sources on topics ranging from economic indicators to government regulations. Every industry has virtually reached a new level of intense activity. The pressures include: Shrinking product lifecycles, increasing production costs, demand for quality and customer service. The business world is in an age of hyper competition.

The Internet has brought new revolution and using this facility has become necessary for their survival in this highly competitive world. The success of a company depends upon the effective use of information technology. The librarian and information professional has a vital role to play in bridging the information gap. The potential of Internet based information sources cannot be overemphasised in this context.

7. ACKNOWLEDGEMENT

FURTHER READINGS

1. Prakash, Vipul Ved. The Internet protocols. *PC Quest*, February 1996, p29-130.

2. Agarwal, Surajit. Beyond Windows. *Dataquest*, 16-31 March 1995, p8-128.

3. Prakash, Vipul Ved. Mailing list with Net. *PC Quest*, February 1996, p6-48.

4. Chitnis, Atul. The Internet is here. *PC Quest*, September 1995.

5. Gulati, Ashish. Netsurfing regardless. *PC Quest*, August 1995. p115-125

6. Net news. *Personal Computer World*. January 1995. p504.

7. *Informatics* (Quarterly Newsletter from NIC). October 1994 & January 1995.

8. Sukhdev Singh. Internet and its resources for the health Information Scientists. CALIBER-95, Hyderabad National convention of INFLIBNET.

9. Exploring the Internet Seminar. STP, Hyderabad, 1995.

10. Cheong, Fab-Chen. Internet agents: Spiders, Wanderers,Brokers and Bots. New Riders Publishing. 1996.

11. Course material on Internet. Satyam Computers. Secundrabad.