

ISU Guide to Internet Services

Eric Behr, Mathematics
Illinois State University
December 1990
Revised September 1992

Due to the growing interest in the Internet, we have prepared a document which lists basic Internet functions available to our faculty and staff.

A lot of important details have been omitted in order to increase readability. If you are already an expert, you will most likely chuckle in quite a few places; if you aspire to become one, please peruse the recommended literature.

Introduction

Internet evolved from a Department of Defense Advanced Projects Agency network, ARPAnet. It is a major world-wide computer network linking academic, commercial, governmental and private computers. It uses relatively sophisticated transmission methods, which gives it an edge over other similar networks.

One of the reasons behind Internet's functionality is the fact that it uses a uniform *communications protocol*, IP (Internet Protocol), as a foundation on which other data transfer applications are based. Any machine which is physically connected to an Internet-related network *and* understands the IP protocol is able to take advantage of Internet services, some of which are described below. A basic set of Internet communications protocols built on top of IP is referred to as TCP/IP. We will use this acronym as a catch-all term referring to Internet compatibility.

Making the connection

First of all, every machine which attempts to use the Internet has to understand TCP/IP. There are several systems in which TCP/IP comes as a standard (all Unix-like systems, including DEC Ultrix, IBM A/IX, Apple A/UX, Sun OS). Other computers have to run a special software gadget. You want to make sure that TCP/IP is in one form or another a part of your system software. For example, a Macintosh will be TCP/IP compatible when it has a MacTCP CDEV in its System Folder. Some IBM PC Internet applications, like WiscWare, come with a TCP/IP driver built-in.

Second, the machine you want to use needs to be physically connected to a network which has a *gateway* to the Internet. This usually means a connection to a major nearby Internet site.

Internet basics

At present, the Internet consists of over 1,000,000 registered hosts; since not all hosts are registered with Internet authorities, the number of individual computers with Internet access is much higher, and is very difficult to estimate. This vast network is connected to similar networks

in Europe, Asia, and Australia. These diverse computers can successfully communicate thanks to two unique Internet features: the standardized set of IP-based

communications protocols, and an addressing scheme which allows one computer to send data to another one in a baton-relay fashion.

Every computer connected to the Internet has its unique *internet address*, which consists of four numbers separated by periods, e.g. 138.87.1.2. The first two numbers denote one of the many mid-sized networks, in this case the Illinois State University campus network. The third one represents an administrative entity within the network; in this example 1 is reserved for Computer Services. Finally, the 2 stands for a specific machine – the IBM RS/6000 computer in Julian Hall.

Fortunately for us, humans, the network is smart enough to let us remember such an address using more natural mnemonics. By convention (not necessity!) those names usually correspond to the administrative network structure. The so-called *name server* computers on Internet have been programmed to recognize the name “ilstu.edu” and substitute the network number “138.87” for it (Illinois State University is an educational institution which has been allocated a unique Internet name “ilstu”). Similarly, computers around us know that “cmp” is the name used within ISU to denote Computer Services machines, and finally “rs6000” is the name which we assigned to the RS/6000. An outside user can now type “rs6000.cmp.ilstu.edu”, and Internet machines will be able to translate that to “138.87.1.2” and route the message accordingly. Note that the separate chunks of the address are listed in reverse order in the human-readable form.

Explosion of Internet popularity in recent years has led to an “address exhaustion” problem – we may soon have to change the underlying mechanism used to address Internet hosts.

Finally, there is a question of addressing messages to a specific user. This is done by means of another standard convention: an individual is known to the Internet by his designation “user@machine”. For example, my login name on the RS/6000 is “ejbehr”, so that my full address is “ejbehr@rs6000.cmp.ilstu.edu”. In some cases a partial address is enough; for example, if another user wants to reach me from the RS/6000 computer, he only needs to say “ejbehr”; if a user on the Music Department network wants to send me a mail message, he ought to be able to use just “ejbehr@rs6000.cmp”.

Reaching users who are not on the Internet may be much more difficult, since other networks usually follow their own addressing schemes. Addressing a letter to someone on a small private network in Malaysia can resemble black magic, but this problem is diminishing as more networks adopt a uniform convention, and routing machines become sophisticated enough to automatically recognize exceptions.

In our case, the main non-Internet network with which we want to communicate is Bitnet. Sending a message to a user on Bitnet is very simple: all you need to do is specify “user@node.bitnet”, e.g. “EBEHR@ECNCDC.bitnet” (now defunct...) The message is passed on to a designated gateway (a computer at UIUC), and then delivered to BITNET automatically. By the way, in 99% of e-mail addresses capitalization is irrelevant, so “ebehr@ECNcdc.BITneT” should be just as good.

Internet services

There are several interesting features of Internet communications; we will concentrate on just a few most important ones. In computerese, they are known as: ftp, telnet, tn3270, smtp and nntp (sounds discouraging... but read on!) You will sooner or later find out about other gadgets: gopher, archie, fsp, POP, and so on.

ftp stands for “file transfer protocol”. As the name implies, it provides a (relatively easy) way of exchanging files between various Internet computers. There are two main flavors of ftp: text (ASCII), and binary. Using one or the other two computers can transfer files at speeds in the range of tens of kilobytes per second, meaning that a large 200 KB article can be sent in half a minute or so. This clearly beats using a 2400 baud modem. The binary option allows transfers of machine-specific executable programs.

Normally, FTP asks for the usual user information: what is the login name and password for the account on the remote machine which you wish to contact. Unless you are a dedicated hacker, you need to know this information beforehand. A variation of FTP which is very popular on the Internet works as follows: most Internet managers have set up a special account, called “anonymous”, on their computers. This particular account usually disregards the password, and contains files that the system manager wants to make public. This way, anyone can use ftp to gain access to those files. The list of currently available anonymous FTP sites on the Internet is 12 pages long; exploring all of them in detail would take a few months.

telnet is probably the most useful Internet protocol. It allows a computer to start acting as a terminal to any other Internet machine. In particular, any microcomputer on ISUnet which runs telnet can serve as an access point to the RS/6000. To use telnet to connect to another computer, you have to be granted login privileges on that machine (there are a few exceptions here as well).

Telnet usually emulates the popular VT-100 terminal. It can be thought of as a very high speed communications program, like ProComm or VersaTerm, except designed to use the network rather than serial telephone connections. Some implementations of telnet combine this functionality with ftp, so you don't have to switch between programs very often.

Some sites set up “anonymous” telnet accounts. For example, the SRI Internet data clearinghouse can be accessed using telnet (to nic.ddn.mil), allowing you to search their records, order documentation, etc.

tn3270 is a modified telnet which provides an interface between the IBM and Internet worlds; it allows a local computer to behave as an IBM 3270 terminal (which differs from VT-100 in that it displays information in various fields on the screen, rather than on consecutive lines). Some IBM computers connected to the Internet expect other machines to connect to them using tn3270 rather than telnet, but many accept both.

smtp stands for “simple mail transfer protocol”, which is used by Internet computers to exchange mail messages. Since SMTP is based on the TCP/IP standard, it doesn't matter whether the machine involved is a PC, a Sun workstation, a DEC VAX minicomputer, an IBM, or a CRAY. A message which is properly addressed should get to its destination in seconds. This of course allows programmers to

customize mail programs to the given environment; for instance, a Macintosh can send mail which contains a MacPaint picture to another Mac, even though the rest of the network doesn't "understand" it (more precisely, it "doesn't care" *what* is being

sent). SMTP generally understands only ASCII data, so sending binary information using it requires additional encoding (such as uuencode or BinHex).

nntp is an acronym for “network news transfer protocol”. It gives Internet computers the chance to participate in information exchange on a mind-boggling scale. Imagine a computer bulletin board with hundreds of thousands of users, several hundred topics, and a few dozen messages in each topic a day. This is precisely what “Usenet netnews” is. One of the nice things about it, compared to other BBS systems, is that you have a good chance to read articles posted by real experts, and not only know-it-alls with inflated egos. It takes a lot of work to sift through all the messages to get at the pearls, but it’s well worth it.

Recommended reading

Brendan Kehoe’s “Zen and the Art of the Internet” is still widely available from ftp sites, even though it has been superseded by a commercial edition. It’s a good place to start.

Most Internet documents are available in the form of RFCs (Request For Comments) which are stored primarily at the SRI-NIC site nic.ddn.mil (see below) and are available through anonymous ftp (get rfc:rfcxxx.txt). If you want to dig into details (and I mean *details*) of the protocols mentioned above, and you don’t mind doing some heavy reading, then you should look at these:

RFC-791	Internet Protocol (IP)
RFC-793	Transmission Control Protocol (TCP)
RFC-821	Simple Mail Transfer Protocol (SMTP)
RFC-822	Standard for the Format of ARPA Internet Text Messages
RFC-854	Telnet Protocol
RFC-959	File Transfer Protocol (FTP)
RFC-974	Mail Routing and the Domain System
RFC-1087	Ethics and the Internet

Interesting places to visit on the Internet

Main Internet Information Center run by SRI International. Use “telnet nic.ddn.mil”

Information about the NSF portion of the Internet (but not only). Use anonymous ftp to nis.nsf.net and to nnsc.nsf.net.

University of California Melvyl library catalog. Use “telnet melvyl.ucop.edu”

Database of geographical data about US (and some foreign) cities. Use “telnet martini.eecs.umich.edu 3000”