The Internet

The Electronic Interstate System of the 21st Century

Alek Komarnitsky
komarnit@tramp.colorado.edu
INFS5820, Spring/1991
Dr. Carroll Frenzel
University of Colorado at Boulder

Abstract

Building the US interstate system in the mid-1900s resulted in sharply lower transportation costs of goods, allowed firms to expand from regional to national markets, and contributed to the economic expansion of the country. The Internet will provide a similar impetus to electronic goods, which is especially pertinent as the US becomes more of a services based economy (where information is the actual good). This paper will give an overview of the Internet: a collection of networks that together comprises over one million computers. The Internet has historically been used primarily for educational and research activities, but in the past several years, it has become more and more commercial. Firms are realizing the competitive advantage of nearly instantaneous electronic transfer of data. Although the development of the Internet is most advanced in the US, there are also a number of international links. This paper will also discuss some issues of concern as the Internet grows into a national (and eventually global) electronic network.

The author predicts that by the 21st century, organizations that do not have access to the Internet will be like one today without phone service.

Contents

Introduction

The Internet is a network of computer networks: a loose collection of computers that all have the capability (at least to some degree) to communicate with each other. That level of communication can vary from simple text-based electronic E-mail to multi-media (text, images, and sound) interactive applications. Similar to the US interstate system, there is a nationwide backbone that provides connectivity nationwide. Regional networks provide electronic transportation within a region. There may also be local networks that actually provide the connectivity to individual organizations. Finally, those organizations themselves have their own internal networks.

History of the Internet

The Internet began as the Arpanet, a project started by the Department of Defense's (DOD) Advanced Research Projects Agency (called DARPA today) in 1969. Although this project was an experimental prototype, it became apparent that it could be used to electronically communicate between the geographically separated and different types of computers used by the DOD. The fact that these agencies often used different hardware and/or software can not be understated, and the ability of the Arpanet to tie these diverse sets of computers systems together is a major factor in its acceptance and popularity in its formative years. The network was based on packet-switching, which is a method of breaking a message into smaller units (or packets) for transmission (which were re-assembled at their final destination). This allowed a single transmission line to be used simultaneously by many users. This was revolutionary at the time, since most computers used dedicated circuits to communicate.

In the 1970's a set of procedures and rules were developed to establish a standard means of addressing and routing packets across the network. These non-proprietary protocols, called Internet Protocols (IP) allowed dissimilar computer systems to converse with each other. Appendix B discusses IP addressing, which (fortunately) allowed the network to scale from (then) several dozen computers to thousands today.

In the early 1980's the Arpanet grew: both in the number of sites directly on the Arpanet, and also from connections to other (sometimes independently formed and operated) networks. In 1983, the Department of Defense spun off their traffic into an entity called the Milnet (which was operated and funded by the Defense Communications Agency). Several connections (currently six) were

established between the Milnet and the Internet - other connections are not allowed. By 1983, the Internet was composed of approximately 200 computers on about 50 networks. These were primarily educational and research organizations (one reason for this is that Internet use was restricted to non-profit usage). At the same time, a number of commercial firms realized the advantages of networking and built private networks; Digital Equipment's EASYnet is an example (which still exists today, but with connections to the Internet).

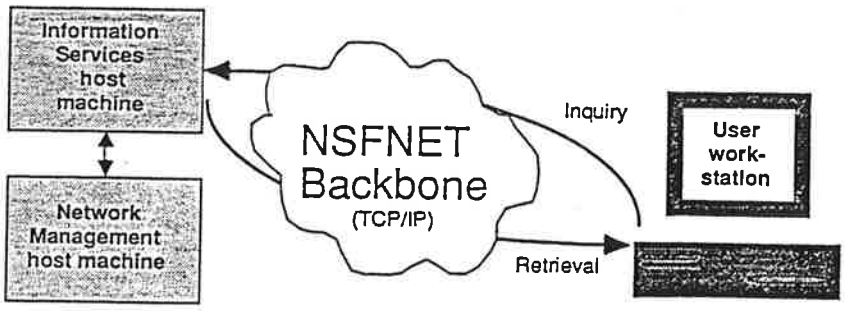
Technology continued to advance, as networking speeds became faster (with 56 Kbps lines) and more reliable. In addition, a set of higher level protocols were standardized to perform certain common functions.

- Telnet allowed one to log into another computer.
- File Transfer Protocol (FTP) was used to transfer files between computer.
- Simple Mail Transfer Protocol (SMTP) was used to send electronic messages back and forth.

Although the above three applications may seem simple, they had to be written so that the same commands could be used on different computers thousands of miles away. Another common application called *news*, an electronic bulletin board, was also gaining in popularity as discussed in Appendix E.

The Arpanet continued to grow, and was rapidly reaching "critical mass" for the scientific community. The present day popularity of the FAX machine can perhaps be similarly credited to the fact that "everyone" has one. Scientists at that time had to travel to the few Supercomputer centers in order to access the most powerful computers. The National Science Foundation, which was funding a number of these supercomputer centers, decided in 1985 to use IP as the standard of communication. NSF built and operated (contracted through Merit corporation) a nationwide backbone network called the NSFNET. In addition, they also encouraged (and funded) regional and campus networks that connected to the backbone. This allowed scientists, from their desktop, to use remote supercomputers thousands of miles away as show in Figure 1.

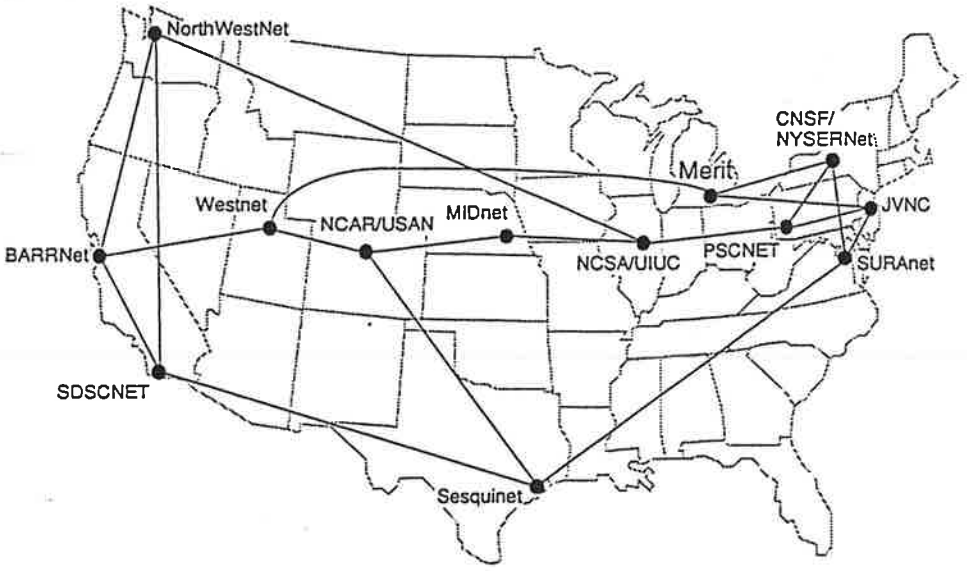
Figure 1: *Schematic of remote access over the NSFNET*



The number of users continued to increase exponentially. In addition, these scientists soon wanted the ability to access their data interactively, rather than downloading it overnight. Finally, more complex applications became available that required transmission of more data in less time. Although connectivity was definitely appreciated at first, it soon became taken for granted. It was obvious that network bandwidth (the amount of information that can be transmitted on the communications medium) was limiting both the number of users and the type of applications that could be used.

Fortunately, technology was improving in the communications field, and T-1 (1.544 Mbps) links became available, which allowed a 26 fold increase in transmission speed versus the 56Kbps. By the end of 1988, NSFNET had 13 networks hubs with 1.544Mbps T-1 links (as shown in Figure 2) which provided nationwide connectivity to over 50,000 computers on over 500 networks.

Figure 2: *NSFNET T-1 Backbone, 1988*



A number of administrative functions were centrally managed. As the Internet grew, this became increasingly unwieldy. For example: E-mail using the existing UUCP style of addressing (Appendix D) was proving unreliable in the rapidly changing network. So perhaps even more important than the technology changes was the hierarchical structure that began to form. Appendix A discusses the geographically/politically based hierarchy of networks, and Appendix C discusses the Domain Name Service. Both of these developments contributed to the success of de-centralized management of the Internet.

Because the Internet was federally funded (and users didn't directly pay for it), use was restricted

to educational and research organizations. As mentioned earlier, some firms had built their own internal networks. However, there was a desire to connect these various networks to allow intra-organization communications. Initially, this was strictly for research purposes, such as HP's computer laboratories in Palo Alto. However, due to demand from users, it soon became obvious that other sorts of support related work would be convenient to the research community. For instance, Sun Microsystems (whose computers are used a great deal on the Internet) set up an E-mail address to handle customer support (*hotline@sun.com*). In addition, operating systems software was made (instantly) available via FTP (although not "officially" by the vendors). Software development firms started doing the same thing. Proper use of the Internet became very blurry.

In response, each network developed a use policy statement similar to the one shown in Figure 3. A number of regional networks that don't receive federal funds do allow (and solicit) commercial traffic. In addition, there are at least two companies (UUNET and Performance Systems International) that "sell" access to the Internet, although commercial traffic should (!) not be sent over networks that don't allow it

Figure 3: *NSFNET Interim Conditions of Use Policy, June/1990*

The purpose of NSFNET is to support research and education in and among academic institutions in the U.S. by providing access to unique resources and the opportunity for collaborative work.

This statement represents a guide to the acceptable use of the NSFNET backbone. It is only intended to address the issue of use of the backbone. It is expected that the various middle level networks will formulate their own use policies for traffic that will not traverse the backbone.

1) All use must be consistent with the purposes of NSFNET.

2) The intent of the use policy is to make clear certain cases which are consistent with the purposes of NSFNET, not to exhaustively enumerate all such possible uses.

3) The NSF NSFNET Project Office may at any time make determinations that particular uses are or are not consistent with the purposes of NSFNET. Such determinations will be reported to the NSFNET Policy Advisory Committee and to the user community.

4) If a use is consistent with the purposes of NSFNET, then activities in direct support of that use will be considered consistent with the purposes of NSFNET. For example, administrative communications for the support infrastructure needed for research and instruction are acceptable.

5) Use in support of research or instruction at not-for-profit institutions of research or instruction in the United States is acceptable.

6) Use for a project which is part of or supports a research or instruction activity for a not-for-profit institution of research or instruction in the United States is acceptable, even if any or all parties to the use are located or employed elsewhere. For example, communications directly between industrial affiliates engaged in support of a project for such an institution is acceptable.

7) Use for commercial activities by for-profit institutions is generally not acceptable unless it can be justified under (4) above. These should be reviewed on a case-by-case basis by the NSF Project Office.

8) Use for research or instruction at for-profit institutions may or may not be consistent with the purposes of NSFNET, and will be reviewed by the NSF Project Office on a case-by-case basis.

Another issue that became important in the late 1980's was security. There had always been a number of computer break-ins, but nothing had affected the Internet as a whole, and users tended to not worry about security. The worm (or virus as some call it) of November 1988, which paralyzed the Internet for several days, changed this perception. The worm was allegedly (the recent court decision has been appealed) started by a Cornell Graduate student (who's father incidentally is the Chief Scientist of the National Security Agency's Computer Security Center). It took advantage of some bugs in the UNIX operating system, and "infected" several thousand computers on the Internet. No damage was done to the computers, but (due to a bug in the worm) these computers (which many people now used as an integral part of their work/lives) became so busy that they were unusable. A "cure" was quickly found within days, but the ramifications from this incident are still felt today.

Another incident at about the same time was a ring of KGB sponsored computer thieves that were electronically stealing military secrets by tapping into the Internet from West Germany - they were apprehended in 1989. A number of things were done to tighten security, including formation of the Computer Emergency Response Team (CERT) which acts as a clearing house for security.

Use of the Internet continued to increase at a rate of almost 25% per month - Figure 4 shows traffic on the T-1 line between CU and NCAR. Although the overall utilization of the total bandwidth on the communications line is quite a bit less than 10%, that is somewhat misleading. Traffic tends to be highest during the working hours, and is very bursty in nature as shown in Figure 5.

Figure 4: *T-1 Traffic between CU and NCAR*

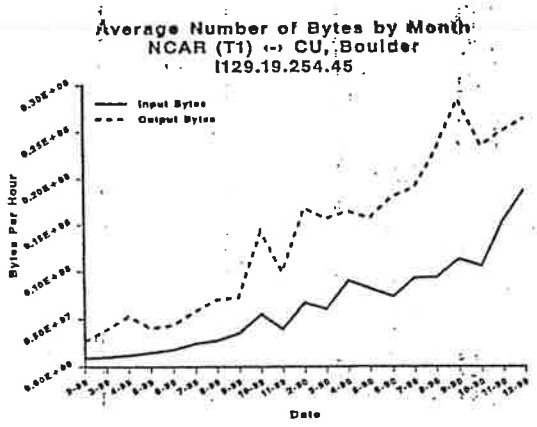
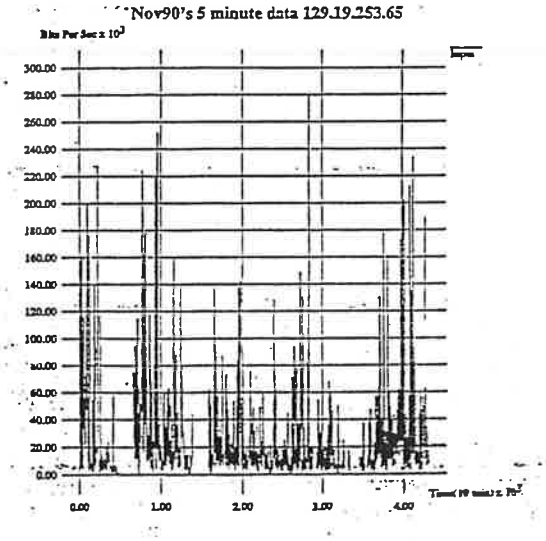


Figure 5: *Bursty Nature of Internet Traffic*



In 1990, work was started in installing T-3 (45Mbps) links for the backbone, a 30 fold increase in speed over T-1. Regional networks, following in NSFNET's footsteps, added more T-1 links.

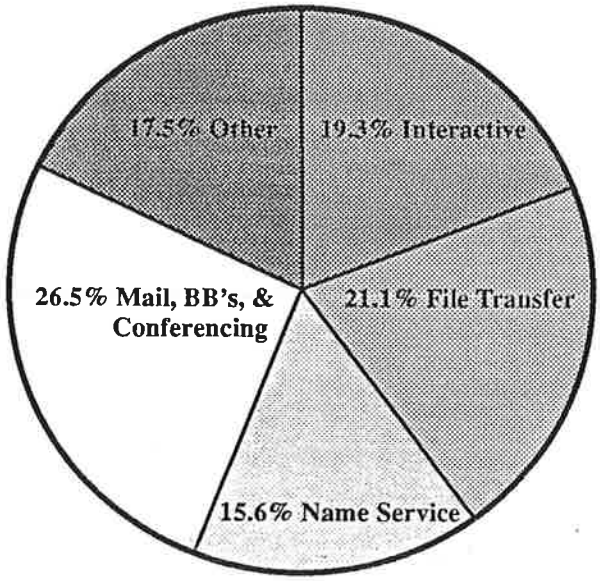
Work was also started in migrating toward the ISO Open Systems Interconnect (OSI) networking model. TCP/IP is just not robust enough to handle the growing size and complexity of the network. The migration to OSI could best be described as slow, since there is a very large base of installed computers using TCP/IP.

Currently, there are an estimated two million computers with perhaps over twenty million users connected to the Internet over six continents.

Future Direction of the Internet

It is interesting to look at the current breakdown by use of the traffic on the Internet today as shown in Figure 6.

Figure 6: *Breakdown of type of traffic on NSFNET, 2nd Quarter, 1990*



Note that over 15% of the traffic is Name Services: i.e. finding out where someone or some data is (interestingly enough, when the Encyclopedia Britannia was put on CD-ROM, the index occupied more space than the data). The Interactive and Other categories will continue to grow, as people

move into more on-line rather than batch-oriented applications. In addition, those types of applications tend to be more bandwidth intensive.There's no doubt that technology will continue to pace development of the Internet. One might think that T-3 (45Mbps) links are more than sufficient, but a number of (single) applications require much higher speeds. Video, animation, image processing are just a few. A simple calculation shows that a 8-bit color Mega-Pixel display running at 30 frames/second will require 30Mbytes of data per second (although data compression can reduce this substantially).

A number of non-technical issues also are being addressed. Probably foremost is funding of the Internet, which ties into who is going to pay for it, and how they will be charged. Some people feel that, similar to the Interstate system, the federal government should pay for it as evidenced by Senator Gore's "High Performance Computing Act of 1991" recently introduced in the 102nd Congress. Funding for network development at the state and regional level remains a problem.

Management of the Internet is also an issue, although de-centralization is working well. This allows each network to concentrate on providing the service to its users. Security and privacy will becomes increasingly important, as proprietary commercial and personal messages are sent across the network. This become even more "thorny" when that electronic information is transmitted across national boundaries. The demand for electronic communications will continue to increase, especially between (in addition to within) organizations.

Although the Internet has historically been used by organizations, it is interesting that individuals are starting to connect to it. The author believes that ultimately many households will want to be connected (similar to the French Minitel system) especially once the appropriate applications are available.

The interstate system was built in the mid 1900's under the auspices of National Defense. In today's world, the "threat" is not so much military strength, but economic competition. Key factors in efficient production and distribution of goods are low cost communication and transportation. A Nationwide, federally funded network that connects all organizations and people will facilitate this.

References

Many of these references were obtained electronically from various FTP archives on the Internet (*uu.psi.com*, *wuarchive.wustl.edu*, *nic.ddn.mil*, *nis.nsf.net*, *uunet.uu.net* and others with help from, of course, *archie* (a database of FTP archives) at *quiche.cs.mcgill.ca*). Although this list is (perhaps too) long, one particular book stands out: *The Matrix: Computer Networks and Conferencing Systems Worldwide* by John Quarterman. Although getting dated, it gives an excellent overview of networking, and also discusses it in great detail. I look forward to reading the next edition, and hope the bozo that checked out the first edition for six months (and just finally returned it) doesn't beat me to it! :-)

Another recommended book is *The Cuckoo's Egg*, by Clifford Stoll. This tells the story of how a 75 cent discrepancy in computer accounting records led to the discovery of a ring of KGB sponsored computer thieves, based in West Germany, who were electronically obtaining US military secrets. Many technical people (including the author of this paper) are not know for their writing skills, but Cliff's book is not only educational, but enjoyable to read and hard to put down.

Finally, if one is interested in keeping up with the top-level changes in the Internet, the NSF Network News newsletter is recommended.

"Internet Resource Guide," NSF Network Service Center, Cambridge, Massachusetts, March 25, 1991.

Slides from Dr. Ken Klingenstein (Director of Computing and Network Services, University of Colorado at Boulder) presentation at the Denver Data Processing Management Association March 21, 1991 meeting.

Klingenstein, Ken, "Beyond the Law," CNS Digit, University of Colorado at Boulder, vol. 26, no. 2, p. 3.

Jacobsen, O. and Lynch, D., "A Glossary of Networking Terms," (RFC1208), Network Working Group, March, 1991.

Milking, G. and Marine, A.,"FYI on Questions and Answers: Answers to Commonly asked "New Internet User" Questions," (RFC1206), Network Working Group, February, 1991.

"High-Performance Computing Act of 1991," introduced by Senator Gore to the 102nd session of Congress, United States Government Publications, 1991.

"NSF Announces Additional Funds: Remaining Nodes will move to T3," NFS Network News, Jan/Feb, 1990, pp. 1-2.

"Dialup/Switched/On-Demand Internet Access," The PSI Connection, Performance Systems International, Reston, Virginia, vol. 2, no. 1, pp. 1,5.

"National Network Expansion," The PSI Connection, Performance Systems International, Reston, Virginia, vol. 1, no. 3, pp. 1,6.

Caroline, Arms, "Using the National Networks: Bitnet and the Internet," *Online*, September 1990, pp. 24-30.

Communications Handbook, Computing and Network Services, University of Colorado at Boulder, September, 1990.

Bowers, K., LaQuey, T., Reynold, J., Koubicek, K., Stahl, M., and Yuan, A., "FYI on Where to Start - a Bibliography of Internetworking Information," (RFC1175), Network Working Group, August, 1990.

"NSFNET Moves toward New Technology," NFS Network News, July/Aug, 1990, p. 1.

"Interim NSFNET Acceptable Use Policy," NSF Network Service Center Press Release, June 1990.

"NSFNET Expansion Announcement," NSF Network Service Center Press Release, June 1990.

Alexander, Michael, "Morris Sentence spurs Debate," *Computerworld*, May 14, 1990, p. 128.

"Running Faster with T3," NFS Network News, April, 1990, pp. 3-4.

System & Network Administration Manual, Sun Microsystems, Inc., Mountain View, California, March, 1990.

Frey, Donnalyn and Adams, Rick, *A Directory of Electronic Mail Networks and Addressing*. O'Reilly & Associates, Sebastopol, California, 1990.

Laquey, Tracy, *Users' Directory of Computer Networks*. Austin, Texas, Digital Press, 1990.

Quarterman, John, *The Matrix: Computer Networks and Conferencing Systems Worldwide*. Bedford, Massachusetts, Digital Press, 1990.

Krol, Ed, "The Hitchhikers Guide to the Internet," (RFC1118) University of Illinois Urbana, Urbana-Champaign, Illinois, September 1989.

"Computer Security: Virus Highlights Need for improved Internet Management," GAO Report IMTEC-89-57, United States Government Accounting Office, Washington, DC, June 1989.

VMS System Manager's Manual, Digital Equipment Corporation, Maynard, Massachusetts, June 1989.

Eichin, Mark and Rochlis, Jon, "With Microscope and Tweezers: An Analysis of the Internet Virus of November 1988," Massachusetts Institute of Technology, Cambridge, Massachusetts, February 9th, 1989.

Stoll, Clifford, *The Cuckoo's Egg*. New York, New York, Doubleday, 1989.

Spafford, Eugene, "Some Musings on Ethics on Computer Break-Ins," Purdue University, Indiana, undated (appears to be early 1989).

Spafford, Eugene, "The Internet Worm Program: An Analysis," Purdue Technical Report CSD-TR-823, Purdue University, Indiana, December 8, 1988.

"Measuring the Size of the Internet," NFS Network News, October, 1988, pp. 1-2.

"NSFNET: The Next Five Years," NSF Network News, July, 1988, pp. 1-2.

"A User's Introduction to the NSFNET (Part II)," NSF Network News, November, 1987, pp. 1,9.

"A User's Introduction to the NSFNET (Part I)," NSF Network News, July, 1987, pp. 5,12.

"Steve Wolff Comments on NSFNET's Progress," NSF Network News, July, 1987, pp. 1,8.

Quarterman, John and Hoskins, Josiah, "Notable Computer Networks," *Communications of the ACM*, vol. 29, no. 10, pp. 932-971.

Rose, Mark, *Interstate: Express Highway politics*, 1941-1956. Regents Press of Kansas, 1979.

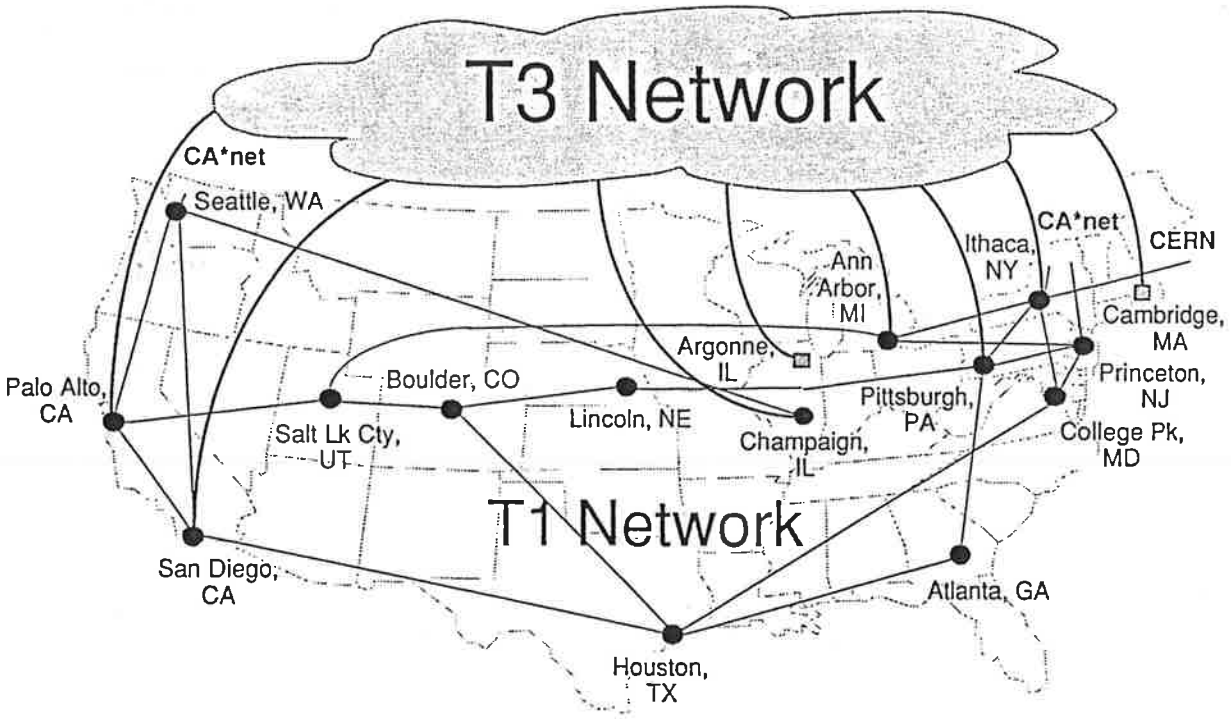
Tucker, Harry, and Leager, Marc, *Highway Economics*. Scranton, Pennsylvania, International Textbook Company, 1942.

Appendix A: The Hierarchical Nature of the Internet

On the Internet, there are at least three distinct "levels" of networks (not counting the one(s) actually inside an organization). This distinction is mostly based on the geographic and/or political considerations, although it should be noted that (most) networks can talk to one another as a peer.

The network that gets the most press is the NSFNET backbone shown in Figure A-1. This is a nationwide set of links, built around key computer centers. It allows connectivity among those central sites, but more importantly, has links to other networks that allow them to route traffic outside of their area of coverage.:

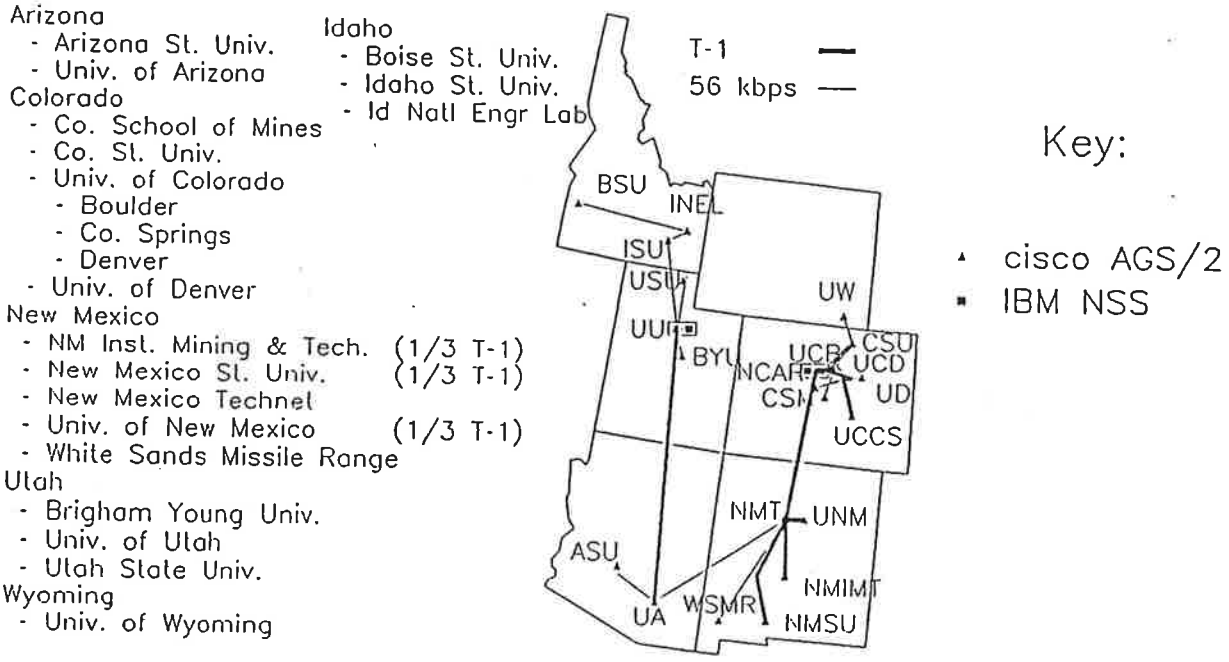
Figure A-1: *NSFNET T-1/T-3 Backbone, 1991*



The NSFNET is managed by Merit Computer Network in Ann Arbor, Michigan. Although the majority of traffic from other networks doesn't traverse NSFNET, a tremendous amount is still transmitted over the backbone. For March/1991, the total number of bytes transmitted over the T1 network alone was almost 700 Billion bytes, and that number doubles (at least!) every year.

Regional networks, such as Westnet (see Figure A-2 below) connect into the NSFNET backbone, but concentrate on providing service to a region. These networks may span multiple states (such as Westnet), or just a section of a state (such as BARRnet, the (northern California) Bay Area Regional Research Network).

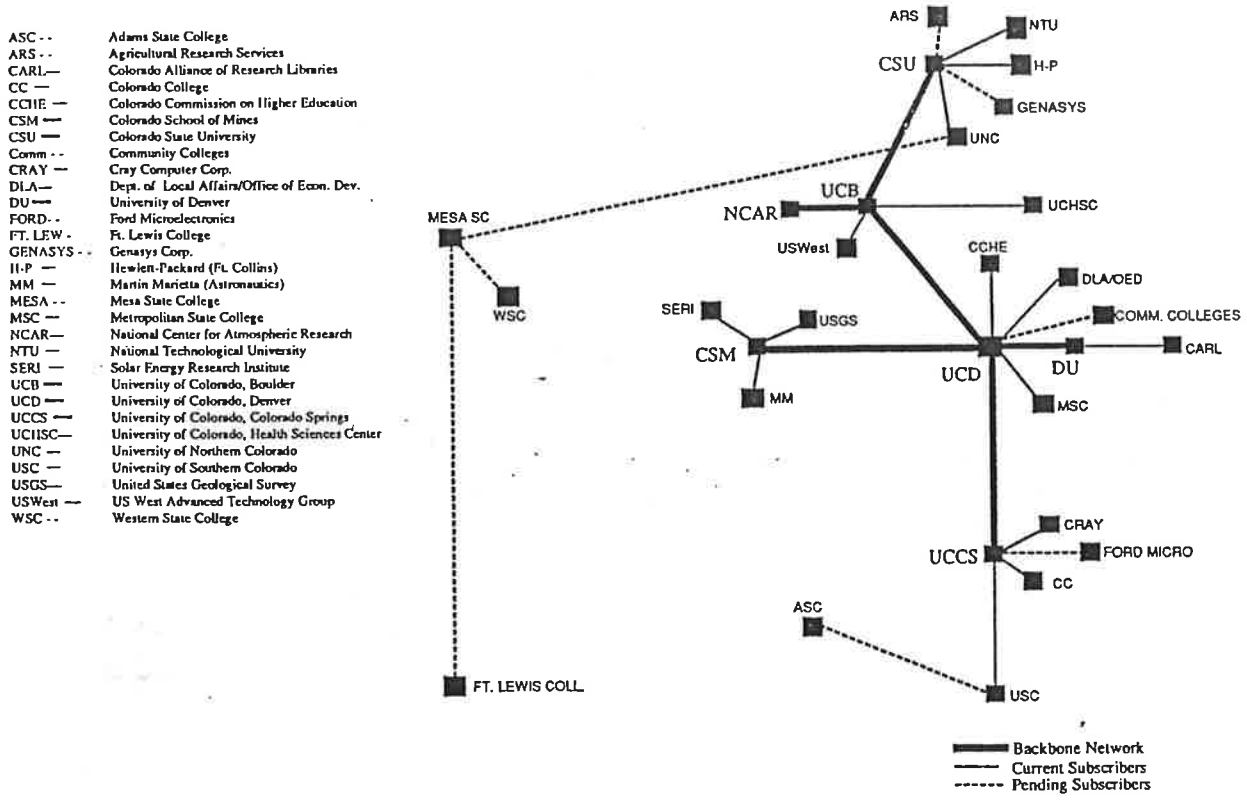
Figure A-2: *Westnet Primary Nodes, February 1991*



Westnet is managed by representatives from the primary universities on its network. Although most of the network is still 56kbps lines, efforts are underway to convert these all to T-1. Figure 4 shown earlier illustrates that traffic within Westnet also continues to increase at a fast pace.

Another level sometimes seen "below" the regionals is a state network. In Colorado, there is the Colorado SuperNet, as shown in Figure A-3. There are also comparable networks in the other states covered by Westnet; NMTechNet is one example.

Figure A-3: *Colorado SuperNet, March 28th, 1991*



The Colorado Supernet is managed at the Colorado School of Mines in Golden, Colorado by Colorado SuperNet, Inc. (a non-profit agency of the state of Colorado). It uses 56Kbps and slower lines, but has a T-1 line to NCAR to connect into the Internet. It offers a variety of very reasonably priced services to companies who don't need or can't afford full-blown network access. The author is currently in the process of obtaining network access for his employer from Colorado SuperNet.

The reader may be confused by the apparent random use of capital letters in the network names. The author is just as confused.

Appendix B: A Brief Explanation of IP Addressing

Computers on the Internet use (unique) Internet Protocol (IP) addresses to communicate with each other. This address is a 32 bit number, divided into four 8-bit fields. Each field is called an octet, and it is usually represented by the decimal equivalent, with periods separating each field. For example, the IP address of *tramp.colorado.edu* is: 128.138.129.4. IP addresses are unique in the world, and are assigned (by class) by the Network Information Center at SRI International.

An organization may have small, medium, or a large number of hosts. Conversely, one can suspect that there will be many, medium, and few organizations that fit this above description. The design of IP took this into account (after the fact), and assigned three main classes of IP addresses:

- Class A addresses have a value between 0-127 in the first octet. The network number of a Class A address is that number itself - i.e. 75. So the remaining three octet fields can be used to specify hosts within that organization - i.e. a total of 16,777,216 hosts. Note that there are only 127 class A addresses, so an organization would have to show true need before getting one of these.
- Class B addresses have a value between 128-191 in the first octet. The second octet is used with the first octet to specify the network number for a total of 65,536 networks (CU's is 128.138). The last two octets are used to specify the hosts: 65,536 are available on a class B network.
- Class C addresses have a value between 192-223. The second and third octet is also used to generate the network number: i.e. 192.9.200. Only 256 hosts are addressable on a class C network, but this is fine for many small companies (and fortunately there are 16,777,216 class C addresses available).
- An astute reader will notice that the range 224-255 is not mentioned above. Class D (multicast) and Class E (experimental) addresses exist, but are not in common use, and won't be discussed.

Although the description of IP addresses may seem technically esoteric, it's critical to the operation of a world-wide network, as it allows a very large and flexible address space. In contrast, DECnet Phase IV (a networking protocol designed by Digital Equipment Corporation) limits networks addresses to a maximum of 63 areas which can each have a maximum of 1023 hosts apiece. Although this may seem adequate for any single organization, this scheme quickly runs out of addresses when computers from different companies connect to each other. DEC is currently upgrading the addressing space with DECnet Phase V, but this is transition is causing a lot of difficulties for their current customer base.

As of March/1991, there were 41 Class A, 4520 Class B, and 24572 Class C network addresses assigned. As discussed above, the number of computers are much greater than this.

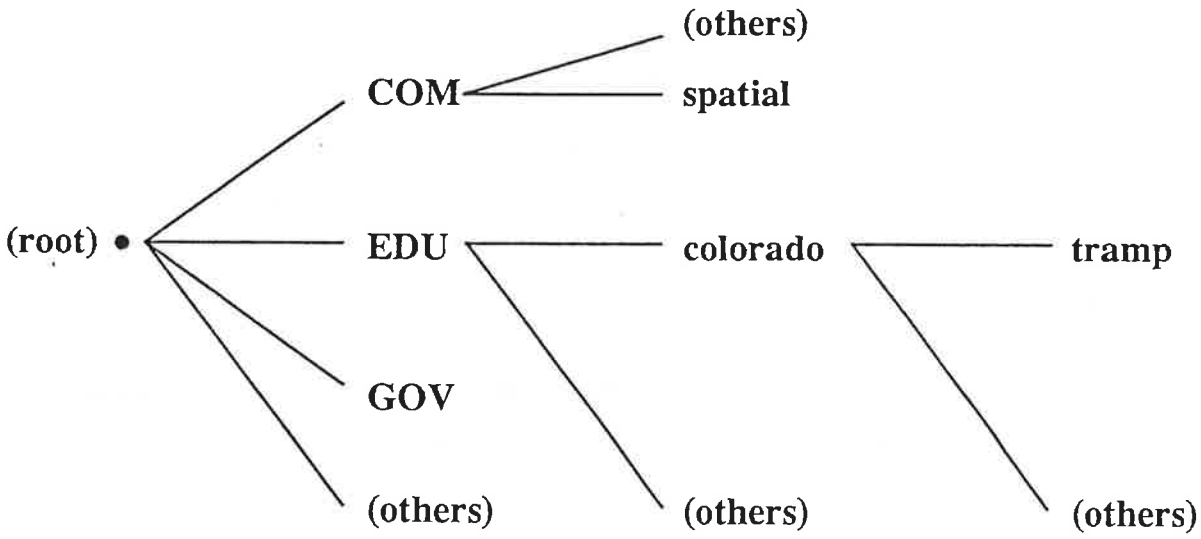
Appendix C: DNS - What is *komarnit@tramp.colorado.edu*?

People who send E-mail (or connect to machines) on the Internet often see something like the title above. What does it mean, why is it that way, and (most importantly) will my E-mail get to this person? This section will (also) be somewhat technical, but the answers to the above questions are important in understanding the Internet.

First, as mentioned in Appendix B, the Internet is a collection of machines, all with unique IP addresses. However, it would be somewhat inconvenient to refer to the machine at CU called *tramp* by it's IP address (which is 128.138.129.4). The solution in the Arpanet days was to simply call that machine *tramp* and associate it with 128.138.129.4. However, one would have to ensure that name was unique, and then insure that any machine that used *tramp* would have this entry in a file called HOSTS.TXT. As the number of machines grew, HOSTS.TXT quickly became unmanageably large. In addition, computers with pre-existing names that joined the net often had previously used names. Finally, if the computer folks at CU wanted to change their local configuration, they had to communicate this to everyone in the Internet community. A better solution needed to be found.

So the Domain Name Service was developed. Basically, this is a hierarchical setup that allows organizations to manage their own (uniquely named) machines, and present a unified front to the outside world as shown in Figure C-1 below.:

Figure C-1: *Outline of the Internet Domain Hierarchy*



The top level, or root, of the Internet is currently maintained at SRI's Network Information Center, which also assigns domain names. This machine (and several duplicates) are called root domain name servers - i.e. they know about the levels below them and which name servers know about them. *EDU* is a top-level domain standing for educational, *COM*: commercial, *GOV*: government, etc. Underneath *EDU* are several universities, one called *colorado* which is a second-level domain. It knows about machines under it: *tramp* in the example given. The name scheme is case-insensitive, although upper case is often used for the top-level domain. Note that although DNS is built on top of IP addresses, there is not necessarily a one-to-one mapping between Network Addresses and domain names (although that is usually the case).

So, when someone wants to connect to the machine called *tramp* (from anywhere in the world), the process goes something like this: First, the users' machine may not know the address of *tramp.colorado.edu*, but it can query a root domain name server to gather information on *edu*. After receiving the name of the top-level *EDU* domain name server, it asks it who is the name server for *colorado.edu*. It then queries that machine for the address of *colorado.edu.tramp*. Finally, once that address is received, a direct connection is established between the users' machine and *tramp*. This may sound overly complicated and possibly inefficient in concept. However, several "short-cuts" are used in the implementation to speed things up and increase reliability such as local caching, backup name servers, etc.

Note that this scheme allows each organization to decide on its own local hierarchy as it sees fit, and then advertise (or perhaps restrict access to) these machines. So if CU adds, deletes, or changes any of its computers on its network, it simply changes the information on its local name server. Although in most cases (such as this), the name server is managed by the local site, that is not necessarily true and in fact not the case for computers which communicate via UUCP rather than direct Internet access - please see Appendix D.

E-mail works in a very similar fashion. Many people believe that their E-mail is sent to some machine called *edu*, then *colorado.edu*, and then *tramp.colorado.edu*. Although this is a correct analogy to how postal mail works, it is not how E-mail works on the Internet (once again in most cases, please see Appendix D). Instead, a direct connection is established from your machine to the target machine (after determining the address as above), and your message is sent directly to the target machines, usually in seconds.

The DNS system allows one to refer to machines in a more logical manner. An extension to this that's currently gaining popularity is assigning geographically based names. For instance, instead of *spatial.com* (a software development firm in Boulder), one might use *spatial.flatirons-west.-boulder.colorado.us*. Many international firms use this style (the ISO two letter country codes are used to specify the top level domain), which also has the nice side-benefit that regular mail can sometimes use the same address.

It should be mentioned that DNS will function with some of the emerging standards coming from ISO: X.400 (message handling) and X.500 (directory services - "white pages").

Many people find DNS confusing at first. However, it is actually more intuitive that other naming schemes we have grown used to, such as phones numbers and zip codes. It is interesting that the phone system (which is basically a set of numbers with no relation to the person at the other end) has been accepted almost universally. One thinks it would be easier to "ask" the phone to call the author referred to by his name, and (autonomously) the phone would determine the "correct" number to call at the time of the call.

There are research efforts underway by the phone company(s) to "assign" people a permanent number that follows them around for life. The author believes a dynamic scheme and/or a more intuitive approach (such as DNS) may prove to be a superior approach.

Appendix D: A Brief Explanation of the UUCP network

This paper has primarily concerned itself with the Internet and tacitly assumed that the computers in question were directly connected to it. Although this is changing, the vast majority of computers do not have direct Internet connections, but instead are still connected via modem with a protocol called UUCP. The primary "advantage" of UUCP is that it is very easy to get on the network. All one needs is a computer, a modem, and a willing site that will establish a connection for it.

UUCP stands for UNIX to UNIX Copy. Computers first networked by establishing dial-up (usually on demand) phone connections via modem to transfer files. Often times, E-mail (and other information) was queued up for later transmission (usually at night when phone rates were cheap! :-). While UUCP "correctly" refers to the protocol used to transfer the data, people often refer to UUCP style addressing. Although DNS has made this (mostly) obsolete, it is still used, and is of interest for historical reasons (and also will illustrate how important DNS is)

In a nutshell, UUCP style addressing requires the sender to specify the path the E-mail must take to reach the receiver. Contrast this with DNS, where one simply specifies the users's address, and the network takes care of getting it there. An analogy to the "regular" mail system may make this more understandable. The author's home address is: 4808 West Moorhead Circle
Boulder, CO 80303-6156 (USA)

One can send a letter to me from Tokyo using that address, and the postal service will insure that (hopefully! :-) the best path is taken to get there. However, if one had to use UUCP style addressing, one might have to say: send this to Tokyo! Honolulu! San Francisco!Denver!Boulder!4808 West Moorhead. Note that the routing path would be different from another place. In addition, this may not only be a less effective routing path, but in fact could be inoperative and/or incorrect. Finally, this requires the sender to know about the entire routing path. As clumsy as this may seem, this is how E-mail used to work (and still does to a certain extent for non-registered domains).

People usually specified their E-mail address as *...{boulder, ncar}!spatial!alek*. This told someone how to get from a "major" site (either *boulder* or *ncar* in this example), but it was left up to the sender to figure out how to get there. It's amazing that it worked as well as it did.

Computers that use UUCP may have their computer registered in DNS as discussed in Appendix C, or they may not be "known" at all to the network. Regardless, the method used to communicate is similar in both situations.

Registered Domains

In this situation, a computer on the Internet acts as, what is called, an MX forwarder. For example, the computer *spatial.com* receives E-mail for a company called Spatial Technology in Boulder, Colorado. However, it is not connected to the Internet. When someone sends E-mail to *spatial.com*, the DNS lookup as specified in Appendix C is followed. Since *spatial.com* is not on the Internet, it can not handle name service, but instead other name servers do this for it. *boulder.colorado.edu* is returned as the place to send the E-mail. After it arrives at *boulder.colorado.edu*, it is delivered via UUCP over dial-up lines to *spatial.com*. At the same time these messages are delivered (or separately), *spatial.com* sends E-mail's to *boulder.colorado.edu* for it to forward to other machines on the network.

Non-Registered Domains

A large number of organizations who have not yet registered their domains fall into this category. The author is intimately familiar with an organization in Louisville, Colorado called Centera Technologies. Their E-mail address is: *username%centera.uucp@flatirons.sun.com*. In this situation, Centera was successful in asking the system administrators of the machine *flatirons.sun.com* (located in Westminster) to act as a E-mail forwarder. So E-mail to this address goes through the standard DNS approach and is sent to *flatirons.sun.com*. Once reaching that machine, *flatirons* on-demand or periodically calls *centera* (or vice-versa) and delivered the E-mail. Although this approach is similar to above, it's important to note that only *flatirons* "knows" about *centera*.

Appendix E: USENET, The Electronic Daily News

USENET is a (very large) network of computers that exchange *news*. *News* is basically a giant electronic bulletin board, on which anyone can add comments to (except as noted below). Articles are grouped in newsgroups, which are a hierarchical representation of the subject matter. For example, *comp.sys.sun* discusses computer issues relating to Sun systems. The subject areas range from the core group of serious ones (*comp.lang.c++, comp.unix.wizards*), local newsgroups (*boulder.general, cu.cs.grads*), humorous ones (*alt.humour, rec.arts.startrek.info*) to the bizarre (*alt.sex.beastiality*).

USENET is one of the most decentralized networks; anyone can read and/or post (with the exception of moderated groups) articles to the net. No one entity pays for USENET, as each host basically handles its own traffic and forwards *news* onto the next host. It is probably for this reason that commercial solicitations are frowned upon, and usually met with "flames" (very negative comments from other users). However, there is a group, *comp.newprod*, for firms to (briefly) announce new products. There are a variety of (somewhat obscure) rules for using USENET that come under the broad category of *net.etiquette*. People that violate these rules are subject to being flamed, especially if they are perceived as using *news* for commercial purposes as shown in Figure E-1. Finally, security can best be described as lax, although it is usually not used in a harmful fashion; please see Figures E-2 and E-3.

Although the two are intertwined, USENET and UUCP are different. One supports *news* but not E-mail, and the other supports E-mail but not *news*. However, in most cases, they co-exist on the same computers and use the same network connections.

USENET was started in 1979 by two graduate students, Jim Ellis and Tom Truscott at Duke University as a method of exchanging information between Duke and the university of North Carolina at Chapel Hill. The originators invited others to join in, and puts the *news* software on the 1980 USENIX tape. USENIX is the oldest and largest UNIX users' group. The software at that time was equipped to handle about a hundred hosts and a few articles per day. It should be mentioned that this "free-ware" concept gets a lot of the credit for software such as *news* today.

As the years went by, *news* gained in popularity. The software to handle it was re-written several times, as the number of groups grew. By 1986, there were 240 groups, posting an average of almost a thousand messages totalling two MBytes a day. Some would say that *News'* popularity has had negative consequences, as the signal to noise ratio (ratio of "good" articles to "junk" articles) became quite low in some popular groups. This led to moderated groups. Users' comments would go to a moderator instead of being broadcast. That moderator would then condense, summarize, or perhaps remove that posting before posting it to the entire net.

Most groups were carried world-wide (although the international links usually don't carry a full *news* feed because of its volume). Regional or local groups were formed for information that pertained mostly to the local area.

USENET flourishes today, although some suggest that its popularity may lead to its demise. Currently, there are almost 2,000 groups posting over 17 MBytes of articles a day. Although the software to handle news has undergone several revisions, it is still somewhat inadequate to handle this information explosion, especially with regards to presenting to the reader what they really need/want to read.

Figure E-1: *USENET response to a survey question*

Fromboulder!csn!ncar!elroy.jpl.nasa.gov!usc!cs.utexas.edu!uwm.edu!bionet!agate!ucbvax!lta.lta.com!david Thu Apr 4 19:35:49 MST 1991
Article: 1768 of comp.windows.x.motif
Path:boulder!csn!ncar!elroy.jpl.nasa.gov!usc!cs.utexas.edu!uwm.edu!bionet!agate!ucbvax!lta.lta.com!david
From: david@lta.lta.com (David B. Lewis)
Newsgroups: comp.windows.x.motif
Subject: re: Question for net.views column in UNIX Today!
Message-ID: <9103302235.AA01027@lta.com>
Date: 30 Mar 91 22:35:10 GMT
Sender: daemon@ucbvax.BERKELEY.EDU
Distribution: inet
Organization: Lewis, Trachtenberg & Associates
Lines: 8

> * QUESTION #2 *
> Is a single GUI standard really necessary?

Anybody considering taking this question from Unix Today! seriously should first check out the comp.misc discussions of March 18-22 on their handling of their Question 1.
And now back to real Motif stuff...

Without going into technical detail, UNIX Today! previously committed a number of breeches in *net.etiquete*. There were several responses that were much more vehement. Although this may seem harsh, it actually leads to a forms of self-policing of the net. Despite UNIX Today!'s miscues, they did receive a number of responses, and they use them in a regular column in their biweekly computer trade journal. USENET allowed them to reach a global audience, and gets replies in a very timely basis - it's questionable if they could do this any other way.

Figure E-2: *Annual (?) April 1st Posting to USENET*

Fromboulder!csn!pikes!mercury.cair.du.edu!mnemosyne.cs.du.edu!uunet!seismo!ukma!wuarchive
!usc!apple!amdahl!walldrug!moscvax!perdue!spaf Mon Apr 8 07:39:15 MDT 1991
Article: 4 of news.announce.important
Xref: boulder news.announce.important:4 news.admin:2162
Path:boulder!csn!pikes!mercury.cair.du.edu!mnemosyne.cs.du.edu!uunet!seismo!ukma!wuarchive
!usc!apple!amdahl!walldrug!moscvax!perdue!spaf
From: spaf@cs.purdue.EDU (Gene Spafford)
Newsgroups: news.announce.important,news.admin
Subject: Warning: April Fools Time again (forged messages on the loose!)
Message-ID: <4-1-1991@medusa.cs.purdue.edu>
Date: 1 Apr 91 00:00:00 GMT
Expires: 1 May 91 00:00:00 GMT
Followup-To: news.admin
Organization: Dept. of Computer Sciences, Purdue Univ.
Lines: 25
Approved: spaf@cs.purdue.EDU

Warning: April 1 is rapidly approaching, and with it comes a USENET tradition. On April Fools day comes a series of forged, tongue-in-cheek messages, either from non-existent sites or using the name of a Well Known USENET person. In general, these messages are harmless and meant as a joke, and people who respond to these messages without thinking, either by flaming or otherwise responding, generally end up looking rather silly when the forgery is exposed.

So, for the few weeks, if you see a message that seems completely out of line or is otherwise unusual, think twice before posting a followup or responding to it; it's very likely a forgery.

There are a few ways of checking to see if a message is a forgery. These aren't foolproof, but since most forgery posters want people to figure it out, they will allow you to track down the vast majority of forgeries:

- Russian computers. For historic reasons most forged messages have as part of their Path: a non-existent (we think!) russian computer, either kremvax or moscvax. Other possibilities are nsacyber or wobegon. Please note, however, that walldrug is a real site and isn't a forgery.

- Posted dates. Almost invariably, the date of the posting is forged to be April 1.

- Funky Message-ID. Subtle hints are often lodged into the Message-Id, as that field is more or less an unparsed text string and can contain random information. Common values include pi, the phone number of the red phone in the white house, and the name of the forger's parrot.

- subtle mispellings. Look for subtle misspellings of the host names in the Path: field when a message is forged in the name of a Big Name USENET person. This is done so that the person being forged actually gets a chance to see the message and wonder when he actually posted it.

Forged messages, of course, are not to be condoned. But they happen, and it's important for people on the net not to over-react. They happen at this time every year, and the forger generally gets their kick from watching the novice users take the posting seriously and try to flame their tails off. If we can keep a level head and not react to these postings, they'll taper off rather quickly and we can return to the normal state of affairs: chaos.

Thanks for your support.

Gene Spafford, Net.God (and probably tired of seeing this message)

Figure E-3: *Follow-up Posting to above message*

Fromboulder!csn!magnus.acs.ohio-state.edu!zaphod.mps.ohio-state.edu!pacific.mps.ohio-state.edu!linac!att!cbnewse!cbnewsd!cbfsb!cbnewsg.cb.att.com!mark Mon Apr 8 07:39:33 MDT 1991
Article: 5 of news.announce.important
Xref: boulder news.announce.important:5 news.admin:2175
Path:boulder!csn!magnus.acs.ohio-state.edu!zaphod.mps.ohio-state.edu!pacific.mps.ohio-state.edu!linac!att!cbnewse!cbnewsd!cbfsb!cbnewsg.cb.att.com!mark
From: mark@cbnewsg.cb.att.com (Mark Horton)
Newsgroups: news.announce.important,news.admin
Subject: Forged warning messages
Message-ID: <1991Mar31.211215.4974@cbfsb.att.com>
Date: 31 Mar 91 21:12:15 GMT
References: <4-1-1991@medusa.cs.purdue.edu>
Sender: news@cbfsb.att.com
Organization: AT&T Bell Laboratories
Lines: 10
Approved: mark.horton@att.com

This is just to point out that the previous message is a forgery. In fact, it's a boring repeat of a forgery done in a previous year. There really are Soviet computers on the net now, but moscvax isn't one of them, as far as I know.

It is true that you should beware of things done on or near April 1.

Such as fake postings from me or Gene Spafford....

Mark Horton

moderator, news.announce.important

A naive reader might ask (like several subsequent USENET postings did) which one is fake? Although this case is pretty innocent, one can imagine that this could cause problems.