GAO
United States
General Accounting Office
Washington, D.C. 20548

General Government Division

B-276887

May 12, 1997

The Honorable Sheila Jackson Lee
House of Representatives

Subject: Internet Census and Use Estimates

Dear Ms. Jackson Lee:

On February 12, 1997, your office requested information relevant to your proposal for a national Internet census. As agreed with your office, our objectives in this letter were to (1) summarize the range of current estimates of Internet access and usage in the United States, (2) describe several key concepts related to defining the "capacity" of the Internet, and (3) identify additional issues where there are questions about the federal role in the development of the Internet.

BACKGROUND

The Internet began as a federal project to facilitate communication and data exchange among research universities and defense contractors. The Internet today has been described as a "network of networks." Networks link computers; for example, a network in a federal agency may provide that agency's staff with access to central data and the ability to share information. The Internet extends that access and ability to other networks by linking them together under common communications protocols. Messages are transmitted from one network to another through the Internet, and may pass through several of the networks en route to their destinations. A single message, in fact, is broken into "packets" that move separately and that may move through different paths to their destination. The packets are routed through the Internet (moving across multiple networks) according to rules set to minimize transmission time. The objective of the common communications protocols is to define how these individual packets can be reassembled at the destination so that the message appears to be intact to the recipient who reads it.

Users gain access to the Internet through several different methods. Some individuals have access through schools or employers. The schools and employers, in turn, generally purchase access through Internet service providers

(ISP), who in turn are connected to "backbone providers." [1] Other individual users purchase access directly from ISPs.

The array of services that an individual user is able to access can be divided into several broad categories, including file transfer protocol (ftp), electronic mail (e-mail), and the World-Wide Web (WWW). By using ftp services, users can electronically move computer files containing programs or data from one computer to another over the Internet. E-mail enables users to exchange messages electronically; these messages may be "posted" on Internet bulletin boards or be part of "news groups."

WWW, or Web access, builds on the other types of Internet access. Using special software or "browsers," users can access "pages" that are stored on distant computers. These pages are formatted or marked in accordance with a defined standard that electronically links them together. These links allow users to jump from one Web page to another. By providing these links that facilitate movement from page to page and by allowing these pages to contain a wide assortment of graphics (which may be used as links themselves), Web services have made it much easier for users to gain access to the information available through the Internet. This increased ease of usage, in turn, has contributed to the explosive growth in Internet use over the last several years.

RESULTS IN BRIEF

Estimates of the current use of the Internet or access to its services vary widely. One Internet source gathered private sector estimates of use and access that were made between July 1996 and January 1997 and reported a range between 5.8 million users and 35 million with access. A subsequent estimate put the Internet population at 47 million, but noted that "relatively few are 'heavy' users." There are several reasons for the range of estimates. Surveys employed different definitions of use and access; for example, some sought to identify those who have used the Internet recently, while others sought to identify those who have access to the Internet that they may or may not use. In addition, studies conducted at different times may have substantially different results because of the rapid growth in Internet usage in the United States. Finally, studies did not have a uniform methodology for estimating Internet usage. However, while the different sources disagreed on the level of Internet use and access, they did agree that regardless of how it may be measured, Internet activity is increasing at a fast rate. This fast growth, in turn, may complicate the task of measuring the level of activity.

By design, the Internet is decentralized, complicating both the definition and measurement of its capacity. And while there have been complaints about Internet

[1] "Backbones" route traffic between ISPs and connect with other backbones. Backbone providers include MCI, UUNet, and Sprint.

congestion, there is no way to conclusively demonstrate the extent to which this is actually occurring or what causes underlie any congestion. The rapid growth in the Internet also makes it difficult to precisely define its capacity. A key question, however, is whether the available capacity is being used efficiently. In addition, new technologies are presently being designed and deployed that are intended to increase the ability of the Internet to transmit messages and data.

While the federal government played a key part in the creation of the Internet, its role has been decreasing during the recent growth in Internet usage. In several areas, however, there is a continued or evolving federal role. The Federal Communications Commission (FCC) is considering whether there is a more efficient way to charge ISPs for phone lines used to provide Internet access to their customers; ISPs and several other businesses providing Internet services and Internet-related products oppose such changes. Debate continues over proposals to restrict encryption technology that can render messages sent over the Internet unreadable to anyone but the intended recipient, with privacy concerns competing against law enforcement and national security needs to intercept and decipher messages. Finally, an interagency task force is working to develop a consistent federal strategy to address policy issues posed by growing Internet use.

FINDINGS

Internet Use and Access

An Internet-based compilation of estimates of the size of the Internet user population noted that "one of the most contentious Internet issues is the size of the user population."[2] It noted two principal reasons for the differences among the competing estimates:

- The rapid growth in Internet usage could lead to misleading comparisons between surveys that were not conducted at comparable times. The compilation suggested that surveys conducted more than 90 days apart from each other should not be compared.

- "Different questions produce different results." Different surveys have used different questions to measure whether a respondent is an Internet user. These questions can differ with respect to the time period in question (e.g., whether a respondent has used the Internet within the last 3 months or 1 year) and with respect to the definition of Internet use (e.g., "logging on," using the WWW, or using any of the other Internet services) or access.

[2]"CyberAtlas" (http://www.cyberatlas.com on Feb. 7, 1997).

This compilation identified 13 studies conducted between July 1995 and November 1996 that sought to estimate Internet usage. The earliest study, by O'Reilly & Associates, estimated that, as of July 1995, the population with Internet access in the United States was 5.8 million. This estimate was based on random phone calls to households. The highest estimates were presented in two studies. In November 1996, Lou Harris & Associates estimated that there were 35.0 million adult Internet users in the United States; Intelliquest offered the same estimate in July 1996.

Subsequently, Intelliquest reported that, during the fourth quarter of 1996, the Internet/online population was approximately 47 million adults. (The online population refers to users of online services, such as America OnLine, which typically provide their customers with Internet access as well as other services.) This estimate was based on random phone calls to households. According to Intelliquest, "this represents a 34 percent growth in the online population from the first quarter of 1996, which was measured at 35 million. Despite this dramatic population growth in 1996, relatively few are 'heavy users', with four and a quarter million people using the Internet and online services 20 hours or more per week."[3]

We have not sought to verify these estimates or compare the differences in the methodologies among the studies.

Internet Capacity

By design, the Internet is decentralized. Originally, this was to provide redundant ability to transmit messages. A consequence of that design is that the capacity of the Internet is difficult to define and measure.

One assessment of Internet congestion listed several commonly cited reasons for congestion: "overloaded Web servers" (computers that make Web pages available to Internet users), loss of packets during periods of peak Internet usage, problems associated with the way that ISPs are linked to the Internet, and capacity limitations of Internet routers and backbones.[4] Another study noted that "two types of Internet-related congestion should be distinguished: congestion of the Internet backbones, and

---

[3]Intelliquest (http://www.intelliquest.com/about/release24.htm on Apr. 22, 1997).

[4]Bob Metcalfe, "NetNow's Statistics Trigger Defensive Responses From Some Corners of the 'Net," Infoworld, February 3, 1997, p. 3. "Routers" are computers that assign individual packets to routes through the Internet that are determined to be the best at the point in time when the assignment is made; since Internet traffic patterns vary over time, packets that comprise parts of the same message may be assigned different physical routes.

congestion of the public switched telephone network when used to access the Internet."[5]

Some components of the Internet, such as backbones, can be described in terms of their capacity to transmit data, and capacity limitations, if any, that exist in this physical component of the Internet can be measured. The impact of physical measures of capacity, such as backbone transmission capability, on users cannot be as easily determined, however, because any congestion perceived by users may have been caused by physical limitations at several different locations.

A second key issue in defining capacity–as it appears to an Internet user—is the efficiency with which the physical capacity to transmit data is used, and this factor also does not lend itself to easy measurement. Messages sent through the Internet are divided into packets to improve the efficiency of transmitting them, using the available backbones and other aspects of the Internet system. This strategy makes it possible to transmit more messages within a given time than would be possible if the messages had to remain intact during transmission. Very roughly speaking, the ability of a user's computer and software to "reassemble" the message (or Web page) from the packets makes the Internet more efficient, and the availability of computers and software acts as a substitute for the greater capacity that would be needed if Internet connections were terminals with no computing ability.

If the available capacity of the Internet is not used as efficiently as possible, there are a variety of possible causes. One assessment says that the "real problem with maintaining the Internet backbone is not traffic capacity. . . . The real problem is keeping track of all the networks that form the Internet." Thus, if one network that is part of the Internet shuts down and restarts, messages must be sent to "routers all over the network so they can make optimal routing decisions based on the state of the network in real-time."[6] Other causes of congestion have also been asserted. For example, some Internet experts have asserted that smaller or inexperienced service providers do not handle Internet addresses efficiently, thereby causing problems for the entire Internet, which must track these addresses.

---

[5]K. Werback, "Digital Tornado: The Internet and Telecommunications Policy," Federal Communications Commission, Office of Plans and Policy Working Paper Series 29, March 1997, p. 52. This paper notes that its analysis and conclusions do not necessarily represent the views of the FCC or individual commissioners.

[6]George Lawton, "Building the Internet Backbone: Take A Virtual Tour of a Busy Network Access Point" (http://www.magpag.com/~kozmando/NAP/comp.html on Apr. 28, 1997).

Current Policy Debates

The federal government had substantial responsibility for the beginnings of the Internet, with both the Department of Defense and the National Science Foundation providing much support to the Internet during its early years. In recent years, that support has diminished as private businesses and institutions have assumed greater responsibility.

The federal government remains involved in the Internet in several areas. The Communications Decency Act of 1996, for instance, seeks to prohibit use of the Internet as a method for disseminating pornography. Other areas where the federal government is considering a continued role include regulation of telecommunications charges for ISPs and regulation of encryption technology.

Many Internet users use telephone lines to access Internet services. Arguing that such use imposes requirements on them that are substantially different from voice communications, regional telephone companies have asserted that they should be compensated for the increased loads put on their systems. The regional telephone companies argue that Internet access not only increases the number of calls made, but that those calls typically are also longer in duration than the voice calls for which their systems are primarily designed. However, the current pricing for phone services in the United States is based on a principle that the party placing a call will pay ("sender pays"). While ISPs commonly have a large number of phone lines, these lines typically are not used for outgoing calls, and ISPs thus effectively pay flat monthly rates for the phone lines rather than usage-based charges. ISPs dispute the contention that the congestion attributable to their operation poses a serious problem. Opponents of proposals to impose use-based telephone charges on ISPs also assert that the flat-rate monthly rates are designed to recover the costs incurred by the telephone companies; if the rates can compensate for local service without usage charges, they equally can compensate for Internet access. The FCC has sought comment on "how it can most effectively create incentives for the deployment of services and facilities to allow more efficient transport of data traffic to and from end users."

Through encryption, a message can be made unreadable to anyone but the intended recipient. Ensuring privacy of communications is often considered essential to commercial use of the Internet: individuals and businesses are not likely to entrust financial transactions to communications that cannot be made secure against interception. As noted in the draft of "A Framework for Global Economic Commerce," a report prepared by an interagency working group, however, ". . . strong encryption not only enables law-abiding citizens to protect better their trade secrets and personal records, it can also be used by criminals and terrorists to hide their activities and

thwart legally-authorized investigations."[7] Exports of certain encryption programs are thus subject to export controls. Opponents of such controls have argued that, since the technology underlying these encryption programs cannot be controlled, such export controls limit U.S. businesses without providing effective means of enhancing U.S. security.

In addition, the draft of "A Framework for Global Economic Commerce" sets out the following recommendations for policy:

"Fostering the Internet as a Non-Regulatory, Market-Driven Medium:
- Establishing cyberspace as a duty-free zone
- Advocating for no new taxes on the Internet
- Allowing electronic payment systems to evolve without premature government involvement
- Encouraging industry self-regulation where appropriate
- Enabling market forces to drive the development of technical standards.

"Ensuring a Transparent and Harmonized Global Legal Environment
- Creating a 'Uniform Commercial Code' for cyberspace
- Protecting intellectual property on-line
- Partnering with industry to safeguard security in the electronic marketplace.

"Allowing Competition and Consumer Choice to Shape the Marketplace
- Maintaining privacy and the integrity of personal information
- Fostering fair competition and striving for interoperability among national telecommunications systems
- Empowering consumers to manage questions of content
- Opposing non-tariff barriers which limit free trade across the Internet, such as content restrictions, discriminatory telecommunications regulations, standards requirements, or anti-competitive compulsory licensing requirements."

As of April 23, 1997, the policy statement had not been finalized.

Federal policy on Internet issues also is among the issues addressed by the interagency Information Infrastructure Task Force. One component of that effort is an Information Policy Committee, chaired by the Administrator of the Office of Information and Regulatory Affairs at the Office of Management and Budget (OMB). That committee is to address information policy issues that must be resolved if the National Information Infrastructure is to be "fully deployed and utilized."

---

[7]"A Framework for Global Electronic Commerce," draft 9, December 11, 1996 (http://www.iitf.nist.gov/eleccomm/glo_com.htm on Feb. 13, 1997).

SCOPE AND METHODOLOGY

Since much of the information about the Internet is available through the Internet, most of the information presented in this letter was collected from Web pages. These pages were identified by several Internet search programs (including "Yahoo" and "Alta Vista") and were accessed between February and April 1997. Footnotes identify the Internet addresses for the Web pages and the dates on which the information was collected. Given the nature of information that is made available on the Web, we could not independently verify the information, nor can we ensure that the information presented remains available at the Internet addresses cited.

We did not seek to verify the estimates of Internet access and use and its capacity that were made available on these Web pages. We did identify key differences in definitions and research methodologies that were included in the Web pages, but did not seek to estimate the effects of these differences on the estimates of Internet users or access.

In discussing the evolving federal role in the Internet, we identified issues about which there has been considerable recent controversy. We selected examples based on our judgment that they illustrated the nature of the federal role in the Internet. We did not seek to provide a comprehensive overview of all issues surrounding federal policies regarding the Internet.

On April 24, 1997, we made a draft of this letter available to OMB for review and comment. The branch chief for Information Policy and Technology, Office of Information and Regulatory Affairs, said that OMB saw no need to comment.

- - - - -

The principal contributor to this correspondence was James McDermott. If you have any questions, please call me on (202) 512-8676.

Sincerely yours,

L. Nye Stevens
Director
Federal Management and
Workforce Issues

(410138)

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are $2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. VISA and MasterCard credit cards are accepted, also. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20884-6015

or visit:

Room 1100
700 4th St. NW (corner of 4th and G Sts. NW)
U.S. General Accounting Office
Washington, DC

Orders may also be placed by calling (202) 512-6000
or by using fax number (301) 258-4066, or TDD (301) 413-0006.

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (202) 512-6000 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

For information on how to access GAO reports on the INTERNET, send an e-mail message with "info" in the body to:

info@www.gao.gov

United States
General Accounting Office
Washington, D.C. 20548-0001

Bulk Rate
Postage & Fees Paid
GAO
Permit No. G100

Official Business
Penalty for Private Use $300

Address Correction Requested