A Note on the Internet

The Internet is a world-wide network of networks, connecting thousands of government, university and commercial computer networks around the globe. It is functionally defined as the network of interconnected and interoperating networks based on the TCP/IP (Transmission Control Protocol/Internet Protocol) standard and a common network addressing scheme (see Section 2). In contrast to other public networks, the Internet does not have a formal institutional or corporate embodiment. The evolution of the Internet is the key to its decentralized structure. This note reviews the Internet's history in Section 1, key elements of Internet technology in Section 2, access options in Section 3 and services available on the Internet in Section 4. Hands-on exercises are included in Section 5 and emerging issues in the development of the Internet are discussed in Section 6.

1. A Brief History of the Internet

The Internet has evolved from a research network to a commercial network. Its history consists of three phases: ARPANET (Advanced Research Projects Agency Network), NSFNET (National Science Foundation Network), and Post-NSFNET developments. In this note we focus on the U.S., as it is here that the Internet was "born" and many other key developments took place. Nevertheless, you should be aware that while more than half of the computers on the Internet are U.S.-based, Internet growth rates outside the U.S. have in fact exceeded those within the U.S. for the past few years. This history is important as it is the reason for the current, highly decentralized structure of the Internet.

1.1 ARPANET

ARPANET, the first computer network based on packet switching technology (see Section 2.1 and “A Note on Computer Network Technology”, teaching note S-OIT-14), was developed in the late 1960s by the Advanced Research Projects Agency (ARPA) to make the U.S. communications infrastructure resistant to nuclear attack. It initially consisted of four host computers (also called hosts): three in

Prepared by Philipp Afèche under the supervision of Haim Mendelson of the Graduate School of Business, Stanford University. Comments by Charles Bonini, Christine Herron, Anne Korin, Daphna Mendelson and Jeff Moore, as well as partial financial support by the Stanford Computer Industry Project, are appreciated.

1 In a networking setting, a host computer is a computer running application programs and storing data.
California (one of which was at the Stanford Research Institute) and one at the University of Utah. These were linked through leased telephone lines, each with a bandwidth\(^1\) (or data rate) of 56 Kilobits (thousand bits) per second. Between 1969 and 1986, the number of hosts grew to about 5000, with the first international connections established in 1973 (to England and Norway). During this period, access to the ARPANET was limited to defense agencies and their contractors, including university-based researchers. Computer scientists were the heaviest users, using the ARPANET mostly to exchange messages via Electronic mail (E-mail) and to access data and programs on remote computers with Telnet (see Section 4.1).

The most noteworthy event in the ARPANET era took place in 1982, when TCP/IP (see Section 2 for further discussion) was established as its standard protocol suite. TCP/IP thus pioneered the open systems approach to networking, which allows very different systems to communicate, and lead to one of the first definitions of an internet as a set of connected networks, and of the Internet as the set of connected TCP/IP networks using a common network addressing scheme. Packet switching, allowing the cost-efficient sharing of network links, and network protocol standardization greatly contributed to the ARPANET's growth. Further factors were the proliferation of Local Area Networks (LANs, networks linking computers in the same or in adjacent buildings on a campus) in the 1980s, the demand for network access from groups outside the defense sector, and the development of regional and special purpose networks requiring greater geographical reach.

1.2 NSFNET

The National Science Foundation (NSF) created the NSFNET in 1986 with the aid of NASA and the Department of Energy to connect the nation's research community to the newly created supercomputer centers. The initial plan was to build on the increasingly popular ARPANET, but due to bureaucracy in the Department of Defense, the NSF decided to bypass the ARPANET and build its own network. The ARPANET subsequently lost significance and was decommissioned in 1990. The NSF grouped the networks that were used to provide supercomputer access into three tiers (see figure 1):

1. The NSFNET Backbone\(^2\),
2. Regional or mid-level networks, and
3. Campus networks.

---

\(^1\) The capacity of a line, expressed in terms of the maximum number of bits it can transmit per second. Current bandwidths are denoted by Kbps, Mbps, and Gbps, which are abbreviations for Kilo (thousand), Mega (million) and Giga (billion) bits per second.

\(^2\) In internetworking, a backbone is a network linking together other networks.
Figure 1: The original NSFNET three tier structure. At the top level is the NSFNET Backbone, to which each mid-level network is connected. The campus networks have in turn connections to a mid-level network and sometimes make additional connections. Neighboring mid-level networks also interconnect directly.

The NSFNET Backbone was the first to connect the supercomputer centers and the mid-level networks. In contrast to other federal agency networks also created in that time period, it was accessible to the academic and research community at large and originally consisted of 56 Kbps lines interconnecting the six NSF supercomputer centers.

The evolution to a competitive market of commercial Internet Service Providers (ISPs) started in 1988, when the NSF signed a five year agreement with a consortium consisting of Merit (a network provider connecting schools in Michigan), IBM and MCI to upgrade the backbone to 13 computers connected by T1 lines (a T1 line has a bandwidth of 1.544 Mbps; see “A Note on Computer Network Technology”) and take over its operation.

In 1990, Merit, MCI, and IBM created Advanced Network and Services (ANS), to which Merit subcontracted the operation of the NSFNET backbone. The NSF was no longer the owner of the network, but rather a customer of ANS. In 1992, ANS installed a backbone consisting of T3 lines (44.736 Mbps) linking 16 computers. One of the main reasons for founding ANS was to start the transition from public to private funding of the Internet.

ANS CO+RE, a profit-oriented subsidiary of ANS, was created in 1991 to offer businesses network access without the restrictions imposed by the NSF’s acceptable use policy (often abbreviated as AUP). This policy came into existence in 1987 and prohibited commercial traffic on the NSFNET backbone and other portions of the Internet that were financed by federal funds. The AUP was initially very strictly interpreted, banning traffic that did not involve research or instruction at non-profit institutions in the U.S., and commercial (profit-oriented) uses not consistent with this overall NSFNET orientation. Exceptions had to be approved on a case-by-case basis. This interpretation eased up over time, as ANS CO+RE and
other commercial ISPs made in effect the NSFNET accessible via their privately funded networks to commercial users (see below). The government also adopted a more favorable attitude toward NSFNET use by businesses, motivated for example by the benefits of obtaining value-added information services via competitive procurement.

Several mid-level or regional networks were established to connect university and other campus networks to the NSFNET backbone. They were generally operated by one or several state agencies as nonprofit organizations. Some received subsidies from the NSF or from state governments. A large share of their funding came from customer connection fees. Over two dozen of these traditional regional networks offer Internet services today.

Among the key developments since the creation of the first mid-level networks are the rapid growth in network usage, the change in their orientation from a technical focus to market awareness, and the emergence of private, commercial ISPs.

The focus of regional networks has evolved from a technical focus in 1985-1988, when the operational aspects of network service provision were of prime concern, to market and service awareness from mid-1988 on, when the user market became increasingly important. Starting in 1987, the proliferation of personal computers and LANs created increasing demand from a much more diverse group of users, e.g., to access libraries, for medical applications, and to foster technology transfer from basic research to marketable products.

The regional networks can be regarded as the precursors of today’s commercial ISPs. Some regionals started to outsource their network operations to commercial network providers and focused on delivering value-added services. The general success of the Internet also spurred demand for access not restricted by the NSFNET’s acceptable use policy. These forces and the lack of federal and state regulation for Internet access led at the end of the 1980s to the creation of the first commercial ISPs that have since rapidly proliferated. Performance Systems International agreed in 1989 to provide and operate the infrastructure for NYSERNet (the New York State Education and Research Network), setting the first example of a commercial provider emerging as a spin-off from a nonprofit regional network.

In order to exchange commercial traffic among their networks without using the NSFNET - at the time subject to a restrictive acceptable use policy - General Atomics (CERFNet), Performance Systems International (PSINet) and UUNET Technologies (Alternet) established in 1991 the Commercial Internet Exchange (CIX), a trade association open to all commercial Internet TCP/IP network providers. The CIX sets a fixed and equal tariff at which all member networks agree to carry each other’s traffic.

Today, hundreds of commercial ISPs, ranging from small semi-professional organizations to telecommunication giants such as AT&T, MCI and Sprint, are offering Internet services. The CIX has presently dozens of members through which numerous businesses can be reached.
The third tier of the Internet consisted of academic and industrial campus networks (internal LANs), connected to one of the mid-level networks. An NSF grant program was available to support colleges and universities to connect to the NSFNET.

The number of connected networks has been growing exponentially over the last few years, as has the number of host computers. Many of the new users are commercial enterprises that could connect because of the increased presence of commercial ISPs and a looser interpretation of the acceptable use policy over time. Many K-12 schools and public libraries have also connected to the Internet.

1.3 Post-NSFNET Developments

Since April 30, 1995, a new Internet structure (see figure 2), conceived by the NSF, has been in effect. It maintains the general 3-tiered hierarchy of national backbone, regional networks and campus networks, but features the following changes:

- The NSFNET backbone was dismantled. The remaining federal agency networks and several so-called Network Service Providers, which are commercial national backbone networks, are interconnected through four Network Access Points, one each in Chicago, New York City, San Francisco and Washington, D.C. In addition to ANS (acquired by America On-Line), current major Network Service Providers in the U.S. include NETCOM, PSI, UUNET Technologies, MCI, Sprint and BBN (Bolt Boranek Newman) (see Section 6.1).

- The Network Service Providers carry the traffic of regional networks which must connect to the Network Access Points to receive NSF funding during a four year phase-out period.

- The NSF will put in place a very high-speed Backbone Network Service (vBNS), operating at least at a bandwidth of 155 Gbps, over the next five years in partnership with MCI, to link the national supercomputer centers and as a testbed for research and development in high-speed computer networking. This backbone is funded by the NSF. It will connect to the other networks through the four Network Access Points and will be the only portion of the Internet subject to an acceptable use policy restricting commercial traffic.
2. Elements of Internet Technology

The Internet is a network of interconnected **LANs, Metropolitan Area Networks (MANs, spanning entire cities but based on LAN technology) and Wide Area Networks (WANs, networks spanning large areas such as countries or the entire globe)**. The resources of the national and regional networks typically consist of a **Network Operation Center (NOC)**, **T1** and **T3** communication lines (also referred to as **circuits** or **links**) leased from telecommunications carriers\(^1\), **routers** (special computers used to interconnect the lines), and networking software to manage the communication tasks among these devices. The NOC consists of hardware and software for monitoring the network, detecting outages and initiating trouble shooting actions. It runs around the clock. Communication **protocols** provide the network functions enabling computers to “talk to and understand each other” on the Internet.

---

\(^1\) Sprint, MCI, AT&T and some other ISPs own their lines.
Key elements of Internet Technology briefly discussed here include packet switching, internetworking, the Internet Protocol and the Transmission Control Protocol, the client-server computing model, the Domain Name System, and encryption technology.

2.1 Packet Switching

Packet switching is the basic communication technology used in the Internet for exchanging data between each pair of communicating computers. It divides data at the sending computer into variable-length packets (also called datagrams in the Internet), which contain in addition to the actual data exchanged by users control information such as the Internet addresses of the source and the destination computers. Packets are sent independently through the network (see figure 3) and reassembled at the destination into the original message intended for the recipient. Bandwidth is only tied up when needed for a packet transmission. This approach takes advantage of the intermittent, or bursty, nature of Internet traffic, thus enabling multiple communicating computers to efficiently share the network. However, because links are shared, packet switching suffers from potentially large delays and lost packets if the network is congested, and requires complex processing for routing packets through the network.

Figure 3: Packet switching in the Internet. Data is transmitted from A to D, A to E, and F to C. The data packets belonging to the same source-destination pair are routed independently, using the address information they carry (represented by the letters written in the data packet symbol). Therefore, they may have to be reordered and reassembled at their destinations (i.e., at C, D and E). Bandwidth is only tied up for the duration of the packet transmission but can otherwise be shared with other transmissions.

---

1 For more details, see “A Note on Computer Network Technology.”
2 This is in contrast to circuit switching, used in the telephone network, where a dedicated physical circuit through the network is set up before any data is sent and it remains active for the entire duration of the connection (see “A Note on Computer Network Technology”).
3 Periods of high data transfer rates are followed by relatively long periods during which no data is transmitted.
2.2 Internetworking

Although the Internet appears to users as a single large network, it is in fact an internetwork, i.e., a network of interconnected networks. The individual networks don't all use the same packet switching technology, as each has been designed to meet its own speed, distance and cost constraints. All types of networks are interconnected using specialized computers, called routers, that pass packets from one network to another. (For more details, see "A Note on Computer Network Technology").

2.3 Internet Protocol (IP)

The Internet Protocol (IP) is one of the two key communication protocols of the Internet. Its main purpose is to make the numerous heterogeneous networks that form the Internet function as a single homogeneous network. This involves defining the format of Internet packets, called datagrams, a common network addressing scheme, and the rules used for routing the datagrams independently from any source to any destination on the Internet. Any computer on the Internet must have IP software, as all Internet Services use it for sending packets across the network. In order to make routing and delivery possible, every host computer on the Internet is assigned a unique numeric Internet address, often referred to as its IP address.

2.4 Transmission Control Protocol (TCP)

Transmission Control Protocol (TCP) is the other key communication protocol of the Internet. It makes the Internet reliable by solving several problems that can occur in a packet switching network. While IP enables any computer on the Internet to send data to any other computer on the Internet, it is quite "negligent" as datagrams can be lost, delivered out of sequence, or delivered multiple times. TCP makes up for these imperfections by resending data when a datagram is lost, reassembling datagrams in the right order at their destination, and eliminating multiple copies. The current version of TCP offers only one class of service, in which packets are delivered as quickly as possible but without any delivery or maximum delay guarantees.

IP and TCP typically work together and are usually referred to as the TCP/IP protocol suite. All Internet Services use TCP/IP to reliably send data across the Internet.

2.5 Client-Server Computing on the Internet

The software for all Internet services is based on the client-server computing model. Initially restricted to the LAN environment, client-server computing has been propelled by the Internet's success to
the world of Wide Area Networks. In this model, each program is classified and operates either as a client or as a server (the term server is used to refer both to the server software and to the computer hosting it). A server program, e.g., a World Wide Web (WWW) server (see Section 4.3), is typically continuously active, ready to accept a request from client programs. Users who want to access a server program run client software only when needed. The client program running on a user's computer, e.g., the Netscape Navigator you are using on a personal computer in the MBA lab, sends a request to the server program, e.g., the GSB’s WWW server, and displays the result, e.g., the GSB Home Page (a special file containing information about the GSB), on the user's computer.

2.6 Domain Name System

While communication protocols identify computers on the Internet by their numeric IP addresses, people use so-called domain names for this purpose. Domain names contain information similar to postal addresses and have a special structure. In the U.S., they typically consist of at least three parts with the following general format:

```
local.organization.domain.
```

For example, your E-mail account is handled by a computer named gsb.stanford.edu. A user’s E-mail address on a computer has the format:

```
user@local.organization.domain,
```

where user may or may not be that user’s real name. For example, people can send E-mail to your personal GSB mailbox at lastname_firstname@gsb.stanford.edu.

The domain is the part of the domain name denoting the organization type. In this example, the domain edu stands for educational. Other domains in use in the U.S. include com for commercial organizations, gov for government, mil for military sites, net for administrative hosts in a network, and org for other non-profit organizations. Most other countries follow alternative schemes, typically appending the 2-letter country code to each domain name (for example, uk for the United Kingdom).

Each domain name corresponds to an IP address. The domain name system (DNS) is the software that translates a domain name to an IP address. It is decentralized, based on the client-server model described above. Each organization operates a domain name server that maintains a list of all its computers along with their domain names and IP addresses. Internet services contact domain name servers as clients when they need to translate a domain name into an IP address. Local and remote domain name servers cooperate to provide the address information of remote sites to the local user.

2.7 Encryption Technology

---

1 To make sure that names are unique, additional parts may be added while maintaining the general format.
In general, encryption technology, also called cryptography, refers to techniques used to make messages that are exchanged between two parties (man or machine) unintelligible to all but their intended users. It is becoming increasingly important for the secure transmission of business records and other confidential information over the Internet. Encryption is based on the following principle. Messages to be encrypted, called plaintext, are transformed by an encryption key into ciphertext. The ciphertext is then transmitted over the network to the intended recipient who deciphers it using the encryption key. While an unintended recipient, or intruder, can possibly access the entire ciphertext, s/he cannot decrypt it easily since s/he is not in possession of the encryption key.

3. Internet Access

Internet access requires hardware and software for physical and logical connectivity.

3.1 Physical Connectivity

Physical connectivity includes the selection and configuration of access lines (also referred to as circuits or links), computers and other equipment such as for example a modem, which allows computers to exchange data over a regular telephone line. The basic access options include dial-up (or switched) and leased (or private) lines. With dial-up lines, access can be established with a phone call when needed for actual Internet use while a leased line provides permanent access irrespective of whether the connection is needed during the entire leasing period or not. The two key selection criteria for access lines are their pricing and bandwidth. In order to choose an appropriate line, customers should estimate their bandwidth requirements by determining what applications they intend to use, the frequency of use, the number of users and so forth. Dial-up connections are very cost efficient but only provide sufficient bandwidth for connecting stand-alone computers or the smallest LANs, offering up to 28.8 Kbps using a modem and a regular telephone line and 64Kbps or 128 Kbps via digital ISDN (Integrated Services Digital Network) connections (for technical details, see “A Note on Computer Network Technology”). Dedicated lines, with access speeds ranging from 56 Kbps to 44.736 Mbps (for T3 lines), are better suited for companies with substantial demand for bandwidth. However, they may not be very cost efficient compared to dial-up lines if the network use is characterized by bursty traffic (see Section 2.1.) Some recent access options, such as Frame Relay (for a detailed description, see “A Note on Computer Network Technology”), offer the flexibility to accommodate high peak data transfer needs in a more cost efficient way than regular leased lines.
3.2 Logical Connectivity

Logical connectivity involves getting an IP address (see Sections 2.3 and 2.6) and setting up various communication protocols. These include:

- Software performing the functions for controlling modems and other devices that serve as network interfaces between computers and communication lines;
- The Internet Protocol (IP) and Transmission Control Protocol (TCP), the software for managing the exchange of data between source and destination computers over the network (see Sections 2.3 and 2.4);
- *Internet services*, which are application programs for using the Internet to exchange messages and to post and retrieve information. These include *electronic mail (E-mail), File Transfer Protocol (FTP), Telnet, the World Wide Web (WWW)* and many others, which are discussed in the next Section.

4. Internet Services

There is a large and growing variety of services on the Internet. We describe the major ones below\(^ 1 \) and include some hands-on exercises on their usage in Section 5.

4.1 Basic services

These were available since the beginning of ARPANET and were essential for the collaboration among researchers at different institutions. They include:

- *Electronic mail (E-mail)* is probably the most popular network application across all user categories. By now you are probably familiar with E-mail through use of the *Microsoft Mail* program at the GSB, which allows you to edit, send, receive and store messages. You can receive messages in your personal mailbox at the address *lastnameFirstname@gsb.stanford.edu*. In addition to sending a typed message to one or more recipients, you can also include a separate text, voice, video, or graphics file in your message. For example, *Microsoft Mail* allows you to attach a *Word, Excel or Powerpoint* file to a message for one of your classmates, and other E-mail programs have similar functionality.

- *File Transfer Protocol (FTP)* is the basic service for transferring, or copying, files efficiently on the Internet from a remote to a local host (*getting a file*) or vice versa (*putting a file*). The use of FTP involves the following steps: *(i)* logging in to (accessing) a remote computer; *(ii)* locating the desired remote file or the remote directory where to put the local file; *(iii)* performing the transfer; and *(iv)* terminating the connection. Many computers support *anonymous FTP*, in which any user on any

---

\(^1\) The Internet Services are grouped here according to the classification used by the Internet Society, a nonprofit organization established to foster exploration of the Internet.
Internet computer can access these files by logging in using the user name “anonymous” and an arbitrary password. Note that while FTP allows you to interactively search for file names through a remote computer’s directories, it is not very user-friendly. It does not describe the file contents, and you can only view or edit files after you transfer them to your local machine.

- **Telnet** permits users to login to a remote computer which can then be used just like a local computer. Unlike E-mail and FTP, Telnet provides fully interactive "remote control" in which the remote computer reacts to the instructions of the remote user and displays information on the local computer’s display. Telnet can also be used to access other Internet services not available on the local host. Exercise 1 in Section 5 illustrates the use of Telnet on an MBA lab computer at the GSB.

### 4.2 Mailing Lists and Bulletin Board Services

While the basic services typically provide for one-to-one communications, mailing lists and bulletin boards are suited for many-to-many communications. Mailing lists are either managed by users or automatically by *LISTSERV*.

Particularly popular is *Netnews* *(for Network News)*, an electronic bulletin board service originally established in 1979 between Duke University and the University of North Carolina, which today comprises thousands of individual discussion groups called *Newsgroups*. The term *Usenet* refers to all sites that maintain and participate in the exchange of Network News. Newsgroups are created and accessed by millions of users around the world. They each have a unique name consisting of character strings separated by periods, much the same as the structure of computer domains. Newsgroups are classified based on the topics they cover. Some of the major Newsgroup categories are:

- **alt** Controversial or unusual topics
- **biz** Business
- **clari** Newsgroups from ClariNet, a commercial news service
- **comp** Computer hardware and software and related subjects
- **misc** Discussions not covered elsewhere
- **news** News about Netnews itself
- **rec** Arts, hobbies, games, recreation
- **sci** Science
- **soc** Social groups and cultural issues.

Not all Newsgroups can be found on every site. The Newsgroups carried by a particular site are determined by the policies of the organization that manages the site. Users can read articles using a *newsreader* program, or post articles either by E-mail or directly, using a special editor. They may put forward an idea, ask a question or respond to somebody else’s post. Copies of new articles are not immediately posted to all Usenet sites maintaining the relevant Newsgroups. Rather, programs on each
site periodically contact other sites to automatically update their Newsgroups. Exercise 3 in Section 5.1 illustrates the use of Network News.

**4.3 Interactive Information Delivery Services**

*Gopher* and the *World Wide Web (WWW, or in short, the Web)* are interactive services providing more sophisticated ways to organize and deliver information than is possible with FTP. Both rely on a network of information servers which are accessed from other computers by a special type of client software, called *browsers* or *navigators*.

Gopher was developed by the University of Minnesota at Minneapolis/St.Paul in 1991. It uses a simple text-based, menu-driven browser, which permits users to search through and read files on remote computers.

The WWW was released in 1992 by CERN in Geneva, Switzerland. In the same year, the National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign released *Mosaic*, a WWW navigator with a graphical user interface. The WWW enhances information delivery in two significant ways compared to Gopher. First, while Gopher separates menu information and documents, the WWW uses a technique known as *hypertext* to integrate menu items and information. As a result, information on the WWW need not be structured in a one-dimensional hierarchy, but can be organized and linked along multiple dimensions. Second, text, graphics, audio, and video can be combined to form *multimedia* documents in the WWW, while Gopher consists of textual information only. Because of these two characteristics, the WWW is also referred to as a *hypermedia* system. The WWW plays a growing role in emerging applications. Thousands of businesses and other organizations make information accessible to the Internet public on Web servers. Mosaic was considered a "Killer Application" because of its immense success. A number of navigators have since come to market, with the *Netscape Navigator (Netscape)*, released last year, now having a 75% market share. They allow access to data not only on WWW servers, but also on FTP and Gopher servers. Netscape also serves as an interface to Network News (see exercise 3) and E-mail. It has a link to a number of third party search mechanisms, also called *search engines*, to facilitate searching the vast amount of information on the WWW (see exercise 5). Exercises 2-5 in Section 5.1 illustrate the use of the Netscape Navigator.
4.4 Directory Services

The enormous number of Internet users makes the development of directory services both a challenge and a requirement. One way to find people on-line is through a product called *Netfind* (developed at the University of Colorado at Boulder), essentially a self constructing database which periodically searches Newsgroups and public E-mail servers to update its content. AT&T is developing a directory service modeled after the phone book. Other services, completed and in progress, include *Netpages*, *Whois*, and *X.500*. Last but not least, the *InterNIC*’s (Internet Network Information Center) *Information Services Reference Desk* also contains directory information. All of these products notwithstanding, it is still difficult to find the address of people on the Internet.

4.5 Indexing Services

Indexing services provide access to indices that are stored and updated on numerous servers across the Internet. *Archie* is a distributed index to the names of all files that can be downloaded using the FTP service and is operated by McGill University in Montreal. *WAIS*, the network of Wide Area Information Servers, is a distributed system for keyword searches of document contents. *Yahoo*, started by two Stanford students, indexes the WWW by topics such as arts, business and economy, government, and science, and also includes a search engine for detailed keyword searches.

4.6 Trends

New services are emerging, fueled by the needs of the growing commercial Internet constituency and by new technologies:

- The immense quantity and complexity of information available on the Internet leaves room for better tools to search for individuals, organizations and information. Some software companies develop “intelligent” software agents, so-called *Knowbots*, to take over our information retrieval tasks. For example, search engines such as *Excite* from *Architext Software* enable users to search for information not only by keyword but also by concept and have additional “intelligent” searching capabilities of the kind we expect to see in the future.

- The advent of commercial users and electronic commerce creates the need for secure electronic payment systems using encryption technology (see Section 2.7). There is also a high demand for the use of encryption and related technologies for signature verification, copyright protection and privacy enhancement.

- New applications will increasingly involve multimedia documents and user-friendly graphical interfaces, thus further increasing the number of users and the traffic each user generates on the Internet.
This will sustain the current high Internet traffic growth rates for some time to come. We observe here the same phenomenon as with computer systems: as the systems get faster processors and more memory capacity, applications developers create new software products that require ever more processing and memory resources and push the new hardware generation to its limit.

- Many Internet applications being designed today build on the services described in this Section by making them more secure, giving them a more friendly user interface or by automating some of the tasks that were previously performed by the user. CommerceNet, the first large scale electronic market, is based on the WWW and Netscape, to which services and standards for directory services, secure transaction enhancements, authentication and payment mechanisms are added.

5. Hands-On Exercises

The following hands-on exercises are not intended as a comprehensive Internet tutorial. Rather, they are designed to familiarize you with some of the services available on the Internet. A wealth of Internet Tutorials can be found on-line.

Exercise 1: Using a remote computer interactively through Telnet

This exercise should be carried out using the software you find on the Macs and PCs in the MBA lab, as the instructions have been tailored to this specific environment. As mentioned in Section 4.1, Telnet enables you to use a remote computer just as if it were local. Of course, you can only login to a remote computer if you have the authorization to do so. You must either have an account on the remote computer (e.g., you may have an account on Stanford University’s leland system which is located at Sweet Hall a few blocks from the GSB) or the remote computer must be explicitly made accessible to your organization or to the general public.

Below are instructions for establishing a telnet session to forsythetn, a Stanford University computer offering information services to the Stanford community.

a) Using Microsoft Windows on a PC:
1. Double-click on the Telnet icon in the Main window. One of two things will happen: If you are prompted for Host Name, go to step 3. Else, a window named Samson for Windows will open. In this case go to step 2.
2. Select Session and New Telnet from the menu. Another window named New Telnet will open.
3. In the Host field enter forsythetn (since you are accessing it from Stanford, you need not specify its full Internet address, forsythetn.stanford.edu) and click on the OK button.
4. Once the connection is established, you are automatically logged in. Once logged in, enter folio at the 
Account? prompt, then your university ID, and finally enter select. This will allow you to browse 
through catalogs of the Stanford University libraries and the Stanford bookstore as well as access 
information on housing, University Policies, public events etc.

b) Using a Mac:
1. Double-click on the MacSamson (Telnet) icon in the Internet Tools window. A window named Telnet 
Session will open.
2. In the Host [port] field enter forsythen (since you are accessing it from Stanford, you need not specify 
its full Internet address, forsythen.stanford.edu) and click on the OK button.
3. Once the connection is established, you are automatically logged in. Once logged in, enter folio at the 
Account? prompt, then your university ID, and finally enter select. This will allow you to browse 
through catalogs of the Stanford University libraries and the Stanford bookstore as well as access 
information on housing, University Policies, public events etc.

5.1 Using the Netscape Navigator

The World Wide Web and multimedia navigation tools such as the Netscape Navigator 
(Netscape) allow you to use the Internet with a graphical user interface unlike the text-based access formerly 
available. The World Wide Web organizes information on pages that can have multiple links to other 
pages. The home page of individuals or organizations serves at least two purposes. First, it is their 
WWW identity, and second, it is a reference point for other pages that a person or organization puts on the 
Web. In the following exercises, you will explore the use of the Network Navigator, both to read Network 
News and to search for and access Web pages.

To start Netscape, simply double-click the Netscape icon on your PC or Mac. If you are doing 
this exercise in the MBA lab, a window named Netscape [Stanford Graduate School of Business: Home 
Page] will open on your screen displaying the GSB home page. If you are not using a computer in the 
MBA lab, Netscape will most likely display another page. The instructions for this exercise are valid for 
the Netscape versions 1.1N and 1.2N, which are available in the MBA lab. For information on how to use 
later versions, ask your Internet Service Provider or the MBA lab consultant.

Each page has a Uniform Resource Locator (URL), which denotes its exact WWW location or 
“address”. The URL is usually displayed in the rectangular field located in the top section of your screen 
just underneath the row of the buttons (from left to right) Back, Forward, Home, etc. We subsequently 
refer to this field as the URL field. If you cannot see the URL field, select Show Location from the Options 
menu option on the Netscape menu bar. The URL field should now be visible. The GSB home page’s 
you started Netscape in the MBA lab, this should now be displayed in the URL field. If this is not the 
case, type http://gsb-www.stanford.edu/ in the URL field and press enter to access the GSB home page.
Whenever you enter a new URL, make sure that you first delete the old URL and that you use the exact spelling and case for the entered URL as you are given. URLs have the following general format:

\[
\text{<transfer protocol>://<Internet address>/<directory and page name>},
\]

where the transfer protocol specifies the rules used for communication between Netscape and the server program involved. The WWW protocol is the hypertext transport protocol (http). As mentioned in Section 4.3, the Netscape Navigator can also be used to access data stored on FTP and Gopher servers, which are based on different protocols. We will see an example for this in Exercise 2.

Recall that in contrast to Gopher, the WWW is based on multimedia documents and the hypertext technique (see Section 4.3). The former is immediately apparent on the GSB home page, as the image-based GSB logo is included on the same document as text. To explore the WWW’s hypertext aspect, click on the GSB logo, or any text that appears highlighted (in blue on this page). Netscape will take you to a page related to the GSB logo or the selected text, from which you can either travel back to the home page or continue your journey, clicking your way through the entire Web.

There are three basic methods for finding pages on the WWW:

1. The first consists of explicitly telling Netscape a page’s URL. It is the most focused but has the disadvantage that you must know a page’s exact URL to be able to access it.
2. In the second method, you use Netscape features to revisit pages without having to repeatedly enter their URL’s. For example, Netscape allows you to mark pages you are likely to revisit in the future using Bookmarks (see the menu bar). It also keeps independently a short history of the pages you recently visited, with functions such as View History under the Go menu option and the Back and Forward buttons that allow you to navigate in both directions.
3. In the third method, a page is accessed indirectly through a reference from some other page. This may be its related home page, a page from a WWW index, or some other document.

Learning to navigate the Web with skill comes very naturally with experience and does not require structured instruction. If you insist on a more methodical approach, you can access a Netscape Tutorial by selecting Handbook from the Help menu option on the menu bar\(^1\). Otherwise, just spend some time exploring the different menu options. The following exercises will help you get started.

---

\(^1\) In Netscape 1.1N, you should instead click on the Handbook button located just underneath the URL field.
Exercise 2: Using Netscape to download a file with a known location from a remote computer

Now that the 1996 presidential race has taken off, wouldn’t it be interesting to review President Clinton’s 1993 inaugural address? The text of his address is contained in the file clinton.in, located on the FTP server ftp.spies.com. You want to download it from this server to your computer. To that end, take the following steps:
2. Choose Save As from the File menu.

You can now read the file using any text editor or word processor.

Exercise 3: Accessing Network News

Netscape also serves as an interface for accessing Newsgroups. To check them out, click the Newsgroups button, located right underneath the URL field. As an exercise, first push the View all Newsgroups button at the bottom of the new page. This will take you to a page titled Newsgroups list. Find the clari.biz.briefs group and read any article by clicking on it. Feel free to search through this and other groups, subscribing to any group you like. Subscribing to a group in Network News is much like marking a WWW page in that it will spare you the effort of retrieving it in the future.

Exercise 4: A few selected World Wide Web sites

Before browsing aimlessly through the Web, here is a selection of URLs of a few interesting or particularly well-designed sites:

- Books by Mail: The world’s (apparently) largest bookstore, with over 1 million titles: http://www.amazon.com.
- And, last but not least, (the URL says it all): http://www.whitehouse.gov.

---

1 This may result in a message like "Before you can view all newsgroups, Netscape must save a copy of the Newsgroup list. If you are using a modem or are on a slow connection, this may take a few minutes... Proceed?" If you get such a message, click the "Yes" button and proceed.

2 Some of them were cited as the best home pages of their kind in the January 1996 issue of Internet World.
Exercise 5: Finding Information on the WWW

Oftentimes, you don’t know what page contains the information you are after, and even if you do you may not have its URL. In this case, you can search the WWW using a number of so-called search engines. You can directly access the infoseek guide search engine by simply clicking in Netscape on the Net Search button located just underneath the URL field. In addition to the search engine appearing at the top of your screen, you will find references to alternative search engines and indexing services, including Excite and Yahoo, at the bottom of the screen page. Use of the search engines is intuitive, and you may want to look at their help screens as well.

Using one or more of these search engines, answer the following questions. Always find the most recent information and indicate your source’s URL and how you found it.

1. What is the average age of the U.S. Web user? What about that of the European Web user?
2. How many networks are part of the Internet? How many of these are U.S.-based, and how many are in other countries?
3. How often is the host name ‘snoopy’ used on the Internet?
4. What are the costs of leasing a T1 line? A T3 line?
5. How many hosts were on the Internet in October 1990?

6. Emerging Issues in the Development of the Internet

6.1 The U.S. Internet Service Provider Market: Status and Trends

Market Size and Growth

According to 1994-1995 Internet Service Provider Marketplace Analysis (published by the Maloff Company, http://www.trinet.com/maloff/), the U.S. ISP market grew from an estimated annualized revenue of $118.8 million in March 1994, to an estimated annualized revenue of $521.41 million (or $1.071 billion including the revenues of America On-line, CompuServe, Delphi and Prodigy attributable to their ISP activity) as of January, 1995. This corresponds to a monthly revenue growth rate of 28%! These numbers are for Internet access services only and exclude revenues from consulting, training, security services, hardware, and “CyberMalls” (Electronic Shopping Malls).

Market Players

The ISP market is presently very competitive with a few large and several hundred medium size and small providers. Traditional regional networks have been around for almost a decade, and they have accumulated substantial expertise in the operational aspects of data networks. The major online services¹,

¹ These provide information content in addition to Internet access.
America On-Line, Compuserve, Delphi, and Prodigy, and the three major common carriers, AT&T, MCI, and Sprint, recently entered the market, bringing a lot of financial resources and expertise in the area of long-distance telecommunications to the game. But at the same time, their traditional centralized structure may be inconsistent with the decentralized structure and informal culture of the Internet. Software companies have recently also shown interest, most notably Microsoft with its Microsoft Network. The majority of providers have been in the business for a very short time, and many new ones, especially smaller local ISPs, are continuing to enter.

Figure 4 shows 1995 revenue estimates for the top ten U.S. ISPs, not taking into account the major on-line services. According to the Maloff report, the rapidly expanding number of smaller providers now account for much of the revenue growth.

Figure 4: Estimated 1995 Revenues of Top Ten U.S. Internet Service Providers (Source: 1994-1995 Internet Service Provider Marketplace Analysis). This listing excludes the on-line services America On-Line, Compuserve, Delphi, and Prodigy.
Market Differentiation

Providers presently differentiate themselves in terms of geographical scope (most operate regionally, a few nationally), their access line speeds (from low-speed dial-up telephone lines to high-speed leased T3 lines) and the range of Internet services they provide (from E-mail only to full access to all Internet services), value-added services (e.g., system configuration, security, consulting and training) offered, the quality of and congestion on their network, and pricing schemes (fixed monthly fees are prevalent, but they are sometimes augmented by hourly and other charges.)

With the emergence of commercial providers, the boundary between backbone and mid-level networks has been blurred. For example, ANS and PSI offer both backbone services to other networks and regional connection services to individual users and organizations. While Network Operation Centers are generally located close to the network customers, the essence of networking makes geography irrelevant. That is, a network can easily grow geographically by extending the reach of its backbone. In deciding whether to compete locally or nationally for users, the revenue effect of serving a larger market and the existence of some scale economies have to be traded off against the cost increase due to additional lines and the possibly more diverse clientele.

In order to remain in the game as competition continues to intensify and telecommunication line costs keep falling, providers will likely have to diversify their services from access only to value-added and information services.

Exercise 6: Internet Service Providers

To get a feel for the variety of ISPs, check out a listing of U.S. and foreign Internet Service Providers at http://thelist.com. From there, visit a few to familiarize yourself with the different services and prices they offer.

Economies of Scale

Due to a number of factors, the average cost of operating network services is decreasing. First, there are considerable economies of scale due to the fixed cost of the Network Operations Center, which represents one of the highest components of a provider's start-up costs but only increases slowly as the network grows. Second, the cost per bit of capacity of a communication line is lower the faster the line. For example (see exercise 7), a T3 line (44.736 Mbps) costs significantly less per bit of capacity than a T1 line (1.544 Mbp.) Third, the variable costs of network operations make up less than half of total costs. Thus, network service provision is characterized by significant scale economies.
Funding

Many regional providers still receive government funds through the NSF and operate on a nonprofit basis. However, NSF funding is scheduled to cease in 1999. Almost all the emerging suppliers are commercial and privately funded. Since they do not get government subsidies, they are free to expand their operations if they find ways to finance them.

Emerging Market Structure

As outlined above, the ISP market is undergoing rapid change. New players are entering the market, new alliances and mergers are being formed and new services are launched at a breathtaking pace. This raises a number of questions regarding the future ISP market structure, including:

• For how long will the ISP market continue to grow at the current enormous rates?
• How long will the trend of strong market entry prevail and how will it affect the mix of national and local ISPs? Will the market soon be saturated with too many providers, turning the fixed start-up costs of network operations into an entry barrier, or will sustained market growth continue to absorb new entrants?
• Will new players enter from other industry segments? Some possibilities include cable TV networks, content providers or software companies that may follow the example set by Microsoft.
• What will be the level of concentration when the Internet Service Provider market reaches maturity? Countering the general trend of numerous players entering the market, the recent acquisitions of BARRNET by BBN Planet, and of ANS by America On-Line, as well as the entry of very large competitors such as AT&T and Microsoft, render the scenario of an industry shakeout with only few major competitors surviving quite plausible.
• Will ISPs increasingly integrate the supply of network transport infrastructure and information content, or rather specialize in one of these two areas?
• What will be the role of the traditional regional networks after government support ceases?

6.2 Congestion and Pricing

Until recently, much of the Internet infrastructure was funded by the federal government. Access was limited to the non-commercial sectors and hence the amount of traffic was modest compared to the present. With the proliferation of commercial ISPs and the development of new bandwidth-intensive multimedia applications, the number of Internet users and the traffic volume they generate is growing faster than the available bandwidth. While telecommunication technology is expanding the available bandwidth, it is unlikely that sufficient capacity can be economically supplied in the near future to keep up with traffic growth. The resulting network congestion, which you probably experienced during the course
of some of the hands-on exercises, is likely to prevail and even become more serious unless congestion is effectively controlled without generally curtailing Internet usage.

Some argue that the cause of Internet congestion is the currently prevalent scheme of connection-based pricing. Under this scheme, a user pays the network access provider a fee that is calculated based on the duration and the capacity of the connection. In the case of leased lines, these are typically monthly charges. In the case of dial-up lines, the customer pays a base fee for a certain number of connect hours, and possibly additional fees when that number is exceeded. (To check out various pricing schemes, you may want to review the ISP listing at http://thelist.com). As a result, the charges users incur are not very sensitive to the actual bandwidth they consume, and there are no mechanisms to prevent them from exploiting the available capacity. Among the approaches proposed for controlling Internet congestion, some involve charging users prices that are more sensitive to their actual network usage than only connection charges (see exercise 7 below.) Others are based on rationing schemes under which the network provider determines the amount of bandwidth each customer can use over a given time period.

Those opposing any form of usage control argue that it runs against the decentralized nature and the exceptional freedom that allowed the Internet to develop into the universal communication medium it is today and that the network has survived so far without any usage control mechanisms. They claim that relying on users’ voluntary efforts, combined with peer pressure and user ethics, should take care of the problem.

Exercise 7: Pricing to Control Internet Congestion
Evaluate the following alternative pricing schemes for solving the problem of Internet congestion:

1. Flat per bit price on bits sent: In addition to connection fees, users are charged a certain price for each bit sent. The per bit price does not fluctuate over time, and is the same for all users and Internet services.

2. Service-dependent per bit price: Same as in 1, except that different prices are set for different Internet Services. For example, a WWW bit may be priced twice as high as an E-mail bit.

3. User-dependent per bit price: Same as in 1, except that different prices are set for different Internet users. For example, businesses may be charged a different per bit price than residential users.

4. Priority-dependent connection prices: A number of priority service classes are defined. Users pay connection charges only, but can select among the offered priority classes. Higher-priority users are charged a higher connection fee than low-priority users. When data packets compete for the same network resource, the higher-priority packets are forwarded before lower-priority packets.
6.3 Impact of Electronic Commerce

The emergence of electronic commerce, especially in the form of electronic markets interconnecting the majority of businesses in many industries, is likely to significantly cut the time and cost involved in acquiring, disseminating and processing information relevant for various kinds of business transactions. It is likely to:

- Shorten procurement cycles through on-line catalogs, ordering and payment, to cut costs through competitive bidding, and to shrink development cycles and accelerate time-to-market through collaborative engineering and product implementation.
- Reduce the incentives for integrating different business functions into one organizational entity, as the required information can now flow very fast between different organizations. This could lead to vertically less integrated firms and the emergence of so-called virtual organizations, i.e., flexible, decentralized systems in which entities with different core competencies temporarily cooperate to produce a new product or service.
- Allow small organizations with limited resources to achieve market presence worldwide as the costs of information dissemination and connectivity are drastically reduced.
- Create more efficient markets and lower prices by intensifying competition.

6.4 Hurdles to Commercial Use

In the course of the last two decades, most large and many smaller businesses have established enterprise networks (also called intra-organizational networks or Intranets) linking their geographically dispersed sites, and also inter-organizational networks with their suppliers, partners and customers. Since commercial use of the Internet was prohibited until the early 1990s (see Section 1.2), this was done by setting up a private data communication infrastructure, based on lines leased from telecommunications carriers and equipment selected from a variety of vendors, and/or by signing up with public data network services with firms like Tymnet, ADP Autonet and GTE Telenet. Intra- and inter-organizational networks are being used as a basis for company-wide logistics systems, manufacturing planning and inventory control, point of sales systems, electronic banking, and so forth.

While the Internet is now open to commerce and an increasing number of businesses are exploring its substantial commercial potential in some fashion, most are hesitant to use it as a replacement for their established intra- or inter-organizational links. This is partly due to security, information retrieval, payment system, and network management and control problems, which apparently still offset the Internet’s cost and connectivity advantages we discussed in the last Section.
Security

Security is one of the major reasons many businesses are skeptical about joining the Internet. When the Internet first emerged as a research network with access restrictions, there was no serious need for sophisticated security mechanisms. As the Internet is now an open system that is accessible worldwide to the general public and electronic commerce, strict security measures have to be sought.

Network security includes physically protecting data against unwanted manipulations, verifying the identity of users that are granted access to data (user authentication), and preserving privacy and confidentiality by making sure that data is only accessible to its intended recipients. Solutions include firewalls, application-layer gateways (e.g., Sprint Plus, ANS Interlock), and virtual private networks (e.g., PSI Secure Stream, ANS Key Ring, UUNET LAN Guardian). Firewalls are created using a pair of routers with a host computer between them. The router on the Internet side lets the host freely communicate with the Internet while the other router prevents unauthorized traffic from reaching the corporate network. An application-layer gateway interacts on behalf of the users with the Internet. Users must authenticate themselves to the application layer-gateway in order to access its services. The greatest security level is achieved by defining a virtual private network, access to which is protected by encryption technology (see Section 2.7). In choosing among these solutions, one has to trade off security against ease of access and use. While technical solutions for network security problems may be available, it will take some time until they have gained the trust of business decision makers.

Information Retrieval

The task of finding information and Internet users was and still is very difficult, as you may have experienced the course of the hands-on exercises. Due to the decentralized nature of the Internet, the information available is rarely well structured and referenced. New concepts and tools for structured, consistent distributed information storage and retrieval are needed.

Payment systems

In comparison to Electronic Data Interchange (EDI) systems, which electronically integrate business transactions with the related payment procedures, the Internet, with a few exceptions, has yet to provide the same capabilities to its potential customers. Several products are in the making, e.g., DigiCash and CyberCash, but established standards have not been formed to date.

Network Management and Control

In the Internet with its decentralized network management and control structure, network congestion and reliability problems (transmission errors, data loss, network outages, and so forth) are beyond the control of corporate network managers. These can only select their access equipment and ISP, but neither they nor any other single institution can possibly track each Internet segment that their business
data will traverse on its way between the source and the destination address. It is possible that advances in network technology will in the future extend users’ ability to remotely control relevant portions of an internetwork that is owned and operated by multiple entities.

In spite of all these problems, the use of Internet technology, both for inter-organizational networks and Intranets is growing. Today, companies such as Federal Express, Hewlett-Packard and Silicon Graphics, just to name a few, view Internet technology as a substitute for their proprietary networks. Further technology advances, driven by the vast promise of the Internet for electronic commerce, will likely provide ways to overcome current and future obstacles.