Telecommunications and the Next Generation Internet for Health Care

Craig Feied, MD

From the National Center for Emergency Medicine Informatics; the Department of Emergency Medicine, Washington Hospital Center; and the Department of Emergency Medicine, The George Washington University School of Medicine and Health Sciences, Washington, DC.

Received for publication May 2, 2001. Accepted for publication June 1, 2001.

Reprints not available from the author.

Address for correspondence: *Craig Feied, MD, Department of Emergency Medicine, Washington Hospital Center, Suite NA-1177, 110 Irving Street NW, Washington, DC 20010; 202-877-7574, fax 202-877-9263; E-mail cfeied@ncemi.org.*

0196-0644/2001/$35.00 + 0
47/1/117952
doi:10.1067/mem.2001.117952

See related articles, p. 286, p. 303, and p. 312.

[Feied C. Telecommunications and the next generation Internet for health care. *Ann Emerg Med.* September 2001;38:293-302.]

THE EVOLUTION OF CONNECTIVITY

Before the advent of the telephone, information exchange was limited, with few exceptions, by the need for face-to-face conversations or the delivery of written documents. Lengthy delays, also referred to as high latency, attended information exchange across any significant distance, and these delays largely defined the communal life of a country in which news of a presidential election could take several weeks to cross the continent. Technological innovations such as the railroad and the telegraph allowed a more speedy exchange of written information, but it was not until the widespread adoption of the telephone that person-to-person interaction at a distance became possible. The telephone eliminated some of the natural constraints of distance and catapulted the world into the first era of rich information exchange.

This ability to exchange information can be called "connectivity"—the ability to give information that we possess to other people and to receive information that we want from other people. When we can exchange information more rapidly, exchange richer types of information, and exchange information with a larger universe of people, we enjoy a greater degree of connectivity. Since the introduction of the telephone, the maximum achievable degree of connectivity has increased dramatically; this has transformed a society that has become increasingly dependent on information transfers for the economic, political, business, and personal vitality of its citizens.

Initially, the telephone created small islands of connectivity with large gaps between them. In the 1870s, the earliest telephone installations connected up to 48 subscribers on a single exchange. All subscribers had to be within a fairly close geographic proximity, and there was no connection between any 2 exchanges. In larger cities, many different companies operated telephone exchanges, and calls could not be connected between competing companies. End-to-end voice connectivity came about only when one telephone company succeeded in obtaining a monopoly, eventually unifying the entire country with a single proprietary standard for telephone connectivity. It is now difficult to imagine life without telephones, but the public was initially slow to accept them. The first branch exchanges were installed in 1877, and it was not until 50 years later that the number of telephone subscribers reached 20 million.[1]

In contrast with hard-wired telephony, mobile cellular telephony was adopted very rapidly. Although at one time they were considered so unlikely a concept that they appeared only in science fiction, 20 million cellular telephones were installed in the first 9 years after their introduction in 1985. At the beginning of 1998, there were more than 50 million cellular telephones in use in the United States. Rapid growth in the number of cellular telephones continues to be fueled by smaller and cheaper portable units; this trend is exemplified by the recent introduction of a cellular telephone built into a wristwatch.

Just as the telephone created a tremendous network of voice connectivity, other devices have piggybacked onto the telephone network to add richness and complexity to that connectivity. The first of these was the facsimile machine, permitting the rapid transmission of a document between 2 standard units that can be attached to any telephone instrument anywhere in the world. When facsimile transmission first became available, few people recognized its value and importance, but the speed and convenience of facsimile transmission eventually became widely recognized. Today, facsimile equipment is nearly ubiquitous and adds significantly to the richness of global connectivity as compared with voice-only connections.

During the 1970s and 1980s, computer systems were also being piggybacked onto the voice communications network. With a computer at each end of the line, a facsimile of a document could be transmitted over great distances, and the editable content of the document could be sent as a computer file. Large numbers of computer enthusiasts subscribed to a tremendous network of public-access file exchange systems known as "bulletin board systems." Because a direct telephone connection between the 2 machines was needed, the system was effective but cumbersome. Like an isolated telephone branch exchange, a bulletin board system was an island of connectivity because it could serve only a few people in an isolated geographic area and could not easily route files or messages to people using other systems in other areas.

In the 1990s, the widespread adoption of transmission control protocol/Internet protocol (TCP/IP), a standard for exchanging messages between computers from different manufacturers, facilitated the development of local-area networks (LANs) of interconnected computers and then of connections between networks to form "internetworks." For the first time, information could be freely exchanged between any 2 computers connected to any parts of a far-flung network of networks, known collectively as the Internet.[2]

The number of Internet-connected computers increased from 0 to 20 million in the first 4 years of its existence, and this exponential growth has continued unabated. As of January 2001, there were nearly 110 million host computers registered as accessible in the Internet Domain Name System (DNS), with many more host systems connected to the Internet but not registered for public access. This level of connectivity is without precedent. For the first time, it has become possible to obtain information on virtually any topic at exactly the place and time at which the information is needed. If the Internet continues to grow at this phenomenal rate, by the end of the year 2003 there will be one computer connected to the Internet for every man, woman, and child in the United States.[3]

UNDERLYING COMMUNICATIONS TECHNOLOGIES

Whether by voice, by facsimile, or by computer connection, modern communication depends on the underlying telecommunications technologies that link buildings, cities, states, and countries. The power of a particular mechanism for information exchange depends partly on the highest speed with which a very small amount of information can be sent, known as the latency, and partly on the amount of information that can be sent per unit of time, known as the bandwidth. An overnight delivery of a shipping crate full of documents has a high latency because it takes an entire day for a single bit of information to be delivered, but it also has a high bandwidth because a huge amount of information can be transferred per day. A flashlight used to transmit Morse code has a low latency because the first bit of information is delivered at the speed of light, but it also has a low bandwidth because very little information can be transferred per day. With the exception of satellite connections, modern telecommunications modalities typically have a low latency, but the amount of bandwidth varies greatly from modality to modality.

VOICE ORIGINS OF DIGITAL COMMUNICATIONS CIRCUITS

Today, many different kinds of information are delivered through telecommunications links, but these links are not optimized for the type of information that they carry because they were originally developed solely to carry human voices. The microphone inside a telephone handset converts a sound wave into a continuously varying voltage, which is known as an analog signal. The sound waves from a normal human voice oscillate up to 4,000 times per second; thus, an analog voltage signal representing that voice will have up to 4,000 cycles per second, or 4 kHz. This 4-kHz signal can be sent through the wires directly as an analog voltage, or it can be digitized and sent through the wires as a series of high and low voltages representing digital ones and zeros, which is known as a digital signal. End-to-end analog connections are rare today; in most cases an analog voice signal goes from the telephone handset through a "local loop" of copper wire to the telephone company branch exchange, where it is converted from an analog to a digital signal and passed along to a destination branch exchange. At the destination exchange, the signal is converted back into analog form and sent through the local loop wires to the telephone handset at the other end.

When a 4-kHz human voice is converted from an analog to a digital signal, it requires 2 digital samples per cycle and 8 levels of measurement per sample; thus, $4\times2\times8=64{,}000$ bits of information per second (64 Kbps). A single 64-Kbps channel is often referred to as a "basic channel" or "B-channel," and most high-bandwidth connections are defined as multiples of this channel. For example, an Integrated Services Digital Network (ISDN) line carries 2 digital B-channels, or a total of 128 Kbps. A leased T-1 line offers 24 B-channels plus one eighth of a channel used for timekeeping, for a total of 1.544 million bits per second (Mbps).

When an "end-to-end" digital circuit is available to connect 2 computers, all of the potential bandwidth of the circuit is available to send and receive information. However, when some part of the circuit must carry an analog signal, some of the potential bandwidth is lost. The computer's digital signal is converted from digital (at the sending end), to analog (at the telephone connection), to digital (at the sending telephone company branch exchange), to analog (at the receiving telephone company branch exchange), and finally back to digital (at the receiving computer). A conversion from digital data to an analog signal and back to digital again is known as modulation-demodulation, hence the term "modem." Similarly, when an analog voice signal is converted to a digital signal and back again, the term "coder-decoder" or "codec" is used.

With a pure digital connection, the 64 Kbps of a B-channel can carry a full 64 Kbps of data, but the same channel used with a modem at each end carries much less data. Modem technology originally permitted only 120 bps through a B-channel but evolved during a 30-year period to support higher transmission rates over the same basic channel. Today, a standard B-channel modem can transmit or receive 38,400 bps when connected to another modem, and data rates can go as high as 56,000 bps in special situations in which at least one end of the circuit is digitally terminated.

HIGH-SPEED DIGITAL LANDLINES

Given the right conditions, a 4-wire copper wire circuit can be made to carry as many as 1 billion bps (1 gigabit/s, or 1 Gbps), but this type of transmission requires high-quality lines and expensive line conditioners and can run only short distances before signal regeneration is necessary. In contrast, light passing along a single fiber of spun glass can carry 2.4 billion bps or more, and the fiber can run 200 miles between regenerators. For this reason, high-speed network traffic today is carried almost exclusively over fiber-optic cable.

Compared with new copper wires, fiber-optic cable is very cost-effective. When fiber cable is made, many fibers are bound together in a single bundle. Only a single fiber strand is needed to make a high-bandwidth connection, and there is a large amount of unused capacity, known as "dark fiber," remaining in most cable runs. Advances in a technique known as "wave-division multiplexing" have increased per-strand bandwidth by making it possible to send and receive data encoded into many different wavelengths of light through the same fiber strand all at the same time. The amount of data that can be carried by a single fiber strand has been increasing rapidly and is likely to continue to increase during the next few decades.

With the same amount of data being squeezed into narrower and narrower parts of the visible spectrum, there is no theoretic limit to the potential future data-carrying capacity of fiber connections. Today, commercially installed synchronous optical network (SONET) systems carry voice and data traffic in a digital format at channel speeds up to 2.4 Gbps. Other optical systems routinely operate at 10 Gbps, and laboratory prototype optical systems now handle up to a trillion bps (1 terabit per second, 1 Tbps) and are capable of carrying 13 million simultaneous telephone calls on a single fiber that is the diameter of a human hair.

In the United States, fiber-optic cables now connect every main telephone company exchange and many local branch exchanges and commercial buildings. Unfortunately, it has not yet proven cost-effective to run new glass fiber-optic cables into every home and office. For this reason, there is great interest in new technologies that can improve data transmission over existing copper wires. The principal "last mile" solutions being deployed today are cable modems that can transmit data over the coaxial cable used for cable television and "digital subscriber lines" (DSL) that use special conditioners to boost the digital speed of existing copper loops belonging to the local telephone system. The bandwidth of these systems depends on the distance the data must travel and the quality of the environment through which the signal must pass. Given the proper conditions, each of these can be a cost-effective solution carrying 1.54 Mbps of data or even more.

COMPUTER CONNECTIVITY OVER CELLULAR TELEPHONE CIRCUITS

When digital data are sent over wireless radio-telephone circuits, there is another added layer of complexity. The radio frequency spectrum is a scarce resource, and current cellular technologies use a variety of tricks to squeeze every possible bit of voice traffic into the limited bandwidth. The first trick used is to simply throw away data; as long as the speaker is recognizable, voice quality can be sacrificed to save bandwidth. A pure analog cellular telephone voice channel can carry only about 48 Kbps rather than the 64 Kbps carried by an analog landline. Other digital and hybrid systems encode and compress a voice signal with algorithms that can deliver a recognizable voice using much less bandwidth. The resulting cellular channels are adequate for voice transmission, but, when used with a computer and a modem, they provide much less bandwidth than a voice-quality landline. Many applications that work well when a laptop is connected through a modem and a landline will fail miserably when the same laptop is connected through a modem and a cellular telephone connection. This bandwidth problem severely limits the connectivity of mobile systems such as those in ambulances and helicopters. Of all the digital systems currently in use, only Code Division Multiple Access (CDMA) systems can deliver usable data rates through a single channel. Unfortunately, CDMA is not a good final solution for mixed voice and data transmission, because adequate data bandwidth is only obtained at the expense of other cellular users who share the channel. The Figure shows the maximum rate at which digital data can be delivered through different types of cellular channels compared with the rate at which it can be delivered through a standard "local telephone loop" landline.

SATELLITE TELECOMMUNICATIONS

Communications satellites may be grouped into 3 broad categories according to the height at which they orbit the earth. The period of an orbit depends on its distance from the earth; the greater the distance is from the earth, the slower the orbit is. At a distance of 22,241 miles above the earth's surface, the orbital period is precisely 24 hours, and the orbit is called "geosynchronous." If a satellite in a geosynchronous orbit travels in exactly the same direction as the earth's rotation, the satellite can remain above the same spot on the earth at all times. Such an orbit is called "geostationary." Because it is so high above the earth, a single geostationary earth orbit (GEO) satellite can transmit to nearly an entire hemisphere. Despite this advantage, GEO satellites are poorly suited to voice communications and other real-time applications because data transmission, signaling, and receipt acknowledgment across this great distance introduce an overall round-trip latency of approximately one-half second.

Figure.
Basic channel bandwidth.

Standard Local Loop Landline
38.8 Kbps with standard modem in analog configuration
56 Kbps with standard modem in optimum configuration
1.54 Mbps if a digital-conditioned line
Analog Cellular Connection
33.4 Kbps with typical modem
48 Kbps maximum with a custom-designed modem
North American Time Division Multiplexing (NA-TDMA)
8 Kbps if a standard modem is used with a cellular telephone containing a codec
16 Kbps if the system permits a digital signal
European Time Division Multiplexing (EU-TDMA)
13 Kbps if a standard modem is used with a cellular telephone containing a codec
33 Kbps if the system permits a digital signal
Code Division Multiple Access (CDMA)
13 Kbps if a standard modem is used with a cellular telephone containing a codec
48 Kbps if the system permits a digital signal

This excessive latency can cause otherwise high-bandwidth GEO connections to communicate at a fraction of their capacity for 2-way or lossless data transfers. Most common data protocols operate on the principle that a copy of each data packet must be kept in a buffer on the sending computer until receipt of an acknowledgment that the packet arrived successfully. The buffer can only hold a limited number of packets, and no new packets can be transmitted until old ones are removed when their acknowledgments are received. Because the round-trip latency is 500 ms, data packets cannot be removed from the buffer for at least that period. The default buffer size in TCP/IP is 32 kilobits. This means that, at any given moment, only 32 kilobits can be in transit and awaiting acknowledgment, and, no matter how many bits the channel can transmit theoretically, it still takes at least half a second for any 32 bits to be acknowledged. Given these conditions, the maximum data throughput rate is 32 kilobits per half second, or 64 kilobits per second. If a 1.54 Mbps T-1 line is connected through a geostationary satellite, the throughput will be constrained to only 64 kilobits per second, which is only 4% of the purchased capacity.

The high power and large satellite dish required to send a signal from the earth up to a receiver at so great a distance also make GEO satellites a poor choice for many applications. GEO systems have found their principal niche for 1-way broadcasts that can tolerate some data loss and therefore do not require packet acknowledgment. Digital radio and digital television are examples of such applications.

Medium earth orbit (MEO) satellites orbit at 6,000 to 12,000 miles above the earth, and, although they move with respect to the ground, they are high enough that each satellite can cover a large portion of the earth and can stay within view of a single ground-receiving station for a long time. A moderately large satellite dish and a large amount of power are required to transmit from the ground to an MEO satellite, but round-trip signal latency is acceptable for voice communications. The International Maritime Satellite Organization (INMARSAT, London, UK) and TRW mobile communications systems (TRW Inc., Cleveland, OH) use MEO satellites for both voice and private network data transmissions at up to 9,600 bps over a standard channel.

Low earth orbit (LEO) satellites have a short signal latency and are close enough (100 to 600 miles above the earth) that handheld cellular telephones can connect to them with a low-power transmitter and a standard antenna. Because they are so close to the earth, LEO satellites must travel about 17,000 mph (relative to the ground) to remain in orbit, circling the globe about every 90 minutes. The first such system, which was known as Iridium, launched 66 active LEO satellites into an orbit 420 miles high and offered satellite telephone service worldwide beginning in the last quarter of 1998. Unfortunately, Iridium proved to be a financial disaster, accumulating $5 billion of debt before declaring bankruptcy and selling the entire satellite system for just $25 million. Another LEO satellite telephony company, Globalstar (San Jose, CA), has been more successful, with 48 active LEO satellites in orbits that cover selected areas of the world.

Prototype satellites have also been launched for another system, known as Teledesic (Bellevue, WA), that proposes to place 288 satellites into orbit 339 miles above the earth, linking them together to create an “Internet in the sky.” Teledesic mobile stations will permit access speeds up to 144 Kbps, and special ground stations could use small roof-mounted satellite dishes to give each home or office a high-speed wireless connection carrying 64 Mbps, a capacity 2,000 times greater than a current analog modem and 6 times greater than most hardwired LAN connections.

The Table shows the time it would take to download the 16-megabyte Internet Explorer 5.5 (Microsoft Corporation, Redmond, WA) Web browser with several different types of connections. This download would

Table.
Time to download a 16-megabyte file.

Connection	Speed	Download Time
Modem (circa 1975)	120 bps	12 d
Modem (circa 1980)	300 bps	5 d
Modem (circa 1984)	1,200 bps	30 h
TRW or Inmarsat satellite modem	9,600 bps	4 h
Digital cellular telephone with modem	13 Kbps	2.7 h
Modem (circa 1992)	28.8 Kbps	75 min
V.34 modem (circa 1994)	38.4 Kbps	56 min
V.90 modem (circa 1998)	56 Kbps	38 min
Dual ISDN	128 Kbps	17 min
Asymmetric DSL	384 Kbps	6 min
Cable television modem	750 Kbps	3 min
T-1 leased line	1.544 Mbps	83 s
Ethernet LAN	10 Mbps	13 s
T-3 leased line	45 Mbps	2.8 s
VDSL digital subscriber line	54 Mbps	2.3 s
Teledesic satellite connection	64 Mbps	2 s
Fast ethernet LAN	100 Mbps	1.3 s
Gigabit ethernet LAN	1 Gbps	0.13 s
SONET OC-48 fiber-optic channel	2.4 Gbps	0.05 s

VDSL, Very high rate digital subscriber line.

have taken 12 days using a modem in 1975 but now takes only one-twentieth of a second using a synchronous optical fiber channel.

GROWTH AND CONGESTION OF THE INTERNET

The power of the Internet to transform society cannot be overstated. In a few short years, this "network of networks" that allows an immediate exchange of data between any 2 computers nearly anywhere in the world has become both ubiquitous and essential to the functioning of many industries.

The Internet as it exists was never planned or developed by any formal group. It grew on the scaffolding of existing telephone and computer connectivity and came into existence mostly as the result of the independent efforts of scattered visionaries working in laboratories at universities and colleges, in the government and military, and in private industry. Two decades of underlying work was largely funded by military and government grants supporting research in a variety of seemingly unrelated areas, much of it aimed at providing distant connections to supercomputers and other shared resources across the country.

The United States has benefited greatly from a position of leadership in worldwide connectivity. Increased production and an improved competitive position in the worldwide economy are just two of the many ways in which the connectivity of the Internet has made a difference. Unfortunately, some of the benefits of this connectivity are being eroded by the very success of the Internet. Other countries are increasing their presence on the Internet and are moving forward in ways that may well erode the commanding lead the United States has enjoyed in technology fields. Private enterprise and governmental agencies are afraid that the country will lose its competitive edge now that the rest of the world has become heavily committed to the Internet.

At the same time, the network bandwidth available for research and development has been nibbled away by private and commercial users. University, government, private lab, and military researchers once enjoyed the exclusive use of a network with bandwidth more than ample to meet the needs of any achievable project. Today, new applications demand more and more bandwidth as research capabilities continue to improve, yet researchers find that their high-speed connections to important resources are slowed to a crawl as Internet traffic continues to rise beyond all expectations. At the current rate of growth, by the time the next generation of applications are released there will be so much Internet traffic congestion that the new systems would hardly be functional at all if constrained to run over the public Internet.

NEXT GENERATION INTERNET AND INTERNET 2

The recognition of certain shortcomings in the current Internet and the desire for continued growth and leadership in technology have sparked several initiatives to underwrite the development of a "new Internet." Just as the current Internet has led to profound and unforeseen changes in our ways of life, a next generation version might permit the development of as-yet unimagined applications that could continue to redefine modern life as the current Internet does today. Two main initiatives are underway in support of many different groups working to bring this future vision to life. The first is known as the Next Generation Internet (NGI) initiative, and the second is known as the Internet 2 (I-2) initiative. These initiatives share many of the same goals but have different constituencies and complementary approaches to those goals.

The NGI initiative, which can be accessed at http://www.ngi.gov/, is a multi-agency federal research and development program with 3 goals: (1) to connect national laboratories and universities with high-speed networks that are 1,000 times faster than today's Internet, (2) to promote experimentation with the next generation of networking technologies, and (3) to demonstrate new applications that require advanced networking to meet important national goals and missions.

In an effort to replicate the dynamics that led to the creation of the current Internet, a 1998 federal investment of US$100 million was allocated as a catalyst for additional investment by universities and by the private sector. The initiative is directed by the National Coordination Office for Information Technology Research and Development, which can be accessed at http://www.ccic.gov/, and the principal agencies involved have been the National Science Foundation, the Defense Advanced Research Projects Agency, the Department of Energy, the National Aeronautics and Space Administration, the National Institute of Standards and Technology, the National Institutes of Health, and the National Library of Medicine (NLM). Most of the project focus is on building a new, Very-high-speed network infrastructure to support national goals. The NGI initiative is a logical outgrowth of the existing High Performance Computing and Communications initiative. All high-performance mission-specific networks now operated by federal agencies are elements of the NGI initiative, which builds on networks such as the National Science Foundation's Very high-performance Backbone Network Service (VBNS).

The I-2 initiative, which can be accessed at http://www.internet2.edu/, is a project of the University Corporation for Advanced Internet Development (UCAID) (http://www.ucaid.edu/), a collaborative effort by more than 180 universities together with their research partners in industry and government. The goal of the project is to establish a distributed knowledge system for achieving ongoing innovations in research, teaching, and learning to develop advanced Internet technologies and new applications that will support the future research and education needs of institutions of higher education.

One I-2 goal held in common with the NGI initiative is the effort to connect universities, national laboratories, and private partners with high-speed networks that are 1,000 times faster than today's Internet, networks that can handle at least 1 Gbps of connectivity for any user or application that needs it. One often-quoted aim is to be able to transmit and receive the contents of the entire *Encyclopedia Britannica* in less than 1 second. Consortium members have established extremely high-capacity network links to a separate national network that has an infrastructure parallel to the commodity Internet and is operated by many of the same common carriers. Members connect to the high-speed backbone through special "point-of-presence" service providers that are known as "gigapops" because they can deliver 1 Gbps of data flow to each connection. Today, it costs approximately US$500,000 per year in commitments and upgrades for a university to join UCAID and become connected to an I-2 gigapop.

The project will demonstrate prototypes for a new generation of high-performance applications that support scientific research, national security, distance education, environmental monitoring, and health care and that need more speed, more bandwidth, more reliability, and more security than can be obtained with today's Internet. It is recognized that most of today's applications would perform better if the existing Internet worked smoothly and reliably within its current design parameters; however, I-2 is not an effort to improve the speed of today's Internet connections or to resolve shortcomings in the existing Internet. Instead, the I-2 project aims to support the development of new classes of applications that by their very nature need substantially higher application-dedicated bandwidth, bandwidth reservation, improved security, guaranteed quality of service, and other advanced features of a new generation network.

PROBLEMS WITH THE CURRENT INTERNET

Improving the current Internet is not a goal of the I-2 project, but, to support advanced "next generation" applications, I-2 will address many recognized problems with the current Internet, including problems with network management, end-to-end control, security, reliability, and the need for embedded functions to support distributed systems. The lack of directory services and of provisions for interrealm authentication will also be addressed. A new addressing scheme, known as Ipv6, will be implemented because the current IP address space is too small to support the number of devices that now need to be connected.

In building the NGI, the most important function that will be added to IP is a set of "internetworking" protocols for managing the end-to-end connection. Central management was intentionally designed out of IP because IP is meant to be deliberately nondeterministic in the sense that each information packet traversing the network "finds its own way" to its destination on the basis of the local conditions it encounters on the way. This makes the network self-healing because individual packets can be independently routed around broken connections or highly congested areas. Unfortunately, because there is no mechanism for sharing information about global conditions along an entire path, each packet can be rerouted only after a problem is encountered. I-2 will improve the sharing of knowledge about distant conditions along a desired route and will permit more sophisticated route management.

The ability to guarantee delivery is another important function that is missing from the current Internet, which uses a delivery service called "best-effort" delivery. Best effort means that all packets are accepted from all senders and are sent on their way whenever possible. If there is too much traffic at a particular router, that router randomly discards packets without any concern for where they are coming from or where they are going. There is neither guarantee of timely delivery nor guarantee of delivery at all. If an expected message or response fails to arrive, the sender and the receiver will negotiate a re-send of the incomplete data, in which case even more packets have to be handled by the network. This is satisfactory for the exchange of e-mail messages but not for a mission-critical system that needs guaranteed delivery with a guaranteed maximum latency.

Along with guaranteed delivery, I-2 aims to improve the end-to-end quality of service of the network by recognizing that, although dependability and reliability are critical quality parameters for all applications, different applications have different networking needs. A text file transfer can tolerate a 10-second delay in delivery, but a command going to a robotic control system in a nuclear power plant cannot safely be delayed. Under I-2, quality of service controls will permit applications to reserve a specific quality of service from the network and to be guaranteed that the reserved latency and bandwidth will be available when needed. Electronic commerce will be able to obtain secure delivery guarantees. Real-time audio and video services will be able to reserve bandwidth and to obtain guarantees of maximum packet loss rate and jitter. Even when an application makes no quality-of-service demands, new "type-of-service" headers can help routers know what type of material is included in each packet. If a Web page request and a fragment of a telephone conversation are competing for the same slot, the router can make intelligent decisions about prioritization. Type of service can actually reduce congestion in absolute terms because preferential delivery of an acknowledgment packet prevents the original material from being re-sent by the sender.

One implication of the availability of quality-of-service controls and multicast capability expected in NGI is that I-2 will be far more hospitable than today's Internet to connecting very large numbers of sensors. The capacity to make large amounts of "public" shared-sensor telemetry available to the I-2 community represents an exciting opportunity to explore new classes of applications. The ability of the network to handle large amounts of data that are irrelevant to most people might be seen as the very definition of today's Internet, but, in reality, it is easy to swamp nearly any real-life network segment by attempting to pass real-time telemetry data for any complex event or series of simple events. I-2 promises to allow such data to exist on the network without interfering with other applications.

As an example of the potential for this kind of sensor-based data, consider that many stores and many IP-enabled soft drink vending machines already use automated inventory systems that are accessible over the Internet. With this sort of sensor data widely available, it may soon be possible to query from any public point of presence and learn in a few seconds where the nearest bottle of a favorite soft drink is available, including both stores and vending machines. Another example of widely deployed IP-enabled equipment is found in the vehicular traffic sensing units that report local traffic conditions on a constant basis. When enough of these have been deployed, it will be possible to receive a constant stream of sensor-provided data about the traffic load at most busy intersections across the country.

PROTOTYPE HEALTH CARE APPLICATIONS OF THE NGI

The NLM has a history of sponsoring health care activities related to high-performance computing and communications and is now funding the development of a variety of health care applications that will use and require the NGI. Areas of particular interest to the NLM include advanced telemedicine, teleimmersion, digital libraries, distance learning, image distribution, and other applications that require the high-speed transfer of large amounts of data. Other funding agencies are also underwriting the development of advanced applications in health care that can serve as demonstration projects for the NGI. Medical applications can qualify as NGI applications without involving large data sets or high-speed networking if they require other features of the NGI, such as guaranteed delivery, guaranteed quality of service, or security of private medical information that is to be transmitted over the public network. Many examples of NGI- and I-2–qualified projects have been funded in the past. Some are from fields other than health care, but they all serve to illustrate the kinds of things that could or should be done in the design of health care projects for NGI.[4]

The Virtual Temporal Bone is a shared virtual reality simulation based on an accurate 3-dimensional (3-D) model of the temporal bone that is designed as a teaching aid for residents in otolaryngology who are learning about the anatomy of the middle and inner ear. The Virtual Temporal Bone project was developed by the Virtual Reality in Medicine Laboratory, the Department of Otolaryngology, and the Electronic Visualization Laboratory at the University of Illinois at Chicago. This is considered an I-2 project because the quality of service and end-to-end bandwidth made available by I-2 will be necessary to share this application between educational institutions. More information is available at http://www.sbhis.uic.edu/vrml/.[5]

A project known as Collaborative Architectural Layout Via Immersive Navigation (CALVIN) is a test bed for applying virtual reality in architectural design and collaborative visualization that uses multiple perspectives in an effort to take advantage of virtual reality in the earlier, more creative phases of the design process, rather than just as a walk-through of the final design. CALVIN is being used to investigate issues in collaborative virtual reality, including network needs, avatar representations, audio and video communication, and unusual user interfaces. This is considered an I-2 project because, for participants to collaborate effectively, virtual environments require high bandwidth and low latency that are not possible on the current Internet. CALVIN is another project of the Electronic Visualization Laboratory at the University of Illinois at Chicago. More information is available at http://www.evl.uic.edi/spiff/calvin. Many other immersive and teleimmersive environments are under development at other centers.[6,7]

Some of the most important projects in health care are those requiring shared access to extremely large data sets that cannot be accessed in a reasonable amount of time without the high-bandwidth connections promised by the NGI. The most well-known of the large data sets in health care is the NLM's Visible Human data set, a complete, anatomically detailed, 3-D digital representation of the male and female human bodies. Transverse computed tomograms, magnetic resonance images, and cryosection images of representative male and female cadavers were acquired at 1- and 0.33-mm intervals to make this data set. The female data set is approximately 40 gigabytes in size and would take 97 days to download through a standard V.90 modem. More than 50 projects and products have been based on the Visible Human data set, including the 3-D Virtual Colonoscopy project, which demonstrates static images and fly-through animations from the Departments of Radiology and Computer Science at the State University of New York at Stony Brook, the Interactive Knee Program from the University of Pennsylvania Medical Center, and "CathSim," an intravenous training system that includes a tactile feedback device.

Another example of large shared data sets is the Human Genome Project of the National Center for Biotechnology Information (NCBI). These are the most extensive databases on DNA and protein sequence data in existence, and they are very heavily used by researchers around the world. The NCBI Web site attracts more than 2 million hits a day. Although the basic search engines serving these data do not require NGI attributes, such attributes may be needed by other applications using large portions of the data set or by applications using it for visualization or rendering.

The Terabyte Challenge is a distributed test bed for research into the management and mining of truly massive data sets. Data mining is the automatic discovery of patterns, associations, changes, and anomalies in large data sets. Developing scalable data mining algorithms, software tools, and applications is a fundamental scientific challenge with important implications for health care and for many other scientific and business disciplines. The Terabyte Challenge project facilitates prototyping of new ideas and software tools and allows them to be tested against both large data sets and distributed data sets. This is considered an I-2 project because the mining of geographically distributed massive data sets cannot be done without high-performance links. The Terabyte Challenge is a collaborative project of the University of Illinois at Chicago, the University of Pennsylvania, and the University of Maryland. More information is available at http://www.nscp.uic.edu.

The term "collaboratory" is a neologism describing a distributed collaborative laboratory. Collaboratories often support teleconferencing as a basic collaboration tool, but, more importantly, they also permit remote access and collaborative or simultaneous access to physical resources at a distant site. Collaboratories need not deal with scientific data; for example, an existing music collaboratory, known as jam session site, actually allows 2 or more musicians to practice together over the Internet. At least one compact disc has been released featuring music that was created and recorded through collaborative telepresence. Another existing collaboratory is a meteorologic site that allows multiple users to communicate in real time while simultaneously displaying weather maps and other streams of meteorologic data that are germane to a discussion. Many medical collaboratories are under development.[8]

Other interesting wide-area projects provide access to and control of special imaging techniques. Examples have ranged from single-photon imaging in cultured cerebellar granule cells to radio astronomy synthesis imaging, a process in which 27 radio astronomy antennas track the galactic center while the data are recorded, merged, and processed in real time to show a visible simulation of the Fourier transforms. The Collaboratory for Microscopic Digital Anatomy makes a unique research tool available over I-2 connections: an intermediate high-voltage transmission electron microscope at the National Center for Microscopy and Imaging Research at the University of California, San Diego, has been modified to function as a computed tomography scanner. The specimen is tilted to a series of different angles for a series of scans, and the resulting data are tomographically reconstructed to give 3-D information about its internal structure. Researchers can control the electron microscope lenses and the 4-axis specimen stage by remote control. The tomographic reconstruction is very computationally intensive; therefore, as the scans are performed, the data are transparently sent out across I-2 to high-performance supercomputers for calculation. The resulting 3-D images are returned to the remote researchers, who possess neither microscope nor supercomputer, but who have enjoyed the use of both to capture 3-D structural images that would otherwise have been impossible to obtain.

One of the most exciting tools for investigating the possibilities of NGI is a visualization tool called the Cave Automated Virtual Environment (CAVE). Created by scientists at the University of Illinois' Electronic Visualization Laboratory in 1992, the CAVE is a multiperson, high-resolution, 3-D graphics video and audio environment. It is a 10×10×10-foot structure that sits in a 35×25×13-foot darkened room and has rear-projected screen walls and a front-projected floor. Inside the CAVE, a user wearing 3-D shutter glasses is fully immersed in a virtual environment. Objects appear to float in space, and the computer tracks a user's position to keep the perspectives correct while the user "walks" around objects. A desktop version of such a display also exists, called an "ImmersaDesk." These visualization capabilities are already available to those who can afford the supercomputing power needed for real-time rendering. In fact, in recent years General Motors has designed automobiles with an automobile simulator that uses the CAVE principles. NGI and I-2 offer the promise that CAVE environments could offload their real-time rendering to supercomputing centers connected by the next generation of network connections.

The National Institute of Health and the NLM recently awarded a series of phase I and phase II contracts for NGI biomedical demonstration projects, including projects related to personal medical records, biomedical teleimmersion using CAVE and ImmersaDesk systems, regional collaborative cancer care, secure medical imaging and teleimaging systems, nursing home surveillance, nomadic data access, cardiology education, rare disease treatment coordination, ultra–high-resolution cell image databases, remote real-time surgical simulations, digital embryology libraries, ambulance-based real-time telepresence, remote treatment planning for radiation therapy, interactive fly-through of 3-D anatomic renderings, and a national network for high-resolution breast imaging archives. These projects are currently underway and promise to help define the need for those aspects of enhanced networking promised by the NGI and I-2 projects.

In summary, from the earliest days of small telephone exchanges until the present time, connectivity has been a driving force for change. The Internet and the World Wide Web arose out of the existing base of telephone connections and computer networks and inherited their bandwidth restrictions but, nonetheless, have brought a previously unimagined richness to connectivity, changing the world dramatically as a result. The NGI is being designed specifically to provide high-bandwidth connectivity on a scale so vast that is difficult to comprehend. It is our task to imagine into existence the richness and complexity of the applications that will exploit this new connectivity.

REFERENCES

1. Naughton J. A *Brief History of the Future: From Radio Days to Internet Years in a Lifetime.* 1st ed. New York: Overlook Press; 2000:1-320.
2. Leiner BM, Cerf VG, Clark DD, et al. A brief history of the Internet. Internet Society Web site. Available at: http://info.isoc.org/internet-history/. Accessed July 3, 2001.
3. Network Wizards. Internet Domain Survey, 2001. Internet Software Consortium Web site. Available at: http://www.isc.org. Accessed July 3, 2001.
4. Shortliffe EH. The next generation Internet and health care: a civics lesson for the informatics community. *Proc AMIA Symp.* 1998;8-14.
5. Mason TP, Applebaum EL, Rasmussen M, et al. Virtual temporal bone: creation and application of a new computer-based teaching tool. *Otolaryngol Head Neck Surg.* 2000;122:168-173.
6. Lanier J. Virtually there. *Sci Am.* 2001;284:66-67.
7. DeFanti T, Sandin D, Dawe G, et al. Personal tele-immersion devices. In: Proceedings of the Seventh IEEE International Symposium on High-Performance Distributed Computing; July 28-31, 1998; Chicago, IL. pp. 198-205.
8. Keating KA, Myers JD, Pelton JG, et al. Development and use of a virtual NMR facility. *J Magn Reson.* 2000;143:172-183.