AN OVERVIEW OF UNIVERSAL INFORMATION SERVICES: CONCEPTS AND TECHNOLOGIES OF FUTURE NETWORKS

Arnold Heiber

Arnold Heiber is manager—UIS Strategic Planning for AT&T Network Systems in Holmdel, New Jersey. He is responsible for developing concepts and strategy for Universal Information Systems. He has a B.S.M.E. from New Mexico State University. He joined AT&T in 1964.

This issue of the *AT&T Technical Journal* addresses the technologies that will play important roles in future networks. In this brief overview, we examine the basis of AT&T Network Systems' view of the future, called Universal Information Services, which predicts a literal fusion between communications and computing.

Introduction

The importance of specific technologies to future networks will not be evident without an understanding of the roles, and therefore the requirements, of future networks. What are the functions that a network can perform? Should future networks be "intelligent" or "dumb"? What benefits to society can be derived, or lost, in making the choice between intelligent and dumb networks? What do we mean by an intelligent network? These questions and many more must be raised and answered before it is clear what technologies are relevant to future networks. Here, we provide some of the answers that form the basis for AT&T Network Systems' view of the future, Universal Information Services.

Although many observations will be made here on the characteristics of future information networks, there are two key observations that permit us to derive an extraordinary amount of information about the nature of successful future networks. The first observation relates the role of intelligent networks to human intelligence. The second observation relates intelligent networks to intelligent machines.

The Purpose of Future Networks

The traditional view of a network as a transport mechanism led people in the telecommunication industry to refer to the "last mile" as the most difficult and costly gap to be bridged in the delivery of bits to the office or residence of the customer. In the information age, the most critical and productive gap to be bridged is the last three-eighths of an inch through the human skull, where, if information has been properly presented to the senses, comprehension is achieved. Information becomes important only when it is comprehended by someone who needs it.

The most important purpose of future networks is to increase information productivity, which refers to all the processes that support

5

the efficient access to, comprehension of, and utilization of required information. The key issue is the connection between the resources of the future information network and the neural networks of the customers. Today an increasing variety of devices and technologies support the rapid production and dissemination of information. Computers produce "output" at an ever-increasing rate. Word processors efficiently produce thicker documents. And faster reproduction machines ensure that everyone has a personal copy. Electronic mail and document distribution systems efficiently distribute this cornucopia of "information." To date, the emphasis has been on information production and dissemination. However, output does not equal insight. Information productivity requires that future networks support a variety of processes that provide efficient generation, access, storage, dissemination, and, perhaps differentiated, presentation of desired information in the forms that make it most easily comprehended and readily employed for the purposes of the recipient.

In the future, as intelligent machines take over the more routine "thinking" processes, networks will be called on to support applications that stimulate the more creative processes that will remain the dominion of humans.

Implications of Applications

It has been interesting in recent years to attend conferences on the Integrated Services Digital Network (ISDN) and see information presented as classical multimedia—the way in which people present information to each other whenever they assemble. At some point in a conference, someone would ask "What applications should the technologies address?" It has been interesting because, whenever people get together, they relate ideas to each other by presenting multimedia: text, data, and images with a voice explication of segments of the text or data or features of the images. This logical, temporal, and spatial relationship of segments of information is the most generic application of a multimedia presentation. It

In This Issue

The papers in this issue reflect a sampling of the activities that are supporting the technological evolution to future networks.

"The Universal Port Concept," by Gary P. Austin, Barry S. Bosik, and Christopher J. Capece, describes work being done to give customers a network interface that permits easy moves and interchangeability of a wide range of terminals and applications at any port on a network.

"The Network Operating System Concept for Future Services," by Wu-Hon F. Leung and Gottfried W. R. Luderer, describes software technologies necessary to support information-productive communications between groups of people.

"Fast Packet Technology for Future Switches," by John J. Degan, Jr., Gottfried W. R. Luderer, and Avinash K. Vaidya, describes one of the fundamental technologies that will be required to evolve future networks into distributed computers for on-demand response to customer needs for service.

Facilitating the creation of new enhanced services is addressed in "Advanced Software Technology Supporting Customer Programmability," by Donald W. Brown, Christopher D. Carson, Warren A. Montgomery, and Paul M. Zislis. The paper describes two promising approaches to specifying services at a higher level of discourse.

A key requirement for realizing the benefits of UIS is increasing bandwidth, and the related response time, in all parts of a network. This is the subject of two papers by Ivan P. Kaminow, "Photonic Multiple-Access Networks: Topologies" and "Photonic Multiple-Access Networks: Routing and Multiplexing."

is used in most meetings and classrooms to convey ideas efficiently.

When people are not together, it would be valuable if multimedia terminals could present stored information in the same manner. For example, if a person moved the cursor to a line of text on the screen, a stored voice might explain the author's concepts just as if the author were there to present the material. In general, recent terminal and network concepts do not yet support this type of interactive inquiry and multimedia response, but some are moving in that direction.

This last point is significant because many of the

6

current network concepts, and thus the constraints that they place on applications, were derived from concepts that relate to the historical voice network. The voice network was designed on the assumption that individuals would be simultaneously present to speak to each other in real time. Recently, of course, telephone answering machines have permitted time-shifting. In ISDN, the concept of *simultaneous voice-data* envisions real-time, simultaneous-presence applications, such as two individuals simultaneously viewing data screens while discussing them in a voice conversation, because the transport mechanism provides two B channels. However, the applications that dominate the greater portion of the knowledge worker's day are those during solitary periods in which information must be comprehended and value must be added to support those briefer intervals in which the value-added information is exchanged with others.

A human being is an integrated voice-image "machine"; it clouds our understanding of applications to speak of "integrated voice-data." Indeed, when most people in the industry spoke of integrated voice-data, they meant multiplexing, not applications that conveyed concepts. Image is the high-information-rate path to the brain, and image requires substantial bandwidth in networks. While a picture may be worth a thousand words, it is important to note that in virtually every application a great many words—in voice and text segments and, perhaps, supporting data—may be required to explain the meaning of various elements of the image.

Feeding the Mind

Having discussed one of the most generic forms in which information must be presented to be comprehended efficiently, let us turn our attention now to the rate at which networks must feed the mind with presentations of multimedia information in order to keep a train of thought efficiently chugging down the track. The mind of a user who has just finished making an inquiry at a ter-

minal should be stimulated by the response in no more than about 100 to 200 ms (milliseconds). After about 300 to 500 ms, the mind may begin to drift. Thus, human beings require huge amounts of network and processing resources for very short periods of time. Anyone who has experienced the short delay in a telephone call carried by a satellite is familiar with how dysfunctional even short delays are to thinking processes. The processing power of today's advanced workstations is a small fraction of what could be consumed in even the most ordinary of applications if the information is presented in the form for most efficient comprehension.

Comprehension is not only individualistic and related to the function of the person who is the recipient of information, but is also frequently enhanced when the information is entertaining. High-quality images, moving images, cartoons, music, and sound effects should have a role in the presentation of information. The related applications will place important requirements on networks.

Connect What?

Future information networks, whatever we may decide should be their functions (and we will describe more of them shortly), provide connections between things. For a moment, let us consider the inclusive set of all things that attach to information networks. These things, whatever may be their purpose or location, may be processors (of one kind or another), may be interfaced to the network by a processor, or may be replaced by what can readily be described as a processor. In other words, we can define the set of things that are primarily connected to information networks as a form of signal processor.

Even the telephone is a rudimentary form of signal processor (and can easily be replaced by more obvious forms of processors). It processes signals into voice, voice into signals, dialing into signals, and signals into ringing. A store-and-forward capability is not a direct interface from the network to memory, but rather to a

7

processor that accesses memory. An information network, whether ancient, current, or future, is a network of signals and is always attached to one form or another of signal processor.

Terminals have evolved to reflect the architecture of generic processors that are now the basic building blocks of their construction. However, the current public networks, other than data packet networks, evolved from networks that were defined before the concepts for generic processors existed. They therefore do not reflect the choices that would create synergy between a modern terminal's internal network and the external network and thus increase functionality while reducing gateway costs. This situation encourages a very costly anarchy in the construction of terminals and in their relationship to networks.

This simple observation has profound implications for our business that become clear when we now ask some critical questions. Given that an information network is always a connection of processors, can a limited set of processor types and related protocols be defined such that the widest range of desired applications might be served efficiently? What kinds of processors, software, and protocols should be defined when multimedia applications must be processed? What is a network that is an interconnection between large numbers of processors, and what functions can it perform? What form should the network take to optimize the opportunities that are presented by the interconnection of huge numbers of processors? Certainly the presence of an interconnection of large numbers of processors immediately suggests such things as distributed computing, multiprocessing, and parallel processing. These uses place a new set of demands on the network, while simultaneously expanding the network opportunities beyond measure. Given a network that can meet the demands, a little imagination suggests using the network for:
- Dynamic reconfiguration of computer systems
- Creation of virtual terminals from components resi-

dent on the network
- Wide-area memory management.

Intelligent Transport for Intelligent Networks

The internal "network" of a computer is a parallel bus that has an information structure in which bits represent priority, address, operations codes (instructions), data field, and error controls. This suggests that a network, as the natural interconnection mechanism between processors, should be able to emulate a processor bus.

Serializing the bits in the backplane (bus) of a computer produces a packet. This is not the traditional data packet, as found in low-speed packet networks, since those packets usually have little more than address and error-control information. When they use the term *packet* in the traditional sense, people are usually referring to packets used to create virtual circuits that are every bit as dumb (lacking in instruction richness) as the circuits in a time-division network. Clearly, if the future network is to gain the opportunities that are presented by the potential interconnection of a multitude of processors, the packets will have to be rich in instructions and move according to priorities at very high speeds.

If the processors and the packets are to handle multimedia applications, they must be modified to handle the timing and sequencing of voice, for example, and the interrelationships between segments of media, as has been described. The flexibility of variable header fields, variable data fields, and variable error control can create a network that can accommodate a wide range of applications and processing characteristics. This structure could be implemented above a fixed-length packet underlayer. Obviously, not all processors need to handle multimedia applications. Note that whether a processor is a terminal, part of a terminal, an enhanced service node, or even inside the network or an appendage is not an important consideration for many applications. However, an intelligent network requires that some of the processors belong to the network, controlled by the distributed operating sys-

8

tem software of the network; otherwise many of the most valuable functions of an intelligent network cannot be realized. The most intelligent of networks is not just a dumb transport provider with processors and databases attached to it. It is a network-scale computer capable of responding to instant needs.

In a sense, the transition between a "voice" network, based on circuits, and the future information network has certain similarities to the choices that designers of the early power networks faced in making the transition from direct current to alternating current. There were advocates, including Thomas Edison, who believed passionately that direct-current power networks were the right answer. Their views were strongly influenced by a single, dominant application: the incandescent light bulb. Nikola Tesla and others saw the importance of transformability to support a variety of applications and to achieve higher transmission efficiency. The key issues in the transition from voice network to information network are

- Transformability—that is, matching the transport, processing, and response needs of the application with the appropriate priority, instantaneous bandwidth, processing instructions, linkage between medium segments, timing/sequence instructions (voice readout), addressing, and error control.
- Efficiency
- Priority of processing

Attaching Terminals and Applications

Customers would like, easily and immediately, to connect, move, and interchange a wide range of terminals and applications at any port on the network without making prior arrangements with the network service provider. Experiments at AT&T Bell Laboratories have verified that a flexible information structure, combined with the ability to allocate bandwidth dynamically and manage delay, can support the common applications in voice, image, and data. Thus, a universal port for terminal access can be defined to meet this customer need. Since most terminal and application functions can be implemented on a limited class of processors, the evolution from terminal adapters that can handle the existing installed base of terminals to a much more efficient interface is assured.

Unified Networking

Communication takes place across concatenated networks. A network within a terminal connects the elements of the terminal to the terminal's processor and an external network interface. That external network may be an on-premises network, which may be connected to a local exchange network, which may be connected to an interexchange network, and so forth. The greater the dissimilarity between each pair of connected networks, the more complex the gateway required for communications. This invariably results in a reduction in communications, functionality, and performance. Of course, the largest number of gateways (modems are an example) are found between terminals and networks. If the number of terminal processor types can be limited, but still kept large enough to support a virtually unlimited range of applications—and it appears that this can be done—then the networks need emulate only a limited class of processor buses, and a unified networking scheme can be implemented that will bring many benefits and economies to the entire industry. In the remainder of this overview, some of these benefits will be described.

Virtual Machines

When a network can be made to appear as a natural extension of the bus of the processor that defines the terminal, the network can act as an integrating mechanism to make any assemblage of resources attached to the network appear as elements of the terminal. Thus, the network can become the instrument for the creation of nearly unlimited varieties of virtual terminals. This efficient sharing of diverse resources can have a

9

profound effect on the cost/performance of customer premises equipment. A network that can minimize access delay and support wideband transfer of information resources can make almost any resource appear to be at any point on the network. Therefore, it is possible for system designers to centralize or distribute resources as design requirements dictate. This is the concept of architectural freedom in future networks that has been presented in Universal Information Services.

Statistical, Software-Defined Networks

A public network is successful when all customers believe that they are being served by their own private networks. Future applications may be delivered by a compounding of multiple subnetworks within the public network. A customer may access an application on a virtual terminal that is defined for that application by a virtual subnetwork. The database may be assembled on another subnetwork. The computer that executes the application may be assembled on another subnetwork. And the customer may use the application to interact with other people on another subnetwork. The public network may be made up of tremendous numbers of dynamically changing subnetworks. To be cost-effective, these subnetworks must be statistical; that is, although they are created as a map of the connectivity, they use the network resources only during the moment that information flashes through the network.

Differentiation of Services

Future information networks must carry a wide variety of services with very widely differing characteristics. In an information network, all bits are not created equal. A bit that represents part of a data character has far greater economic value than a bit from a digitized voice stream. And both bits have far greater value than a bit that represents one of the gray scales of one color of one pixel of an image. Further, bits in a data communication require greater accuracy than voice or image bits. Person-to-person voice is delay-intolerant traf-

fic, whereas person-to-storage voice calls are delay-tolerant. As more and more voice traffic uses store-and-forward services, the networks can take advantage of the differing characteristics of services. High-priority traffic should pay a premium, as does priority mail. In packet networks, time, which is one of the most valuable commodities, can be sold in priority services. Distributed processing in networks is not feasible without a system of priorities. The primary mechanism for differentiating services with different priority, accuracy, bandwidth, and instruction requirements is packet networking.

Connection versus Connectionless

With the introduction of competition after deregulation, some manufacturers promoted the digital private branch exchange (PBX) to the business community as a solution to both their voice and data problems with the simplistic argument that both voice and data could be represented by bits. Of course, a digitized voice PBX is not a data switch, and it was not too long before customers figured out that something was amiss. This created a gap in the market that was soon to be filled by local-area networks supplied by non-traditional communications vendors. The first of these, called the Ethernet® local-area network (Ethernet is a registered trademark of Xerox Corporation), attempted to reduce a network to nothing but a broadcast space (ether) in a piece of coaxial cable by making the communication connectionless and placing all contention resolution and other network functions in the terminals or adapters.

In a connectionless network it is relatively difficult to know what is going on inside the network, but on the premises, where needs can be met cost-effectively with an overabundance of bandwidth and where the network is not required to make money through charges for its services, it can make sense. In public networks and, in particular, future networks, which must provide a wide variety of services and recover costs rapidly in order to keep network services up to date with rapidly advancing

10

technologies, connectionless services do not, in general, make sense. There are those who put forth arguments for connectionless networks on the basis of avoiding call setup delay; however, a packet network with the ability to maintain statistical, virtual private networks is all that is required to avoid call setup delays.

Responding to Customer Needs

Historically, public networks have been slow to respond to customer needs. If response to needs remains coupled to major changes in a very large and costly installed base, future networks will continue to lose market to more responsive solutions. The response problems of large traditional networks result from the following causes:

1. *Software.* Network features and services are currently implemented in switching-system generic software programs by centralized software development groups working for the switch manufacturer. Once a need reaches an important enough level on a priority list that a design effort is started, two to three years will pass before a product emerges. Software-related delays have received the most attention. They are serious, although they are not the longest.
2. *Hardware.* Many network services, such as the ISDN basic rate service, are implemented with service-specific hardware in the network. To develop hardware to the point where it can offer merely limited connectivity to a target market can take more than six years.
3. *Standards.* In many cases, it will take more than a decade to reach agreement on an important interface standard, particularly in international standards bodies, because attempts by some participants to achieve self-serving results delay proceedings. The end results are rarely determined by customer needs and are frequently compromises that serve no interest well.

In addition to the above sources of delay, there are the following compound sources of delay:

1. *Hardware/software.* When network services are implemented with hardware-specific designs, instead of being determined by parameters with general-purpose hardware, the software development environment becomes destabilized. Even greater delays result.
2. *Network/terminal interface.* Terminal manufacturers are reluctant to develop terminals for network services that are not widely deployed. Thus, substantial network development and deployment may be required before terminals appear to support a service.

A network that cannot respond to customer needs in a reasonable period of time will become by default merely a transport network. The response problems of today's networks are great, but not insuperable. We now possess the technologies and the concepts to evolve networks to wonderful new machines.

Ideally, a network should provide its services instantaneously to meet the needs of the application. Thus, the network itself should provide the characteristics determined by the application, rather than "omniscient" technologists and market planners. Applications software should be developed close to the applications by those most familiar with the customer needs. This means that the largest body of software developers will work on network applications software. However, it also requires the development of network operating system software with stable interfaces to the applications development environments. This desirable software structure for the network can result if the transport fabric is both general-purpose and consistent with processing structures.

The Fusion Point

There exists a nearly magical point at which the difference between communications and computing disappears. That point is reached when the network is a priority-controlled, instruction-rich, high-speed packet fabric. Such a network can emulate a processor bus between distributed processors on the network. It can provide all the basic transport functions. It can establish

11

virtual circuits with bandwidths to match the requirements of the broadest range of the most common applications. It can provide on demand services that are determined by the application. It can provide statistical, virtual private networks to support the creation of virtual machines. It can support a terminal interface that accommodates the most common applications and terminals. It can support multimedia applications and the information structures that interrelate segments of voice, image, and data on a logical, temporal, and spatial basis to make information understandable. And, above all else, it will permit the development of a viable network software environment that will permit networks to achieve their intellectual (and revenue) potential.

The Network Computer

The transport fabric of the future network is the bus of a new kind of computer. It is not just a transport medium with computers and databases attached to it. Nor is it just a computer on a grand scale; it is a new machine, a distributed multimedia computer that encourages the creation of new software structures, processor architectures, and terminals. The potential of public networks has barely been scratched.

Public Networks as an Information Marketplace

Beyond the fusion point, there exists a public network that is the largest-scale computer and the gateway to an abundance of information resources that can be assembled in any combination to serve customers' needs. This network computer also can assemble the charges and bill the customer, then distribute payments to the participants in the provisioning of any service. Thus a marketplace is created. The marketplace gives immediate access to a very large market for service creators and tremendous and immediate rewards for those who serve the public well. There will be a great incentive to create new and better services. This large and immediate market will result in competition among the value-added

resources that are attached to the network. It will create cost economies for customers. And it will generate benefits to society that far exceed the benefits that can be derived by constraining networks to a dumb transport role and minimizing the cost of a delivered bit.

Evolution

There are many possible branches of evolution for future networks, but some are much more powerful than others. It is actually easier and less costly to take the path toward the general-purpose network computer that was envisioned as Universal Information Services than to continue the current practice. It is easier and less costly to deploy general-purpose hardware than to develop, deploy, and attempt to maintain hardware-specific service solutions. The great bonus is the productive software environment and the growth of applications that results when a great body of programmers who are close to the problems become the sources of the solutions.

A great gap, and one of the greatest business opportunities, exists between terminals that are based on widely available, inexpensive, powerful processors and public networks that still reflect their historic foundations.

The Greater Good

When we make choices about the nature of future networks, we are making choices that will have the most profound impact on our society. In an information-age economy, the capabilities of public networks will directly affect the efficiency of our industries and the quality and availability of education. Those who choose to limit the capabilities of public networks will forgo the benefits of the most flexible of all processing machines and the efficiency of an information marketplace that optimizes all the value-added resources that compete in it.

Of course, there will be more intelligent and less intelligent networks. It is quite likely that those nations that make the wrong choice for the future of their networks may never recover because of the time and ex-

12

pense involved in changing a national network. Further, it may be that those societies that do not have a free and open exchange of information—and do not maintain good relationships with other nations participating in interconnection of "high IQ" networks—will fall behind at an ever-increasing rate. Intelligent networks will nourish the minds of humans and the minds of humans will nourish intelligent networks. The most powerful network computer, with the widest selection of attached resources, including human resources, will become a strategic national resource. It is information productivity that will determine the successful participants in an information-age economy.

13