

A DISTRIBUTED HOSPITAL INFORMATION SYSTEM

A distributed hospital information system is operational at the University of California at San Francisco Medical Center. The system uses a fiber-optic local-area network developed at APL.

INTRODUCTION

The problems and costs of processing information in hospitals have attracted the attention of private industry, the executive and legislative branches of the federal government, the Department of Health and Human Services, and the Department of Defense. Hospital information systems have been analyzed, developed, implemented, and evaluated for over two decades at a cost of over \$100 million. Despite this large effort, there is still a need for more acceptable alternatives to the present systems.

Two different models, or "architectures," have been proposed as the basis for hospital information systems: the centralized system model and the distributed system model. Centralized systems incorporate a set of hospital software applications in a single large computer sharing common files. Several commercial systems of that kind are available. Almost all the documented experience and evaluation to date have focused on the centralized approach.

Distributed information systems employ several computer systems, possibly of different design, and use a communications subsystem to solve a common problem. During the last few years, technology has progressed to the point where it seemed feasible to design and implement a distributed hospital information system and to document the experience.¹⁻⁵ This paper describes such a system at the University of California at San Francisco Medical Center. The local area communications network technology for this project has been developed and installed at the Medical Center by APL. It represents a generic approach to problems common to many hospitals.

PROBLEM AND ENVIRONMENT

The hospital information system design remains a challenge because of a combination of characteristics of the hospital environment. There is significant complexity in the "standard" financial and administrative functions and for the specialized, sometimes unique, clinical functions. Billing, cost reimbursement, and accounting are enormously complex. Charge capture is tightly coupled to clinical service functions. Clinical registration and clinic scheduling are more complex than standard reservation systems (as in airlines) since resource utilization is much less

predictable. The complexity of clinical functions has led to the development of stand-alone minicomputer systems for ancillary services and patient management to provide specialized functional support not available in single centralized systems.

The hospital environment presents a broad range of heterogeneous functions that must be automated, as a result of the diversity of clinical services provided. Furthermore, the computer needs of these services are quite different. Clinical laboratory and intensive care units involve on-line instrumentation and real-time data processing; radiology involves text processing; pharmacy requires maintenance of large, volatile files.

This diversity of function and associated data must be integrated to provide useful and correct applications support. A subset of patient data is common to all functions, including patient identification, demographics, financial information, status, location, and clinician. Multiple entry and redundant copy maintenance of this information is labor-intensive and will invariably result in errors. Furthermore, integration of functional area data is necessary for clinical management, decision-making, and billing.

Systems must be flexible, extensible, and easily modifiable to support the specific needs of individual hospitals and be able to incorporate new developments, growth, or change.

COMPARISON OF MODELS

Centralizing the computing resource solves many of the problems discussed above. In particular, centralization avoids multiple entry of identical information and redundant processing and storage of this information. Also, the various functional modules, implemented on a single machine with common software and data base, can all access any of the data needed.

However, centralization imposes certain undesirable constraints. Each participating department does not have autonomous control over its applications. Modifications of a single application may ripple through many parts of the system, complicating maintenance and restricting change. Hardware changes or additions present difficulties due to the system complexity. Thus, centralized systems are not flexible and are not easily modified or extended.

Some vendors do provide facilities for users to define interfaces flexibly (screens, reports), but software modification is often difficult or impossible.

Many medical centers currently have decentralized, but not (fully) integrated, systems. The components of such systems tend to include various ancillary systems, separate financial and administrative systems, and sometimes even a separate information system. Integration of such systems can be achieved but is often expensive in terms of new applications and specialized communication software development. Also, information transfer may not be rapid; typically, tape or batch transfer modes are used, and thus multiple entry and maintenance of common information become a problem. In addition, each computer-to-computer interface is regarded as a new project involving new communications rules (protocols) and software. This approach is both inefficient and difficult to maintain. However, since the heterogeneity and complexity of clinical functions have spawned several specialized systems over which departments need and want to have control, there is a growing demand for integrating these distributed systems.

The distributed approach allows modular implementations of each of the departments' application systems, allows each department to retain control of its own system, and permits easy hardware and software upgrades.⁶ However, the distributed approach generally requires the maintenance of a highly heterogeneous grouping of hardware and software. There is often redundancy of information between distinct systems.

A communications system is needed to support functional integration and one-time capture and distribution of common information. Fragmentation of patient clinical information into several physically distinct data bases poses data access problems for both active patients and for archived information. The overall system design must address these problems.

The basic obstacle to exploring the distributed hospital information system approach has been the development and application of communications support between diverse hardware and software systems that permits system and applications designers to solve the problems of common information distribution and multiple-system access to remotely stored data. This requires establishment of communications rules known as protocols. Since the distributed system is usually highly diverse, any given hardware/software configuration may require N sets of protocols for talking on N communications links. This requires massive amounts of work to add new systems, and it becomes extremely difficult to develop and maintain $N(N-1)$ communications programs.

One solution to the communications problem is the use of a local area communications network.^{7,8} This greatly simplifies the implementation and maintenance problems, since there is only one interface (to the network) for each hardware/software configura-

tion. But such a network does require that more depth be given to the protocols to permit the fair and accurate use of the network and adequate network support for applications development in the distributed system. This can put more of a burden on the host systems, or the protocols can be shifted onto a network interface unit between the host system and the network that would handle much of the protocol support that is required.

The local area communications network approach was adopted for this project. The University of California at San Francisco Hospital developed a set of distributed applications and served as a test bed facility for the system. APL developed a fiber-optic-based network, that contains microprocessor-based interface units to handle the protocols involved in communicating over a network.

PROJECT OBJECTIVES

The University Hospital has been developing the applications basis for a distributed hospital information system for several years and has acquired several component minicomputer systems. The hospital has 550 beds, and annually admits 20,000 inpatients and has 160,000 outpatient visits. The demonstration project began in February 1981. The objectives of the project are

1. To demonstrate the basic feasibility of the technology by integrating four heterogeneous ("incompatible") minicomputer systems serving major functions of the hospital;
2. To exercise the network as much as possible by performing functions at the application level that require communications among the four systems;
3. To document the costs, benefits, utilization, and problems associated with the use of the network.

INITIAL APPLICATIONS ENVIRONMENT

The initial demonstration project involved four minicomputers (hosts) interfaced to achieve a level of functional integration that forms an extensible core for a comprehensive information system. Figure 1 and Table 1 identify the four application systems and the general system configuration. It is completely heterogeneous with respect to machine/vendor models, operating systems, application software, and local data bases or file access methods. The Patient Identification/Registration System and the Radiology/Medical Record System are located in the Ambulatory Care Center on the north side of Parnassus Avenue. The Clinical Laboratory System is located on the fifth floor of Moffit Hospital on the south side of Parnassus. The Outpatient Pharmacy System is located on the twelfth floor of Moffit Hospital. The distance from the Ambulatory Care Center to the fifth floor of Moffit (Fig. 2) is about 400 meters. As indicated in Fig. 1, the logical and physical interconnections among the host computers are accom-

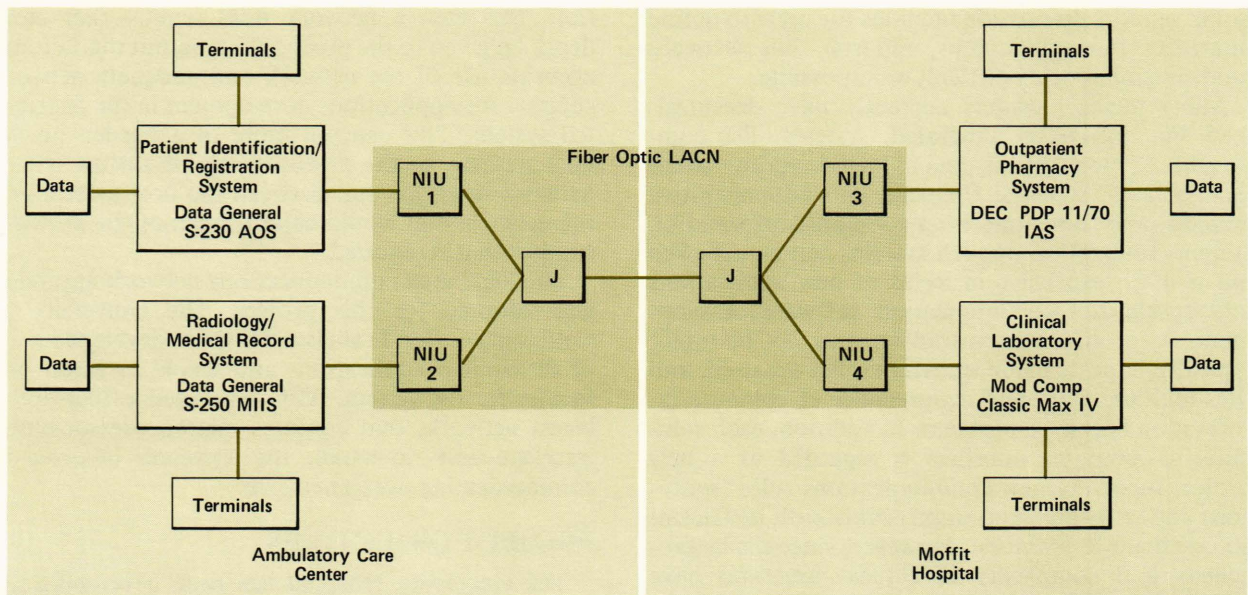


Figure 1 — The initial system implementation integrates four different hospital computer systems by means of the APL local area communications network. The hospital computers are located in two different buildings. Each computer is connected to a network interface unit, which provides the data communications services used to build the distributed hospital applications.

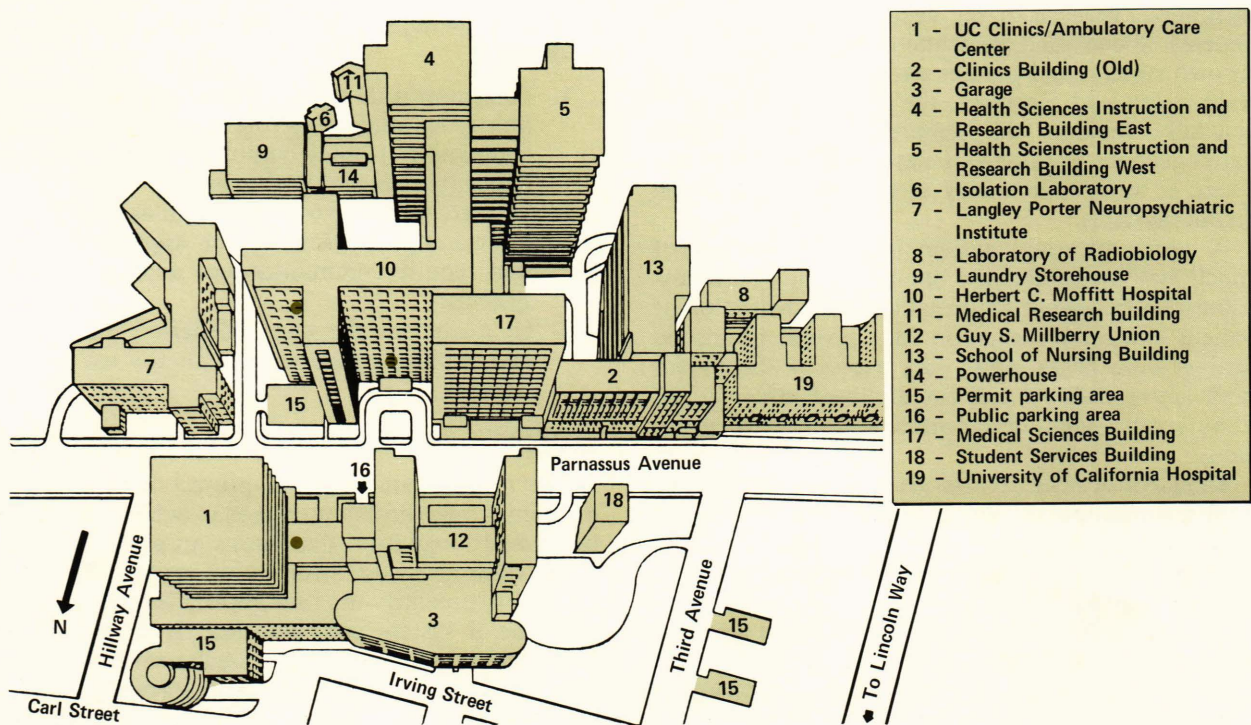


Figure 2 — Map showing the UC Clinics/Ambulatory Care Center and Moffitt Hospital. Computers in the initial system are located in the basement of Building 1 and on the fifth and twelfth floors of Building 10.

plished by means of the local area communications network. Each host computer system retains its own terminals and its own local data base. Although the network can support terminals directly, currently all remote access from a given terminal is through that terminal's defined host.

The Patient Identification System is the core of all systems in the hospital. Its function is to provide a

patient with a unique medical record number that can be consistently used at any time or any place in the medical center. Thus, the Patient Identification System file consists of a large number of patients (750,000), with a small amount of information per patient.

The Registration System contains records on patients who have open accounts with the hospital. No

Table 1—General characteristics of demonstration host computers.

CPU	Data General Eclipse S-230	Data General Eclipse S-230	Mod Comp IV	PDP 11/70
Core	512 kb	512 kb	512 kb	768 kb
Operating system	AOS	MIIS	MAX IV	IAS
Applications in demonstration	a. Patient identification b. Patient registration c. Inpatient reservation (future) d. Census (future)	a. Radiology b. Medical records c. Pathology (future) d. Operating room (future)	Clinical laboratory	a. Outpatient pharmacy b. Inpatient pharmacy (future)
Hours of operation of demonstration applications	23 hours per day	23 hours per day	23 hours per day	10 hours per day, outpatient 23 hours per day, inpatient
Users of CPU other than demonstration application	None	None	None	Cardiac Catheterization Laboratory research
Location of CPU	A-57	A-57	Moffit 5	Moffit 12
Number of application terminals, printers, and peripherals	a. 35 OWL CRT's b. 2 character printers c. 1 embossing terminal d. 2 196-Mb disc drives e. 1 800/1600 disc drive f. 1 line printer	a. 43 OWL CRT's b. 7 character printers c. 15 megadata word processing d. 2 196-Mb disc drives e. 1 800/1600 bpi tape drive f. 2 line printers	a. 35 CRT's b. 4 character printers c. 10 laboratory instruments including real-time analog d. 2 256-Mb disc drives e. 1 800/1600 bpi tape drive f. 1 line printer	a. 13 CRT's b. 5 character printers c. 1 line printer d. 1 196-Mb disc drive
Application software	UCSF/Fortran	Purchased (CMS-MIIS)	Purchased (CHC-Fortran)	Purchased (custom Fortran)

patient can receive any services in the hospital unless an account has been opened for that individual. Thus, the Registration System contains a smaller number of patients than the Patient Identification System (approximately 130,000), but considerably more information per patient, primarily regarding insurance and payment information.

The Clinical Laboratory System is a general-purpose system for managing all laboratory data related to chemistry, hematology, serology, and blood bank tests. All laboratory requests are entered into the system, and all results are processed through the system. Most laboratory results are entered through direct in-

terfaces with clinical laboratory instruments. The remainder are entered manually. Reports are distributed via hard copy printed output, as well as through cathode ray tube terminal displays. Individual patient reports, nursing unit summary reports, and patient cumulative summaries are provided as hard-copy output. Input is generated automatically for the hospital billing system, and management statistics are also provided.

The Radiology System contains functions for patient log-in and arrival, X-ray film tracking, patient flow control, management statistics, X-ray reporting, pathology coding, and billing. X-ray reports re-

main available for on-line display and may be retrieved by diagnostic code. The American College of Radiology coding scheme is used.

The Outpatient Pharmacy System was implemented during 1981. All prescriptions filled at the pharmacy are entered into the computer. Prescription editing, drug-drug interactions, and allergy checks occur automatically. Patient drug profiles are displayed on-line, and prescription refills are facilitated by the system.

Table 1 describes the initial characteristics of the host computers. Four distinct operating systems are involved. The AOS machine is a multiprogramming, multitask system. IAS is multitasking, but each task is a program. MIIS (a MUMPS variant) is a time-sharing system, with each program executing independently. MAX IV is a time-sharing system that checks each executing program periodically for status changes. Each system handles communications line inputs differently. The initial interface to each computer is a 9600-baud serial (RS-232) link.

INITIAL SYSTEM

Local Area Communication

Network Components

The network components consist of network interface units, fiber-optic links, and "junction boxes" to interconnect these links. The interface unit is a multiprocessor consisting of two Z-80 microprocessors that communicate through a shared memory. One-half of the unit contains the hardware and protocol software required to interface to and communicate with the local computer, terminals, peripherals, or other network subscribers. This is called the "host" half. The other half contains the hardware and software necessary for interfacing to the fiber-optic bus and communicating with other units. This is termed the "network" half. Each junction box provides up to six full-duplex fiber-optic ports, permitting interface units to be linked to the network and permitting extension of the network by links to other junction boxes. In the junction box, an optical signal arriving at a receiver at any given port is distributed to the transmitters at all other ports.

The use of fiber optics as the communications medium on the network is not based on the bandwidth requirements for the system. Actually, the hospital network is running at 0.5 megabit per second, far below the cable rating of 25 MHz-km, and also far above the likely utilization level. Fiber optics was chosen for the potential bandwidth, low noise, and relative security. In hospital environments, noise can be a significant problem due to the presence of fixed and portable radiology equipment. Furthermore, fiber-optics cable is easy to install in existing facilities and is not a significant factor in the system cost. Hence, it seemed advantageous to use a fiber-optic network despite any reservations about its novelty compared with a more traditional medium such as a coaxial cable.

The source network interface unit disassembles messages into smaller pieces called packets. These packets are reassembled into messages at the destination. Packets of variable length up to 256 bytes are broadcast to all units except the source. Each unit samples packet address headers received and ignores all packets not addressed to it. Recognition of packet addresses and aborts due to packet collisions is achieved by the interface unit hardware. Packet sequence and data integrity are maintained by use of a balanced synchronous data-link protocol. Complete messages are delivered across the interface to hosts.

Initial Network Services

The logical design of the protocols within the network interface units was guided by the International Standards Organization Reference Model,⁹ with the major emphasis on providing as much communications support within the unit as possible. The software in all the interface units is identical. This software was developed using a high-level language in modular structures associated with the levels in the International Standards Organization Reference Model. These modules were mapped onto the dual processors, with the network-side processor performing the physical, link, network, and transport layers, and the host-side processor performing the session layer and the interfacing between the host and the interface unit. A description of the protocol functions may be found below.

The APL network provides a variety of services, including virtual circuit, single message delivery, matched query/response, and acknowledged subgroup broadcasts. It also handles several network control messages, which can be used to change or measure the performance of the network. Virtual circuits are used to provide host-port to host-port delivery of messages, with messages arriving in sequence. Each message is directed to a particular user process by means of a header known to both the host and the network interface unit.

One major category of service is query/response session management. When a user (e.g., program) wishes to send a query message to a destination host, the session layer is invoked to manage the exchange. Three fields from the query are saved as a query identifier; this character string is placed in the header of every returning response message that matches the query. At user request, there may be a time-out placed on the responses. If a time-out occurs, an error message with the query identifier is returned to the originating host. The session layer permits multiple outstanding queries on the same connection between interface units. When each response message is received, the query identifier is placed in the header by the interface unit and the message is delivered to the host.

Currently, communication is largely between the Patient Identification/Registration System and each of the others. Common patient identification and registration information is exchanged among these

systems. This is done to distribute to, and improve the consistency among, all of the local data bases.

Whenever a change is made to a patient's identification or registration information, the updated record is sent over the network using the broadcast message service. A broadcast message is delivered to all other nodes on the network (Radiology, Pharmacy, and Clinical Laboratory). If sometime later the patient's information is needed at the pharmacy system, for example, the local data base is checked, and the information is likely to be present. Occasionally the information may not be present for a number of reasons, such as a host being down for maintenance, purged or lost records, or no prior visit to the ancillary unit subsequent to network installation. In this case, the query/response service is used. The host sends a query along a "permanently" established connection to a host that has the information needed (usually the Patient Identification/Registration System), which then sends a response back to the originating host. Matching of queries and responses is done within the network, which answers with an error message if a response does not appear within a certain time interval. For example, if a patient's information is not available locally to the pharmacy system, it sends a query to the Registration System, requesting the data on the patient, which system then sends the information back via a response message to the pharmacy system. Network processing of the query/response includes error checking, flow control, time-outs, and matching of multiple responses to the query.

PROJECT HISTORY AND SYSTEM EVOLUTION

History and Experience

The first year of the demonstration project has now been completed. The first two network interface units were delivered in August 1981 and were installed to connect the Patient Identification/Registration computer to the Radiology Medical Records computer.

The third unit was delivered in December 1981 and was installed in a production mode for the Outpatient Pharmacy System in January 1982, thus creating a three-computer network. The fourth unit was installed and tested in February 1982 for a new Clinical Laboratory System that became operational for the hospital in May 1982.

Four network interface units, two junction boxes, and approximately 500 meters of dual-strand fiber-optic cable constitute the network components. Use of the network has been limited to the key functions of synchronizing patient identification and registration information among the four computers.

One of the most important goals was to reduce the application level effort required to communicate to multiple computer systems in a hospital information system. The relatively small effort (4.6 to 7.3 man-weeks) required for each of the systems to use the

network is well within expectations. Each of the computers has gained access through its network interface unit to all other computers currently on the network and to those that will be added in the future.

Utilization of the network is currently extremely low compared to network capacity. During the first six months of use, when there were only two computers on the network, there were approximately 70,000 broadcasts (about 400 per day) and 8,000 query/response pairs (about 45 per day). The peak transaction rate occurs between 2 and 3 PM on Tuesdays, with an average of approximately 120 transactions during that hour. The transaction rate is insignificant for a fiber-optic bus that currently operates at 500,000 bits per second. With the current four computers and the planned addition of several more computers over the next 18 to 24 months, the transaction rate will obviously increase. The number of broadcasts from the Patient Identification/Registration System will remain relatively stable, since each broadcast goes to all computers. However, the number of query/response pairs will increase rapidly as more applications and more functional uses of the network are added. It is not anticipated that the fiber-optic bus media will ever be stressed. However, other points of stress may become significant. The current computer-to-network interface unit links are standard RS-232 at 9600 baud; it is conceivable that a particular link will become saturated. This will necessitate going either to a parallel interface or to multiple RS-232 ports between a unit and its host (or possibly adding a second unit for a given computer). Another issue that may eventually affect performance is the limited buffer capacity in the interface unit. Since complete messages are delivered across the host interface to avoid the need for host reassembly, long messages may degrade performance. Until higher volume of network use occurs, we can only speculate on these potential limitations.

The Broadcast Problem. The major problem revealed by the first year's experience relates to the use of the broadcast transaction. The relative importance of this transaction type was not fully appreciated until we had experience with several applications. Broadcasting the registration of all new patients, both inpatient and outpatient, to all data bases and creating a record on each data base at that time is a highly desirable feature. This is true even if many patients never receive the clinical service provided by that particular application. Having the patient record already established by the time of the first entry of a lab or X-ray request, for example, eliminates the need to query the Registration System for that record. Not only does that reduce delays in the operation of the applications, but it reduces the use of the network and the risk that the record will be unavailable through the network because of registration downtime or other hardware/software failures. These benefits far outweigh the problem of having to purge unused patient header records from the appli-

cation data bases. In view of the importance of this transaction type, the software design at both the application level and the network interface unit level was found to be inadequate to (a) ensure delivery of the broadcast, (b) monitor the broadcasts to ensure that they were successfully delivered to each computer, and (c) provide mechanisms to rebroadcast in case of failure. A new assured broadcast service has been designed to overcome these limitations, and the units were all upgraded with the service in July 1982.

Technical Issues of Assured Broadcasts. In a broadcast service, a message transmitted by one host computer is expected to be received by all the other computers in the network or at least by those hosts selected to be in the particular broadcast subgroup. The broadcast service is particularly well suited to a contention bus network, since the communications medium is common to all network devices and every device can easily be made to receive anything that is transmitted to the bus. However, a major problem is that broadcasts are not acknowledged. Acknowledgment of broadcasts would require receipt and coordination of many acknowledgments from the destination nodes. The acknowledgments might collide, and the sending device would have to resend the message until all destinations acknowledged receipt. Furthermore, the sending device would have to know either how many acknowledgments should be expected, or what destinations were expected to return acknowledgments. This acknowledgment coordination problem is not easy, especially on a bus where control is distributed and all network devices have equal bus access rights.

In the University of California network, the standard broadcast service has been enhanced by employing the existing reliable datagram service to provide assured and acknowledged broadcasts. In the datagram service on that network, a resource identification field is provided at the header of the message. The field identifies the destination device or process (e.g., PHA for the pharmacy system). The message is guaranteed to be delivered error-free to the destination, assuming there are no insurmountable problems such as a dead network or host device, or a break in the fiber-optic cable. If the datagram cannot be delivered, the sending host is notified by means of an appropriate error message. Because a virtual circuit is not created for datagrams, it is possible that a message may be duplicated at the destination. However, a message will never be lost. This level of datagram service should be contrasted with the less reliable type of service provided by most commercial local networks, in which datagrams are similar to letters dropped into a mailbox: there is a good chance that the message is delivered, but there is no guarantee.

The assured broadcast service provided by our network is based on sending datagram messages to each destination in the broadcast group. Because each datagram message has assured delivery when the network and hosts are functioning properly and the

sending host is notified if a problem prevents delivery, the broadcast built upon the datagrams will be reliable. However, because under certain rare circumstances messages may be duplicated, the application must be idempotent; that is, if a message is received twice, the application should be unaffected. Database updates resulting from patient identification broadcasts are idempotent.

The assured broadcast begins when the network interface unit receives a message from the host computer with an assured broadcast indication in the header. The unit then consults a table that has a list of all the resource identification names of the destination processes. The message is sent as a datagram to the first resource. If the message cannot be delivered after retrying a certain number of times, the sending host computer is notified. The message is then sent to the next destination on the list, and so on, until the message has been sent to all destinations. At that point, a *broadcast complete* message is returned to the sending host notifying it that an attempt has been made to send the message to all destinations. Even if the message has not been delivered to any destination, the broadcast complete message is returned. Thus, the datagram-supported broadcast service provides a guaranteed and acknowledged broadcast capability, and the sending host is notified if the message cannot be delivered to a particular destination. This level of service should be contrasted with the standard unacknowledged broadcast, where resends are not even attempted in the event of a checksum error, and the sender never knows who did and who did not receive the message.

Current Extensions

Planned growth of the hospital information system is illustrated in Fig. 3. An additional Data General Eclipse S-250 was installed in 1982. The existing medical record system was transferred to that machine, and a summary-time-oriented Medical Record System began operation on the network. The system is expected to place a greater load on the network than do current applications. Because it keeps a redundant data base rather than trying to assure consistency, queries across the network are used to obtain current data unless a remote system is unavailable. An operating room system that handles scheduling, management control, patient and surgeon times, and billing is also being implemented on the machine. A pathology system has been installed on the other S-250, which continues to host the radiology system. A census system was introduced in the fall of 1982.

A clinical display system for nursing unit and physician use is currently being installed on a Data General S-140 computer located in Moffit Hospital. Initially, eight nursing units in the intensive care unit will be serviced by this system, which may grow to support 12 to 18 terminals. Additional display systems will be brought up as needed. They do not maintain a local patient data base, but rather query several

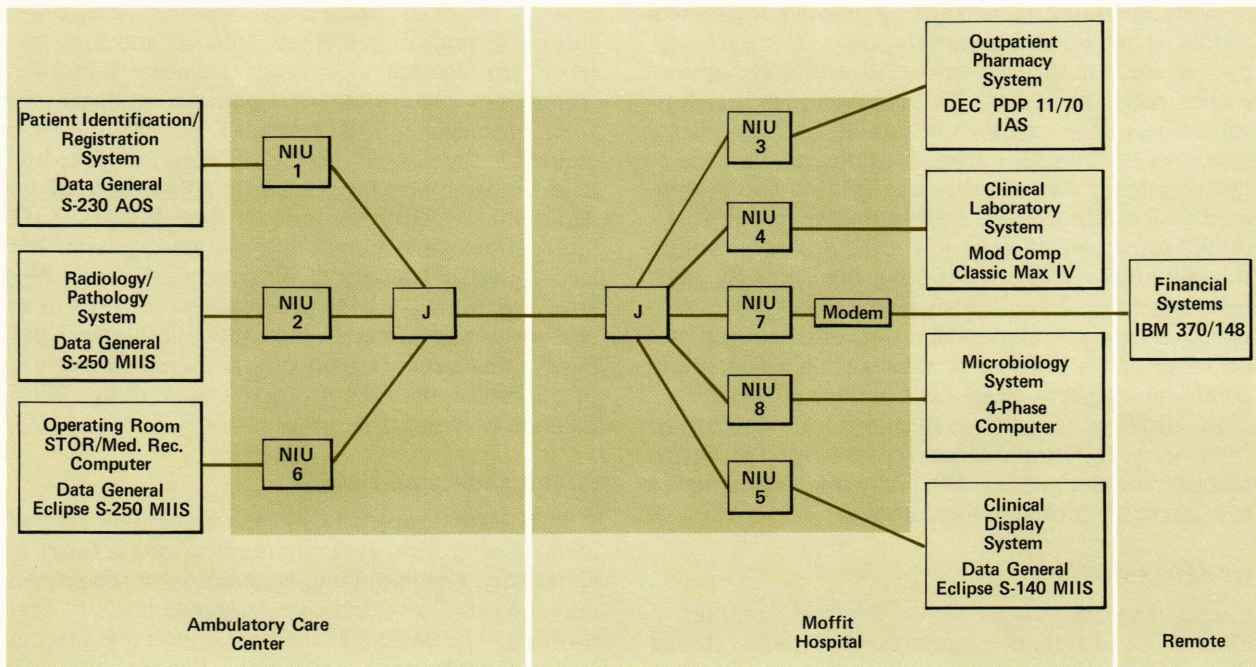


Figure 3 — Currently, the initial demonstration is being extended to include three more computers and several new software systems. Also, software systems have been moved to new computers as part of the reconfiguration. The new systems are Clinical Display, Summary Time-Oriented Medical Record (STOR), Pathology, Operating Room, Microbiology, and Financial.

other systems for patient clinical information, which is viewed as a series of “pages.” These response pages are simply displayed on the clinical display terminal that originated the request; they are formatted by the source systems as standard displays. Thus, responses consist of several messages of up to 2 kilobytes each, which the network matches to the query and delivers to the clinical display system. Several such query/responses may be concurrent. Thus, the traffic across the interface may saturate a single port link.

Plans for 1983 include integrating the financial systems on an IBM 370/148 located about five miles from the Medical Center. This will be accomplished either by use of a modem on a network interface unit or by using a local IBM Series I computer, which would be linked to the remote system.

THE APL LOCAL AREA COMMUNICATIONS NETWORK

In this section, the APL local area communications network is described in greater detail. There have been a number of recent efforts to design improved communications support for diverse computer resources through a local area network system.¹⁰⁻¹⁴ Each approach reflects an architecture keyed to a given set of objectives. The APL network is a new approach to local area networking that is aimed at improving network flexibility and extending the range of services to subscribers.

The APL network has a broadcast packet switch network design that uses a fiber-optic contention bus

and “smart” interface units to support high performance data communications among heterogeneous terminals, computers, and digital devices.

Several local area network topologies, including star, ring, and bus, have been proposed and developed. Because of the relative benefits in a distributed control environment,⁸ attention was restricted to bus architectures. Existing local area network implementations with bus architectures may be characterized by one or more of the following operating features:

- Coaxial cable based
- Electromagnetic intrusion or extrusion or both
- Head-end driven “bus”
- Active bus devices (for example, radio frequency repeaters)
- Central or master controller of bandwidth allocation
- Low-level protocols in net interfaces; high-level protocols in subscribers

Potential problems of coaxial-based networks include intrusion and radiation of electromagnetic signals, difficulty in installing large coaxial cables in existing facilities, lack of physical security of the transmission medium, and lack of efficient full-duplex communication due to a physical bus structure. Time-division multiplex schemes using a master controller for time-slot allocation and networks with a head-end driver are vulnerable to catastrophic single-point failures. Broadband coaxial systems present radio frequency cable plant maintenance problems.

A major limitation of existing network implementations is the limited communications software support resident in the network relative to that required within each subscriber. As long as individual host subscribers must deal with "dumb" or unreliable networks for data exchange, the potential benefits of heterogeneous digital system integration and distributed processing may be largely unrealized. Furthermore, provision of only low-level protocols within the network will likely involve a high cost for each new subscriber. Even if high-level protocols are standardized, a costly implementation effort would still be required to connect new hosts and to maintain communications software for extended periods.

In addition to a general interest in developing fiber-optic communications technology, the design and development of the APL network were initiated to address the problems cited above.

Design Objectives

The network design is intended to support a number of different objectives. Foremost among these is expansion and integration of user services in an environment involving network and computer system development, laboratory activities, and office support functions. A particularly desirable communications service is network support of simultaneous data transfer between any two subscribers.

A further objective is to provide secure, reliable, broadband data communications free from electromagnetic intrusion and radiation that may be easily installed in existing facilities.

Data communication must be reliable in two different senses. First, the integrity of bits, bytes, packets, and messages must be assured with an error-detection, -correction, and -reporting capability. Second, without resorting to component redundancy, the network must not be vulnerable to single-point catastrophic failures.

For growth, maintenance, and configuration flexibility, the network should be an open system architecture implemented with modular and structured design principles. In addition, the engineering design of network interface devices should support the progressive transfer of communications functions from hosts into the network.

System Overview

A network system satisfying the design objectives outlined above will be reliable, secure, fault tolerant, and capable of concurrent service to many diverse systems.

Each node of the network is logically fully connected to all other nodes via a programmable network interface unit and one or more active junction devices. A logical bus structure and contention access method with collision resolution form the basis for signaling data across the net. Also, the architecture supports simultaneous data transfer services between subscriber pairs.

A full range of communications support including short impromptu message broadcasts and long-term scheduled processing sessions between subscribers (sometimes called datagram and virtual circuit services, respectively) is provided. Communications control is distributed among network interface units to avoid or minimize effects on subscribers of network interface unit or junction box failures. Communications protocol functions ranging from low-level (physical transfer of bit streams) through high-level functions (such as process session control) are performed as required in each network interface unit. Distributed network control is achieved through use of a common network protocol system design within each network interface unit.

Architecture and Design

The term "duplex bus communications service" identifies our local area communications network architecture. The initial implementation of the physical network (see Fig. 4) consists of several levels of organization. The devices at the lowest level are host central processing units, terminals, or peripherals. Devices are connected to "network interface units," which are in turn connected to "junction boxes." The junction box gates all incoming data on a given port to all other ports, but not back to the port from which the data arrived, i.e., a one-to-all-others broadcast. (Therefore, with an appropriate connection management protocol, the network could handle true full-duplex sessions between any pair of nodes. See Fig. 5.)

Each network interface unit competes with all others for use of the common communications channel. By listening on the channel until no signal is detected, a network interface unit can determine when to begin transmitting data. Another unit may also attempt to transmit at approximately the same

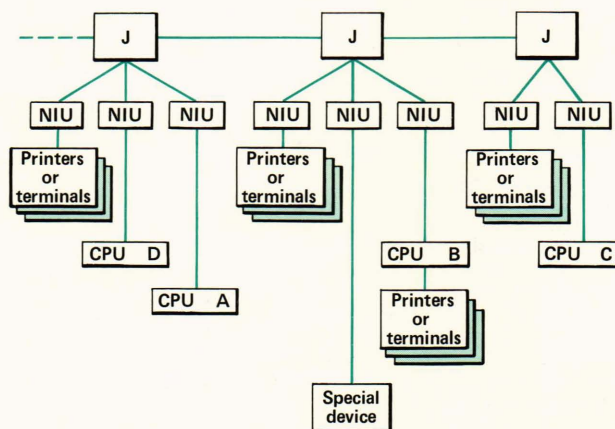


Figure 4 — A local area communications network provides the physical and logical connectivity and the data communications support needed to allow various computers and other devices to exchange data. The APL network provides junction boxes, J, allowing branching of the fiber-optic bus and intelligent network interface units to support computers and devices.

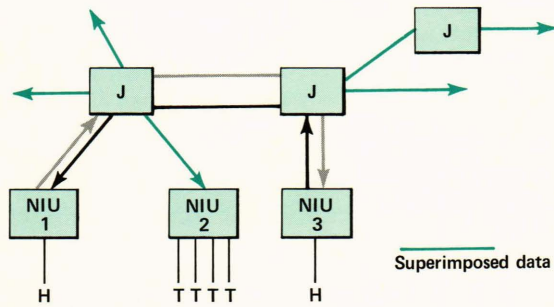


Figure 5 — The architecture of the APL local area communication network can support simultaneous exchange of data between two network interface units by using dual-strand fiber-optic cables. Junction boxes, J, distribute signals inbound on a port to all outbound fibers on other ports but not back to the originating interface unit.

time. This may result in the superposition of two or more messages on the common channel. Such an event, called a “collision,” is detected by the network interface units, which issue a network abort. This abort is achieved by flooding the network with a pulse of light that directs all receiving network interface units to disregard the packet. Each sending unit then awaits a randomly determined interval before attempting retransmission in order to avoid recurrence of the collision.

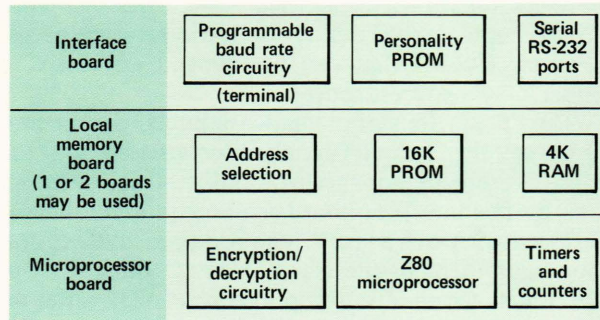
In the scheme presented above, each network interface unit contends with all others for access to the shared communications channel, which is like a contention bus. The combination of listen-while-talk for access and collision detection logic is known as the Carrier Sense Multiple Access with Collision Detection algorithm.

Message packets of variable length are broadcast to all other network interface units. Each unit samples received packet address headers and ignores all packets not addressed to it. Messages are disassembled into packets at the source network interface unit and are reassembled at the destination. Recognition of packet addresses and aborts is achieved by interface unit hardware. Packet sequence and data integrity are maintained in the network by use of a balanced synchronous data-link protocol within the network.

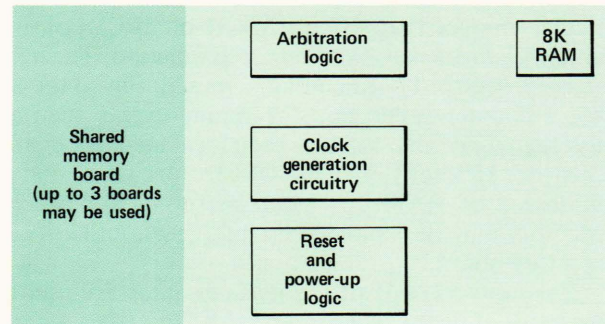
The architecture offers local reliability; all network interface units on either side of a failed junction box can continue to communicate with all units that are linked by operative junction boxes—there is no critical node. The network architecture is modular and expandable so that network interface units and junction boxes may be easily added and subscriber devices may be easily reconfigured.

Physical Design. The network components consist of network interface units, junction boxes, and fiber-optic links. The network interface unit is a multiprocessor consisting of two Z-80 microprocessors that communicate through shared memory. One-half of

(a) Terminal/host microcomputer module



(b) Shared memory module



(c) Network microcomputer module

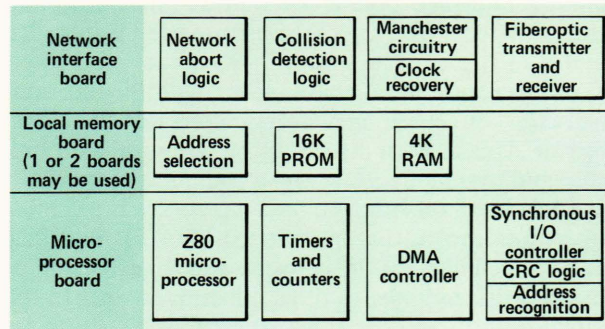


Figure 6 — Block diagrams of the two microcomputer modules and the shared memory module, showing the allocation of hardware functions in the network interface units.

the unit contains the hardware and protocol software required to interface to and communicate with the local computer, terminals, peripherals, or other network subscribers. This is called the “host” half. The other half of the unit contains the hardware and software necessary for interfacing to the fiber-optic bus and communicating with other units. This is termed the “network” half.

Each half of the network interface unit is a distinct microcomputer (see Fig. 6). Packets of data must be transferred between the halves of the unit, however, and this occurs through a two-port shared memory. The two Z-80 microprocessor data buses converge at the shared memory, and an arbiter controls all accesses to it. Thus pointers to data, rather than the

data themselves, are exchanged between host and network microprocessors. The shared memory consists of up to three boards of static random access memory (RAM) that can be mapped anywhere in either Z-80's address space.

The "host" half of the network interface unit consists of the host/terminal interface board, the host/microprocessor board, and a local memory board. The interface board may contain parallel interface logic (such as for Digital Equipment Corporation equipment or the Naval Tactical Data System) for a host access link, or up to four RS-232 serial interface ports (e.g., for terminals). One personality programmable read-only memory (PROM) per port is provided to tailor each port to its subscriber's particular characteristics (e.g., ASCII or EBCDIC), including a break-key baud rate step function. The host microprocessor board contains, in addition to the Z-80, a counter/timer chip for timing events such as message time-outs. Finally, each local memory board provides 16 kilobytes of PROM for program, and 4 kilobytes of RAM for local buffering and stack. Memory may be expanded by inserting another local memory board.

The network half of the network interface unit is made up of the network interface board, the network microprocessor board, and a local memory board. The fiber-optic transmitter and receiver are located on the interface board, as well as the abort and carrier-detect logic. Clocking information is included in the bit stream. The second Z-80 resides on the network microprocessor board, along with the link control chip, direct memory access interfaces for transferring data in both directions between the data link chip and shared memory, and a counter/timer chip.

Each junction box provides up to six full-duplex fiber-optic ports that permit (a) network interface units to be linked to the network and (b) extension of the network by links to other junction boxes. In the junction box, an optical signal arriving at a receiver at any given port is distributed to the transmitters at all other ports. Signal distribution within the junction box is done at transistor-to-transistor logic levels. The junction box must be powered; it is not a passive tap. The network architecture has been designed, however, in anticipation of future developments in fiber-optic passive taps and couplers that would provide an equivalent capability.

Each full-duplex fiber-optic link currently consists of dual step-index fiber cable with two standard connectors at each end. The two fibers carry signals in opposite directions. The fiber connectors mate to connectors on the transmitter and receiver modules.

Logical Design. Capabilities of the initial APL local area communications network include full logic connectivity of diverse subscribers; distributed control of access to network bandwidth; efficient, correct, and secure information exchange; high throughput with various modes of data exchange; and a nominal process-to-process management ability. The logi-

cal architecture has been guided by the International Standards Organization Reference Model,⁹ with major emphasis on developing comprehensive host-to-host communications support within each network interface unit.

A few definitions are useful for discussing logical design. A *connection* is a logic circuit between two endpoints on the network. Connections may be either negotiated by a host with its network interface unit or pre-loaded into a network unit on power-up. A *message* is the basic unit of information on the host-network access link. There are four types: (a) the data message contains the actual information that is to be transferred across a connection; (b) a network control message is used to change or query status information in any network interface unit; (c) error messages are sent by the unit to its host if a time-out or other problem in message delivery is determined; and (d) upon network power-up, an initialization control message is sent to the host. A *packet* is the basic unit of information on the fiber-optic bus and is limited to 256 bytes. The terms *input* and *output* are used with respect to the host or terminal associated with the network interface unit. For every protocol layer on the network (that is, between network units) there is an equivalent protocol for the host-network access link.

Certain terms describing communications services have become widely used. The datagram has been described above; generally, datagrams do not have assured delivery, error detection, or notification if not delivered. Also, there is generally no assurance that datagrams will be received in the order sent. In the APL local area communications network, reliable datagram service is provided, including error detection, retransmission, and notification if repeated delivery attempts fail. Broadcast service has also been discussed above, along with the enhancements for assured broadcast needed in the University of California application. Virtual circuit service is typically described as analogous to a telephone call. There is usually a set-up phase for circuit negotiation and establishment and a close-down phase for termination. While the circuit exists, messages are transferred in guaranteed sequence with error detection and correction (e.g., retransmission of packets with errors) transparently to applications using the circuit. (The analogy with telephone service breaks down for the error issue.) Also, after establishment, messages sent do not need addressing, although for local networks logic addresses are usually provided. Thus, when a virtual circuit has been set up, resources for communications are guaranteed for the duration of the circuit. For many computer-to-computer communications, which are often bursty with long intermessage times, the overhead of virtual circuits is not appropriate, and datagram service will suffice. For terminal-to-computer usage, virtual circuits are needed since both sides must have availability of mutual communications. The mapping of software functions onto the hardware is shown in Fig. 7.

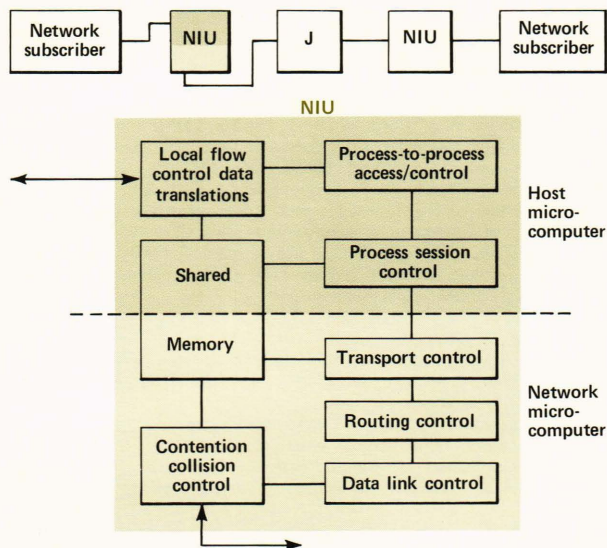


Figure 7 — Block diagram of software functions, showing the mapping onto the hardware modules. Functions corresponding to sub net protocol layers 1 through 4 of the International Standards Organization Reference Model are performed in the network half of the interface unit. Layer 5 functions and the host or terminal interface software are located in the host half of the interface unit.

The **physical layer** (layer 1) protocol contains the hardware and procedures for controlling access to a data link. In the case of the network, the data link would be the dual fiber-optic bus. Thus the fiber-optic transmitter, receiver, and cable fall in the domain of layer 1. Also included are the Manchester logic for encoding clocking information on the serial data stream, and the abort and collision logic used by the Carrier Sense Multiple Access with Collision Detection algorithm.

The **data link layer** (layer 2) protocol is responsible for delivering a packet over a physical link, for detecting and possibly correcting errors, and for handling network-to-network flow control.

The **network layer** (layer 3) protocol provides a network connection that is used to exchange messages. It handles packet assembly and disassembly, and switching and routing in the case of multiple links (e.g., a gateway). Layer 3 makes use of layer 2 to provide sequenced delivery of packets within a message.

The host-network interface unit access link layer 3 has an analogous responsibility for message assembly and disassembly. Upon output from the host, this protocol builds messages in shared memory from each port. Because each slot in memory is the same size as the maximum length packet, this layer distributes multi-packet messages among several memory slots and then links them. Upon input, the protocol has the responsibility for delivering received messages to the appropriate destination port and for freeing up memory after successful delivery.

The **transport layer** (layer 4) protocol assumes the burden for the sequenced, error-free, end-to-end delivery of messages across a connection. In this implementation, this service is supported by layer 3.

The **session layer** (layer 5) protocol provides various levels of service to coordinate dialogue between cooperating processes. It permits multiple outstanding queries on the same connection and keeps track of the returning responses. When each response message is received, the query identification is placed in the header and the message is delivered to the host.

In addition to coordinating query/response dialogues, the session layer handles network control messages that are used to modify or query the information in various network interface unit status tables. It also converts the connection number that is used by the sending unit into the connection number recognized by the receiving unit. The connection number indicates to the host which connection the message is associated with, and the host then can distribute the message to the appropriate process.

The access link (layer 5) gives the host access to the network (layer 5). Thus, upon output to the network, this protocol layer finds a connection for each message. Upon input to the host (when a response message is received, for example), the layer 5 protocol finds the appropriate destination port and associates the connection with that port for the duration of message delivery.

The **presentation layer** has the task of presenting information in a form meaningful to the applications process. This might involve character set conversion, encryption, data compression, and virtual terminal services.

The network interface unit assumes the burden of processing at various protocol layers. Figure 7 indicates the mapping of the layers onto software modules and the residency of these modules in the hardware. Note that the network protocols through layer 4 (transport) are located on the network side, and the higher level network protocols, as well as all the access link protocols, are located on the host side.

System Properties

For the initial implementation, the operating characteristics were chosen as follows. The total address space for each APL network is 8 bits, thus permitting up to 255 physical address nodes and a broadcast. Each network interface unit may serve up to four subscriber ports, each of which may serve multiple processes. The network operates reliably at 0.5 megabits per second in half duplex or simplex modes. This characteristic speed resulted from a decision to use off-the-shelf components and a simple design for the network interface board. Current work is aimed at increasing bus speed to 10 megabits per second. Based on the family of fiber-optic components chosen, the maximum distance between junction boxes may be up to a few kilometers. The maximum network packet size was chosen as 256 bytes based on

an analysis of optimal packet size; the maximum message size is not limited since full messages are not buffered in network interface units.

CONCLUSION

The joint University of California/APL project has been running consistently ahead of schedule. The first goals have been achieved successfully, and no major problems have been encountered. It is a statement of the suitability of the approach that very minimal host software development has been required of hospital personnel. The network has been operational continuously since initial installation in August 1981. A flexible distributed hospital information system based on a generalized local area communications network has been demonstrated and is supporting continued growth and new applications.

REFERENCES

- ¹S. G. Tolchin, R. L. Stewart, S. A. Kahn, E. S. Bergan, G. P. Gafke, D. W. Simborg, M. G. Chadwick, and Q. E. Whiting-O'Keefe, "A Prototype Generalized Network Technology for Hospitals: Initial Implementation," *Proc. 15th Hawaii International Conf. on System Sciences* (1982).
- ²S. A. Kahn, S. G. Tolchin, R. L. Stewart, G. P. Gafke, and E. S. Bergan, "Design and Implementation of a Fiber-Optic Contention Bus Network," *Proc. 15th Hawaii International Conf. on System Sciences* (1982).

- ³S. G. Tolchin and R. L. Stewart, "The Distributed Processing Approach to Hospital Information Processing," *J. Med. Sys.* **5-4** (1981).
- ⁴R. L. Stewart and S. G. Tolchin, "A Distributed Processing/Fiber-Optic Hospital Information System," 4th Annual Symp. on Computer Application in Medical Care (Nov 1980).
- ⁵M. L. Zeichner, P. J. Brusil, and S. G. Tolchin, "Distributed Processing Architecture for a Hospital Information System," 3rd Annual Symp. on Computer Applications in Medical Care (Oct 1979).
- ⁶Q. E. Whiting-O'Keefe, D. W. Simborg, and S. G. Tolchin, "The Argument for Modular Distributed Hospital Information Systems," 5th Annual Symp. on Computer Applications in Medical Care (Nov 1981).
- ⁷S. A. Kahn, R. L. Stewart, S. G. Tolchin, and S. J. Healy, "Functional and Logical Description of a New Fiber-Optic Contention Bus Network," IEEE COMPCON (Sep 1980).
- ⁸D. Clark *et al.*, "An Introduction to Local Area Networks," *Proc. IEEE* **66-11**, 1497-1516 (1978).
- ⁹"Data Processing—Open Systems Interconnection—Basic Reference Model," ISO/TC 97/SC16/N5 37/(DP 7498), American National Standards Inst., New York.
- ¹⁰N. B. Meisner and J. L. Segal, "Dual-Mode Slotted TDMA Digital Bus," *Proc. 5th Data Communications Symp.*, pp. 5/14-5/18 (1977).
- ¹¹R. M. Metcalfe and D. R. Boggs, "Ethernet: Distributed Packet Switching for Local Computer Networks," *Commun. ACM* **19-7**, 395-404 (1976).
- ¹²R. J. Carpenter *et al.*, "A Microprocessor-Based Local Network Node," *Proc. IEEE COMPCON*, pp. 104-109 (1978).
- ¹³N. Abramson, "The ALOHA System," *AFIPS Conf. Proc.* **37**, pp. 281-285 (1970).
- ¹⁴G. T. Hopkins, "A Bus Communications System," *Mitre MTR-3515* (Nov 1977).

ACKNOWLEDGMENT — This project was funded in part by Grant No. 1 R18 HS04196 from the National Center for Health Services Research.