Report Number: WUCS-92-52

1992-12-31

Project Zeus: Design of a Broadband Network and its Application on a University Campus

Jerome R. Cox Jr., Michael E. Gaddis, and Jonathan S. Turner

This is a report of the results of the initial step in a plan for the design, deployment and operation of a high speed campus network at Washington University. The network is based on ATM switching technology that has been developed here during the last several years. This network will support ubiquitous multimedia workstations with high-resolution graphics and video capabilities, open up a wide range of new applications in research and education. It will support aggregate throughputs of hundreds of gigabits per second and will be designed to support port of 100 MB/s is now in operation. The next... **Read complete abstract on page 2.**

Recommended Citation
Cox, Jerome R. Jr.; Gaddis, Michael E.; and Turner, Jonathan S., "Project Zeus: Design of a Broadband Network and its Application on a University Campus" Report Number: WUCS-92-52 (1992). *All Computer Science and Engineering Research.*
https://openscholarship.wustl.edu/cse_research/614

Project Zeus: Design of a Broadband Network and its Application on a University Campus

Jerome R. Cox Jr., Michael E. Gaddis, and Jonathan S. Turner

Complete Abstract:

This is a report of the results of the initial step in a plan for the design, deployment and operation of a high speed campus network at Washington University. The network is based on ATM switching technology that has been developed here during the last several years. This network will support ubiquitous multimedia workstations with high-resolution graphics and video capabilities, open up a wide range of new applications in research and education. It will support aggregate throughputs of hundreds of gigabits per second and will be designed to support port of 100 MB/s is now in operation. The next phase of network implementation will operate at 155 Mb/s port rates, with higher rates introduced as the demand arises and as economics permits. We propose to move this technology quickly into a production setting where the objectives of network use and network research can be pursued concurrently.

Project Zeus: Design of a Broadband Network and its Application on a University Campus

Jerome R. Cox, Jr., Michael E. Gaddis and Jonathan S. Turner

WUCS 92-52

December 31, 1992

Department of Computer Science
Campus Box 1045
Washington University
One Brookings Drive
St. Louis, MO 63130-4899

Abstract

This is a report of the results of the initial step in a plan for the design, deployment and operation of a high speed campus network at Washington University. The network is based on ATM switching technology that has been developed here during the last several years. This network will support ubiquitous multimedia workstations with high-resolution graphics and video capabilities, opening up a wide range of new applications in research and education. It will support aggregate throughputs of hundreds of gigabits per second and will be designed to support port interfaces at up to 2.4 Gb/s. An experimental prototype network with port rates of 100 Mb/s is now in operation. The next phase of network implementation will operate at 155 Mb/s port rates, with higher rates introduced as the demand arises and as economics permits. We propose to move this technology quickly into a production setting where the objectives of network use and network research can be pursued concurrently.

Project Zeus: Design of a Broadband Network and its Application on a University Campus

Jerome R. Cox, Jr., Michael E. Gaddis and Jonathan S. Turner

1. Motivation and Objectives

During the last several years there has been a growing recognition that *Asynchronous Transfer Mode* (ATM) cell-switching technology will form the basis of next generation communication networks. One attractive aspect of ATM technology is its inherent scalability, both in the total throughput a network can support and the port data rates. While much of the early focus in ATM was on public network applications, most people now agree that in the next few years, the demand for these new networks will come from computer-based applications needing higher bandwidth than current shared-access LANs are able to deliver. LAN and workstation vendors are recognizing the need to introduce switching within campus networks to expand network capacity and the range of applications, and are now moving aggressively to introduce products to fill this need.

Washington University has been deeply involved in the development of ATM switching technology, a technology based on small, fixed-length packets called *cells*, and the application of this technology to medical imaging. We have completed a project sponsored by Southwestern Bell and NEC America aimed at demonstrating technical feasibility of the underlying technology and providing some initial applications. We are now working with several industrial partners to create commercial implementations of the technology, to apply that technology throughout the university community for the benefit of users and to answer several pressing system questions that can only be addressed in an operational network environment. Thus, our objectives are threefold: 1) to collaborate with industry in the transfer of the technology for all the components needed to construct an ATM network; 2) to make available to users an advanced network with thousands of high performance workstations supported by the necessary hardware and software; and 3) to provide a realistic testbed for communications research addressing questions concerning network congestion, routing, planning, interoperability, interworking, remote visualization and techniques for operations and management.

Figure 1 illustrates the concept behind the proposed ATM network. The system will consist of several switches on each of the university's two campuses. The "hilltop" campus, at the left, is separated from the

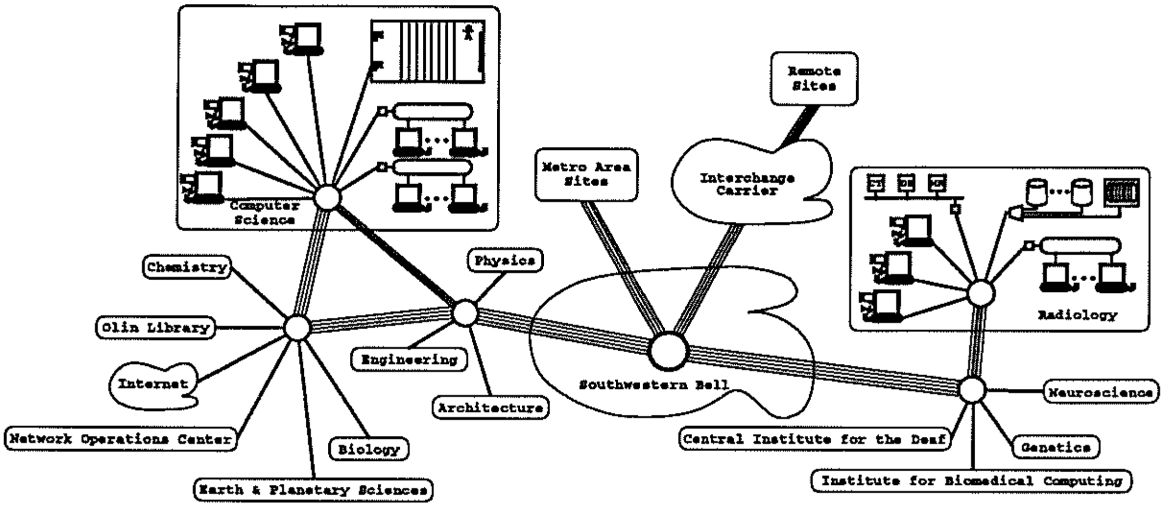

Figure 1: ATM Campus Network

"medical" campus, at the right, by 2.5 miles line-of-sight or 7.5 miles of single-mode fiber in Southwestern Bell's public network. The network's switches will be connected by transmission links operating at speeds of 155 Mb/s, 620 Mb/s and 2.48 Gb/s. Each switch would support up to several hundred interfaces, at a variety of port speeds. Most ports will operate at 155 Mb/s, but higher speed ports will be supported as the need arises. These interfaces will be connected directly to multimedia workstations and central compute servers or indirectly through shared access LANs such as Ethernet or FDDI. Video will play a central role in the network, allowing access to centrally stored video information through the network, two-way or multipoint video conferencing and remote classroom instruction using video.

The network will include connections to remote sites using either dedicated or switched channels provided by the Internet, the emerging National Research and Education Network, the local exchange carrier or interexchange carriers. In particular, connection to new broadband services planned by Southwestern Bell would make possible classrooms, medical offices and hospitals, all at locations more convenient to their clientele, and all linked to the university or the medical center by video, high resolution image transmission and shared databases.

2. Creating the Network Components

Project Zeus [5] is organized in three phases. Phase 0 was primarily experimental, demonstrated the feasibility of the core technology, provided a basis for a more complete design and serves as a testbed for application development. This phase of the project was begun in 1988 and completed in 1992.

Phase 1, begun in January, 1992, will run through 1994 creating all the key components of an ATM campus network including extensive support for application development. The completed phase 1 network will be an operational system supporting a variety of users in key departments within the university.

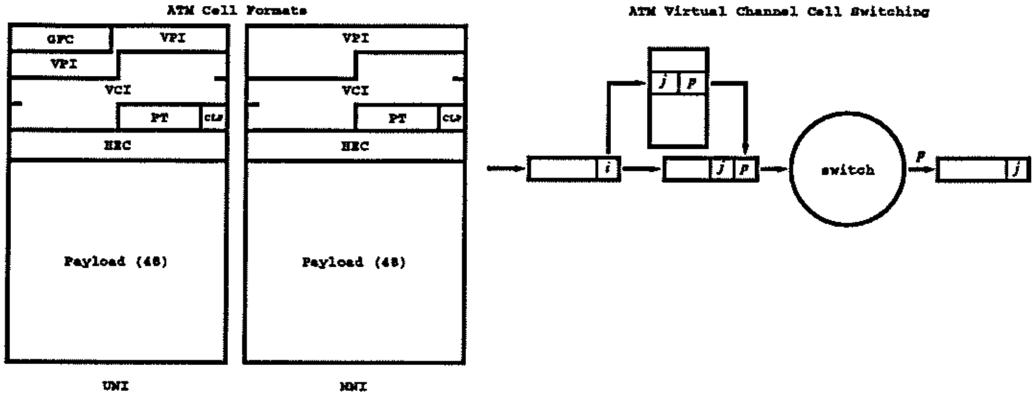

Figure 2: ATM Cell Formats and Cell Switching

During phase 2, which will run from 1994 through 1996, we plan to expand the range of interfaces that can be used to access the network, construct components for larger scale networks and reduce the cost of key network components. The phase 2 network will support users in all departments of the university.

2.1. ATM Network Technology

Project Zeus will take advantage of the emerging technology for *Asynchronous Transfer Mode* (ATM) networks. ATM networks provide virtual channel oriented cell switching, an internationally standardized form of packet switching in which user data is carried in small, 53-byte, fixed-length blocks called cells, where the first 5 bytes of each cell includes a label that identifies the user channel to which it belongs. Communication over an ATM network takes place over virtual channels which are typically established when some user application is initiated. When a virtual channel is established, a route is selected, and subsequently all cells transmitted on that virtual channel are forwarded along the selected route. Since the data objects transmitted by users typically consist of large amounts of data, it is necessary to fragment user data objects into cells and reassemble them at the receiver. This is accomplished using a simple fragmentation and reassembly protocol and supporting hardware in a host's network interface. Host software, is typically free to work with data units of arbitrary size, limited only by internal buffer space. All cell level processing can be left to the interface hardware.

Figure 2 illustrates the essentials of an ATM network's virtual channel oriented cell switching. As shown on the right side of the figure, the cell's multiplexing label is used to select an entry from a routing table; the selected entry includes the number of the switch output to which the cell is to be forwarded and a new multiplexing label. The figure also shows the ATM cell formats used at the *User Network Interface* (UNI) and the *Network-Network Interface* (NNI). There are two multiplexing options in ATM networks, *Virtual Paths* and *Virtual Channels*. Cells belonging to different virtual paths are distinguished by their *Virtual Path Identifier* (VPI) and those to different virtual channels are distinguished by their *Virtual Channel Identifier* (VCI). Most often, user connections are implemented using virtual channels. Virtual paths are essentially bundles of virtual channels. Two hosts can use a virtual path to multiplex many individual application streams together, using the VCI to distinguish these paths. The network does not interpret or modify the VCI fields of cells on virtual path connections, so the hosts can setup new virtual channels on an established virtual path without having to request them from the network. In this paper we don't generally distinguish between virtual paths and virtual channels; the reader may assume that statements made about virtual channels also apply to virtual paths. Other fields in the five byte header of the ATM cell include a *Generic Flow Control* field (GFC), a *Payload Type* (PT), a *Cell Loss Priority* bit (CLP), and *Header Error Control* field (HEC). The *Payload* field carries user information and is 48 bytes long [2, 4].

A key objective of ATM network technology is to ensure consistent performance to users in the presence of stochastically varying traffic. This is necessary to provide adequate performance for many high speed applications (such as video), which require guaranteed throughput. The objective is attained by selecting virtual channel paths according to the anticipated traffic and allocating the necessary network resources. For such guaranteed throughput applications, users must specify, at virtual channel setup, the amount of network resources they require. It also means that if the required resources are not available, a user's request could be refused by the network. The resource specification allows statistical sharing of bandwidth along virtual channel paths; users may specify their peak and average data rates, as well as a maximum burst size. Using this information, the network allocates resources in such a way that almost all information bursts are delivered intact. Soft resource specifications are also possible; users can specify minimum and maximum bandwidth requirements, or even no bandwidth specification at all, and the network will accommodate such requests to the best of its ability. In this case, the network can also adjust the resources allocated to established virtual channels in order to avoid blocking new virtual channel requests.

Conventional communication networks focus on point-to-point communication. Multicast communication, in which information from a single source is distributed to multiple receivers, is a natural generalization that is essential for efficiently supporting video distribution applications and useful in a variety of other applications as well. Figure 3 illustrates the basic concept of multicast virtual channel switching: a network consisting of a collection of switches (shown by circles) and concentrators (shown by trapezoids), configured to support two multicast virtual channels, with sources (shown by small squares) at the top of the figure. At each switch involved in a particular multicast virtual channel, the incoming cells are replicated, assigned new virtual channel identifiers and forwarded on selected outgoing links. The configuration of a particular virtual channel is specified in the switching systems' internal control tables and can be modified through control messages.

As in the case of ordinary virtual channels, the data rate of a multicast virtual channel is completely flexible, as is the number of endpoints. The number of endpoints can change dynamically during the lifetime of a virtual channel with new endpoints added and removed over time. One-to-many virtual channels are included as a special case of a more general communication model that is described in reference [8]. Briefly, a general multicast virtual channel supports transmission and reception at all the participating endpoints. For each endpoint, transmission and reception can be independently enabled, allowing a very wide range of connection configurations. The bandwidth resources associated with such a multicast virtual channel is viewed as a common bandwidth pool that is shared by the participating endpoints in whatever fashion they choose. Coordination of transmission by the different sources in a multicast virtual channel is left to the sources. The network merely monitors the total bandwidth usage and ensures that it does not exceed what has been allocated.

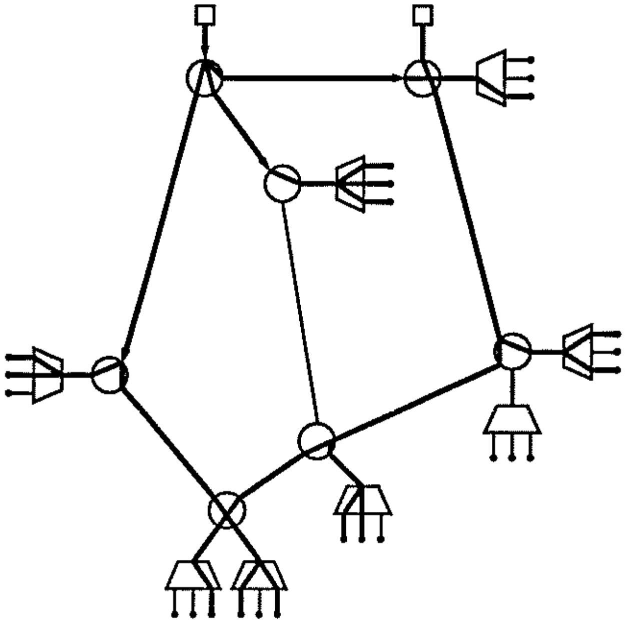

Figure 3: Multicast Virtual Channel Switching

2.2. Phase Zero

Phase 0 demonstrated the feasibility of the multicast ATM technology and applied it to several high speed applications. Specific achievements include the design and construction of a network consisting of four 16 port broadcast packet switches, design and implementation of an ATM video interface, an ATM Ethernet interface and a physician's workstation to support medical imaging applications.

Figure 4 shows the configuration of the network that has been constructed for phase 0. One of the four switches in the network is located in the Applied Research Laboratory of the Washington University School of Engineering and Applied Science (WU EN), one at the Electronic Radiology Laboratory of the Mallinckrodt Institute of Radiology in the Washington University Medical Center (WU MC), one at the Southwestern Bell Advanced Technology Laboratory in downtown St. Louis (SB ATL) and the fourth at Southwestern Bell Technology Resources, Inc. (SB TRI) in St. Louis County. These sites are connected using single mode fiber links provided by Southwestern Bell. The network supports interfaces to a *broadband terminal* supporting video, a *physician's workstation* for medical imaging applications, and Ethernet.

A total of eight phase 0 switches have been built to our specifications by SynOptics. Four are in service in the network, one is a spare and three have been purchased by collaborators who are interested in experimentation with ATM switching systems (Bellcore, Sun, SynOptics).

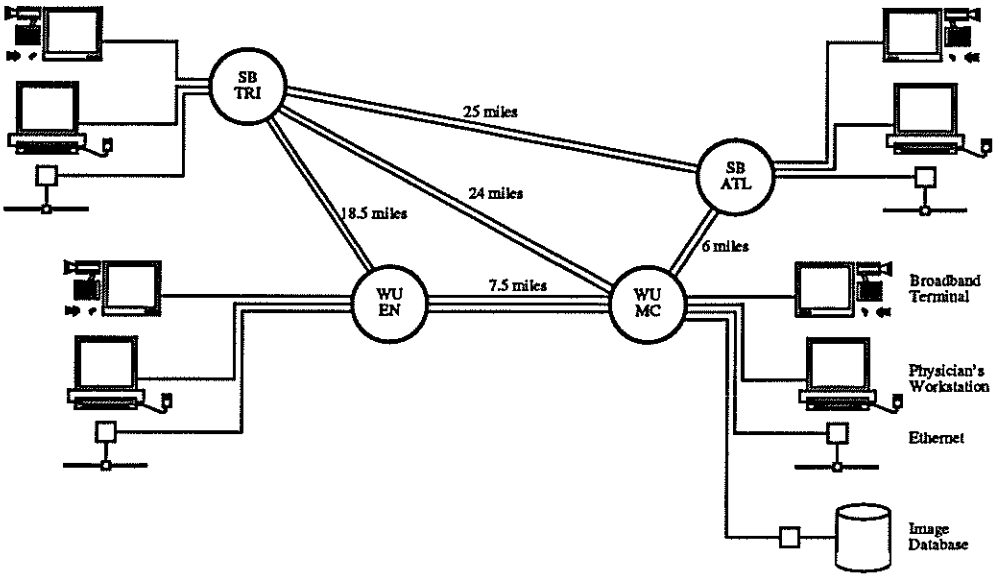

Figure 4: Phase 0 Network Configuration

2.2.1. Phase 0 Switch Architecture

Reference [12] describes a broadcast packet switching system that can support a wide variety of different applications, including video distribution, LAN interconnection and voice/video teleconferencing, all of which require multicast connections. The overall structure of the system is shown in Figure 5. Data is carried between switches in the form of ATM cells over shielded-twisted-pair or fiber-optic transmission links. The *Port Processors* (PP) provide cell buffering and perform link level protocol functions, including the determination of how each cell is routed. The core of the system is a switching network comprising a *Copy Network*, a *Routing Network*, and a set of *Broadcast Translation Circuits*; these are described below. The *Control Processor* (CP), is responsible for establishing connections, including both point-to-point and multipoint connections, as well as overall system control. The *Switch Module Interface* (SMI) provides an interface between the CP and core of the switch.

When a cell enters the system, it is reformatted by the addition of several new fields containing information needed to process a cell within the switching system. In the case of point-to-point cells, the added fields include an outgoing link number which is used to route the cell through the switch and an outgoing VCI. In the case of multipoint cells, they include a *Fanout* field (FAN) which specifies the number of outgoing links that must receive copies of the cell and a *Broadcast Channel Number* (BCN), which is used in a second stage address translation.

The switching network contains three major components, a *Copy Network*, a set of *Broadcast Translation Circuits* and a *Routing Network*. When a multipoint cell having *k* destinations passes through the Copy Network (CN), it is replicated so that *k* copies of that cell emerge from the CN. Point-to-point cells pass through the CN without modification. The function of the Broadcast Translation Circuits (BTC) is to assign outgoing link numbers to the copies of multipoint cells. The Routing Network (RN) delivers cells to

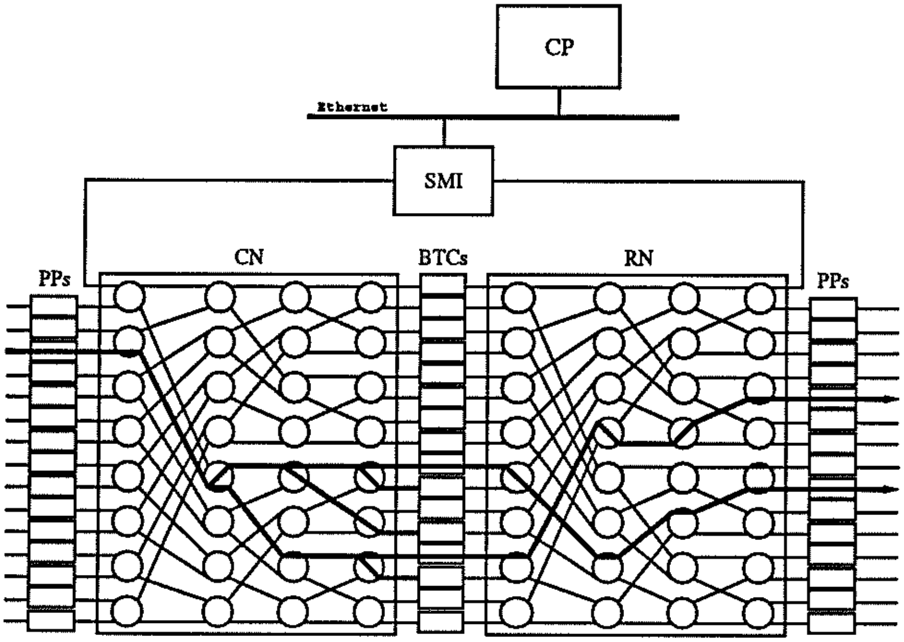

Figure 5: Phase 0 Switch Architecture

the proper outgoing PP, based on the address information given in routing field. The topology shown in the example is a delta network. However, other topologies such as a Beneš topology may also be used.

The Copy and Routing Networks are made up of Packet Switch Elements (PSEs) that contain internal buffers capable of storing several cells. A cell may pass through a PSE without being buffered at all if the desired output port is available when the cell first arrives. Indeed, in a lightly loaded network, a cell can pass through the CN and RN without ever being buffered. In addition to the data path between switch elements, there is a grant signal used to implement a simple flow control mechanism. This prevents loss of cells due to buffer overflows within the fabric. The entire network is operated synchronously, both on a bit basis and a cell basis--that is, all cells entering a given stage do so during the same clock cycle.

The structure of the Copy Network (CN) is the same as that of the RN. The CN's function is to make copies of multipoint cells as they pass through, as illustrated in Figure 6. When a cell passes through the incoming port processor, the virtual channel identifier in the cell's header (8 in the example) is extracted and used to perform a table lookup as illustrated in the figure. In the example, this yields a FAN field of 5 and a BCN of 9. At the first stage, the cell is routed out the lower port. This is an arbitrary decision--the upper link could have been used at this point. At the second stage, the cell is sent out on both outgoing links and the FAN fields in the outgoing cells are modified. The upper cell generates three copies and the lower one two.

In general, a node in the copy network will replicate a cell if its current FAN value exceeds one-half the number of CN output ports reachable from that node. The FAN values are split as evenly as possible, with an arbitrary decision being made as to which port gets the "bigger half" in the case of an odd FAN value. Point-to-point cells are routed through the CN arbitrarily, taking the "path of least resistance."

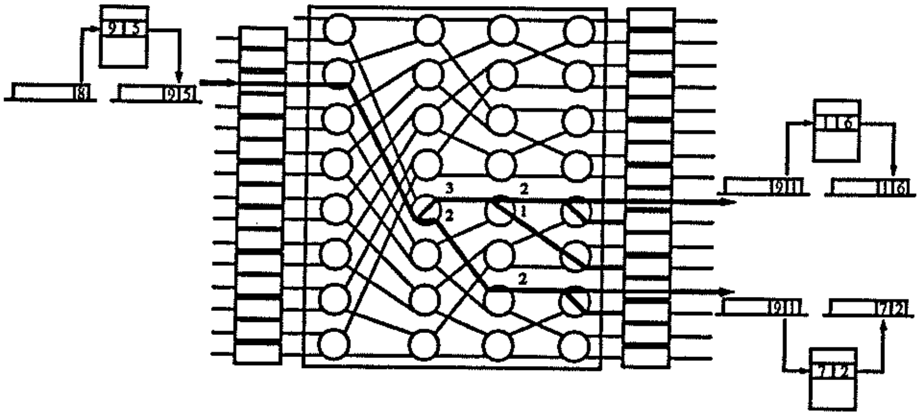

Figure 6: Copy Network Operation

When a broadcast cell reaches a Broadcast Translation Circuit, the BCN is used to select a new routing field from the BTC's internal table. This information is then used by the RN to guide the cell to its final destination. In the example, this routing translation is shown for two of the five copies created by the copy network. The first copy is sent to output 6 and will be assigned a logical channel number of 1 on the outgoing link. The second copy is sent to output 2 and will be assigned a logical channel number of 7.

The phase 0 switch supports external links operating at 100 Mb/s using the ATM cell formats. The internal data paths are eight bits wide and the system operates with a clock rate of 25 MHz. All the crucial components are custom integrated circuits fabricated in 2 μm CMOS. There is a total of five distinct chip types, one for the PSEs, one for the BTCs and three for the PPs. These chips range in complexity from about 45,000 to 220,000 transistors. The phase 0 switch has a 16-port switching network (Figure 7) with 15 external ports. The copy and routing networks are each implemented on a single switch fabric board (Figure 8), the BTCs are packaged on another and the SMI on yet another. The PPs are packed four to a board, so the switch as a whole requires eight printed circuit boards. In conjunction with the early work on

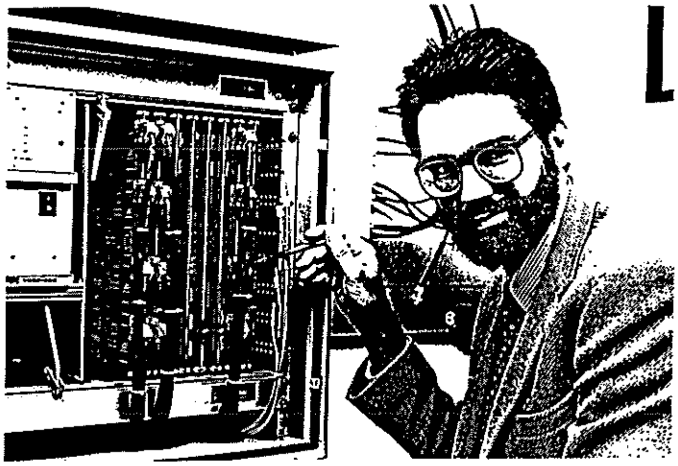

Figure 7: View of Switch Including Front Panel

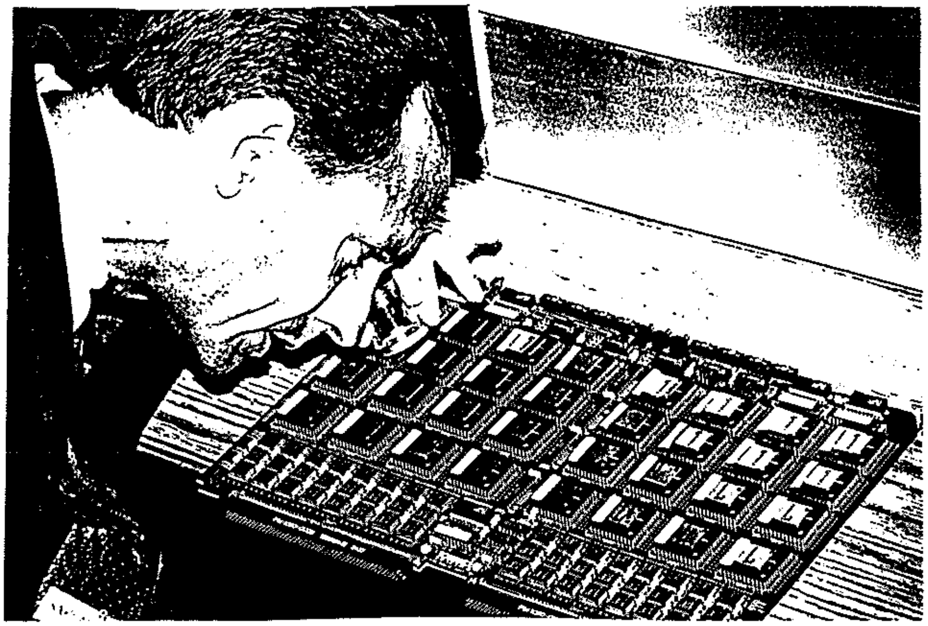

Figure 8: Switch Fabric Board

switching hardware, researchers at Washington University undertook the task of developing a complete signaling system consisting of communication protocols and network control software to build ATM networks that are both scalable and efficient.

2.2.2. Network Control Software

The protocol suite consists of two connection management protocols. The first, called the Connection Management Access Protocol (CMAP), is an access signaling protocol (analogous to Q.931 in narrow-band ISDN and Q.93B in ATM) used to signal for ATM connections at the interface between the ATM network and network clients. CMAP supports dynamic, multipoint, multiconnection calls. CMAP is dynamic, in that clients and connections (in a multiconnection call) may be seamlessly added or dropped during the lifetime of a call without affecting the other clients or connections in the call. CMAP supports the concept of "most" generalized ATM multipoint connections, in that a CMAP ATM connection may have any number of transmitters and receivers. Furthermore, each client accessing a CMAP connection may also change his access mode (e.g. from transmitter only to transmit and receive) during the lifetime of the call/connection. The second connection management protocol, called the Connection Management Network Protocol (CMNP), is a protocol that is utilized inside the ATM network for control processing entities to coordinate the allocation of network resources and to set up connections between clients (typically initiated by CMAP). CMNP is analogous to SS7 in narrow-band ISDN and B-ISUP in ATM. CMAP and CMNP are described in detail in [3, 6, 7]. A *Call Model*, which describes the network as viewed from the perspective of a CMAP client was also developed in conjunction with the connection management protocol suite. The Call Model is described fully in [8].

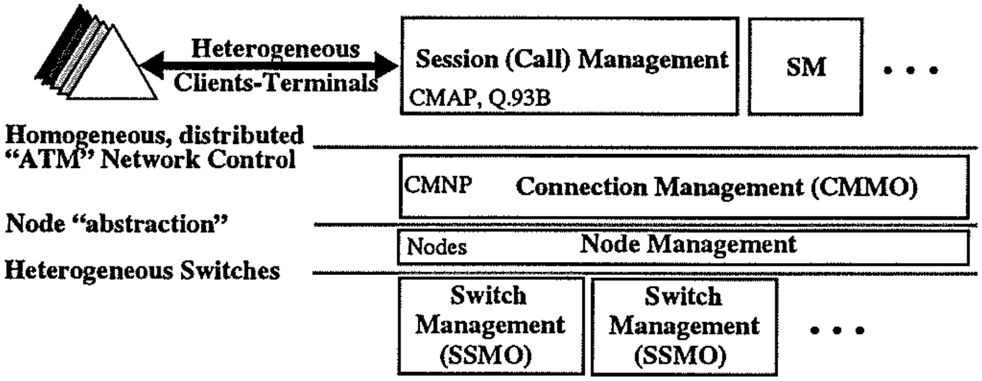

Figure 9. Simplified View of Software Architecture

The protocol suite also consists of a number of lesser protocols that are used in the signaling and software system. A reliable datagram protocol called the Broadband Packet Network (BPN) Reliable Datagram Protocol (BRDP)—analogous to SSCOP in ATM—is used inside the network and at the network access point to provide a reliable data transport substrate for CMAP and CMNP to operate over. The Simple and Efficient Adaptation Layer[1] (SEAL, also referred to as AAL5 [1]) is used as a segmentation and reassembly (SAR) protocol to break BRDP messages up into an ATM stream of cells. In addition to the above protocols, the software system has a number of protocols that were designed to support switch specific control operations, namely, the Switch Fabric Update Protocol (SFUP), the Control-Processor to Switch Module Interface protocol (CP-SMI) and the Packet Processor protocol (PP).

The network control software now exists as a hybrid system composed of the main elements of our early design (described in [7]) as modified to comply with the architectural requirements for separation of switch-specific control software from switch-independent control software (according to Bellcore's Information Network Architecture (INA) principles) and finally, as influenced by our corporate sponsors with regard to requirements for commercial grade software systems.

The software system is described conceptually in Figure 9. This architecture attempts to solve many of the important problems that confront ATM network control engineers. First, we encapsulate heterogeneous network switching equipment interfaces and their control into a software component we call the Switch-Side Managed Object (SSMO). The role of the SSMO and its accompanying "Application Programmers Interface" (API) is to hide the switch specific aspects of ATM switch control and to abstract the generic aspects of ATM switching into a common interface. With the SSMO interface, it is envisioned that multiple vendor switches may be seamlessly controlled by the same connection management software. With this approach, the problem of integrating multivendor, heterogeneous ATM networks may be more easily solved. This concept emerged from Bellcore's INA project and has been implemented on the prototype switching system developed at Washington University. The SSMO's API is based on the emerging CCITT standards with regards to a standard B-ISDN Management Information Base (MIB).

---

[1] The initial development of SEAL was done through a joint effort by Tom Lyons and Allyn Romanow (SUN Microsystems), Steve Deering and Bryan Lyles (Xerox-Parc), Mike Gaddis and Rick Bubenik (Washington University).

The next software layer, referred to as the Node Management layer (NM), allows the NM to group multiple SSMOs into a "Node" and present that aggregation to the connection management layer as a single SSMO. This way, a single network control processor can control multiple, distributed switching systems. This approach provides a flexible and cost efficient way to manage network control processing resources and is particularly effective in ATM LAN networks where the "remote" switches may be small concentrators. Requiring a control processor for each concentrator will needlessly raise the cost of such a network. Nonetheless, the NM architecture allows the Node to consist of just one switch, and, if so, network configurations with a single network Control Processor for each switching system may be supported (including a network that consists of only one switch). Since the NM and SSMO APIs are the same, the connection management layer may be layered directly over the SSMO, eliminating the NM layer in the single switch case.

The next layer is the Connection Management Managed Object (CMMO). This layer is constructed around the CMNP protocol to provide a homogeneous "most generalized" ATM connection management layer—analogous to a "bearer" control layer in B-ISDN standards terminology. The CMMO API can support virtually any ATM multipoint connection conceivable. The API was developed in collaboration with Bellcore based on CMNP principles and influenced by B-ISDN standards (although, in the later case, little work existed to guide us). The CMMO is responsible for end-to-end connection set-up (whereas the SSMO can only support connection segments within the boundaries of its Node). The CMMO sets up these end-to-end connections by communicating with peer CMMOs that control neighboring Nodes and propagating the end-to-end request through the network's mesh topology.

The final layer is the Session Management (SM) layer (it may also be called the call management layer). The SMs interface to connection management through the CMMO API. The CMMO API is constructed in such a way that the it can support multiple SMs. Each SM may compete freely for network resources, it is the CMMO's responsibility to resolve contention for network resources. Session Managers may coordinate their actions but are not required to do so. The access protocols, Q.93B and CMAP, would each be implemented as an SM. This architecture provides considerable flexibility for the network managers to support multiple (simultaneously supported) access protocols and, perhaps even more importantly, new complex services.

2.2.3. Application Interfaces

In the phase 0 network, there are several application interfaces. The first is an audio/video interface for a *broadband terminal.* The broadband terminal consists of a commercial workstation equipped with a video-in-a-window card, together with an ATM audio/video interface as shown in Figure 10. The interface card takes analog audio video from any source, modulates the audio into a band above the video and mixes the two, then digitizes the resulting signal using 8-bit samples at a sampling rate of approximately 10.74 MHz, packetizes the resulting digital stream and transmits the cells on the outgoing ATM link [10].

On the receiving side, data is extracted from cells and placed in a synchronization memory. At this stage cells are resequenced using a sequence number inserted in the packet by the sending interface. Data is read from the synchronization memory and converted to analog. The audio and video portions are then filtered and the audio is demodulated to baseband. The audio signal is then sent to the audio input on the workstation and played out on the workstation's built-in speaker. The video is connected to the video input on

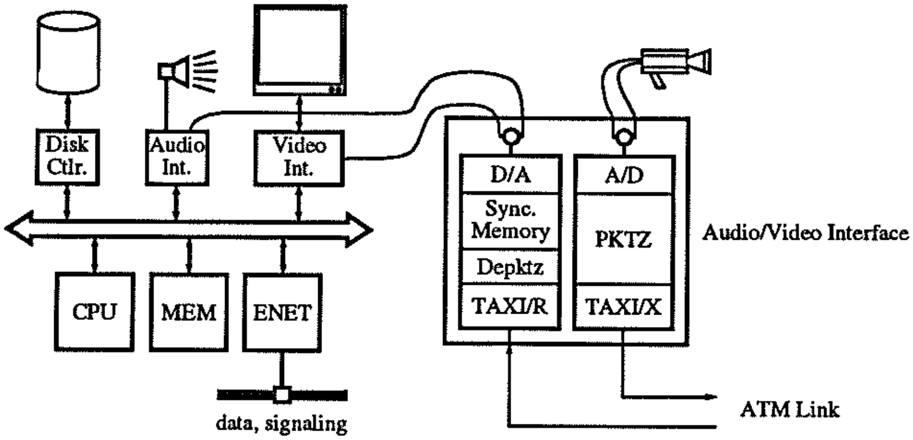

Figure 10: Broadband Terminal

the workstation's video-in-a-window card, allowing it to be displayed on the workstation screen. Data and signaling information are not carried directly on the ATM link. These pass through an Ethernet interface and are forwarded to the ATM network by an Ethernet portal.

Software called VideoExchange has been developed to support the switching of multiple video calls with both point-to-point and multipoint connectivity. Thus, multiparty video conferences are possible without additional equipment, special call setup or operator assistance. The switching takes place in a small fraction of a second and artifacts from cell loss and clock asynchrony are essentially imperceptible.

The physician's workstation shown in Figure 11 is the second application interface used in the phase 0 network. It has been developed on a NeXT platform and incorporates four plug-in boards within the NeXT Cube enclosure. Two are standard boards, the CPU and NeXTdimension supporting the Mach operating system, the NeXTstep user interface programming environment, display Postscript and drivers for monochrome and color displays. Two other boards have been developed especially for Project Zeus [11]. The first of the two, the ATMizer, carries out segmentation and reassembly of cells, does virtual channel translation, prepares and checks cell headers and provides an interface between the ATMizer and the fiber or twisted-pair link to the switch. An embedded processor (M68030) handles segmentation, reassembly and error detection. The second of the two boards developed especially for Project Zeus includes a JPEG [13] coder and decoder for studio quality video, a CD-quality stereo audio channel and a channel for high-resolution images.

Demonstration software has been developed on the NeXT to request and display 12-bit 1024 × 1024 pixel images from a parallel disk array image server, to manage full-duplex 640 × 480 pixel 30 Hz video, to display medical signals such as an electrocardiogram and to retrieve and display various medical records. These multimedia capabilities allow the physician to view and interact with all the information that might be found in the medical record and in addition conduct video conferences with the other physicians, health care providers or patients.

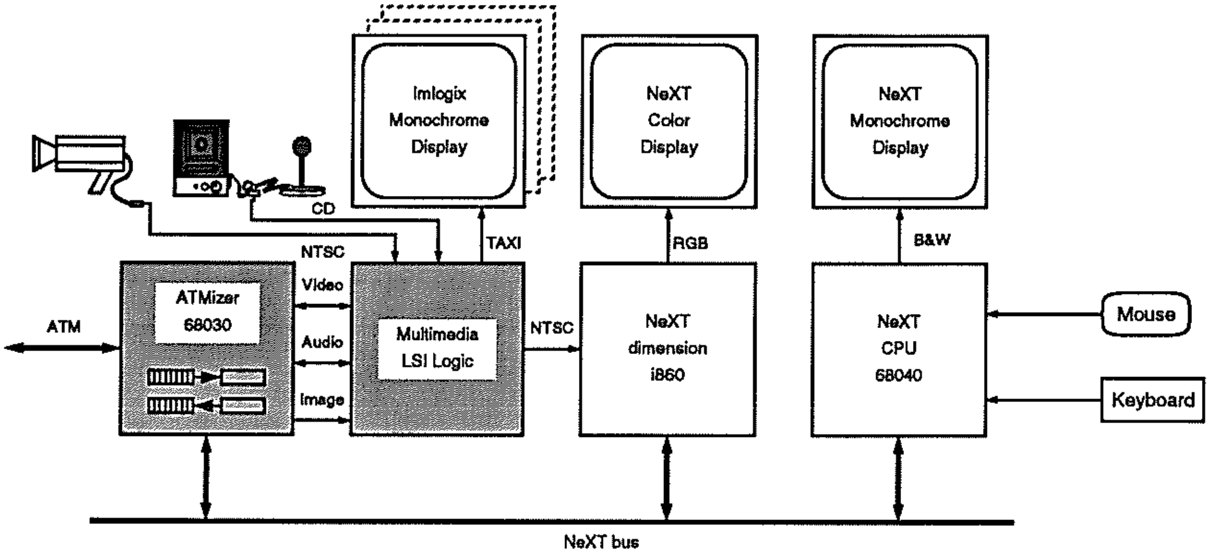

Figure 11: The Physician's Workstation

2.2.4. Internetworking

Internetworking is supported in the phase 0 network using a specialized interface device called an *Ethernet Portal* which is illustrated in Figure 12. The portal transfers Ethernet packets from the local Ethernet, fragments them into a series of ATM cells and reassembles the received ATM cells at a remote Ethernet. The phase 0 implementation does not include any filtering capabilities, but simply transfers the Ethernet packets onto a preconfigured multicast virtual channel. Received ATM cells are reassembled and repeated onto local Ethernets. Separate, commercially available bridges or routers are placed between the portal and the actual Ethernet if filtering is required.

As shown in Figure 12, the portal consists of a commercial EISA bus PC class computer with an ATM interface board including a dual ported memory through which packets are transferred. Fragmentation and reassembly takes place in the memory through a cooperative process involving the Ethernet Controller, the microcontroller on the interface board and the PC's processor, which handles buffer management and makes minor header modifications to ATM cells.

2.3. Lessons Learned

We have learned a number of lessons from our experience with the phase 0 prototype network. Most of our experiences lead us to be optimistic about the future of ATM networks, but a few experiences suggest areas that may require extra care.

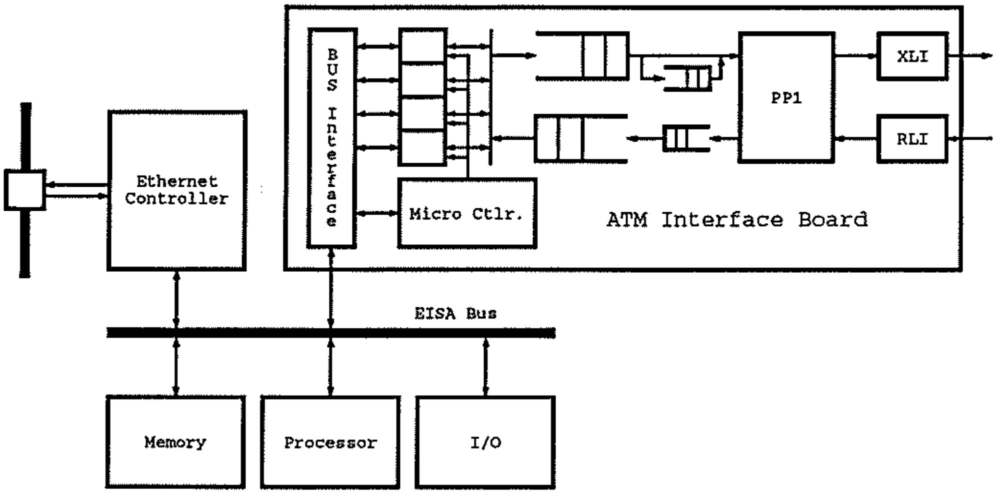

Figure 12: Ethernet Portal

2.3.1. Hardware Experience

The feasibility of multicast ATM networks has been clearly demonstrated. Even fully-loaded switch ports can participate in multicast calls. In our network, cell loss is rare and cell delay minimal.

The custom integrated circuit technology embodied in the five chip types we developed is not particularly demanding. There are many opportunities for cost reduction, even through simple repackaging. The eight phase 0 switches cost about $2,500 per port to fabricate. By simply repackaging the custom chips, the manufacturing cost could be cut to under $1,000 per port. Volume production would produce additional savings in the per-port manufacturing cost, but marketing, administrative and engineering costs will offset these savings.

Redesign for higher performance and increased density is straightforward. This prospect encourages us to believe that within a few years, switches with scalable architectures like our phase 0 switch will be sold for under $1,000 per port. In fact, we estimate that a 16-port 155 Mb/s switch can soon be manufactured with a shielded-twisted-pair interface for $200 per port. A switch with a fiber-optic interface would double the manufacturing cost because of the present relatively high cost of optical devices.

We were surprised by the time required for thorough testing of the phase 0 switches. The complexity of ATM hardware is clearly greater than Ethernet hardware, but even so, testing all modes of operation for a switch took days to weeks. The development of specialized switch test equipment to speed this process seems prudent.

2.3.2. Software Experience

The software system is now operational but refinement and enhancements will continue through 1993—94. The general call model implemented by the system provides needed flexibility in video and voice conferencing and proves to be only slightly more complex than a point-to-point call model with multipoint extensions. Commercialization of the software is planned.

By placing call processing in software rather than in hardware, substantial flexibility and generality has been achieved without sacrificing throughput. In fact, the software architecture supports call processing at very high thoughput. Transactions are batched and shared-memory queues facilitate the use of multiprocessing. Furthermore, the software architecture can support the coming generation of multiprocessors for additional improvements in performance.

The experience that Washington University has gained in developing the connection management protocol suite has been invaluable in our efforts within ATM signaling standards bodies. Washington University researchers have played a substantial role in defining the course of development of the ATM Forum's access signaling protocol (an enhanced version of CCITT's Q.93B protocol). Many concepts from CMAP have influenced design aspects and base requirements of Q.93B. Perhaps more importantly, our knowledge of the internal framework of multipoint connection management (through the design of CMNP) has helped us gain valuable experience about what is "doable" and what is not.

Finally, we believe that interworking of ATM switches should be easy at the hardware level but may be extremely difficult at the software level. Unless the same well-defined, switch-independent control operations are implemented by each vendor's switch encapsulation software, we anticipate lingering problems in multivendor networks.

2.3.3. Application Experience

Very simple methods allow recovery from cell loss in both audio and video transmission. These same methods accommodate transmitting and receiving clock differences and are effective, although somewhat complicated by JPEG coding.

JPEG coding of video produces studio quality images at 15 to 20 Mb/s. This makes possible the delivery of multiple channels of video on a single 100 Mb/s link. However, the cost of multiple CODEC chip sets is presently a severe deterrent to more than one video window. In fact, the audio and video hardware already occupies a considerable fraction of the space and cost budget in our multimedia board in the physician's workstation. A redesign of the ATMizer and multimedia boards to reduce complexity is planned, but a custom host interface chip may be necessary to reach desired space and cost goals.

Not surprisingly, audio problems were the most stubborn ones we encountered during our application experiments and demonstrations. Audio feedback occurred regularly unless lapel microphones were used.

The wide dynamic range of the CD-quality sound in the physician's workstation was clearly superior to the limited dynamic range available in the broadband terminal.

3. Future Plans

Phase 1 of Project Zeus began in January of 1992 and will run through 1994. Its primary objective is to create, with the help of our industry sponsors, the key components needed to establish an operational ATM campus network, develop a set of core applications that use the network and develop support tools to facilitate future application development. The specific goals for phase 1 are listed below.

- Develop an economical switch configuration supporting up to 128 ports at 155 Mb/s each, and an inexpensive concentrator that can be located in a wiring closet within close proximity of desktop workstations.
- Enhance the multicast connection management software and add signaling interfaces to workstations. Design and implement basic network management software.
- Design and implement a workstation interface capable of using a full 155 Mb/s link and equipped with cell-pacing circuitry. Include support for both coded and uncoded video.
- Design and implement a multiport ATM router that forwards IP datagrams across multiple fixed virtual channels and provides Ethernet and FDDI connectivity.
- Establish basic interoperability between the campus network and a public network ATM switch in order to support connections to off-campus sites. This will require development of an ATM/SONET interface to the public network.

During 1994 through 1996, in phase 2 of Project Zeus, we plan to expand the range of interfaces that can be used to access the network, construct components for larger scale and higher speed networks and reduce the cost of key network components. In addition, we plan to extend access to the network to departments all across the university. The specific goals of phase 2 are listed below.

- Develop a core switch supporting 256 ports at 620 Mb/s each, together with an inexpensive concentrator with 16-48 user ports. Improve the functionality and performance of the connection management software.
- Design and implement compatible ATM interfaces for multiple workstations, including support for HDTV video and multiple coded video channels.
- Extend interoperability with public network switches to include signaling, as well as basic cell transfer. Add 620 Mb/s SONET interfaces to the campus network.

- Design and implement internet processors capable of processing fragmented packets without reassembly. These would support both IP datagrams and MCHIP packets [9].

We have, and will continue to depend on government sponsors for the support of basic and applied research. In addition, in both phases 1 and 2, we will depend upon our industry sponsors for financial support and for technical help with network components that they have licensed and developed from Project Zeus technology.

4. Conclusion

The phase 0 Project Zeus network is in operation, providing demonstration of all necessary components. Several industry partners have been recruited and additional partners, particularly in the area of applications are being sought. Phase 1 of Project Zeus is now underway with 155 Mb/s switches to be installed at a handful of sites around our campus in 1993.

Five application areas have been chosen to explore possibilities that may transform daily practice in the selected areas and, at the same time, carry out experiments that are useful in understanding future bandwidth requirements. Application development is underway in electronic radiography, in optical sectioning microscopy, in preservation of images from planetary databases and in visualization in art and architecture.

We believe that Project Zeus creates a unique combination of advanced communication technology and realistic broadband applications that are important to the development of the ATM market and to the understanding of its strengths and weaknesses.

Acknowledgments

The basic and applied research that formed the foundation for project Zeus was sponsored in part by NSF, (NCR-8600947) Bellcore, Italtel, NEC and BNR. We thank our sponsors, Southwestern Bell and NEC America for their generous sponsorship of Phase 0 of Project Zeus. Phase 1 of Project Zeus is now underway and we are grateful to its current sponsors, Southwestern Bell, SynOptics, Ascom Timeplex, Bellcore and Washington University. The work described has been the result of extraordinary efforts by the staffs of the Applied Research Laboratory and the Electronic Radiology Laboratory, and many of the faculty and students in the Departments of Computer Science and Electrical Engineering at Washington University. We also thank Louis Antonopoulos for the preparation of the figures.

References

[1] ANSI T1S1.5. "AAL 5 - A new High Speed Data Transfer AAL," Standards Project: T1S1.5 AAL-ATM. November 4 - 8, 1991.

[2] ATM Forum (Richard Breault and David McDyson, eds.). "ATM UNI Specification (V. 2.0)," June 1, 1992.

[3] Bubenik, Richard, John DeHart and Michael Gaddis. "Multipoint Connection Management in High Speed Networks." In *IEEE Infocom '91: Proceedings of the Tenth Annual Joint Conference of the IEEE Computer and Communications Societies*, pages 59-68, April 1991.

[4] CCITT, Study Group XVIII. *Recommendations Drafted by Working Party XVIII/8 (General B-ISDN Aspects).* Report R 34, June 1990.

[5] Cox, Jerome R. and Jonathan S. Turner. *Project Zeus: Design of a Broadband Network and its Application on a University Campus.* Washington University, Department of Computer Science, technical report WUCS-91-45, July 30, 1991.

[6] DeHart, John, Michael Gaddis and Richard Bubenik. *Connection Management Access Protocol (CMAP) Specification.* Washington University, Department of Computer Science, Technical Report WUCS-92-01, Version 2.1. February 1992.

[7] Gaddis, Michael, Richard Bubenik and John DeHart. "A Call Model for Multipoint Communication in Switched Networks." In *ICC '92: Proceedings of the IEEE International Conference on Communications*, pages 609-615, June 1992.

[8] Gaddis, Michael, Richard Bubenik and John DeHart. "Connection Management for a Prototype Fast Packet ATM B-ISDN Network." In *Proceedings of the National Communications Forum*, pages 601-608, October 1990.

[9] Mazraani, Tony and Guru Parulkar. "Specification of a Multipoint Congram-Oriented High-Performance Internet Protocol," Infocom 90, June, 1990.

[10] Richard, W. D., P. Costa, and K. Sato. "The Washington University Broadband Terminal", Special Issue on High-Speed Host/Network Interface, *IEEE Journal on Selected Areas in Communications* (in press).

[11] Richard, William D., Jerome R. Cox, Brian Gottlieb, and Ken Krieger. *The Washington University Multimedia System*, Washington University, Department of Computer Science, Technical Report, WUCS-93-01, January 1993.

[12] Turner, Jonathan S. "Design of a Broadcast Packet Switching Network." In *IEEE Transactions on Communications*, 36(6):734-743, June 1988.

[13] Wallace, G.K., "The JPEG Still Picture Compression Standard," Communications ACM, Vol. 34, No. 4, pages 31-44, April, 1991.