SuperComputing-94 — Networking Technologies Summary

D. P. Koester
School of Computer and Information Science and
The Northeast Parallel Architectures Center (NPAC)
Syracuse University
Syracuse, NY 13244-4100
dpk@npac.syr.edu

NPAC Technical Report — SCCS 708

22 November 1994

NORTHEAST PARALLEL
ARCHITECTURES CENTER
AT SYRACUSE UNIVERSITY

Abstract

This technical summary examines the networking technologies presented at SuperComputing-94 that are used to implement some Scalable Parallel Processor (SPP) interconnection networks, input/output (I/O) networks, computer clusters, and traditional local/metropolitan/wide area network (LAN/MAN/WAN) capabilities. Most technologies discussed here relate in a broad sense to clustering computers into SPP architectures for (heterogeneous) distributed computing, although it is conceivable that distributed computing may extend over WAN distances. Notes from the *Gigabit Networking* Roundtable and the *Future of the Internet* Panel are also included in this paper.

1 Introduction

Networking offerings at SuperComputing-94 could be classified into the following categories:

- Massively Parallel Processor (MPP) interconnection networks
- Scalable Parallel Processor (SPP) interconnection networks
- Input/output (I/O), channel, and computer cluster interconnection networks
- local/metropolitan/wide area networks (LANs/MANs/WANs)

The focus of this technical summary is on the networking technologies used to implement some SPP interconnection networks, input/output (I/O) networks, computer clusters, and traditional LAN/MAN/WAN capabilities. Most technologies discussed here relate in a broad sense to clustering computers into SPP architectures for (heterogeneous) distributed computing, although it is conceivable that distributed computing may extend over WAN distances. While MPP interconnection networks are of interest to those developing algorithms, this paper does not address these proprietary networking architectures.

1.1 Networking Technologies

The networking technologies offered by vendors at SuperComputing-94 included, but were not limited to:

- Scalable Coherent Interface (SCI)
- High Performance Parallel Interface (HiPPI)
- Fibre Channel
- Asynchronous Transfer Mode (ATM)
- Fiber Distributed Digital Interface (FDDI)
- Ethernet (10 megabit-per-second and 100 megabit-per-second)

HiPPI, ATM, FDDI, and 10 megabit-per-second Ethernet were offered as networking technologies in the SCinet '94, the SuperComputing '94 show net. SCinet '94 was advertised as the most advanced show net anywhere — a viable claim, because this show network did offer true gigabit-per-second (1.6 gigabit-per-second per link) local area networks built upon Dual-HiPPI. A diagram of SCinet'94 is provided in figure 1. This network could be compared to the show net of Interop+Networld in Atlanta in September 1994, where HiPPI saw little use as a networking technology and claims of gigabit networking technologies generally were ATM-based or Frame-relay-based, often operating only at OC-3 data rates (155 megabit-per-second). Gigabit-per-second data rate claims for these networks generally arise from the opinion that Broadband Integrated Services Data Network (BISDN) technology and fiber-optic-based Sonet technology will eventually scale to gigabit speeds.

1.2 Data Communications Environments

The networking technologies listed above operate across a wide range of environments defined by both the distance and required bandwidth. The relationships of distance to bandwidth for the following computer interconnection technologies are described in figure 2:

- buses (ie., SBus, MicroChannel, PCI, VME, etc.)
- I/O (ie., Small Computer Serial Interface (SCSI), etc.)
- channels (ie., hardware interconnections, etc.)
- LANs
- MANs
- WANs

In general, distances over which data is carried for these technologies increases as you look to the right in this diagram, while communications bandwidth increases as you look to the left.

Figure 3 illustrates how various networking technologies would be superimposed on the environment hierarchy presented in figure 2. Networking applications requiring the high communications throughput of buses has special interest to computer clustering applications where sufficiently high speed and low latency communications are required to provide shared memory capabilities between workstations. Adding the intelligence of networking protocols can provide shared resources for I/O interfaces and hardware interconnection channels. Meanwhile, the throughput potentials for local area networks is increasing to meet that of dedicated hardware interface channels, and with ATM, the only limitations for wider-area networking is the availability of bandwidth to interconnect sites. The availability of adequate bandwidth for enterprise networks within organizations was discussed at the Roundtable entitled *Gigabit Networking: State of the Art and Applications*, and is reported below in section 3.

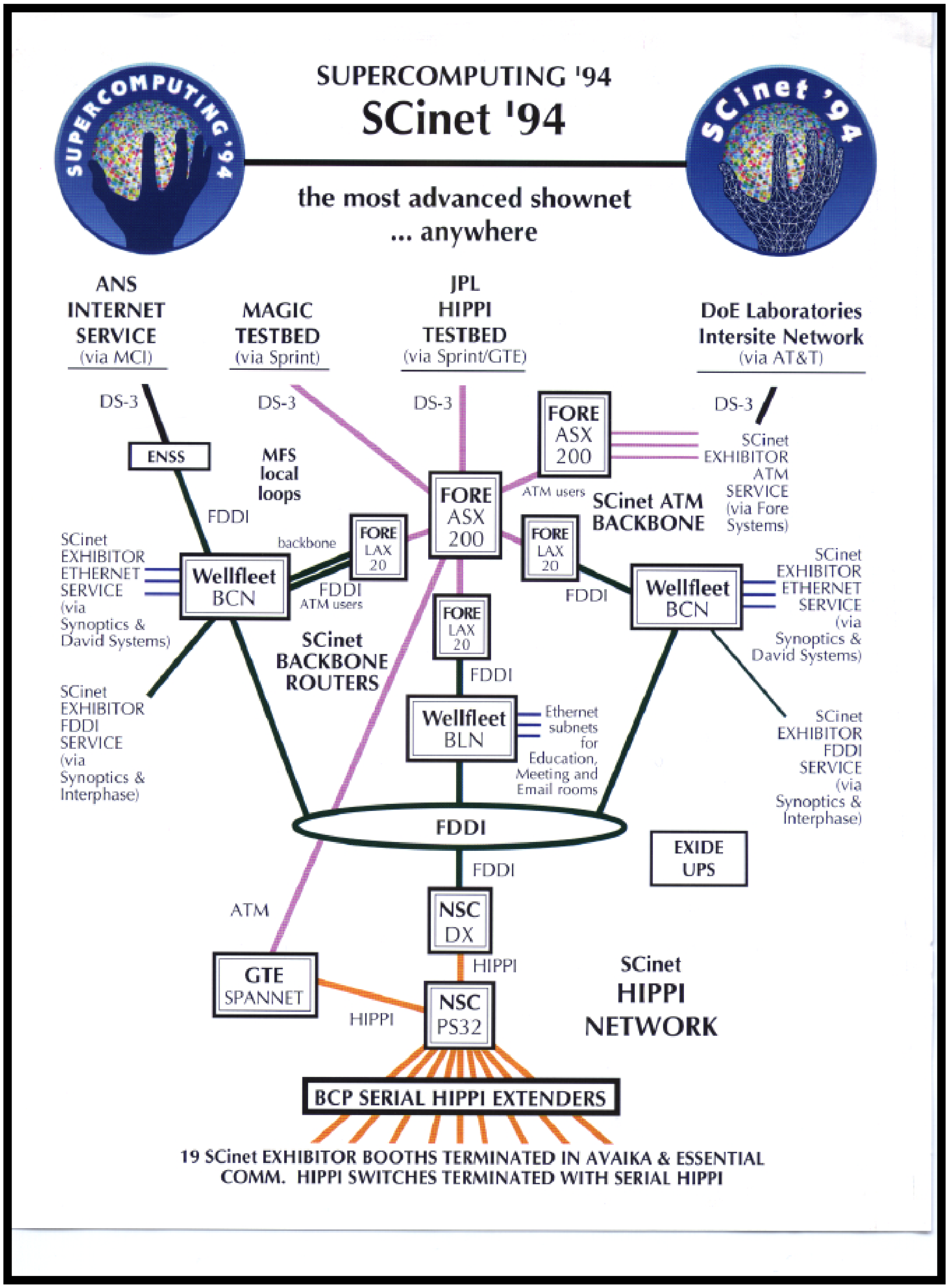

Figure 1: SCinet'94 (Scanned from a handout)

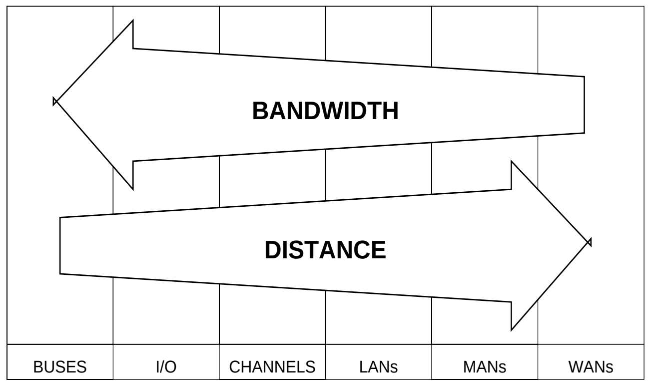

Figure 2: Bandwidth versus Distance for Computer Interconnection Technologies

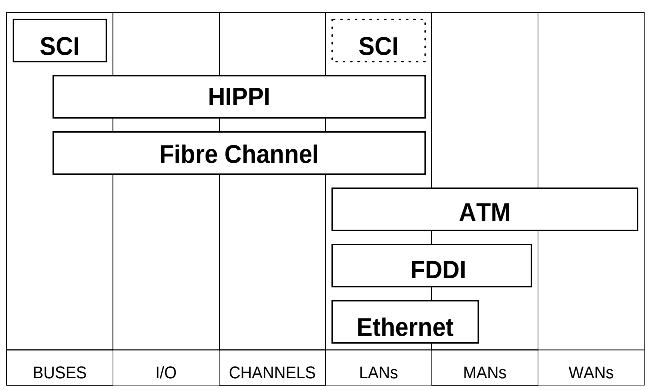

Figure 3: Networking Strategies

1.3 Overview

This summary of the networking technologies offered at SuperComputing '94 is organized as follows. In section 2, I briefly describe each of the networking technologies listed above, with special interest to the environments where the technology would be used effectively. In section 3, I briefly describe some of the highlights of the Roundtable on *Gigabit Networking: State of the Art and Applications*, while in section 4, I cover some of the high-points from the Panel *The Future of the Internet*. In the last section, I discuss the present state of networking in the supercomputer community.

2 Available Networking Technologies

Networking technologies offered by exhibitors at SuperComputing '94 included, but were not limited to:

- Scalable Coherent Interface (SCI)
- High Performance Parallel Interface (HiPPI)
- Fibre Channel
- Asynchronous Transfer Mode (ATM)
- Fiber Distributed Digital Interface (FDDI)
- Ethernet (10 megabit-per-second and 100 megabit-per-second)

In the introduction, the relationships of these technologies to distance and bandwidth were described. Figure 3 illustrates that there is significant overlap in the LAN environment for all of these technologies.

All networking products discussed in this paper, claim to have some LAN utility — even those networking technologies that are more closely associated with point-to-point communications. These capabilities are due in part to a common paradigm used in the protocol stack that controls how applications access the networking hardware. Vendors for SCI, HiPPI, Fibre Channel, and ATM technologies have all implemented a software protocol access paradigm that provides two distinct ways for an applications to access a network interface. One way for an application to access the network interface is to include direct access to the UNIX device drivers for the interface through an application programming interface (API). The other part of the interface access paradigm provides transparent integration of the networking technology to existing applications through the Internet protocol (IP) socket-based protocol stack, and provides transparent, but sometimes

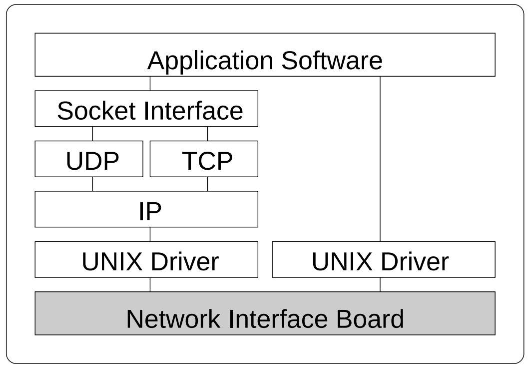

Figure 4: Dual Protocol Paradigm for Interfacing Applications to Network Hardware

crude, LAN capabilities. Figure 4 illustrates this two option paradigm. The most significant aspect of this dual protocol paradigm is that existing applications can utilize the networking technology without having to undergo costly modifications to the software. Utilizing the transparent access to the interface through the IP stack causes additional overhead when compared to the API that directly accesses the device drivers. There is additional overhead involved in implementing the transmission control protocol (TCP)/IP protocol, especially if it is not required because lower layer protocols ensure data reliability.

Multi-computer clustering is an important application that uses these various networking technologies to interconnect multiple computers into a *(poor-man's)* scalable parallel processor. Considering the obvious fact that *as interprocessor communications latency decreases and bandwidth increases, more classes of applications and algorithms can be efficiently implemented on parallel processors* — those technologies with software protocols and hardware designs that provide low latency and ample bandwidth will provide the most robust clustered computing environments. Latency and bandwidth are important to computer scientists designing algorithms and software engineers porting applications; however, system cost is also a factor that must be considered.

Throughout the remainder of this section, I will describe these six networking technologies considering

- the applicable environments for the networking technology
- software and hardware latency
- bandwidth
- relative cost per port for a network

This paper will provide only a general discussion of latency and cost. Discussions with exhibitors did not provide detailed information in all areas. In many instances, additional information has been promised to be provided at by mail — many vendors have promised to send me detailed white-papers on their networking products.

2.1 Scalable Coherent Interface (SCI)

SCI technology is based on the ANSI/IEEE SCI standard 1596-1992 and is designed to provide low latency and high bandwidth interfaces for workstation clustering. In particular, this technology can provide $1\mu s$ latency and, consequently, can efficiently support both shared memory and message passing applications between a variety of workstation types. Bandwidth on individual links is one gigabit-per-second, and the standard can support either ring or switched communications topologies. The SCI standard can be used to create a virtual bus, so that workstation internal bus protocols can be extended between (heterogeneous) clusters of computers.

I found only one company offering SCI products at SuperComputing '94 - Dolphin Interconnect Solutions, a company with main offices located in Norway. Dolphin offers SCI interface cards to interconnect SBus-based or VME-bus-based workstations. In additional, this company offers bridging solutions between:

- VME-SBus
- VME-PCI
- VME-ATM

The SBus product offered by Dolphin has software that provides the options either to use low latency UNIX device drivers accessed directly by applications, or to use the internet protocol (IP) socket-based protocol stack for existing applications. This relationship is described in greater detail above, and is illustrated in figure 4. Due to the fact that this technology can be used for either low-latency bus interconnections or transparent transmission control protocol (TCP)/IP interconnections, figure 3 shows this networking technology as being able to provide LAN connectivity as well as bus connectivity. Dolphin offers a mailbox on the card that can be used for low latency message exchanges, which requires application program modifications to directly send data from memory to the interface card via direct memory access (DMA).

No pricing information was available for this networking technology, however, interconnections are made with relatively low cost shielded twisted pair cables. At these bandwidths, this physical interconnection media surely limits the distances between clustered workstations. Dolphin is examining using ATM-based communications to extend the distances between clustered workstations. If computers are clustered in a ring topology, latency could suffer as a function of the distance (number of network hops) from transmitter to receiver, which is a function of their relative locations on the ring. The SCI protocol can be used in a switched architecture, however, no information on switch availability or pricing was available. In general, this is a limited use networking technology, only to be used in high-performance workstation clusters.

2.2 High Performance Parallel Interface (HiPPI)

HiPPI-based communications can be used for a wide range of applications ranging from

- point-to-point input/output (I/O) data connections
- channel devices
- network peripherals
- workstations

Many supercomputers have connections to separate I/O subsystems using HiPPI links. Cray research has all user network connections in a separate device that is connected to its supercomputers via a point-to-point HiPPI link. Disk farms and tape robots can be connected either as channel devices directly to supercomputers or some devices are HiPPI-network capable and connected to a HiPPI switch for access by multiple supercomputers or workstations. The Sony booth had a terrific example of a HiPPI-based application, were data was moved from a tape system through a HiPPI network at 32 megabytes-per-second for real-time display of a digitized segment of a movie on a 2k $\times$ 1k monitor.

HiPPI-based networks were the surprise of the trade-show, especially serial HiPPI connections. HiPPI has the potential bandwidth of 800 megabits-per-second, while Dual-HiPPI can provide 1.6 gigabits-per-second. Serial HiPPI protocols have freed standard HiPPI from a 25 meter maximum cable-length limitation, so HiPPI links can now be placed over dual-mode fiber optic cable with significantly greater cable run lengths. HiPPI networks were prolific on the trade show floor, with a subset of SCinet '94 being the first all fiber-optic multi-gigabit-per-second computer network that included:

- 21 HiPPI switches with over 136 gigabits-per-second switching capacity connected by over 50 gigabits-per-second of fiber transmission capacity connecting 130 computers and peripherals
- 31 HiPPI fiber links
- the first public demonstration of serial HiPPI ports on a HiPPI switch

- demonstrations of interoperability between various vendor's HiPPI interfaces and switches

HiPPI appears to be a networking technology that primarily has a home in the supercomputing community. There are no other networking area that actually has applications that can utilize 800 to 1600 megabits-per-second throughput at the present time. In many cases, these speeds are not fully obtained because of the capacity of end-user systems, such as workstations and frame-buffers. At Interop+Networld this September in Atlanta, I saw **no** HiPPI networking, which possibly was strongly influenced by the fact that switched serial HiPPI is such a new technology. Conventional HiPPI has 25 meter maximum cable lengths, limiting wire-based HiPPI to the machine room, and consequently limiting conventional HiPPI as a general-purpose networking technology. Meanwhile, serial HiPPI has the potential for more realistic cable-run lengths, but switched serial HiPPI was demonstrated only for the first time at this show. Technologies generally do not show up at Interop or Networld until there is some degree of stability in the technology and some degree of interoperability between vendor's equipment has been established. Moreover, HiPPI may never become a mainstream networking technology, because of the limited number of applications that require the extensive bandwidth. Moreover, SuperComputing conferences tend to highlight the state-of-the-art in computing and networking technologies, with Interop+Networld conferences tending to highlight interoperability.

HiPPI is a networking technology that is in competition with some networking technologies, while complementing other technologies. HiPPI is in direct competition with Fibre Channel technology. As illustrated in figure 3, these two technologies are attempting to operate in the same environments. Meanwhile, networking technologies, like ATM compliment HiPPI, because ATM offers protocols designed for metropolitan and wide area networks. HiPPI can provide gigabit-per-second throughput for peripherals or workstations in switched LAN configurations, however, the protocol is not suited for wide-area data transfers. ATM will eventually be able to provide gigabit-per-second bandwidths available today with HiPPI, but at greater distances, even over WANs. Fibre Channel and ATM technologies will be discussed below in subsequent parts of this section of the paper.

There are several technology problems that illustrate the downside of HiPPI technology. Maximum throughput for interfaces and switches often can only be reached with large sized data packets. Multiplexing of HiPPI signals between switches appears to be a problem, because no data stream buffering is possible at the high HiPPI data rates. Network management for HiPPI switches must be performed manually, with no automated network interface capabilities. Lastly, present workstations can handle the communications throughput of a HiPPI interface but there is little workstation left over to do other work.

The cost of HiPPI is comparable to Fibre Channel, for those applications that require the bandwidth available with this technology. Conventional HiPPI uses heavy multi-wire shielded twisted pair cables with 25 meter length limitations, while serial HiPPI uses fiber optic transceivers, incurring similar costs to FDDI or ATM over glass. Nevertheless, HiPPI will never be a cost-effective LAN technology. HiPPI interfaces and switches are often more expensive than other shared media and switched media LAN technologies, due to the economy of scale available in the manufacture of more mainstream networking technologies. Also the raw speeds of HiPPI links require high-speed components, often requiring components manufactured from GaAs — which significantly increases the cost of HiPPI technology.

HiPPI has matured significantly as a networking technology and is no longer a high-speed point-to-point protocol, once referred to at a roundtable as *RS232 on steroids*. Rather, at SuperComputing '94, HiPPI was demonstrated to have the capabilities to connect networked peripherals and networked workstations with supercomputers. HiPPI has direct competition with Fibre Channel technology, although other networking technologies complement the capabilities of HiPPI.

2.3 Fibre Channel

Fibre Channel is being developed to resolve fundamental problems inherent to channel technologies, and as a result, extend the benefits of channel technology into the local area network. Fibre Channel technology addresses the distance limitations of conventional HiPPI and other channel technologies and the address space limitations of these technologies. Fibre Channel has scalable bandwidth/media options that permits the use of shielded twisted pair for 133 megabit-per-second links and multi-gigabit-per-second rates over fiber optic links for distances of up to ten kilometers. Fibre Channel has the capability of carrying both channel and network protocols simultaneously on the same media, by encapsulating IP signals in a manner illustrated in figure 4. Lastly, Fibre Channel uses 24 bit wide addresses so that over 16 million nodes can be included on a network.

Those computer vendors that are providing backing for Fibre Channel are:

- Hewlett-Packard
- IBM
- Sun Microsystems
- Silicon Graphics Incorporated (SGI)

Hewlett-Packard is offering their HP 9000 Computational Cluster with Fibre Channel as one cluster interconnection option. The HP 9000 Computational Cluster has up to eight

HP series 700 workstations packaged in a single cabinet with high-performance networking support. In addition to Fibre Channel, the Computation Cluster can be purchased with ATM, switched FDDI, or HiPPI as the interconnection network. IBM was also demonstrating Fibre Channel technology and distributed an *IBM Technology Brief* that discusses the complementary features of Fibre Channel and ATM technology. Adapters are available for workstations from Sun Microsystems and from SGI. Anchor Communications is a networking company that has been developing commercial Fibre Channel products for the last several years, and marketing Fibre Channel products for the last year.

Fibre Channel appears to be a technology designed specifically to overcome the shortfalls of conventional HiPPI. Meanwhile, the HiPPI standards are being expanded in an attempt to compete with Fibre Channel technology. Serial HiPPI has been developed to provide similar gigabit-per-second throughput over long distances (10 kilometers) over fiber. HiPPI can encapsulate data from TCP/IP applications, and HiPPI has the capability to double the address space to match the 24 bit addresses of Fibre Channel. As the supercomputing industry readjusts to the end of the Cold War and reduced government spending for research, it will be interesting to see if both Fibre Channel and (serial) HiPPI can survive.

2.4 Asynchronous Transfer Mode (ATM)

ATM technology is a networking technology designed to permit telephony, video, and data communications on the same network using statistical multiplexing to provide *bandwidth on demand.* ATM technology differs drastically from HiPPI and Fibre Channel in the area of statistical multiplexing multiple data streams onto a single physical signal to obtain maximum utilization of the bandwidth. ATM uses small 53 byte cells with a 48 byte payload to provide the potential for rapid switching that will maintain the constant-bit-rate characteristics and timing of isochronous signal sources. The small, constant cell size permits processing of the cell in parallel within a switch — the entire cell can be moved bit-parallel through a switch. This bit-parallel feature in ATM cell processing means that data rates can increase to 10+ gigabits-per-second without requiring expensive GaAs technology as required in other giga-switch technologies.

ATM is being developed with very strong backing from the telecommunications industry to support telephony and video for wide-area connections, although companies like Fore Systems and SynOptics have effectively pushed ATM technology into the LAN environment. While there may be overlap in the LAN environment for ATM and HiPPI/Fibre Channel, these are complementary technologies. All supercomputer vendors have ATM interfaces for the user to interface to their hardware, in the same manner that they have used Ethernet and FDDI networks for some time. As ATM-based enterprise networks arise, all supercomputer vendors are prepared to interface their hardware to users via ATM. Moreover, NCube has included specialized ATM output interfaces to support real-time video from their newest computer, the NCube 3. This computer will support video distribution over AAL1 or AAL5, where ATM Adaptation Layer (AAL) type 1 is for constant-bit-rate data and AAL type 5 is for limited overhead available-bit-rate data. Hewlett Packard is offering a version of their Computational Cluster with an ATM-based LAN as the interconnection network. HP has incorporated a Fore Systems ASX-200 switch in the rack with eight workstations.

ATM technology is an area of personal expertise, nevertheless, there was much to learn from discussions with exhibitors at the SuperComputing '94 trade show. I saw a demonstration of ATM congestion control at the Digital Equipment Corporation (DEC) booth. They demonstrated 135 megabit-per-second throughput on a 155 megabit-per-second OC-3 link between two Alpha-based workstations connected with a single DEC ATM switch. Without congestion control, throughput was less than 40 megabits-per-second. Due to the queuing/buffering architecture within the DEC switch, one would expect no degradation of isochronous signals, because these time sensitive signals are carefully separated in the buffering/priority architecture of the switch. A single buffer-memory space is divided into partitions for each signal type and priority. At present the layout of buffers is static, although future versions of buffer software will permit dynamic allocation of buffer memory. Discussions with IBM revealed that they also recognize that reserved bandwidth signals must be processed separately from workstation generated available-bit-rate data streams to avoid data loss.

The SCinet '94 included two Fore Systems ASX 200 ATM switches and a GTE SPANet switch. The Fore Systems switches were used to connect the MAGIC Testbed and the JPL HiPPI Testbed. The GTE SPANet switch was used as a gateway between the ATM and HiPPI networks. There were three Fore Systems LAX 20 legacy network interconnection devices to route FDDI traffic over ATM. A scan of the network diagram is provided in figure 1.

I believe that ATM will be the workstation networking technology of the future, due in part to the backing of the technology from the telecommunications industry. ATM networks collapse the concept of LANs/MANs/WANs because the same protocol is used throughout the entire network, eliminating the requirements for routers or gateways. The transition to ATM-based LANs for all workstations is being hindered because of the cost to replace legacy networking technologies. Existing workstations have some network technology (by definition) and costs are between one and two thousand dollars to add a workstation to an ATM-based LAN. The cost-point for transitioning to ATM technology must be compared to the cost of maintaining the legacy networks and especially the costs (in both dollars and degraded performance) for additional routers if the number of networked workstations are increased. Standards-based interoperability between ATM networks and legacy Ethernet and FDDI networks are not here today, but can be expected to arrive in the next several years. Proprietary solutions exist from Fore Systems and SynOptics for interconnecting legacy networks to and through ATM-based LANs, but these vendor's propriety protocols are not interoperable.

While HiPPI and Fibre Channel technologies have limited utility as LAN technologies because of the limited maximum connection distances, ATM has no such distance limitation. Rather, the limitations of ATM in the LAN environment are the overhead encountered with the small cell size. As available bandwidth permits users to readily access computing resources across larger physical distances, ATM-based networking will overcome the resistance induced by the cost to upgrade networking capabilities in existing workstations. The requirements for wider-area access to computing resources will spur the wide-spread deployment of ATM-based networking and promote the implementation of ATM on desktop workstations. Further discussions on the use of ATM at Sandia National Labs for large-scale enterprise networks is provided in section 3.

2.5 Fiber Distributed Digital Interface (FDDI)

FDDI is a token-passing counter rotating fiber-optic-based ring that has the capacity to interconnect up to 1000 devices over a 100 kilometer path. Due to the maximum distance runs of FDDI, it is often used as a MAN technology in addition to being used as a LAN technology. FDDI standards specify LAN-to-LAN packet-switched backbones, in addition to the counter-rotating ring configuration operating at 100 megabits-per-second. While the token-ring version of FDDI is the most well known, switched FDDI has the most potential to support workstation clusters. Token-ring FDDI still has some life remaining as a LAN technology — its peak data rate is useful in moving large data files from servers to end-users. However, latency is high in a FDDI ring (you must wait for the token) — so high that token-ring FDDI generally performs worse in computer clusters than TCP/IP communications over Ethernet (for correctly configured FDDI and Ethernet networks). By collapsing FDDI rings to a single switched connection, the problems with latency are removed. A version of the HP Computational Cluster is offered with switched FDDI as the interconnection network.

Due to the fiber optic-based design of FDDI, there is a question of network performance and cost. The laser diodes cause FDDI implementations to be much more expensive that Ethernet-based networks. FDDI throughput performance is much greater than Ethernet, but the cost/performance ratio must be examined as a function of one's specific application. Early ATM-based LANs used physical connections based on the TAXI standard developed for FDDI, but per-port-costs generally were less for ATM when compared to switched FDDI. The electronics were simpler for the interface cards. With the advent of serial HiPPI and Fibre Channel, there is strong competition from these technologies for significantly higher data rate LANs, and ATM will offer strong competition for both LAN and MAN environments. I believe that FDDI will be supplanted by ATM as a networking technology long before 10 megabit-per-second Ethernet — mainly for cost considerations. Existing investments in legacy networks will remain in place for extended periods of time, although I see little (cost effective) reason to expand legacy FDDI networks. For high-performance LAN environments, serial HiPPI is much faster for similar investments in fiber-optic-based technology. ATM should also be comparable for high-performance LAN environments with similar investments, and ATM should have a significant cost/performance advantage in the MAN environment. Lastly, ATM can be used transparently throughout either the LAN, MAN, or WAN environments.

2.6 Ethernet

Venerable 10 megabit-per-second Ethernet is a networking technology that still has some remaining life, because of the momentum of its massive legacy status and because of its cost/performance ratio. Most workstations have built in Ethernet adapters and the cost to add an Ethernet adapter to a personal computer (PC) is generally significantly less than $100. Ethernet adapters cost at least an order of magnitude less than any fiber optic-based networking technology. IBM is attempting to reduce the cost of their 25 megabit-per-second ATM cards that use twisted-pair wire for interconnections to compete with Ethernet interfaces. However, that technology presently costs at least five times that of Ethernet adapters, with only marginal increases in performance. If market share for low-data-rate twisted wire ATM interface cards increases significantly, then the cost for this technology will decrease. Unfortunately, there will always be the additional costs of switches for ATM networks over and above the costs of network interface cards. Switches are not required with shared media like 10 megabit-per-second Ethernet, although switch costs must be included in higher speed Ethernet products.

Higher-speed versions of Ethernet have been developed, although there was almost no sign of support for these high-speed LAN technologies from exhibitors at SuperComputing '94, although network product vendors at Interop+Networld the previous September provided an extensive offering of high-speed switched Ethernet products. Many vendors at SuperComputing '94 responded that they will support such networking technologies, although I was regularly asked why I was interested in high-speed Ethernet technologies. Exhibitors attempted to direct me towards technologies that offered higher speed, ie., serial

HiPPI, Fibre Channel, and ATM, or more robust capabilities, ie., the extensibility of ATM to the LAN, MAN, and WAN environments.

3 The *Gigabit Networking* Roundtable

This section of this report on networking technologies at SuperComputing '94 highlights some of the comments made at the *Gigabit Networking: State of the Art and Applications* Roundtable. Many of the comments reported here appear as sound-bites — single, notable catch-phrases. Given the short time for presentations by the members of the roundtable, the talks often made only one or two points and were presented in the same manner as sound-bites. Those presenting at this roundtable where highly technical, knowledgeable people in the field of high-performance networking to support supercomputing. All comments reported here are from the panelists unless otherwise specified.

- Serial HiPPI has enabled this high speed networking technology to move out of the machine room, and move onto the desktop.
- ATM will be the WAN technology in the near future, and ATM will drive other technologies out of the WAN market within the next five years. ATM will continue to increase its market share in the LAN market as global networking access increases and users want faster access to resources located throughout the world.
- The next generation of research must examine ways to reduce both hardware and software latency in networking products. At present, latency in distributed computations is the most significant issue — ATM is presently only 50% faster than Ethernet — entirely due to hardware and software latency within the workstation.
- Sandia National Laboratory is concentrating its networking investment on ATM technology. Sandia has strong interest in heterogeneous computing over the wide area. Sandia has found that software latency is one of their biggest problems, and they are examining ways to optimize the entire software stack to minimize latency and improve performance. Sandia is looking for ATM switches with high capacity, gigabits-per-second, to handle their network traffic. Sandia claims that it has not been difficult to obtain the massive bandwidth to interconnect their enterprise networks, but the cost has been high. At present Sandia is paying $500 per month per mile for connectivity from commercial sources. The only thing that could be worse than those high costs would be to have to develop and maintain those communications capabilities in-house.
- The single killer application of the future is the *lawyer application*. Legal concerns may minimize data compression for medical images, or anything other than end-

user entertainment. A legal case already has been filed because image compression destroyed the phenomenon of interest on a medical image.

- A representative of the Ford Motor Company discussed their corporate enterprise network(s). Ford presently has world wide TCP/IP, SNA, ... , etc. networks with the interest in consolidating to one networking technology. Ford has a supercomputing center in Dearborn, MI equipped with (among other computers) massively parallel computers from MasPar. Ford has made a commitment to build *world cars* and has instituted a corporate policy to design automobiles over networks sharing data and using available remote resources. Designers in Australia have access to the most available computing power, due to their work times in the global day. (Authors note — remember that Ford is a private company, competing in a world economy, using state-of-the-art networking and distributed computing. Ford's networking plans appear as ambitious as Sandia's.)

- HiPPI is RS-232 (a serial technology) on steroids. We can get gigabits-per-second to a workstation today. Eventually, we will be able to use the workstation to do something with that data. Presently, the entire workstation is required to support a gigabit-per-second interface. Even more capable workstations are required if we are to use such large amounts information in constructive calculations or transaction processing.

4 The *Future of the Internet* Panel

This section of this report on networking technologies at SuperComputing '94 highlights some of the comments made at the *Future of the Internet* panel. Many of the comments reported here also appear as sound-bites — single, notable catch-phrases. Given the short time for presentations by the members of the panel, the talks often made only one or two significant points. The demographics of the panelists included politicians, businessmen, and networking experts. All comments reported here are from the panelists unless otherwise specified.

- Presently there is vertical integration in our telecommunications infrastructure, ie., telephony, video, and data. We need horizontal applications, that combine all aspects of multimedia. Presently, telephony, video, and data exist in totally separate user architectures: we have separate media bringing our telephone service and video entertainment into our homes and businesses. High bandwidth data communications must also be provided over separate communications lines or even separate media. We need a communications format, available to all, that combines voice services, both video broadcast and video transmission, and high-speed data lines available in a multimedia format to a customer base that includes all households and business - large or small.

- The technology of the Internet is the enabler, but the Internet is all about people. It was predicted that 98% of American homes will have networking technology by 2000-2020. (I believe that number is somewhat high, given the age, education, and earning demographics proposed for this country in that time frame. I believe that a substantially smaller percentage of Americans will have any interest, have the time, understand how to use, and be able to afford the technology in their homes. NYNEX places the percentage of homes that it plans to cost effectively connect onto the Information Superhighway at approximately 60%.)

- Given the potential of an expanded version of the Internet, it is conceivable that the last 500 years may be referred to as the *gray ages*.

- The Internet will be able to support:
  - Virtual Organizations
  - Virtual Communities
  - Virtual Universities
  - Virtual Governments
  - ⋮

- Two years ago, no one understood the interest in and the potential for virtual communities, but Mosaic has proven to be a matrix upon which to build virtual communities. Mosaic has grown 10,000 fold in the last 18 months.

- The Internet is no longer a technical revolution, but rather a social revolution. People are the killer applications of the 1990s.

- Today's tools, grep, ftp, Mosaic, etc., are **not** good enough — especially if there is a desire to mandate the Internet for nearly all homes in America.

- Information will be mined for personal and corporate use by smart agents, aka., *knowledge-robots or knowbots.* But information will take on the characteristics of property, and have economic value. As a result, enhanced encryption techniques will be required to replace public key encryption in order to simplify key management and to enhance reliability.

- There was a request from those representing the Clinton Administration, Laura Breeden (DoC/NTIA) and Jonathan Gill (The White House), to emphasize the development of Internet tools for electronic Town Halls.

- A question was posed on the viability of the concept of a *Virtual University*. The panel readily agreed that the *Virtual University* will not replace the right-of-passage that a young person endures by permitting a student to take classes from his/her home, but rather the *Virtual University* should be an opportunity for students to experience the greatest intellectuals by being able to have a virtual presence at classes that includes viewing the lecture and being able to directly participate in class discussions.

5 Conclusions

Supercomputer networking has a broader ranger of requirements than general corporate enterprise networking. While the requirements for user access to workstations via (X-windows-based) client/server applications exist in both networking communities, the supercomputing community has requirements for significantly greater network bandwidths to support clustered computing applications and massive input/output requirements. To support supercomputing data requirements in the bus environment, SCI networking can provide $1 \mu s$ latency at gigabit-per-second network speeds to support shared memory applications on a workstation cluster. Serial HiPPI and Fibre Channel are competing to provide switched gigabit-per-second networking capabilities for the I/O and channel environments. ATM offers the potential for seamless networking within the LAN, MAN, and WAN environments. It is important to remember that **all** of these technologies are still immature. They each appear technically promising, but will all of these technologies have the financial backing to mature? Will there be a demand for SCI technology as more capable workstation clusters are developed? Will both HiPPI and Fibre Channel survive? Will standards-based ATM user-network interface (UNI) and network-network interface (NNI) protocols be implemented for true multi-vendor interoperability? Will the cost for ATM interfaces and ATM network services drop to a point where ATM will be the user network of choice?

My opinion is that all of these technologies will survive, but various technologies will be effected by different market forces. SCI, HiPPI, and Fibre Channel will most likely survive as vertical market items — high cost specialty components used in a limited market. While these technologies survive and prosper, their cost/bandwidth ratio will remain high, and only limited numbers of companies will be able to generate revenue with these technologies. Meanwhile, ATM will expand its market share until this technology commands a horizontal market status. ATM will be massed produced to support the telecommunications and video-entertainment industry, and the benefits of inexpensive custom circuitry will be available to the computer networking community. A significant problem that ATM must face is the inability to develop the technology rapidly enough to support the massive thirst for bandwidth that is developing within world-wide corporate enterprise networks, the National

Information Infrastructure (NII), and the Worldwide Information Infrastructure (WII).

When comparing the opinions of the highly technical members of the *Gigabit Networking* Round-table to the more secular members of the *Future of the Internet* Panel, opinions often were nearly contradictory. There was little disagreement between the technical members of the roundtable, or between the members of the panel; however, there was significant differences between groups. The opinions of technical people were that networking technology still had far to advance in order to be able to proliferate throughout our day-to-day lives. Meanwhile, those people involved in government policy areas declared that the technology issues were of less importance than the social issues of using the technology as the means to empower all Americans and provide equal opportunity for advancement. Technically oriented people expressed the view that there must be continued revolutionary improvements before there will be the scalable network technology available to provide the bandwidth to support true multimedia applications into all homes. The non-techies expressed the concern that the applications do not exist today that will permit the majority of Americans to have a positive experience from Internet connectivity. I believe that the hardware issues can be solved for cost effective large-scale proliferation of bandwidth in the 2000-2020 time frame, but I am less confident that the applications issue can be resolved because it will require overcoming many problems that exist at the most difficult of all levels, the user-machine interface.