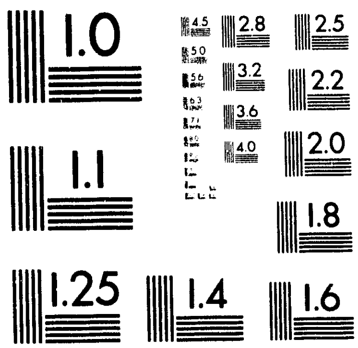

1.0
4.5
5.0
5.6
6.3
7.1
8.0
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

1 of 1

SAND--93-1686C
Conf-9310131--2

Delivery of Very High Bandwidth with ATM Switches and SONET

Steven A. Gossage

KEYWORDS

Asynchronous Transfer Mode, Synchronous Optical Network, Fiber Distributed Data Interface

ABSTRACT

To deliver high bandwidth, a ubiquitous inter-/intra-building cable plant consisting of single mode and multimode fiber as well as twisted pair copper is required. The selection of the "glue" to *transport* and *interconnect* distributed LANs with central facility resources over a pervasive cable plant is the focus of this paper. A description of the traditional problems that must be overcome to provide very high bandwidth beyond the narrow confines of a computer center is given. The applicability of Asynchronous Transfer Mode (ATM) switching (interconnection) and Synchronous Optical NETwork (SONET) (transport) for high bandwidth delivery is described using the environment and requirements of Sandia National Laboratories. Other methods for distributing high data rates are compared and contrasted. Sandia is implementing a standards based foundation utilizing a pervasive single mode fiber cable plant, SONET transport, and ATM switching to meet the goals of gigabit networking.

Introduction

Sandia installed its first inter-building optical fiber trunk cable (144 fiber, 50 micron core, multimode, AT&T ribbon) in September, 1981. At that time, optical fiber service to the computer customer community was dominated by point to point circuits to support distributed DEC VAX machines at 1 Mbps over distances of 2 - 3 km. Rapidly, the benefits of bandwidth caused an explosion in the demand for more optical fiber cable to serve distributed computers. As distributed host computing sites evolved into local area networks based on workstations, intra-building fiber distribution was focused directly at the desktop. Today, Sandia is completing a single mode inter-building trunk system. A multimode and single mode distribution to the desktop is being installed.

Approximately two dozen prototype optical fiber to the desktop distributed local area network connections to support the full 10 Mbps bandwidth of Ethernet were installed in 1988. In late 1990, a much more pervasive distributed Ethernet system based on smart hubs supported by both unshielded twisted pair and fiber to the desktop with a total connectivity of over 200 workstations was in place. Today, that system is beginning the conversion from Ethernet to the Fiber Distributed Data Interface (FDDI) standard. In August of 1991, the central site network was converted to FDDI. At that time, Sandia operated the largest FDDI ring built with technology from Network Systems Corporation. The first high performance local area networks served by distributed routers over DS3 (44.736 Mbps) circuits were installed in 1991. By July, 1993 the DS3 (44.736 Mbps) circuits were multiplexed onto a SONET aggregates of 155.52 Mbps (OC-3) and 622.08 Mbps (OC-12).

There are two major computing sites within Sandia (Albuquerque, New Mexico and Livermore, California, sites separated by approximately 1100 miles). The design philosophy proposed here addresses the requirements of both wide area and local area network high bandwidth delivery. The wide area signaling rate increased by a factor of greater than 10^3 over a decade. Today, the two labs are connected by two pairs of AT&T BNS-2000 ATM switches over DS3 trunks.

Sandia's goal is to provide a very high performance communications and computing environment so that no customer is "too far away" to enjoy the maximum available performance. To that end we are researching emerging technologies such as ATM service to the desktop. Plans are in place to extend desktop fiber distribution at a rate of two buildings per year. A distribution infrastructure composed of ATM switches coupled to multiplexed SONET transport will be used to provide *gigabit* capacities to distributed customers.

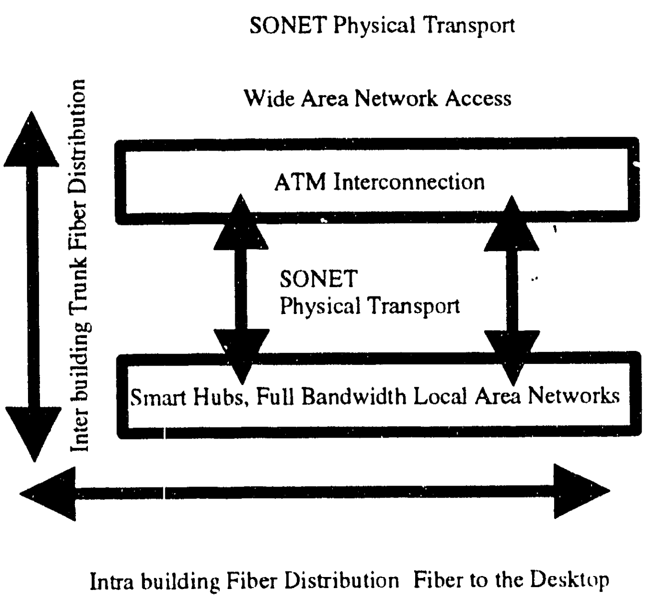


Figure 1. Structure for High Bandwidth Delivery

In the last decade, Sandia has evolved from a facility where asynchronous terminals accessed supercomputers at rates generally below 9600 bps to a facility where hundreds of workstations have connections to optical fiber and access remote supercomputers at rates in the 10's of Mbps. This transition has been driven by the rapid evolution in distributed computing resources from minicomputers to workstations. The transition has been *enabled* by optical fiber cable, advances in local area network technology and standards, as well as by evolution in switching technologies from basic circuit switches to ATM.

Environment at Sandia

The components of Sandia's high bandwidth (45Mbps and beyond) delivery system include: high performance central facilities, distributed, high performance local area networks, inter- and intra-building cabling, and the electronics necessary to distribute bandwidth among the LANs and the central facility. The central facilities include the Secure Supercomputing Network, major mainframes (CRAY Y-MP, etc.), file storage, and specialized peripherals (color printers and copiers, etc.). Ethernet and FDDI are the primary LAN technologies. The inter- and intra-building cable plant consists of high performance unshielded twisted pair (UTP) copper cable to virtually all desktops, and multimode (MM) fiber and single mode (SM) fiber to high-end customers. The electronics used to deliver bandwidth between the distributed LANs and the central facility include an array of routers, ATM switches, and SONET multiplexers **(See Figure 1.)**.

The physical dimensions of the Albuquerque campus are large when a LAN "ruler" is applied. There are many LANs that are separated by more than 2 km of cable which is beyond the capability of Ethernet or FDDI without repeaters. The total number of LANs is greater than 245 requiring a multitude of interconnection and segmentation options. The performance of the Wide Area Network (WAN) components linking the New Mexico and California campuses is critically important. Characterizing the realizable throughputs obtainable from various hosts and workstations connected through a variety of communications/networking elements provided key data for planning the transition to SONET transport. The 11dB budget of FDDI is frequently exceeded in both sites, especially on cross campus, desk to desk links. In many ways, the two lab campuses more closely resemble the Metropolitan Area Network (MAN) model and "ruler".

The next generation inter-building trunk cable plant is being installed using 8.3 micron core SM fiber, the telecommunications standard of the '90's and beyond. The new wiring standard for offices includes the highest quality unshielded twisted pair, 62.5 micron core MM fiber, terminated to support the FDDI standard, and unterminated SM fiber to the desktop **(See Table 1.)**. Both Black (open) and Red (classified) voice, data, and video communications are supported by the new wiring standard being installed pervasively at Sandia. The communications/networking philosophies are strongly influenced by the requirement to properly protect and segregate information.

BUILDING	MEDIA (MM,SM,UTP)	TERMINATED (Y/N)	No. of CONDUCTORS	SERVICE CLASS
880 (Installed)	MM	Y	4 EA	BLACK DATA, VIDEO
	SM	N	4 EA	BLACK DATA, VIDEO
	UTP	Y	4 PR	BLACK VOICE
	UTP	Y	4 PR	BLACK DATA
980 (Installed)	MM	Y	6 EA	BLACK DATA, VIDEO
	SM	N	4 EA	BLACK DATA, VIDEO
	UTP	Y	4 PR	BLACK DATA
PROPOSED BLACK STANDARD	MM	Y	2 EA	BLACK DATA, VIDEO
	SM	N	2 EA	BLACK DATA, VIDEO
	UTP	Y	4 PR	VOICE
	UTP	Y	4 PR	DATA
PROPOSED RED STANDARD	MM	Y	2 EA	RED DATA, VIDEO
	SM	N	2 EA	RED DATA, VIDEO
	UTP	Y	2 PR	RED DATA

Table 1. User Outlet Configurations

Traditional Problems with High Bandwidth Delivery in the World Outside Computer Rooms

High speed data delivery on a large scale between buildings and offices has always been a difficult problem. Common problems include: 1) Severe distance limitations (both inter- and intra-building) due to physical drive capabilities or timing imposed by interface specifications; 2) Nonstandard or proprietary interfaces which are very difficult to extend or to interconnect to a standard system; and 3) Nonscaleable high speed network topologies.

Distance, attenuation, and bandwidth limitations can be largely overcome by pervasive utilization of optical fiber cable. The cornerstone facility cable plant, both inter- and intra-building, must provide large numbers of conductors (MM, SM, and twisted pair) distributed ubiquitously. It is crucial to install a flexible cable plant topology with large capacity to reduce the impact of difficult network topologies. System planners should investigate the effects of large numbers of rings mixed with point to point circuits to test the potential capacity of their cable plants. Such investigations amplify the utility of multiplexing and standards integration. Systems are nonscaleable with respect to bandwidth if more optical conductors are required to deliver more bandwidth instead of the more elegant method of utilizing untapped bandwidth within existing conductors. Nonscaleable transport electronics can only deliver more bandwidth through replication versus the option of increasing individual system performance. A multiplexed transport system such as SONET can add backbone capacity without changing individual user access speed.

A Structured Approach to Delivering Large Bandwidth

The pivotal component of the system is the inter- and intra-building cable plant. To provide virtually unlimited bandwidth, Sandia has committed to a major cable plant upgrade. All trunks between communications centers will have access to between 36 (Area V) and 408 (Area I, Central Site) single mode fibers. This cable plant is capable of supporting Gbps networks. Inside the buildings, fiber to the desktop is being provided on a prioritized basis to customers at the forefront of computing capacity. The MM facility will support 600-800 MHz systems easily, while the SM cable plant awaits the arrival of desktop interfaces with Gbps capability. The inter-building trunk should be complete in about one year **(See Figure 2.)**. For the widely dispersed computer sites on the Albuquerque campus, single mode cable was the best choice for today and for the future. The selection of optical fiber cable technology, MM versus SM, for the inter-building cable trunk system required detailed analysis of the benefits of each technology. The length of a typical circuit in Albuquerque is over 2 km and significant growth is being experienced in circuits that must cover 3.5 to 10 km. Single mode technology offered the greatest available bandwidth at the lowest possible attenuation, thus providing an optimal solution to servicing a growing population of long distance customers while eliminating the need for repeaters. Single mode cable is also best suited for aggregating (multiplexing) many sources of bandwidth together. Bandwidth aggregation is a key application of the inter-building trunk system Single mode fiber cable technology has clear advantages over multimode cable in the Albuquerque campus trunk environment. For the desktop intra-building distribution system, the combination of MM and SM with large numbers of fibers was the best choice. Desktop workstation LANs are still dominated by shared media technologies that are well suited to MM fiber. Conduit space is generally more readily available to serve the multiconductor, high bandwidth needs of a Laboratory computer user in the intra-building distribution system than in the inter-building conduit. The greater availability of conduit and conductors for the intra-building distribution is a critical distinguishing feature when comparing desktop and inter-building cable plant decisions. Recognizing that LAN bandwidth requirements are increasing, unterminated SM cable is also being installed to the desktop. Initially, desktop fiber will serve the customers

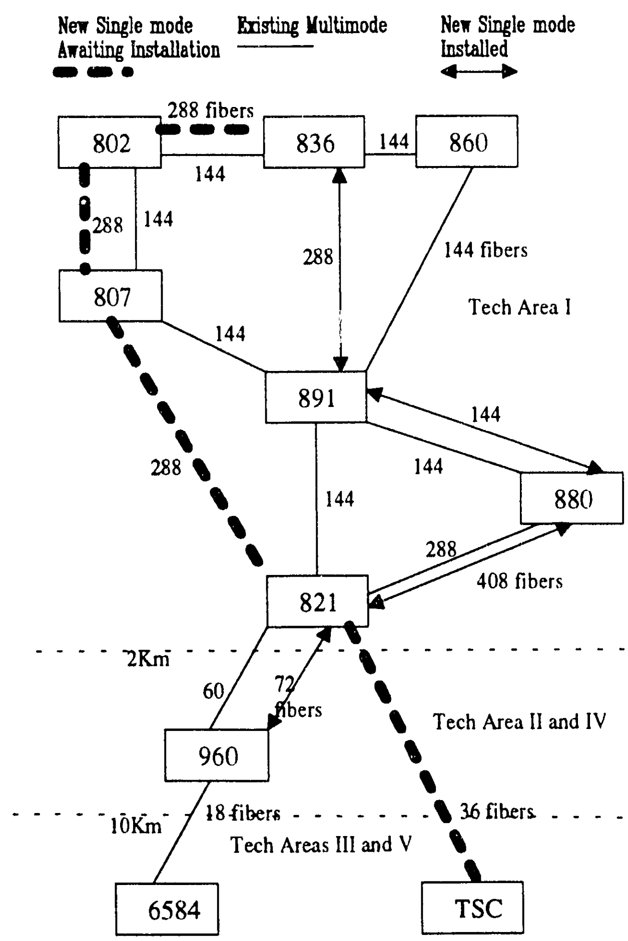


Figure 2. Optical Fiber Interbuilding Trunks

(approximately 20% of the total) with leading edge requirements for bandwidth, but eventually desktop fiber will be used by the majority. Local area networks (LANs) utilize the improved cable plant with both fiber and twisted pair Ethernets. The highest speed LANs are built around smart, manageable hubs connected in a star wired topology. Most of the LANs at Sandia are Ethernet with a migration to FDDI expected when the intra-building cable plant becomes more pervasive and the cost per FDDI connection drops to approximately 20% (or less) of the workstation cost. **(See Figure 3.)**.

High Bandwidth Delivery: Electronics and Requirements

ATM switches and SONET distribution were chosen to interconnect distributed LANs and the central facility to create an integrated high bandwidth delivery system. ATM technology addresses the crucial issue of *interconnection* while SONET solves the *physical transport* problems. The requirements for transporting and interconnecting high bandwidth pipes include: 1) Very high bandwidth capacity, from 10's of megabits to Gbps, 2) Scaleable architecture so the bandwidth increases can be done in a modular, evolutionary fashion, 3) Efficient use of the cable plant so that multiplexing / concatenation methods efficiently use single mode cable bandwidth by increasing the *modulation rate* instead of the *number of fibers* required, 4) Virtually unlimited loss budget and distance to support even the most remote systems, 5) Very high reliability to maximize delivery system availability, 6) Standards oriented transport to assure interoperability, maintainability, future enhancement, and growth in a multivendor environment, 7) Effectiveness in local and wide area network applications; Minimal transport latency, 8) Flexible architectures to accommodate a range of reliability, performance and physical plant constraints, 9) Integrated system management for maintenance and reliability; and 10) Capacity to flexibly interconnect 100's of end points without loss of performance.

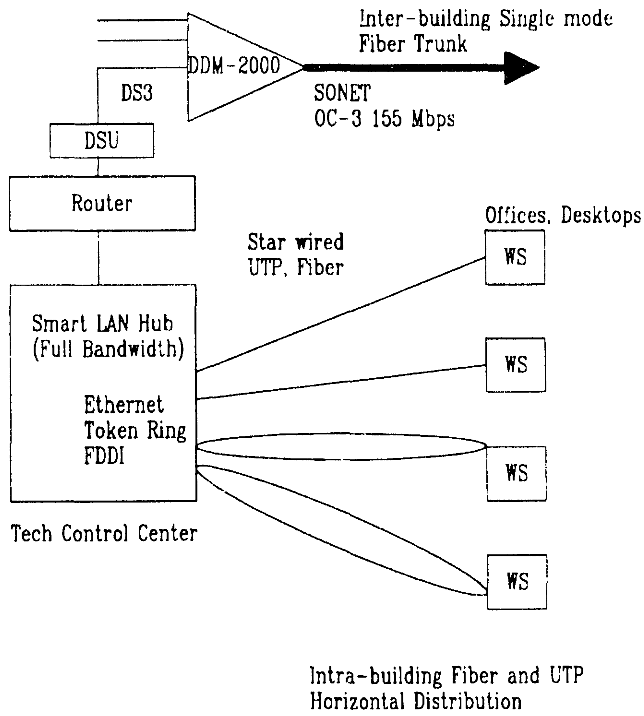


Figure 3. Distributed High Performance LANs

Sandia has network bandwidth requirements that range from a few megabits per second (Mbps) into the gigabit per second (Gbps) range. Ethernet and FDDI can effectively supply 1-100Mbps for LAN applications. Visualization of scientific/engineering data analysis that utilize CRAY channels, massively parallel machines, and frame buffers require 100's of Mbps to 1 Gbps. Aggregation of the above applications drives the system transport bandwidth requirements into the 10's of Gbps. Sandia must supply this wide range of bandwidth (3 or 4 orders of magnitude range) to widely dispersed customers as effectively (with regard to cost and performance) as possible. The current LAN standards, Ethernet and FDDI, and the emerging de facto standard for high speed computer channels, HiPPI, cannot efficiently support the wide range of bandwidths and the dispersed customers. In addition, those standards exact a huge price in the number of fibers required for a sizable implementation. For high bandwidth transport, SONET is the most suitable technology **(See Figures 4, 5, Tables 2. and 3.)**.

Implementation Scaling and Cable Plant Impact

Tables 2. and 3. illustrate the tremendous cost of other high speed distribution methods to both the cable plant and system bandwidth. Ethernet, FDDI, and HiPPI all suffer from the combined limitations of low effective utilization of fiber bandwidth and high utilization of cable conductors. These issues are particularly difficult in the Sandia environment where LAN traffic must be kept separated to meet security requirements. There are also performance issues to be considered. Tests conducted at Sandia indicate that four to five workstation pairs are capable of filling the total bandwidth of an FDDI ring. In a high performance environment, multiple FDDI networks would be required. Unfortunately, the cable plant / bandwidth cost is too great.

The figures and tables clearly show why Sandia must have a scaleable transport architecture. No Laboratory facility can afford enough optical fiber to allocate 2-4 single mode trunk fibers to each distributed high speed application. There are simply too many distributed applications. It is clear that graceful scaleability is crucial in this environment. SONET allows the bandwidth to be fully exploited without increasing the number of fiber pairs required. The transport fabric can greatly increase its capacity without any visible upset to the customer community.

Transport Method	Number of LANs	Number of Fibers	Total number of LANs Supportable/ Major Bldg. w/o Routers*
Ethernet	N	4N	6
FDDI	N	4N	6
HiPPI	N	4N	6
SONET	N	4	48

Assumptions: All links are redundant; 24 fibers / major building; * 100% utilization of the Sandia trunk;
Table 2. Scaling of LAN Transport Technologies

Transport Method	Total LANs Supportable with Routers*	Relative Impact on the Cable Plant**	Bandwidth Utilization Factor %***
Ethernet	36	4	0.4
FDDI	12	4	4
HiPPI	6	4	44
SONET	288 Ethernets, 96 FDDI	1	100

*Assume 6 Ethernets per router or 2 FDDI LANs; **Relative Impact = Number of Fibers Required / Lowest Number of Fibers Required; ***Bandwidth Utilization Factor = Bandwidth Used / Maximum Bandwidth Supportable

Table 3. Cable Plant Impact of LAN Transport Technology

For example, if Ethernet was used as a transport backbone where six Ethernets were supported by each router, then the entire backbone cable plant of 24 fibers would be used up by only six routers (6 routers X 4 fibers/ routers = 24 fibers) to support only 36 Ethernets (6 routers X 6 Ethernets /router). Furthermore, Ethernet with a signaling rate of 10 Mbps utilizes only 0.4% of what could be supported by SONET transport at 2.488 Gbps.

Efficient Use of Bandwidth

Ethernet, FDDI, and HiPPI backbones only scale by replication, thus exacting a large price from the cable conductor plant. SONET, on the other hand, never requires more than 4 fibers and evolves through technology upgrades to utilize increasing amounts of available single mode cable bandwidth. One of the major features of single mode optical fiber is its huge bandwidth (10-100 GHz-km). Since bandwidth is valuable, it should be efficiently utilized by squeezing the most out of each conductor rather than applying more underutilized conductors to deliver aggregate bandwidth.

Interface Standard	Maximum Unrepeatered Distance (km)	Fiber Type	Power*
Ethernet	5	Multimode	0.05
FDDI	2	Multimode,	.2
	20	Single mode	2
SONET OC-3	20	Single mode	3.1
HiPPI	20	Single mode	22
SONET OC-48	40	Single mode	100

*Power = Bandwidth in Gbps * Distance in Km

Table 4. Power to Deliver Bandwidth Over Large Distances

Table 4. compares the relative power of each of the transport technologies to deliver bandwidth at large distances. SONET is the only high speed transmission standard defined and available for rates from Mbps to Gbps. It should be noted that LAN interface conversions to SONET at OC - 3 (155.52Mbps) are now available. A standard has been proposed for mapping FDDI into SONET OC-3c. More work is being done to develop HiPPI extensions to SONET via ATM technology. Additionally, several projects are underway to convert FDDI to ATM where SONET could again be the transport.

Sandia Transport Backbone Implementation

Sandia has chosen to implement a SONET high speed transport system utilizing a Digital Access Cross Connect System (AT&T DACS III - 2000) and Digital Data Multiplexers (AT&T DDM - 2000) **(See Figure 5.)**. The DACS provides for rapid reconfiguration of circuits and enhanced manageability. The DDM - 2000 provides the actual SONET transport. The DS3 rate is currently well matched to present workstation to workstation (memory to memory) transfer rates. Typically, 14 - 31 Mbps are achievable via an FDDI to router interface and those rates fit easily within the DS3 payload.

With SONET transport, one high speed network facility can deliver voice, data, and video since interfaces exist to convert each of those inputs into the SONET standards. No other high speed interface method (FDDI, Ethernet, HiPPI, etc.) can provide that level of bandwidth performance and also the versatility to support every major type of communications input. For the first time in our facility, SONET transport provides greater bandwidth on the transport/interconnection pipes than the available bandwidth on the distributed LANs. Bandwidth between LANs is no longer constrained by the physical transport or cable plant. A number of interesting results are anticipated. SONET transport equipment overcomes all of the traditional limitations of high bandwidth transport.

ATM Interconnection System

To complete the system, some method is needed to interconnect the LANs, the central facility, and the WAN environments. In the Sandia environment, the best method is a high performance switch that utilizes an Asynchronous Transfer Mode (ATM) switching fabric. A switching platform, well integrated with SONET physical transport, was chosen over alternatives such as a hierarchy of routers glued together with LAN technologies. The combination of contention type access and very large switched backplane bandwidth provided by ATM technology is a powerful and efficient way to access precious resources such as supercomputers.

There are three key elements driving the choice of interconnection method. They are bandwidth, switching, and switch technology. Recall the model shown in **Figure 4.** where many separate full bandwidth (up to 100 Mbps, FDDI) distributed LANs are interconnected to each other and to the central facility with SONET. It has been shown that five or fewer pairs of workstations can saturate a single FDDI ring and that a single host to host transfer can take over 65% of the available ring bandwidth. In the Sandia environment which contains hundreds of LANs and nearly a thousand workstations, significant aggregate bandwidths are easily and commonly generated. FDDI does not offer enough bandwidth to serve as the interconnection mechanism for high bandwidth environments especially when the LANs being connected equal the bandwidth capacity of the backbone interconnection method.

The choice to switch rather than route or bridge provides access to very large aggregate bandwidths, allows customers to contend for bandwidth, and avoids single processor bottlenecks and other shared single media problems. In a router or a bridge, a single processor must see all of the traffic coming off a LAN. If multiple LANs are interconnected to a single router, then that single processor can become the bottleneck. In a switched interconnection system, one path or processor is not required to see all of the network traffic, thus avoiding the limitation of shared single media.

ATM advantages include low latency, scaleability for capacity and bandwidth, and support for multimedia (voice, data, and video) in a single platform. Just as optical fiber cable has been hailed as a "virtual pipe" for its ability to handle all types of traffic with a single media, ATM may well be the "virtual interconnection fabric" for the same reasons. LAN technologies are severely limited in their ability to support more than data. FDDI-II will offer the capability to deliver isochronous traffic (voice and video) by allocating multiple Wide Band Channels (WBC) of 6.144 Mbps. The single shared media problem remains and now video and voice will be supported at the expense of high speed data. An ATM switch, on the other hand, acts as a massively parallel interconnection system where many paths are active concurrently and no single processor or media acts as a bottleneck. A massively parallel ATM switch is a scaleable architecture. Cell relay systems in the 2.5-40 Gbps range are realizable now. Bandwidth in the 10 -100 Gbps range will follow. Already prototypes from Fujitsu can demonstrate capacities of 40 Gbps with access speeds of 155 Mbps. Other companies including AT&T, GTE, TRW, Hitachi, Toshiba, Sumitomo and others have introduced their very high performance ATM platforms.

Conclusions

ATM and SONET technologies provide significant performance advantages for the interconnection and transport of LAN traffic. Low latency, scaleability, and high bandwidth are advantages of ATM and SONET that must be applied to the next generation of gigabit LANs. ATM switches and SONET transport will be the enabling technologies.

Acknowledgments

The author wishes to thank Roger Adams, Terry Bisbee, Michael Vahle, Lawrence Tolendino, Nicholas Testi, Jim Schutt, Helen Chen, Richard Swerdlow, and John Naegle for their contributions to the clarity and content of this document. This work was performed at Sandia National Laboratories supported by the United States Department of Energy under contract DE-AC04-76DP00789.

References

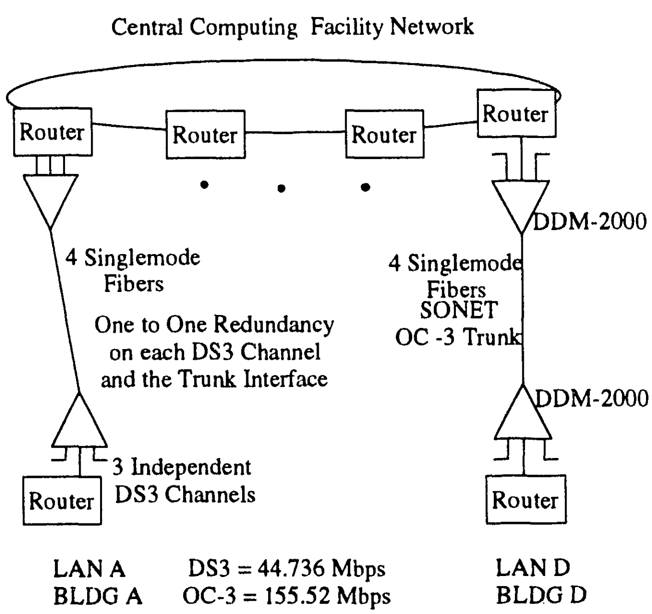
1. Gossage, Steven A. and Nicholas Testi. "Design and Implementation of an FDDI Cable Plant and Network," NEXUS XXIIII Proceedings, Network Systems User Group Annual Conference, Grosvenor Resort, Orlando, Florida, (September 11-14, 1990), 53-60.

2. McQuillan, John M. "Cell Relay Switching," Data Communications, (September 1991), 58-69.

3. Miller, Calvin M., Stuart C. Miller, and Ian A. White. Optical Fiber Splices and Connectors Theory and Methods. Optical Engineering Series Vol. 10. New York: Marcel Dekker, Inc., 1986, Pp. 163-164.

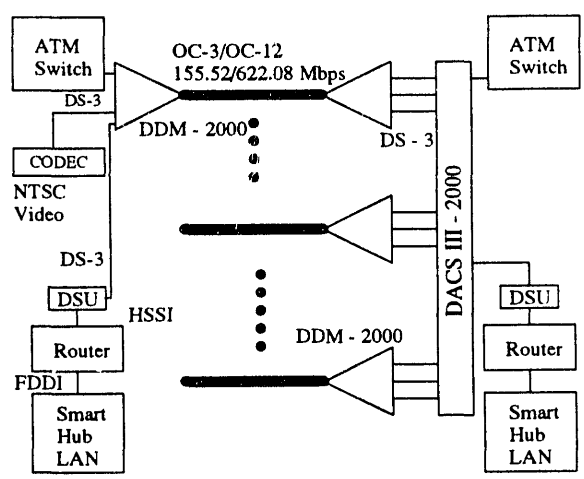
4. Testi, Nicholas and John H. Naegle. "Interpreting and Achieving Performance in the TCP/IP FDDI LAN," Business Communications Review, Vol. 22 No. 3, (March 1992), 16-21.

5. _______. "ATM Sets Speed Record," MAN Newsletter, Vol. 3 No. 1, (January 1992), 18.



Assumptions: Fully redundant OC-3 Connection
Four Distributed Customers With
Independent Networks

Figure 4. Hypothetical Host to Customer Interface Using SONET As A LAN Transport



ATM--Asynchronous Transfer Mode
CODEC--COder DECoder
DDM--Digital Data Multiplexer
DSU--Data Service Unit
DS-3--Digital Signal 3, 44.736Mbps
HSSI--High Speed Serial Interface
NTSC--National Television Systems Committee
OC-n--Optical Carrier n
SONET--Synchronous Optical NETwork

Figure 5. SONET Distribution

DATE

FILMED

10/26/93

END