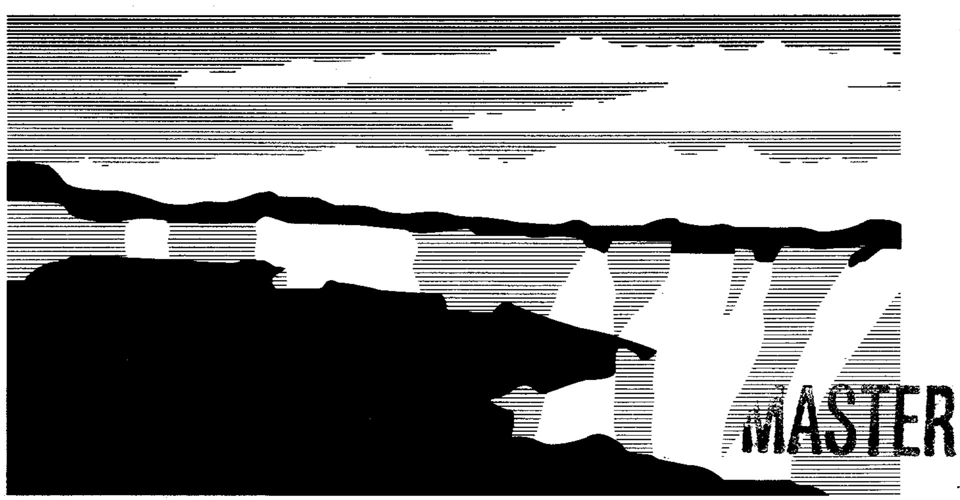
Conf-9505116--1

LA-UR- 94-3994

Title: Gigabit LAN Issues - HIPPI, Fibre Channel, or ATM?

Author(s): Donald E. Tolmie

Submitted to: International Conference and Exhibition on High-Performance Computing and Networking, May 3-5, 1995, Milan, Italy



Los Alamos
NATIONAL LABORATORY

Gigabit LAN Issues – HIPPI, Fibre Channel, or ATM ?

Don Tolmie
Los Alamos National Laboratory

Abstract

Computer networks that operate in the gigabit per second speed range are becoming very important for interconnecting supercomputers, clusters of workstations, and other high end equipment. HIPPI is the current interface of choice, while Fibre Channel and ATM are emerging standards. These systems are examined as to their backgrounds, advantages, and shortcomings.

Interconnect cultures

Often a simple name on an interface does not give a very good idea of what application area the interface was intended for. One way is to look at the background of the people who designed the interface – what sorts of interfaces have they developed in the past. The interface world can be divided into five separate cultures. These are listed below followed by examples.

- Backplane
  VME, Futurebus +, SCI
- Peripheral I/O Channel
  SCSI, IBM Block Mux, IPI, Fibre Channel
- Local Area Network (LAN)
  Ethernet, FDDI, HIPPI
- Metropolitan Area Network (MAN)
  FDDI, IEEE 802.6 (DQDB)
- Wide Area Network (WAN)
  Telecommunications, X.25, Internet, SONET, ATM

The cultures can be differentiated by the problems they are trying to solve and the building blocks that were used in the solution. They can also be categorized by "Control" and "Trust". For example, in a backplane environment, a single user has complete control, and gives complete trust since usually the card being installed can read or write anywhere in the system, possibly crashing the system. Compare this to the wide area telecommunications network -- here a single user has very little control (does your cross-country phone conversation go through St. Louis, Chicago, or Dallas?), and shows the lack of trust by using protocols with lots of checking and firewalls.

Figure 1 gives a brief comparison of some of the differences between the computer LAN culture and the telecommunications culture. For example, on Ethernet, Fiber Distributed Data Interface (FDDI), or Token Ring LANs, once a user has arbitrated for, and obtained, the media, he transmits a packet at the full bandwidth of the media. In comparison, the telecommunications culture is oriented around multiplexing and aggregating many low bandwidth connections on a single physical media. The LANs mentioned also send data in a simplex "ship and pray" datagram mode, with responses returning in separate datagrams. In comparison, a full-duplex telecommunications circuit must be established before any information is transferred between the end nodes. While the LANs have used variable sized packets on the media, constant cell sizes, or fixed frequency bandwidths, are used in telecommunications.

	Computer	**Telco**
Data flow	Full b/w for short time	Multiplex many low b/w
Switching	Simplex datagrams	Duplex circuits
Info unit	Variable packet size	Constant cell size
Overload	Slow everyone down	Deny service, toss data
Charging	Buy switching hardware	Rent allocated b/w
Goal	Don't corrupt data	Keep channels busy

Figure 1. Comparison of computer LAN and telecommunications cultures

Another example is charging and data reliability. On a backplane, peripheral I/O channel, or LAN, once you buy the equipment the vendor does not care if you use it at 1% or 90% of the total bandwidth. But in these environments the data must be delivered correctly, or if an error occurs, this should be a very rare exception and should be flagged. In contrast, the end user does not directly buy telecommunications network central office equipment, but instead rents bandwidth. Hence, the goal for the telecommunications providers is to keep the channels as full of data as possible since this is where they get their revenue. To achieve the full bandwidth, the providers may oversubscribe the links, depending upon statistics over a large user population to avoid dropping data. Since the telecommunications WANs were developed primarily for voice traffic, dropping small amounts of voice data is not catastrophic; but dropping packet data is catastrophic.

If an interface is developed for a particular "culture", it is not unrealistic to use the interface for a different culture, but it may not be optimized for that culture. It is highly unlikely that an interface developed for a backplane would be appropriate for a WAN, and vise versa. If someone touts a particular interface as the "best for all applications", you should be very cautious – it is probably wishful thinking.

HIPPI

The High-Performance Parallel Interface (HIPPI) was developed in American National Standards Institute (ANSI) Task Group X3T9.3. The work started in 1987, and today HIPPI is the interface of choice for high-end applications.[1] The HIPPI physical layer specifies an 800 or 1600 Mbit/s (100 or 200 Mbyte/s) simplex interface for distances of up to 25 meters using copper twisted-pair cables. Other HIPPI documents define the packet format, mappings for upper-layer protocols, control of physical layer switches, and a fiber-optic extender.[2]

The documents defining HIPPI include (1) the physical layer (HIPPI-PH, ANSI X3.183-1991), (2) a framing protocol that defines the packet format (HIPPI-FP, ANSI X3.210-1992), (3) a mapping to IEEE 802.2 to support communications protocols such as TCP/IP over HIPPI (HIPPI-LE, ANSI X3.218-1993), (4) a definition for physical layer crossbar switches (HIPPI-SC, ANSI X3.222-1993), and a fiber-optic extender for distances up to 10 kilometer (Serial-HIPPI, not an ANSI standard). The ANSI documents are also being processed as international standards.[3]

Additional documents being developed in X3T9.3 include (1) mappings to the IPI-3 generic command sets for magnetic disks and tapes, (2) a mapping allowing upper-layer HIPPI protocols to use a Fibre Channel physical interface, and (3) a mapping between HIPPI and ATM.

HIPPI came from the high-speed local network culture. Attempts to include features to better support direct disk and tape I/O were expressly omitted. HIPPI is intended as a memory-to-memory interface, and is used in this mode with high-end striped disk and tape systems. Figure 2 gives a brief summary of HIPPI, and a comparison to Fibre Channel.

HIPPI ADVANTAGES

HIPPI is the current interface of choice, largely because it was the first standard at close to the gigabit speed. It came to fruition quickly because of a "keep it simple" goal, and a well-focused direction in the standards committee that avoided adding lots of bells and whistles. Some of the advantages of HIPPI include:

- It is simple, elegant, and easy to understand.
- It has a good physical level flow control. The flow control even works with very long links by the addition of extra buffering at the receivers (approximately 1 kilobyte per kilometer of distance).
- A good tester was developed early on which allowed vendors to test implementations in-house so that interconnection with other vendor's equipment was usually a plug-and-play.
- A variety of products with HIPPI interfaces from a fair number of vendors currently exist. Many are second generation designs, incorporating improvements from earlier designs.
- HIPPI crossbar switches are available from multiple vendors.
- HIPPI specific integrated circuits are available. Even so, some vendors find that small scale integration parts are more suitable due to the simplicity of the physical interface and limitations of the HIPPI specific ICs.
- HIPPI to SONET adapters are available for very long distance links using telephone network facilities.

	HIPPI	Fibre Channel
Data rate (Mbit/s)	800, 1600	100, 200, 400, 800
Striping	No	Yes (future)
Transmission	Parallel	Serial
Media	Twisted-pairs	Fiber, Coax, Twisted-pairs
Architecture	Simplex	Duplex
Components	Off-the-shelf	Custom silicon
Data Classes	Datagram	Datagram Multiplex w/ACKs Connection w/ACKs
Options	Speed only	Lots of options
Fabric	Crossbar	Crossbar, Ring, Tree
First Delivery	1988	1993

Figure 2. HIPPI and Fibre Channel comparison

HIPPI SHORTCOMINGS

HIPPI is not without limitations and shortcomings. Perceived shortcomings include:

- It is not a mass-market item, the number of applications that require the bandwidths are not that numerous. Hence the price is higher. It is questionable whether competing gigabit/s technologies, e.g., Fibre Channel or ATM would be any cheaper.
- It does not support speeds slower than 800 Mbit/s. Slower speeds would help make it more of a mass market item.
- It does not support multiplexing. If you transfer a megabyte over a HIPPI channel as a single entity then it will take at least 10 milliseconds. During this time the channel cannot be used for any other communications.
- HIPPI does not support time-critical or isochronous data.
- The HIPPI specification limits the distance to 25 meters (82 feet) with copper twisted-pair cable. Serial-HIPPI defines a fiber-optic extender than is useful for distances up to 10 kilometers, but it is an added expense.
- The cable is somewhat bulky and stiff.
- The cable connector is large and somewhat fragile.

Fibre Channel

Fibre Channel is an emerging standard, also from the ANSI X3T9.3 Task Group.[4] The Fibre Channel work started on it in 1988, one year after HIPPI started. The first Fibre Channel documents are just now being completed, and the first products being delivered. Fibre Channel supports burst data rates of 100, 200, 400, and 800 Mbit/s. As the name implies, it is based on serial transmission over optical fibers, whereas HIPPI was based on parallel transmission over copper wires. The first products are being developed at the 200 Mbit/s speed, higher speed products will follow shortly. Figure 2 gives a brief summary of Fibre Channel, and a comparison to HIPPI.

You may see Fibre Channel referred to with different rates, for example, 133, 266, 531 or 1062.5 Mbit/s. These rates are the serial stream signaling rates that include the 8B/10B encoding and other overhead. The corresponding rates for the user data portion of the serial stream are 100, 200, 400 and 800 Mbit/s respectively.

The Fibre Channel Physical and Signaling Interface (FC-PH) document has completed development and is in the review process. The FC-PH draft document is available from Global Engineering as ANSI X3.230-199x. Other Fibre Channel documents under development include:

- FC-EP, enhanced physical layer with support for isochronous, stripped physical layers (e.g., running three FC-PH physical layers in parallel for three times the bandwidth on a single transfer), and other things left out of FC-PH.
- FC-IG, implementation guide with state diagrams for FC-PH and a collection of folklore and helpful hints.
- FC-SB, mapping to single-byte command code sets, i.e., IBM Block Mux command sets
- FC-FP, mapping to HIPPI upper-layer protocols
- FC-ATM, mapping to Asynchronous Transfer Mode (ATM) transport
- FC-LE, mapping to IEEE 802.2 for support of communications protocols such as TCP/IP
- SCSI-FCP, mapping for SCSI protocols to use Fibre Channel physical layer for higher speed SCSI devices
- SCSI-GPP, mapping for SCSI generic packetized protocols
- IPI-3 Disk and Tape, revisions to the existing IPI-3 standards to include running over the Fibre Channel physical layer
- FC-FG, fabric generic requirements
- FC-SW, crosspoint switch fabric
- FC-AL, arbitrated loop
- FC-DF, distributed fabric
- FC-DS, directory services

Fibre Channel advantages

Fibre Channel came from the "mainframe I/O channel" culture, and it should provide an excellent solution for that application. Fibre Channel's success at penetrating the LAN environment remains to be seen. Fibre Channel is considerably more complex than HIPPI, but it also includes many more features. It will be interesting to see if this extensive set of Fibre Channel features turns out to be a boon or a bane. Advantages of Fibre Channel include:

- Very versatile; can do almost anything.
- Supports multiplexing of 2 kilobyte frames of different information transfers.
- Supports dedicated switched circuits.
- Supports datagram service, i.e., best-effort transfers without acknowledgments.
- Supports broad range of speeds with common integrated circuits.
- Defines a variety of interconnection fabrics, including (1) a crossbar for highest throughput, and highest cost, (2) distributed switching elements for maximum flexibility and ease of growth, or (3) arbitrated loop (a ring architecture with no extra fabric elements required) for lowest cost and lowest performance.
- The switch definitions allow for easy mixing of speeds in a single system. For example, a switch can simultaneously interconnect 200 and 800 Mbit/s end nodes. As you may expect, speed mixing is not supported on the arbitrated loop.
- Fibre Channel may support time-dependent isochronous data, e.g., voice or video, in future releases. Striping across multiple physical channels is also planned as a future enhancement.

Fibre Channel shortcomings

In trying to be "all-things-for-all-people", Fibre Channel included an extensive set of options, which some people label as "bells and whistles". Only the future will tell if this large option set made Fibre Channel stronger by being useful for a large set of applications. The large set of options could also be a detriment, resulting in vendors having difficulty making interoperable products, i.e., the set of options used by one vendor are not compatible with another vendor's equipment. Fibre Channel may have tried to do too much in one interface. Other perceived shortcomings of Fibre Channel include:

- The development process has taken a long time; Fibre Channel may miss its window of opportunity.
- Integrated circuits supporting Fibre Channel are just becoming available, and they may be made obsolete by later changes in the specification before it is an approved standard.
- Fibre Channel does not currently support time-dependent, i.e., isochronous, data. There are plans within the ANSI committee to add this support, but when it will be developed, and when it will be available in integrated circuits, are open questions.
- All of the options and capabilities resulted in a specification that is quite difficult to read and understand. The complexity will also make interfaces difficult to implement, check out, and verify against other vendor's interfaces.

ATM

In this context, ATM stands for Asynchronous Transfer Mode, not automatic teller machine. ATM came from the telecommunications community, and defines a protocol for sending information in 53-byte cells.[5,6]

Note that ATM is not a physical level interface. SONET (Synchronous Optical NETwork) is the physical layer interface most often mentioned with ATM. SONET is a point-to-point interface supporting data rates from 51 Mbit/s to gigabits per second.[7] SONET does not support switching by the end users, ATM will provide this function. Hence, SONET is comparable to a leased line, and ATM to a dial-up connection. The SONET speeds on optical interfaces are designated as OC-*n*, where the serial speed on the link is about 51.8 Mbit/s times *n*, i.e., OC-3 is about 155 Mbit/s, OC-48 is about 2.4 Gbit/s. As with Fibre Channel, these rates include overhead; the actual user data rates are about 75% to 85% of the signaling rates.

The 53-byte ATM cell was designed for carrying many separate voice traffic connections over a single physical media. The ATM cell is composed of a 5-byte header with routing, control, and checking information, and a 48-byte payload. Adaptation layers, called AAL1 through AAL5, define the nature of the information in the payload. For example, AAL1 is intended for constant bit rate data, e.g., voice or video, while AAL5 is intended for packet data that has no specific timing requirements.

ATM standards documents are being developed in ANSI, in ISO, and in the ATM Forum.

ATM ADVANTAGES

ATM was designed for wide area telecommunications networks, but there is also a lot of interest in using ATM technology in LANs. Some of the advantages would include:

- ATM already has good support for mixing time-dependent data, e.g., voice and video, with packet data.
- ATM is independent of the underlying physical media, but is most often mentioned in conjunction with SONET. This combination supports a wide range of speeds from megabits per second to gigabits per second.
- It is easy to mix equipment with differing speed interfaces, e.g., OC-3 and OC-12, in the same system.
- By using the same technology as the telecommunications industry, larger volumes of common components should result in lower prices.
- Bridging between LANs using ATM and wide area networks (which also use ATM) should be simpler than converting between two dissimilar standards.
- WANs presently have extensive network management tools and these tools may be available in an ATM based LAN.
- There is a lot of interest and momentum behind the ATM work – if it fails it will not be due to the lack of talented people working on it, or the lack of effort.

ATM SHORTCOMINGS

ATM was not specifically designed for LAN usage, and hence has some shortcomings when used in that environment. Perceived shortcomings include:

- Vendors are building and delivering products before the standards and problems have been solved. This is largely a result of "over hyping". If too many troubles or delays occur, then there may be a backlash against ATM.
- Poor flow control can cause serious lost data problems in LANs. The ATM Forum has selected the rate-based paradigm, but the details are still being hammered out. Interoperating implementations from multiple vendors may be some time in coming.
- The loss or corruption of a single 53-byte ATM cell may result in the retransmission of a much larger entity, e.g., 16K byte packet.
- The early ATM equipment for LAN usage supports only Permanent Virtual Circuits (PVCs). This requires bandwidth to be dedicated on the basis of "might be needed sometime". Support for Switched Virtual Circuits (SVCs) is being developed, but is not available in early equipment.
- Setting up an SVC may take a fairly long time, e.g., milliseconds.
- The common speeds supported today with ATM are 100 Mbit/s and 155 Mbit/s. There is some equipment supporting 622 Mbit/s starting to appear, but higher speeds have yet to be tapped. The next higher speed used in the wide area networks will be 2.4 Gbit/s; it remains to be seen if this speed, or an intermediate speed, will be used in LANs. Hence, true gigabit speeds are

not available now, and probably not in the near future.

- Splitting large packets of data, e.g., 1M Byte, into many 53-byte cells for transmission seems intuitively wrong. Experience has shown that the fewer times you "touch" the data, i.e., the less the overhead, the faster things run.

Summary

Local area network speeds are increasing to keep up with the new generation computing equipment, and gigabit per second speeds are becoming a reality. Switches are replacing shared media at the higher speeds, and fiber optics are changing the error characteristics. Standards are becoming more important with few customers willing to invest in proprietary solutions.

HIPPI is the current interface of choice for high-speed LANs, but it is being challenged by Fiber Channel and ATM. There are advantages and shortcomings for each of these interfaces and a potential user will need to examine their requirements carefully in order to select the most appropriate technology. It is unlikely that a single technology will be the best for all applications.

ACKNOWLEDGMENTS

The Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36. The author's work was performed under auspices of the U.S. Department of Energy.

REFERENCES

1. D. Tolmie, J. Renwick, "HIPPI: Simplicity Yields Success," *IEEE Network - The Magizine of Computer Communications*, January 1993, pp 28-32.

2. D. Tolmie, "High Performance Parallel Interface (HIPPI)," in *High Performance Networks – Technology and Protocols*, A.. Tantawy, ed., Kluwer Academic Publishers, Norwell, Massachusetts, 1994, pp. 131-156.

3. Approved standards are available from the American National Standards Institute, Attn: Customer Service, 11 West 42nd Street, New York, NY 10036, (212) 642-4900, or from Global Engineering, 15 Inverness Way East, Englewood, CO 80112-5704, (800) 854-7179. The HIPPI documents are also available via anonymous FTP from the X3T11/hippi directory of ftp.network.com.

4. M. Sachs, "Fibre Channel," in *High Performance Networks – Technology and Protocols*, A.. Tantawy, ed., Kluwer Academic Publishers, Norwell, Massachusetts, 1994, pp. 109-129.

5. R.Handel and M.Huber, *Integrated Broadband Networks – An Introduction to ATM-Based Networks*, Addison-Wesley, Wokingham, England, 1991.

6. M. Zeug, "Broadband Integrated Services Digital Network (B-ISDN) Standards," in *High Performance Networks – Technology and Protocols*, A.. Tantawy, ed., Kluwer Academic Publishers, Norwell, Massachusetts, 1994, pp. 183-204.

7. G. Shenoda, "Synchronous Optical Network - SONET," in *High Performance Networks – Technology and Protocols*, A.. Tantawy, ed., Kluwer Academic Publishers, Norwell, Massachusetts, 1994, pp. 205-228.

Don Tolmie joined the Los Alamos National Laboratory in 1959 as a Technical Staff Member, and has been involved with networking of supercomputers since 1970. His current task is defining the next generation computer network to support higher speeds and visualization, and working with vendors to provide the appropriate products. He has been involved in the computer interface standards activities for over 13 years, and for five years chaired ANSI Task Group X3T9.3, responsible for the High-Performance Parallel Interface (HIPPI), Intelligent Peripheral Interface (IPI), and Fibre Channel (FC). Tolmie received a BSEE degree from New Mexico State University in 1959 and an MSEE degree from University of California, Berkeley, in 1961. He can be reached at the Los Alamos National Laboratory, MS-B255, Los Alamos, New Mexico, 87545; Internet, det@lanl.gov.