5 ATM

ATM is specified for OC-3 and OC-12. Particularly at OC-12, there is a very substantial throughput loss due to overhead from cell headers (53 byte cells), giving an effective rate of approximately 450 Mbps - a long way short of 1 Gbps. Several critical standards, including P-NNI for dynamic routing of cells between different manufacturers' switches, and Available Bit Rate (ABR) support, are not yet complete. in particular, ABR causes existing ATM hardware to be obsolete. At the InterOp Engineering conference in March'95, Raj Jain [8] discussed the possibility that ATM may turn out to be like OSI - an idea about which there was great enthusiasm, but little knowledge, and for whose application niche there turned out to be a simpler, more effective alternative.

In particular, there is 100 Mbps Ethernet, which is substantially cheaper, perceived as simpler, and has almost as much throughput when used with a switch as 155 Mbps ATM for local area connections. The 48 byte cell size (with 5 bytes of header) was chosen for support of 64 kbps voice channels. The Congestion Loss Priority bit to be set to indicate which cells could be dropped is fine for a voice circuit - but if you have to drop one cell in a data stream, then you had better drop all the cells from that frame, since the higher level protocols will detect the error and retransmit large quantities of data. HIPPI and Fibre Channel have large, variable sized frames, much better suited to data transmission. Video transmission also creates very large frames of data compared to the 53 byte cell size.

ATM is being developed both by the ITU and the ATM Forum - implementors have to incorporate the features proposed by both. The ATM Forum has at least 500 interested members, all of whom have an interest in a particular feature. ATM is being asked to support a much wider range of functionality than has been attempted before - local networking, wide area networking, private networks, public networks, voice, video, data - none of this allows it to be simple. Hence ATM switches and host interface cards have to be complex. In 1990, there was widespread enthusiasm for (although very few implementations of) OSI CLNS. It was a mandatory requirement for government purchases in the USA and the UK. It was too complex, and too slow, so TCP/IP replaced it.

Hence my suggestion that there are several other protocols - HIPPI, Fibre Channel, PPP over Sonet/SDH- better suited to the support of TCP/IP over gigabit networks than is ATM.

References

[1]J. Renwick, A Nicholson "IP and ARP on HIPPI" RFC 1374

[2] V. Jacobson, R. Braden, D. Borman " TCP Extensions for High Performance" RFC 1323

[3] Fibre Channel - Physical and Signalling Interface (FC-PH), Revision 4.3, X3T11/Project 755D/Rev 4.3

[4] P Walford, E Frymoyer "Providing Interoperable High Speed Connections with Fibre Channel Technology" Connexions, Vol 9, No 3, pp. 24-34, March 1995

[5]Y. Rekhter "IP and ARP on Fibre Channel", Internet Draft, 1995

[6] A Aldridge, CG Omidyar "Introduction to SDH/SONET" IEEE Comm Mag Vol 31 No 9 pp. 30 - 33, Sept.'93

[7] W Simpson "PPP encapsulation over SONET/SDH" RFC 1619

[8] R. Jain, "ATM Networking: Issues and Challenges Ahead" Engineer Conference E10 InterOp, March'95

To find out more

Read the comp.protocols.cell-relay Usenet newsgroup.

Read C Partridge "Gigabit Networking", Addison-Wesley ISBN 0-201-56 333-9

Author Information

Anne CM Wilson is the Principal Consultant for Chernikeeff Networks. Chernikeeff does not manufacture data communications hardware. It buys it in large quantities then designs, installs and supports networks with it for commercial and academic customers with large, high volume network requirements.

3 Fibre Channel

This is an ANSI standard X3T11[3]. For a detailed discussion of all of the possible profiles, topologies, traffic flow models and error recovery procedures, refer to [4].

When used for data networking (rather than as a mass storage interface) Fibre Channel is either used point-to-point between pairs of hosts, each having an interface card, or the host interface cards are connected to one or more Fibre Channel switches, for any to any connections. At the end of April, Fibre Channel cards and switches are available for various workstations including IBM RS6000s and HP 700s at 266 Mbps. Ancor have announced jointly with Sun host interface cards and switches operating at 1 Gbps which are expected to be available in June'95.

Unlike HIPPI, Fibre Channel supports both connection oriented (Class 1) and connectionless (Class 2) operation, and either may be used for TCP/IP traffic. The maximum MTU size is 62580 bytes. The connection oriented model is very like that of HIPPI, being a dedicated path, end to end, through the switching fabric, with very short (10 - 40 υsecs) setup and tear down times. Fibre Channel addresses are at least 48 bits long, and may be as long as 60 bits. Connectionless operation allows for acknowledgment between switch ports (rather than end system ports), and for flow control back to the end system. Again, RFC 1323 [2] is relevant. There is an Internet draft, "IP and ARP on Fibre Channel" [5].

Ancor make a router which can have several Fibre Channel interfaces, ethernet interfaces, and an FDDI interface. None of the high-end router manufacturers supports Fibre Channel on their routers at the time of writing.

3.1 Implementations

A customer was able to demonstrate data transfer rates of 128 Mbps using TCP/IP over Fibre Channel between two IBM RS6000s over quarter speed (266 Mbps) Fibre Channel interfaces and switches. The bottleneck at this point was the disc access rate of the RS6000s.

4 PPP on Sonet/SDH

Sonet (Synchronous Optical Network,USA) and SDH (Synchronous Digital Hierarchy, Europe) are being widely installed and used by the large telcos as the replacement for PDH (Plesiochronous Digital Hierarch) signalling and control of their large Fibre trunk networks. There is a hierarchy of specified speeds, of which OC-3/STM-1 (155.52 Mbps) is the lowest CCITT designation. There is also OC-1 (51.84 Mbps) designated for Sonet only. Sonet/SDH provides the potential for high reliability, given reasonable design, since it specifies counter-rotating rings in the manner of FDDI. It also specifies operations and maintenance parameters, for supervision and protection switching through multivendor networks. An introduction to SDH/SONET describing how PDH is carried in it can be found in [6].

Table 1: Sonet/SDH signal hierarchy

ANSI	Optical signal	CCITT	Data rate (Mbps)	Payload rate (Mbps)
STS-1	OC-1		51.84	50.112
STS-3	OC-3	STM-1	155.52	150.336
STS-9	OC-9	STM-3	466.56	451.008
STS-12	OC-12	STM-4	622.08	601.344
STS-18	OC-18	STM-6	933.12	902.016
STS-24	OC-24	STM-8	1244.16	1202.688
STS-36	OC-36	STM-12	1866.24	1804.032
STS-48	OC-48	STM-16	2488.32	2405.376
STS-96	OC-96	STM-32	4976.64	4810.176

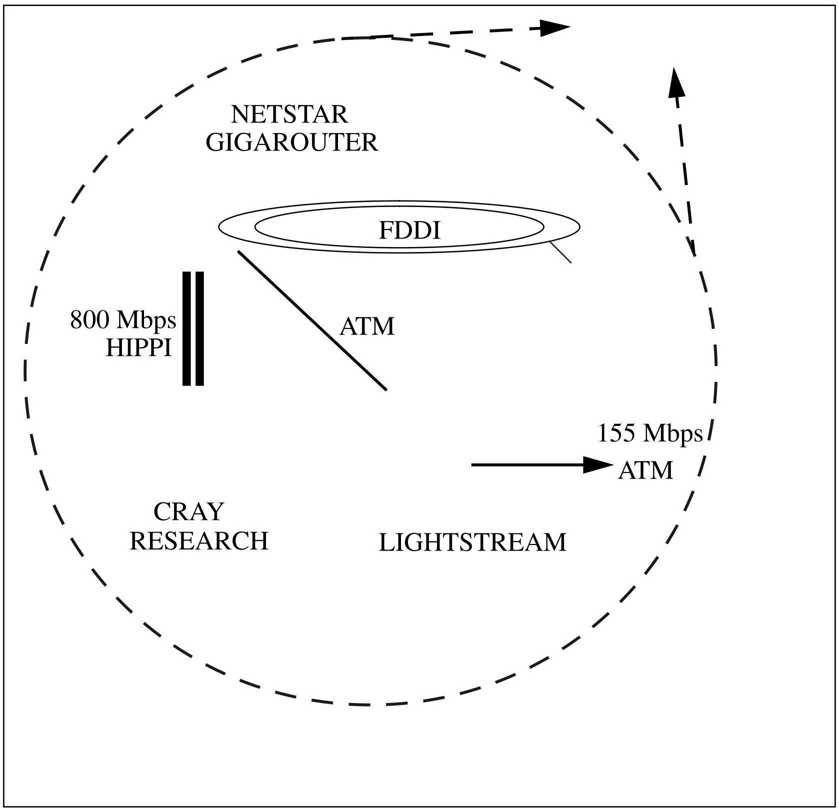
RFC 1619 [7] specifies PPP encapsulation over SDH/Sonet. The interesting speeds here are OC-12, OC-24, and above. Looking at the comparison between the data rate and the payload rate it is obvious that before any data framing is introduced, that less than the full advertised bandwidth is available. It is proposed that running PPP directly over Sonet/SDH would impose less overhead than running ATM over Sonet/SDH, which is the only scheme being widely implemented for wide area networking at OC-3. It would be a point to point implementation between routers, similar to the style in which most of the existing large high bandwidth TCP/IP networks in service have been constructed using T1 or T3 circuits. The routing functionality which is required would be carried out using the existing interior and exterior routing protocols. Features like robust multicast support and mobility are already available, interoperable between manufacturers, and being further developed. Unlike ATM, TCP/IP is very widely used and understood in its characteristics both for local area network operation and for very large international public and private networks.

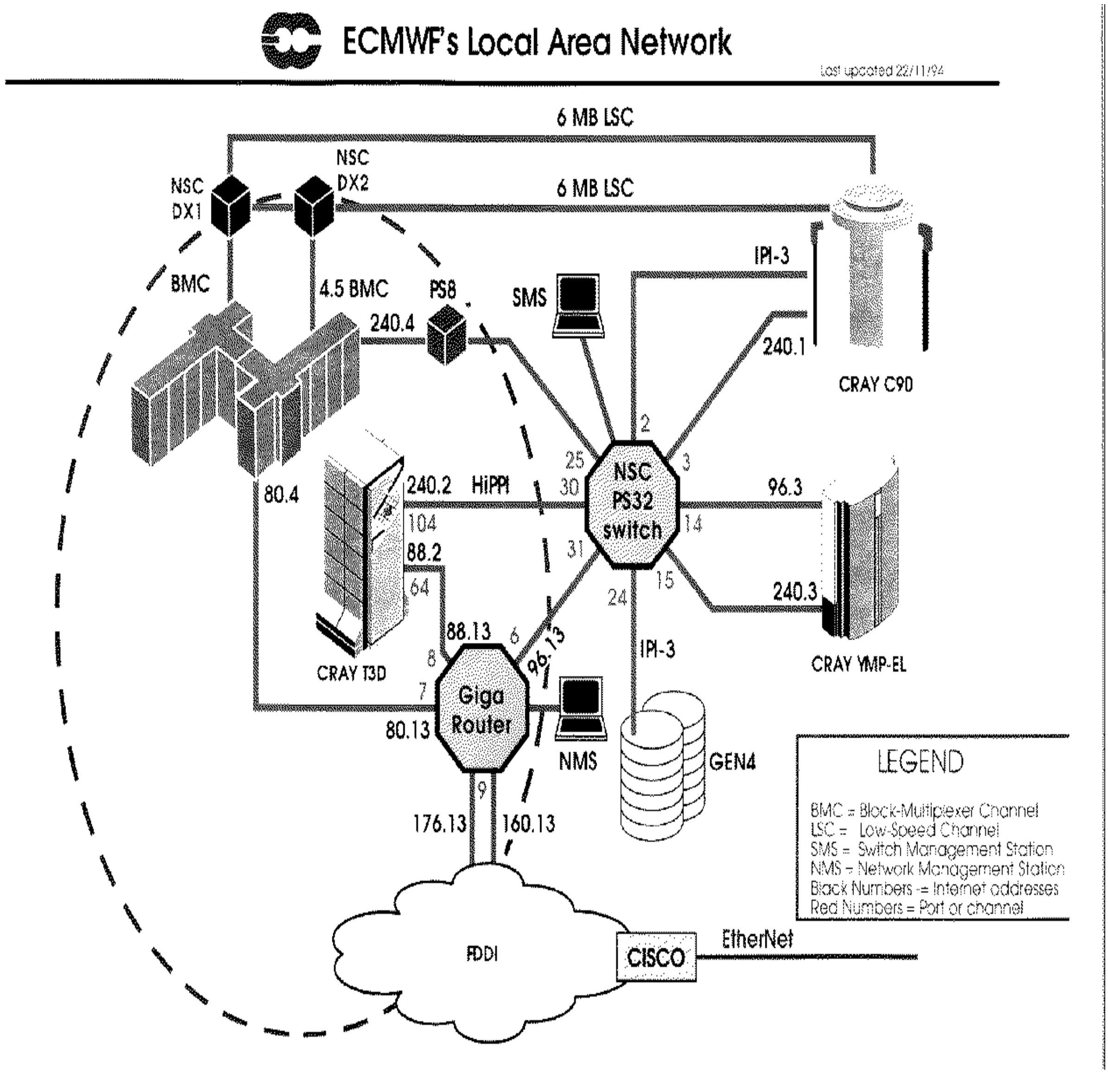
Given the PPP encapsulation, and a thorough knowledge of the existing Sonet/SDH chip sets, the other requirement is for a router capable of supporting, say 6 OC-12 or 4 OC-24 streams. For campus networks and MANs where unstructured fibre is available (in at least five cities in the UK) this would avoid the requirement for ATM. None of the high-end router manufacturers supports this on their routers at the time of writing.

The vBNS (very high-speed Backbone Network Service) supercomputer network implementation was announced on April 25th (just as this paper was being written). Netstar routers with HIPPI, FDDI and ATM interfaces are being used by MCI for site supercomputer connections to Lightstream 2020 ATM switches.

vBNS (very high-speed Backbone Network Service)

Detailed topology at each site




ECMWF's Local Area Network
Last updated 22/11/94
6 MB LSC
6 MB LSC
NSC DX1
NSC DX2
BMC
4.5 BMC
PS8
240.4
SMS
IPI-3
240.1
CRAY C90
2
25
3
NSC PS32 switch
30
240.2
HiPPI
96.3
14
104
88.2
31
15
240.3
64
24
80.4
88.13
6
96.13
IPI-3
CRAY T3D
8
7
Giga Router
80.13
NMS
GEN4
CRAY YMP-EL
9
176.13
160.13
FDDI
CISCO
EtherNet
LEGEND
BMC = Block-Multiplexer Channel
LSC = Low-Speed Channel
SMS = Switch Management Station
NMS = Network Management Station
Black Numbers = Internet addresses
Red Numbers = Port or channel

TCP/IP for gigabit networks

Anne CM Wilson <awilson@chernikeeff.co.uk>

Abstract

This paper discusses HIPPI, Fibre Channel, and PPP over Sonet/SDH as potential technologies for carrying data using the TCP/IP protocol at gigabit speeds. It explains why ATM is not a suitable candidate as a high speed data networking protocol.

1 Introduction

Many sessions of this conference are about using the Internet - about intellectually or socially or commercially interesting applications for this great mesh of connected computers and wires and routers and switches. Yakov's paper yesterday was part of a session discussing the routing issues - how to determine which path(s) data from applications is going to take when passing over the network. This paper is about some of the possible framing technologies which could be used to carry the TCP/IP protocol at much faster speeds than are currently in widespread use. One of the technologies, HIPPI, has been in use in the supercomputing arena for several years, but only recently have there been TCP/IP routers supporting it. Another, Fibre Channel, is only just becoming available at 1 Gigabit per second speeds. The third, PPP over Sonet, only exists as a specification. ATM is not a gigabit technology, and there are reasons for pessimism about its use as such for transporting data.

2 HIPPI (High Performance Parallel Interface)

HIPPI was designed as a "firehose" for data transfer at 800 Mbps or 1600 Mbps. The standard, and many implementations, use 50 pair, twisted-pair cables. For a full-duplex circuit, two of these connections are required. There is an implementors' agreement describing Serial-HIPPI, which allows for extension from the HIPPI interface over fibre for distances of up to 10 Km. The network running at the Supercomputing '94 exhibition in Washington DC last November included several examples of Serial-HIPPI extenders. Integration of the Serial-HIPPI specification directly onto interface boards will enormously ease the physical installation of HIPPI.

HIPPI is a point-to-point interface, so connections between more than two HIPPI devices are made using a crossbar switch or a router. Host interfaces are available for workstations from DEC, HP, SGI, and Sun as well as for Cray Research, Convex and IBM mainframes.

RFC 1374 [1] describes TCP/IP over HIPPI. The maximum IP packet size is about 64K bytes. The logical address space is 12 bits for any one HIPPI switched domain. After transmission across some other media using TCP/IP, the 12 bit address space can be reused. RFC 1323 [2], which describes window scaling, round trip timing, and wrapped sequence protection modifications to TCP for high speed operation is relevant. Since HIPPI is strictly connection oriented, and allows for data exchange between only one source destination pair at any one time, hosts use very short connection times, enough for a very small number of (possibly very large) packets. The setup times can be kept within the 20 υsec range.

2.1 Implementations

The European Centre for Medium Range Weather Forecasting in Reading, England, have a NetStar HIPPI router installed which allows for TCP/IP routing between several Cray Research machines and an IBM ES/9000. They have achieved data transfer rates of 57 Mbytes per second (about 450 Mbps), limited not by the router but by the available bandwidth on the source and destination machines. A modified form of ttcp was used for testing.

ECMWF diagram on page 2