

Fermi National Accelerator Laboratory

FERMILAB-Conf-92/288

A Comparison of High-Speed Links, Their Commercial Support and Ongoing R&D Activities

H.L. Gonzalez, E. Barsotti

*Fermi National Accelerator Laboratory
P.O. Box 500, Batavia, Illinois 60510*

S. Zimmermann

*FNAL/Universidade Federal do Rio Grande do Sul
Brazil*

M. Nomachi, O. Sasaki

*KEK, National Laboratory for High Energy Physics
Japan*

October 1992

Presented at *IEEE Nuclear Science Symposium and Medical Imaging Conference*

Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

A Comparison of High-Speed Links, Their Commercial Support & Ongoing R&D Activities

H.L. Gonzalez, E. Barsotti, Fermi National Accelerator Laboratory (FNAL), U.S.A *

S. Zimmermann, FNAL/Universidade Federal do Rio Grande do Sul, Brazil **

M. Nomachi, O. Sasaki KEK, National Laboratory for High Energy Physic, Japan ***

I. INTRODUCTION

Technological advances and a demanding market have forced the development of higher bandwidth communication standards for networks, data links and busses. Most of these emerging standards are gathering enough momentum that their widespread availability and lower prices are anticipated. The hardware and software that support the physical media for most of these links is currently available, allowing the user community to implement fairly high-bandwidth data links and networks with commercial components. Also, switches needed to support these networks are available or being developed.

The commercial support of high-speed data links, networks and switching fabrics provides a powerful base for the implementation of high-bandwidth data acquisition systems. A large data acquisition system like the one for the Solenoidal Detector Collaboration (SDC) at the SSC can benefit from links and networks that support an integrated systems engineering approach for initialization, downloading, diagnostics, monitoring, hardware integration and event data readout. The issue that our current work addresses is the possibility of having a channel/network that satisfies the requirements of an integrated data acquisition system.

In this paper we present a brief description of high-speed communication links and protocols that we consider of interest for high energy physics data acquisition systems. The protocols to be covered are High Performance Parallel Interface (HIPPI), Serial HIPPI, Fibre Channel (FC) and Scalable Coherent Interface (SCI). In addition, the initial work required to implement an SDC-like data acquisition system is described.

II. HIGH PERFORMANCE PARALLEL INTERFACE

The HIPPI standard consist of five standards that address primarily the lowest two layers of the ISO model [4,5]. The HIPPI-PH is an approved standard that describes the physical layer. The Framing Protocol draft, HIPPI-FP, defines a standard framing protocol to facilitate the development of computer communication protocols and how to access HIPPI. The Switch Control draft, HIPPI-SC, defines the methods to control a HIPPI physical layer switch. The Link Encapsulation draft, HIPPI-LE, defines the method to transport packets conforming to the IEEE 802.2 Logical Link Control over HIPPI. Lastly, the HIPPI-IPI draft defines the means for an IPI

peripheral device to interface to HIPPI. A brief description of HIPPI-PH is given on the next two paragraphs.

HIPPI-PH is an ANSI standard (ANSI X3.183) that specifies the physical layer for HIPPI [4]. The driving philosophy of the HIPPI standardization was to provide a high performance point to point communication link. HIPPI supports transfer rates of 800 and 1600 Megabits/s over twisted pair cable, with data paths of 32 and 64 bits, respectively. To support duplex communications, two HIPPI links are required.

The HIPPI protocol is simple with synchronous data transfers. In the HIPPI protocol, the source requests a connection to a destination, if accepted, one or more packets can be transferred. A packet is made up of one or more burst and a burst can be from 1 to 256 words. Flow control is provided by the ready signal supplied by the destination for every burst that it can accept. Data integrity checks are provided by byte parity for each word transmitted and a length/longitudinal redundancy checkword for every burst.

Table I shows some of the companies supplying HIPPI hardware. There are many more HIPPI hardware products available or under development. Among these are I/O interfaces for mainframes and workstations, storage devices, test equipment and frame buffers.

Product	Description	Company
Chipsets	HIPPI Protocol ICs	BIT, AMCC
HIPPI-VME	VME Duplex HIPPI	CHI Systems
RIO OPT8260(4,5)	Adapter for VME RIO 8260	CES
Ultra's HIPPI Adapter	UltraNet Switch Adapter With HIPPI I/O	Ultra Network Technologies

Table 1: Companies Providing HIPPI Hardware

A. HIPPI Switch

The HIPPI-SC addresses the issue of physical HIPPI switches. Commercial switches are available from Network System Corporation (i.e. PS8 and PS32) and from Input Output System Corporation (i.e HIPPI 4x4). The PS8 switch has been used on a CERN experiment [9].

B. Serial HIPPI

An ad-hoc serial HIPPI implementation group has developed a point to point serial transmission protocol that conforms with HIPPI-PH. The specification is based on a 1.2 Gigabit/second (GB/s) link using Hewlett-Packard's 20B/24B encoding/decoding scheme. A single fiber link supports the 800 Mb/s version, two links support the 1600 Mb/s. Hewlett-Packard (HP) supplies a transmitter and receiver chip set

* Work Performed under the auspice of the U.S.A DOE.

** Work partially supported by CNPq, Brazil.

*** Work performed under KEK funding.

(HDMP-1002 and HDMP-1004) that meets the serial HIPPI specifications. Commercial support of serial HIPPI includes the BIT Technologies chip set (B3030, B3031 and B2040) that interfaces to the parallel HIPPI ICs and the HP chip set.

III. FIBRE CHANNEL (FC)

The Fibre Channel standard is being developed by the ANSI task group X3T9.3. A great deal of the current work has gone into the physical interface and version 3.0 of the document. The FC standard is organized in five levels: FC-0, FC-1, FC-2, FC-3 and FC-4. For this brief presentation, we will limit the scope to the FC0 to FC3 levels. The first three levels are included in the FC-PH document which covers the physical and signaling interface.

FC-0 specifies the physical link in terms of the transmission media and the transmitter/receiver electronics. The specifications of this layer is broadly defined, including transmission rates of 133, 266, 531 and 1062 Megabaud, with optical and electrical implementations, giving a broad range of cost/performance possibilities.

FC-1 is the transmission code which uses the 8B/10B to bound the maximum run length of the code, maintain DC balance, provide word alignment and some error checking.

FC-2 is the transport mechanism of FC. The physical model of FC is based on N-ports, with inbound and outbound serial links. A node can have multiple N-ports. The node to node connection can be point-to-point or through a fabric (switch). To perform its functionality, FC-2 uses frames, sequences and exchanges. Frames are the basic unit of communication and are classified as data frames or link control frames. A sequence is a set of related data frames transmitted unidirectionally with response frames if applicable. An exchange is built from concurrent sequences between two N-ports.

FC-2 specifies three classes of services. Class 1 establishes a dedicated connection, class 2 is a connectionless service with acknowledge and class 3 is a connectionless service without acknowledge. An intermix class is defined which allows the transmission of class 2 or class 3 packets during a Class 1 connection.

A list of some of the companies that have announced FC products is provided in Table 2.

A. Fibre Channel Switch

The standardization process of FC includes a level FC-F for the specification of the requirements imposed on a fabric that supports the FC standard. We have knowledge of two companies developing FC switches, Ancor Communications and Canstar. Ancor's switching network is a space/time division switch and is scalable to thousands of inputs/outputs (labeled as F-ports by the FC standard).

IV. SCALABLE COHERENT INTERFACE (SCI)

SCI is an IEEE standard [1], IEEE1596, that provides point to point communication at 1000 Megabytes/sec (MB/s) over a sixteen bit differential data path or 1000 Mb/s over coaxial or

fiber optics cable. The serial version of SCI is supported by the HP chip set used on serial HIPPI.

<u>Product</u>	<u>Description</u>	<u>Company</u>
S2030/S2031	TX/RX FC Transmission & Signaling protocol	AMCC
S2034	Open Fibre Control	AMCC
G-Taxi Chipset	Serial TX/RX, Mux/Demux with 8B/10B Enc/Dec	VITESSE
HOLC-0266 Module	TX/RX 266MBaud (FC-0)	HP
OLC266 Module	TX/RX 266MBaud (FC-0)	IBM
RCV1201-1.2/1.5	Light to Logic RX	BT&D
XMT1300 -1.2/1.5	Logic to Light TX	BT&D
HFM1400/HFM2400	RX/TX Set, 1GBaud Modules	Honeywell
VHSCI	FC V1.6 Protocol ASIC	ANCOR Commun. Inc
S6001	FC Evaluation Kit	AMCC
FC VHSCI Kit	Evaluation & Develop. FC Standard Interface Hard.	ANCOR Commun. Inc

Table 2: Fibre Channel Commercial Products

The unit of communication is the packet which includes an address, a command and data as needed. To improve the use of the media, SCI uses split transactions where the request for service is separated from the response. From the user point of view, all SCI transactions look like read/write operations. As part of the specification, SCI supports cache coherency among the memory distributed on the nodes, a feature that is very attractive for multiprocessor applications and event building.

Each node on an SCI network implements at least two link ports, an outgoing one and an incoming one. In general, ports on SCI are connected in a ring architecture. To support a ring architecture, buffering of incoming packets is required, such that a packet passing by (addressed to another port) can be stored while a packet from the node itself is being transmitted. A high performance SCI system will use a switch to route packages from source to destination.

The SCI protocol has been implemented in VLSI by Dolphin Technologies in the Netherlands. A data acquisition system for LHC at CERN using SCI has been proposed and simulated [13].

V. LINKS COMPARISON

A detailed comparison between HIPPI, FC and SCI is not easily done, given the scope of each of these standards. To simplify this task we provide Table 3.

VI. FC HARDWARE DEVELOPMENT AND TESTS

The primary goal of our current research and development is to test and experiment with commercial hardware and software to decide its feasibility for the SDC data acquisition system. We will perform tests with a high-speed switch designed at KEK [14] and hopefully a FC commercial switch. Based on the industry's interest in FC, we believe that wide support for FC will be available in two to three years. As part

of this work we are interested in performance parameters (i.e. reliability, throughput, ect) that are useful for system simulations.

	HIPPI	FC	SCI
Data Rate (Mbit/sec)	800,1600	100, 200 400, 800	1000MByte/s 1000Mbit/sec
Components	Off-the-shelf	Custom VLSI	Custom VLSI
Architecture	Simplex	Duplex	Duplex (Bus)
Transmission	Parallel, Serial	Serial	Parallel, Serial
Media			
Elect. Parallel	Twisted Pair	No	Backplane, Shielded Cable.
Elect. Serial	Coax	Coax, STP	Coax
Optical (Mode)	Single	Single, Multi	Single
Serial Encoding	20B/24B	8B/10B	17B/20B
Data Classes	Datagram	Datagram Multiplex/w. Ack Connection/w. Ack	Packets RD & WR transactions
Fabric	Crossbar	Crossbar, Ring Tree	Crossbar, Ring Packet Switch
Cache Coherency	No	No	Yes

Table 3: Characteristics of HIPPI, FC and SCI

A. Fibre Channel Modules

The initial tests will be based on Ancor's VME adapter boards (VME CIM 250) with 266 Mb/s links. A SUN3/260 driver for this board is being developed at Lawrence Livermore National Laboratory (LLNL). For our application we will use a ported version of the driver to the VME167 board under VxWorks. Some work on the ported driver has been done at LBL [14].

At Fermilab we are on the initial phase of developing a FC auxiliary card for a FRC [8] using Ancor's VHSCI ASIC. The same FC protocol IC is used on the VME CIM 250.

B. Test With KEK Switch

As part of our collaboration with KEK, three Ancor VME adapter boards will be tested with the KEK high-speed switch [12]. The test setup is shown in Figure 1.

The KEK switch does not use resynchronization logic at the input or elastic buffers to keep transmitter-switch and switch-receiver synchronized at all times. The switch performance on data acquisition depends on receiver resynchronization time after a configuration change. To test this, the KEK switch configuration will be changed when the transmitter is sending synchronization patterns. The induced phase jump should cause loss of synchronization on the FC receiver.

In addition, the test setup will be used to perform basic event building functions. Parameters will be collected and included on data acquisition system simulations.

D. FC Based Data Acquisition System

The last part of our preliminary work, if funding permits, will be to use the VME CIM 250 boards, the FRC and a CTX

250/1000 Network (FC fabric) to implement a small scale data acquisition system. The hardware configuration is similar to the one shown in Figure 1, but the KEK switch is replaced by the Ancor's FC switch. Note that in this case the routing capabilities of the switch are used and the switch does not have a VME port. This system has the general features of the proposed SDC data acquisition system [11]. As before parameters will be collected and included on simulations of the data acquisition system.

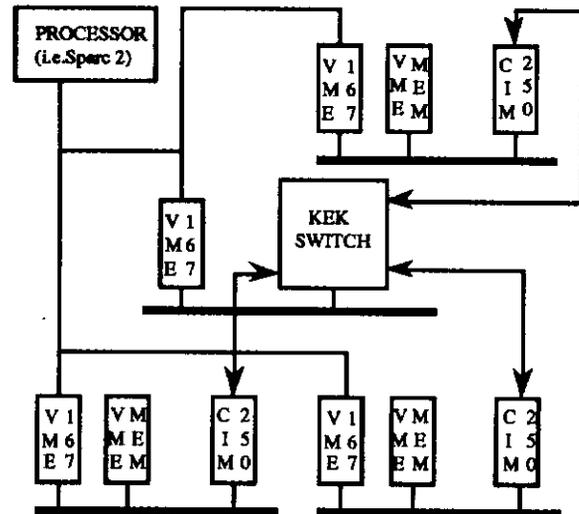


Figure 1. KEK Switch Test Setup

VII. REFERENCES

- [1] SCI-Scalable Coherence Interface, P1596/D2.0 18Nov91 Draft.
- [2] D.B. Gustavson, "The Scalable Coherent Interface and Related Standards Projects", IEEE Micro, Feb. 1992, pp. 10-22.
- [3] Fibre Channel - Physical and Signaling Interface, ANSI X3T9/92-007, Rev 2.2, Jan. 1992.
- [4] ANSI, "High Performance Parallel Interface- Mechanical, Electrical and Signaling Protocol Specification(HIPPI-PH)" Amer. Nat'l Std. X3183-1991.
- [5] ANSI Drafts for HIPPI-FP, HIPPI-SC, HIPPI-LE and HIPPI-IPI,
- [6] Serial HIPPI Specification, Rev 1.0, Serial HIPPI Implementations Group, May 1991.
- [7] Gustavson D.B., "The Scalable Coherence Interface and Related Standards Projects", IEEE Micro, Feb. 1992, pp. 10-22.
- [8] S. Zimmermann, "FASTBUS Readout Controller (FRC) Card For High Speed Data Acquisition", IEEE NSS and Medical Imaging Conference 1991, Santa Fe, N.M., Vol. 2, pp. 794-798.
- [9] A. van Praag, "HIPPI Developments for CERN Experiments, IEEE on NS, vol 39, No 4, 1992, pp. 880-885.
- [10] E. Barsotti, A. Booth and M. Bowden, "Effects of Various Event Building Techniques on Data Acquisition System Architectures, Fermilab, AIP Conference Proceedings 209, pp. 82-101, Santa Fe, NM, April 1990.
- [11] SDC Data Acquisition Technical Proposal, April 92.
- [12] O. Sasaki, M. Nomachi, T.K. Ohska and H. Fuji, "A VME Barrel Shifter for Event Reconstruction From up to 3Gbps Signal Train", IEEE NSS at Orlando, Florida, Oct. 1992.
- [13] Application of the Scalable Coherence Interface at LHC, CERN/DRDC/91-45, Oct 91.
- [14] B. Greiman, Lawrence Berkeley Laboratory, private talk, "Use of Commercial Switches In SDC Event Builder".
- [15] D. Tolmie, Los Alamos National Laboratory, private talk.