

ORGANISATION EUROPEAN POUR LA RECHERCHE NUCLEAIRE

CERN EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH

PCI - HIPPI Interface Modules

T. Anguelov, G. Antchev, G. Georgiev, S. Piperov, I. Vankov, INRNE-BAS, Sofia, Bulgaria

R. A. McLaren, A. van Praag, CERN, Geneva, Switzerland.

O. Orel, IHEP, Prodvino

D. Gillot, A. Guglielmi, Digital Equipment Corporation, Joint project office, CERN

CERN/ECP 96-14, 18 September 1996

Summary

Interface modules between PCI local bus and HIPPI are described. The modules are intended to aid the implementation of the high performance computer network. The aimed maximum throughput of the interface module is 100 MBytes/sec while sustained data transfer rate depends of the particular system performance.

Presented at the SOZOPOL-96 Workshop of Relativistic Nuclear Physics, Sozopol, Bulgaria, 30 September - 6 October

Introduction

PCI - the next generation local bus

An introduction to the HIPPI

Hardware

Achieved results

Acknowledgements

References

Introduction

The High Energy Physics (HEP) experiments require complex Data Acquisition Systems (DAQ). Very often the DAQ is a multilevel systems, allowing filtering and data processing for a huge amount of input data to an acceptable for off-line analysis output data. The size of the events, delivered from the accelerators in the HEP experiments, ranges from Megabytes to Terabytes and the event rate is up to hundred of kilohertz. This big amount of data implies fast data moving between DAQ's levels.

The major importance is the throughput of the DAQ in the upstream direction. Using standard interfaces and powerful workstations is acceptable approach to build up cost effective DAQ. The data transport between different levels in the complex DAQ uses the high bandwidth interfaces as FDDI, FASTBUS, FCS, SCI or ATM.

One of the fastest interfaces, intended to meet the requirement of the high speed data transfer is the High Performance Parallel Interface (HIPPI) [1]. Originally developed for use in the supercomputer area, now the HIPPI is widely used in other applications. An interface module from HIPPI to the recently introduced high speed local bus as Peripheral Component Interconnect (PCI) [2] will be useful to create powerful DAQ based on HIPPI.

PCI - the next generation local bus

Most of the computer companies use different I/O buses in their computers thus making the add-in modules for one machine incompatible to others. In 1992 Intel introduced the bus named Peripheral Component Interconnect. More than 168 computer companies as Apple, DIGITAL, IBM, Intel, NEC and others have adopted this bus as internal bus of their computers. Now the PCI dominates in the computer industry. Many OEMs offer PCI compatible products as video-controllers and interface modules for Ethernet, SCSI, FDDI.

The PCI is a high performance local bus architecture intended to meet the requirements of today's computer systems. In technical aspect PCI is powerful bus architecture with many advances features:

- Synchronous bus with clock frequency 33 or 66 MHz.
- 32/64 bits address/data path.
- Peak transfer rates as high as 528 MBytes/sec.
- Multiplexed address-data lines.
- Dynamic data width sizing by means of byte enables strobes.
- Double handshake mechanism.
- Sixteen type bus operations.
- Hidden (access based) central arbitration mechanism.
- Full Multi-master capability.
- Exclusive access using a locking mechanism to an accessed device.
- Cache coherency support.
- Hierarchical (tree likes) architecture based on PCI-to-PCI bridges.
- Auto configuration through a configuration space register set.
- Predictable target latency.
- Defined bus utilisation by means of a master latency timer.
- Fundamental new approach of bus driving by using of the reflected wave.
- Even parity for address/data and command/byte enables lines.
- Expansion ROM for device specific initialisation and system boot.
- JTAG/Boundary scan capabilities.

In the application aspect the PCI is moving in multiple dimensions - Figure 1.

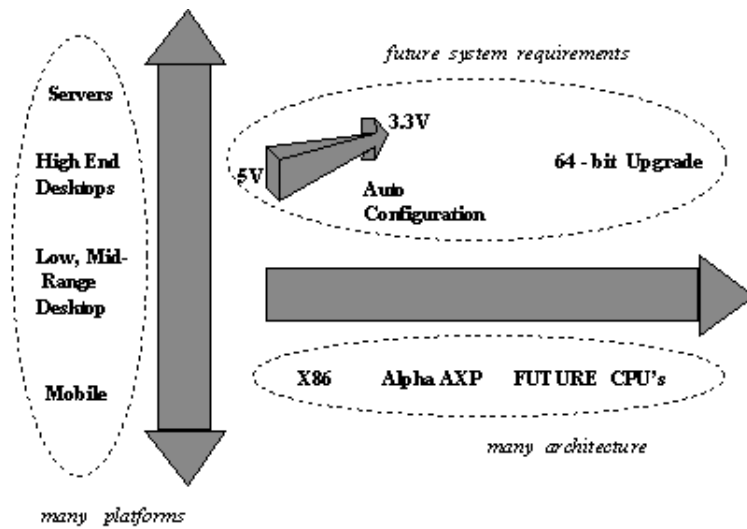


Figure 1. PCI Applications

- Universal bus for all platforms from low-end PCs to high-end workstations
- Processor independent 32 or 64 bit local bus.
- Forward and backward compatibility of 32 and 64 bit add-in boards.
- Auto configuration enabling "plug and play" add-in board insertion.
- Operation at 5V, 3.3 or both.

The PCI data transfer mechanism is burst oriented, that is a train of data words during one bus transaction is moved between the initiator module and the target module. The PCI specification does not limit the burst length, however an infinite data transfer is not possible because of:

- System memory limitations.
- Presence of other bus initiators that might arbitrate the bus.
- Software programmable and hardware implemented bus utilisation timer.
- Type of the bus operation.

There are two generic types of bus master devices. A device that very often needs a fast access to the bus and transfer small data block is called "latency sensitive master". Device that needs the bus access occasionally but transfer large block of data is called "high throughput master". PCI tolerates both types of devices - latency sensitive devices and high throughput by means of hardware implemented latency timer and software controllable type of bus operation. This makes a PCI ideal solution when different type of bus mastering interface modules are connected to the system. Figure 2

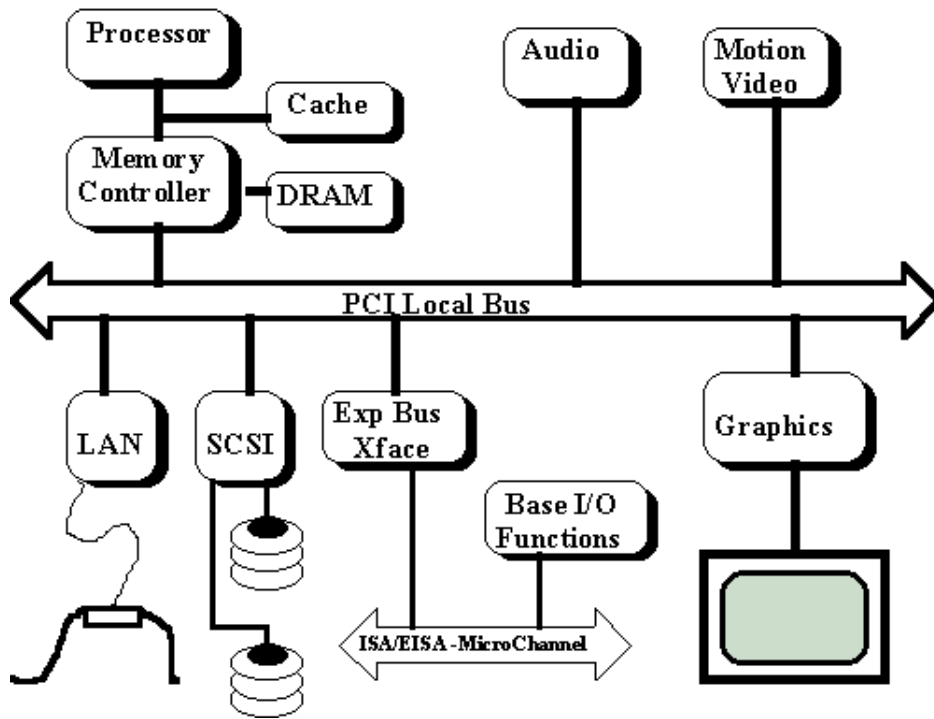


Figure 2. PCI based system configuration

A PCI based system may dynamically tune itself according to some conflicting requirements as the fast access to the system devices and the large data transfer to the system memory.

An introduction to the HIPPI

The HIPPI standardisation was started in 1987 at American National Standards Institute (ANSI). The HIPPI standard includes HIPPI-PH (Physical layer), HIPPI-SC (Switch Control), HIPPI-FP (Framing Protocol) and HIPPI-LE (Link Encapsulation). All of them are ANSI Standards.

The HIPPI-PH specifies mechanical, electrical and signalling parameters of the HIPPI over 25 meters cable of the twisted pairs carrying differential ECL level signals. The HIPPI is unidirectional, synchronous channel with fixed clock frequency of 25 MHz. The theoretical maximum throughput of the HIPPI is 100 MBytes/sec for 32 bit and 200 MBytes/sec for 64 bit version. Separate links can be interconnected through HIPPI switches in HIPPI networks having a very high throughput. The HIPPI switches are commercially available from NCS, Essential Communications and Gigabit Labs..

The HIPPI is a point to point link from a data transmitting device "HIPPI SOURCE" to a data receiving device "HIPPI DESTINATION" - Figure 3.

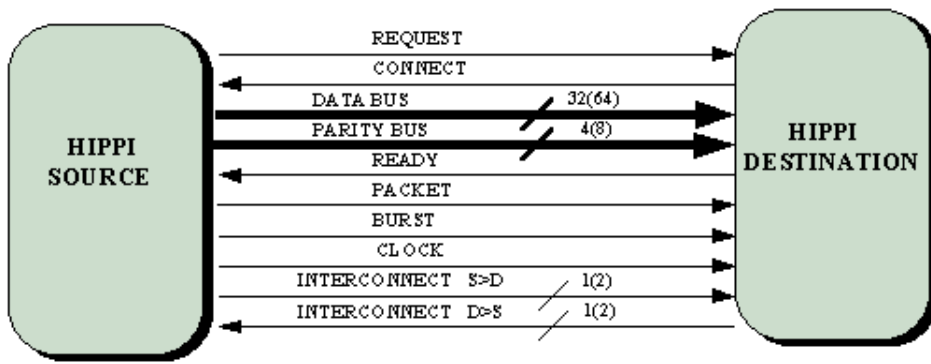


Figure 3. HIPPI Interface signal summary

The protocol is based on the relationship between the "REQUEST" and "CONNECT" and a look-ahead data flow control by the "READY" pulses. The physical framing hierarchy is: HIPPI connection, HIPPI packet and HIPPI burst - Figure 4.

The HIPPI connection consists of one or several packets. Upon requesting a connection the source asserts the I-Field, a 32 bit word, used as routing information when the connection is established through one or more HIPPI switches.

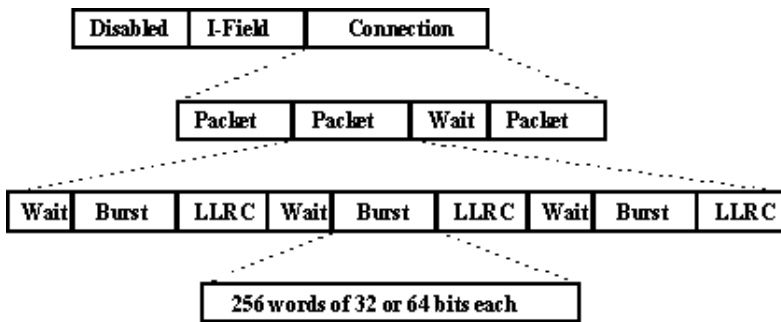


Figure 4. Physical framing hierarchy

The packet consists of one or several bursts. The HIPPI burst is a train of up to 256

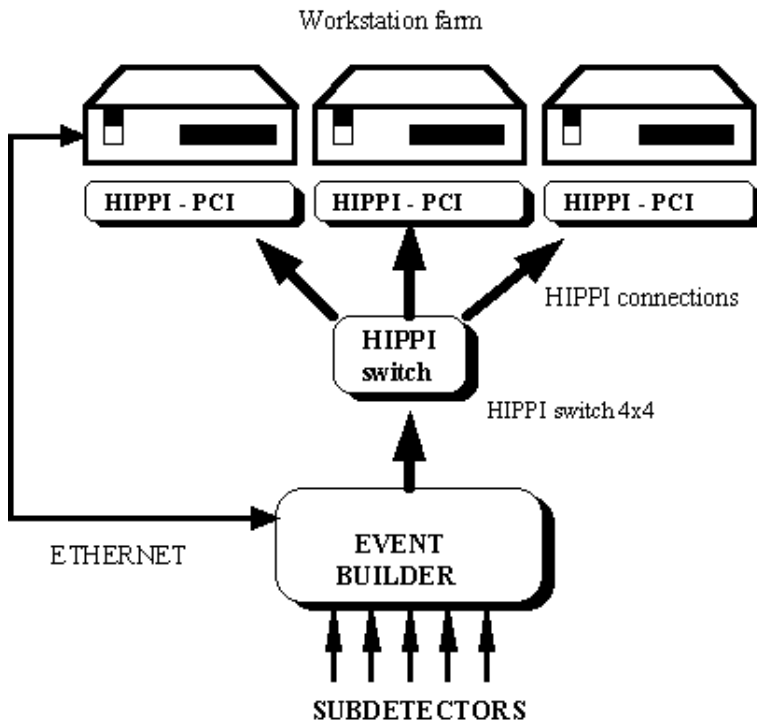


Figure 5. The second level of the DAQ for NA48 experiment

words of 32 or 64 data bits each and 4 or 8 byte parity bits. Every burst is followed by a length-longitudinal redundancy checkword (LLRC) which provides a proper data integrity for eventual error detection and correction.

Most of the DAQs for the HEP experiments need the movement of data only in one direction - from previous to the next DAQ level. In such case the use of a simplex HIPPI segment is relevant. The first application of the HIPPI based computer network is the second level of the DAQ for NA48 experiment at CERN [3]. The second level of the NA48 DAQ is an array of powerful workstations receiving data from an event builder - Figure 5.

The HIPPI switch is used as a distributor, forwarding large data blocks (approximately 180 MBytes) to the currently available workstation while other workstations are processing data, already transferred to them.

HIPPI can be used to build a local area network (LAN) with various configurations. A simple network, configured as "star" has a HIPPI switch and number of network stations connected to the switch - Figure 6.

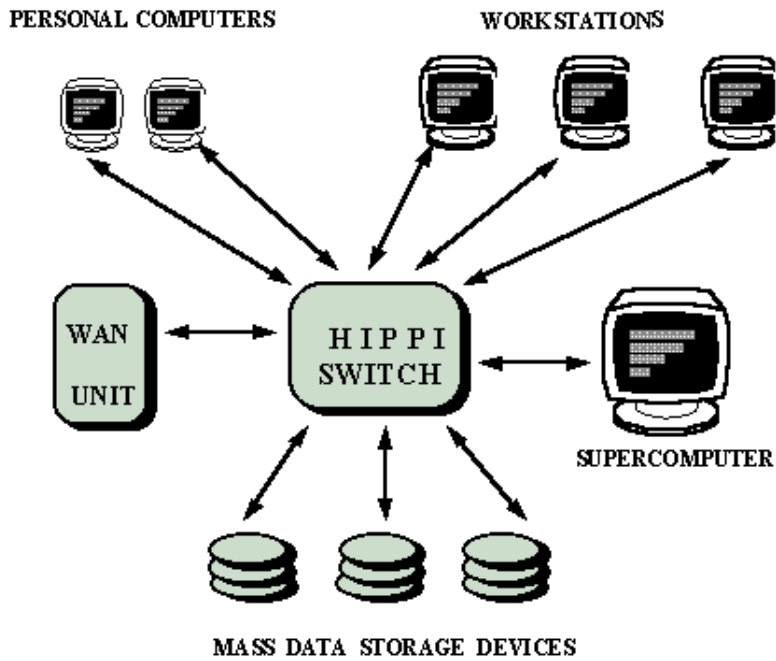


Figure 6. HIPPI based LAN

LAN applications requiring a full duplex use two HIPPI channels in both directions. The total throughput of the network is a sum of the particular throughput of all network segments.

Hardware

Two interface modules were developed - a HIPPI "Destination" and a HIPPI "Source". The "Destination" module block diagram is depicted on Figure 7. The incoming data is transferred to the system memory by means of a Direct Memory Access (DMA) transactions. The advantage of DMA is that the system is free to perform other tasks on the same time the DMA occupies the PCI bus. For large data transfer the DMA requires on-board Scatter-Gather Table Memory (SGTM). The SGTM contains the correspondence between the physical and logical addresses of the system memory buffer, reserved for the current HIPPI connection. The DMA engine automatically retrieves the physical address of the next memory page in the case of page segmented system memory. This mechanism allows large data blocks to be transferred to the system memory without the time consuming interrupt procedure.

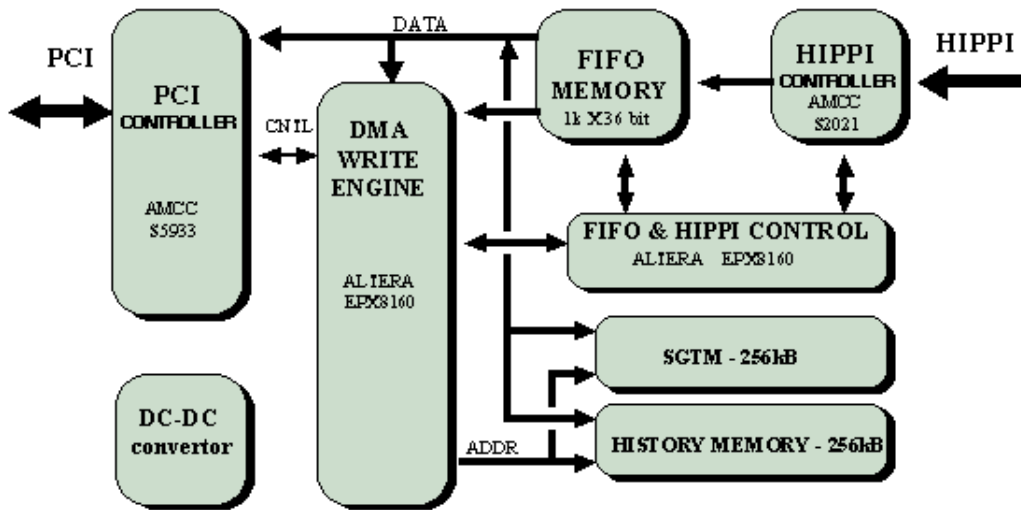


Figure 7. HIPPI "Destination" block diagram

A FIFO memory is used as an elastic buffer of data words. It is clocked by the PCI clock on the input port and by the HIPPI clock on the output port. The HIPPI protocol information and eventual HIPPI data errors are stored into the History Memory. If HIPPI data packets contain one or more parity errors, they can be located in the system memory after the end of the current HIPPI connection. This mechanism avoids the delay due to interrupt handling if the protocol information and data errors would be reported to the system by an interrupt request. The interrupt request breaking the data transfer is generated only after the HIPPI connection end or after a serious interface error. The HIPPI "Source" module block diagram is shown on Figure 8.

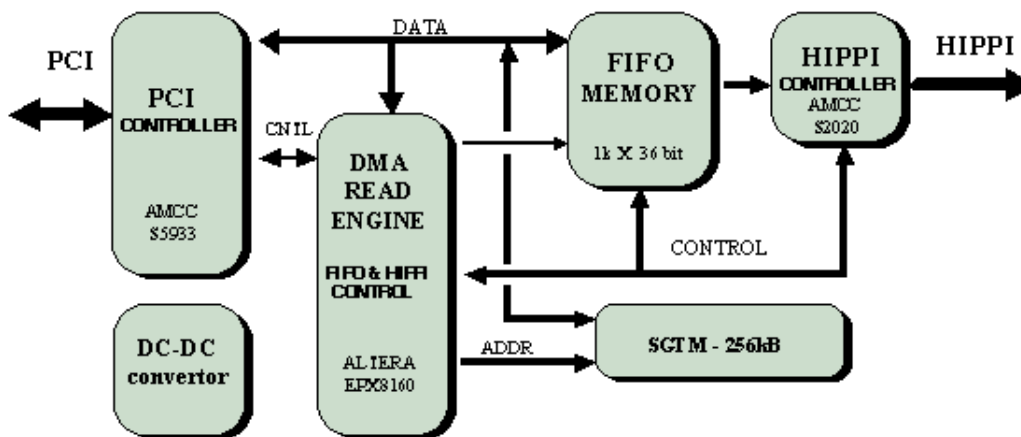


Figure 8. HIPPI "Source" block diagram

In addition to the physical addresses, the SGT M contains the HIPPI packet delimiters and connection end as well as the memory buffer size. The DMA engine moves data from the system memory to the HIPPI until the end of memory buffer is reached, or if the HIPPI connection has to be terminated. In the "Source" and in the "Destination" the AMCC device S5933 is used as PCI controller. It is a general purpose PCI controller with flexible internal structure intended to satisfy various user requirements to the PCI interface.

Achieved results

The speed measurements were carried out on Digital workstations Alpha 400, Alpha 600 and VME module AXBVME 160. The listed throughput in Table 1 was measured on the HIPPI port of the modules without taking in account the software overhead.

	DEC Alpha 600	DEC Alpha 400	AXBVME 160
"Destination"	91 MBytes/sec	72 MBytes/sec	65 MBytes/sec
"Source"	20 MBytes/sec	20 MBytes/sec	20 MBytes/sec

Acknowledgements

The Autor thanks the National Found for Scientific Research of Bulgarian Ministry of Education, Science and Technology for sponsoring this project.

References

1. PCI Local Bus Specification, Rev 2.0, April 1993, PCI Special Interest Group, Oregon, USA
2. HIPPI - Mechanical, Electrical and Signalling Protocol Specification REV 8.1, June 24, 1991
3. J-P. Matheys, et al. "Data Transfer and Distribution at 70 Mbytes/s", CERN/ECP 93-7.
Presented at IEEE RT 1993

This is one of the CERN High Speed Interconnect pages - 10 December 1996 - Arie van Praag