

Integrating High Performance VME, PCI and CPCI processors into CPU clusters for Data Acquisition and Control Systems

L. Vivolo, M. Weymann, F. H. Worm

Creative Electronic System,
70 Route du Pont-Butin,
CH 1213 Petit-Lancy/Geneva,Switzerland
email: ces@ces.ch, web: <http://www.ces.ch>

Abstract

Data acquisition and control systems using a large number of embedded VME processors have a long tradition in High Energy Physics. More recently, CPCI is in some cases considered as an alternative to VME. CES has developed a processor board architecture optimized for high-throughput deterministic bus operation, which can be used with minimal adaptation on both backplanes. The RIO3 8064 (VME) and RIO3 4065 (CPCI) are to a large extent software compatible which helps to develop software solutions almost simultaneously in both domains. Both boards couple the CPU bus directly to the backplane bus (VME or CPCI) and to two independent PCI busses. This twin-bus architecture allows to separate data flows in a similar way than VME/VSX architectures did in the past. The MFCC 844x, a PowerPC based PMC module is ideally suited to handle complex I/O protocols or to build multi-CPU clusters coupled by the carrier boards local PCI bus rather than the backplane bus. The CES PVIC allows to interconnect distant PCI segments (e.g. a VME based processor cluster and a desktop workstation) using both memory mapped access and DMA mechanisms. With its *backplane driver* CES provides an ideal tool to integrate CPUs interconnected by PCI, VME, CPCI and PVIC into a homogeneous, network-oriented environment taking full advantage of the high bandwidth and low latency features of PCI and backplane busses.

Keywords: PowerPC, VME, PCI, CPCI, PVIC

1 Introduction

The vast volume of data generated by physics detectors requires computing power beyond the reach of any existing single computer. On the other hand an ever higher bandwidth is required for accommodating the data flow between the digitizing electronics and the final data analysis and storage stages. Looking at the dataflow through a typical data acquisition system of an HEP experiment we can distinguish different requirements for DAQ processors.

- In a **front-end stage**, sub-detector data are stored temporarily. A large fraction of the data may be rejected at this stage. Accepted data are forwarded to the following stages. Critical requirements here are low latencies and high I/O performance.
- In an **Event-building stage** data arriving in parallel on a large number of input channels are grouped together into complete events. This may imply covering relatively large distances and requires very high integrated throughput. Candidates are switched networks (e.g. ATM) or high-performance bus-to-bus links like PVIC [5].
- In the **analysis stage** complete events are processed. The key requirement here is CPU power. Proposed solutions include processor farms in which a large number of processors work in parallel. Each processor works on a single event at a time, the next event arriving is given to the first free processor in the farm.

In the following we try to show how combinations of VME or CPCI processor boards [1][2], PowerPC based PMCs [4] and PVIC [5] may be used to implement the functionality required by the DAQ system of a HEP Experiment.

2 The Processor Boards

The new generation of CES real time processor boards were designed to meet the requirements outlined above.

- A powerPC CPU kernel based on G3 and G4 processors for maximum CPU power.
- Three 64-bit (1 backplane + 2 PCI) busses equipped with DMA engines for maximum I/O performance.
- Direct coupling of the backplane bus to the CPU (not via PCI) for minimal access latencies.

The following discussion centers on the VME processor, the RIO3 8064 [1], figure 1 shows its block diagram. All features not related to VME are identical in the CPCI version, the RIO3 4065 [2].

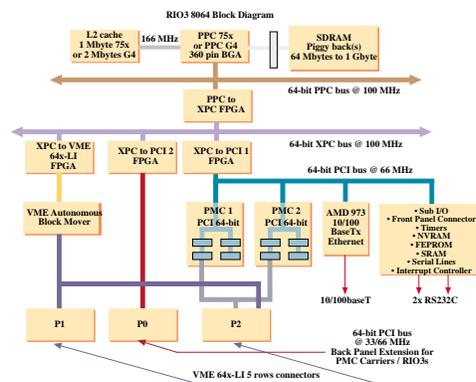


Figure 1: The RIO3 8064 block diagram.

The CPU subsystem consists of PowerPC 750 (G3) at 466 MHz or PowerPC 7400 (G4) at the maximum available speed, coupled to a 2Mbyte L2 cache running at 166 MHz and to a large system memory (up to 1 Gbyte) designed to support a bandwidth of 400 MByte/s.

Three different 64-bit wide I/O busses are coupled to the PowerPC bus through the intermediate *XPC bus* which is essentially a synchronous copy of the PowerPC bus, optimized for burst transactions.

- The **local PCI bus** connects the onboard PCI devices (10/100 Mbit/s Ethernet, Serial lines, EPROM, NVRAM, ...) and two 64-bit onboard PMC slots.
- The **PCI extension bus** provides an independent 64-bit, 33 MHz PCI bus that is routed to the RIO3s P0 connector (J3 on RIO3). Section 3 shows how bus can be used to connect additional PMCs or to interconnect RIO3s without loading the system bus.
- The **system bus (VME or CPCI)**. The RIO3 8064's VME interface follows the VME64X/LI specification, incorporating 2eSST transfers (including broadcast), CR/CSR implementation for geographical addressing and Live Insertion hardware support. This allows to build VME systems in which individual boards in a crate can be replaced without shutting down the complete system, substantially reducing the overall system-down time. A64 addressing capabilities (master/slave) allow to map all internal resources between RIO3s on the same VME backplane.

The CES proprietary bridges used for the coupling also implement DMA engines which allow to transfer data between any of the three I/O busses and system memory without loading the CPU. A sophisticated priority scheme allows high-priority DMA channels to pass between blocks of background DMA transfers, cutting down latency for high priority messages.

3 PMC carriers

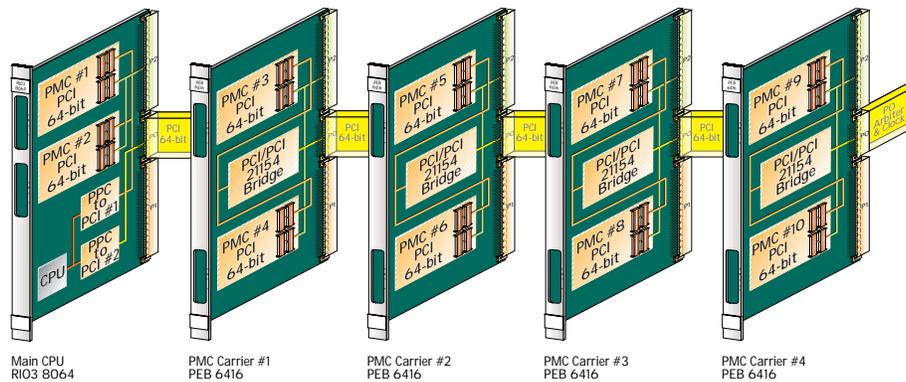


Figure 2: Controlling up to 10 PMCs by a single RIO3 8064 .

Figure 2 shows how a single RIO3 can control up to 10 PMCs. PMC carrier boards (PEB 6416) are connected to a PCI backplane implemented behind the VME P0 connector. Each carrier board can hold up to two PMCs which are isolated from the P0-PCI bus by a PCI-PCI bridge. This configuration is interesting if either the dataflow from each PMC is relatively weak, or if at least one of the PMCs on each carrier is an MFCC (s. Section 4). A possible configuration consists of a combination of one MFCC and one acquisition PMC (e.g. ATM) on each PEB carrier. The MFCC processes the acquired data locally, leaving the RIO3s CPU free for overall control and communication with other system levels. Processor farms could also be constructed by connecting several MFCCs to a single RIO3.

4 PowerPC based PMCs

A large number of I/O subsystems used in data acquisition require the implementation of complex protocols. The *Multifunction Computing Core* (MFCC 844x [4]) is a PMC that provides a PowerPC computing core together with a large *Front End FPGA* suited to implement complex application specific I/O functions. A second FPGA implements a PowerPC to PCI bridge that provides all features necessary for efficient communication with the host system. The PowerPC can be used to pre-process data already on the PMC level, giving the MFCC key role in the construction of PCI data acquisition front-end. The MFCC family includes:

- The **MFCC 8441**: PPC603ev@300MHz, 32 MByte DRAM. A *Front End Adapter* ('nose') can be designed to meet application specific connector and signal conditioning requirements. Some standard 'noses' are available (FND 6451: 64 lines TTL single-ended. DIO 6452: 32 lines differential, ...)
- The **MFCC 8442**: PPC750@366/466MHz or G4, 64 MByte DRAM, 1 MByte L2 cache. This card is optimized for computation (64-bit PPC bus). It has a Front End FPGA, but no 'nose' - Front end signals are routed to the VME P2 connector.
- The **MFCC 8443**: same as MFCC 8442, but with a 64-bit PCI interface.

5 Integrating CPU clusters - the BpNet backplane driver

The MFCC does not have console or network interfaces of its own. It is coupled to its host only via its PCI bridge which renders its memory dual-ported between PCI and CPU and provides two sets of FIFOs and a DMA controller. However, porting an operating system like LynxOS or VxWorks to the MFCC certainly implies an IP emulation across these resources. When studying how to integrate MFCCs into a multi-CPU environment running an operating system CES decided to aim at a tool with a much wider functionality. In resulting *BpNet* system is based on *channel* connections, IP emulation is just one particular application using the underlying channels. POSIX Semaphores and shared-memory mechanisms are other applications.

Channels are bi-directional links between two processes or threads. A channel is established if two processes connect to the same port which may either be an explicit address (host number, CPU number) or a symbolic label. The number of open channels in a system is basically unlimited.

A single BpNet system can connect up to 16 nodes. The nodes may reside on the same PCI bus, or on two different PCI busses linked by a *backplane* bus which can be either VME, CPCI or PVIC. An real Acquisition system using the BpNet is described in [3].

6 PVIC

The PVIC [5] allows PCI based processors to be connected into clusters of up to 15 nodes spanning up to 200 meters. While preserving the full PCI throughput for block accesses, the transparent mapping of remote PCI addresses minimizes latencies and drastically reduces the software overhead. An integrated DMA controller allows complex data transactions with minimal CPU load. Broadcast and multicast cycles, interrupt dispatching, mailboxes, a mirrored memory and global semaphores provides hardware support for efficient inter-processor communication. PCI based processor boards may be connected to each other and to PCI based desktop workstations on three different distance scales: below 1 meter (e.g. within a VME crate, flat cable, GTL+), up to 20 meters (PECL differential copper connection), and up to 200 meters (850 nm multimode fiber).

7 Conclusions

The new generation of PowerPC based VME or CPCI processors designed by CES are well suited to match the requirements of modern experiments in High Energy Physics. They offer a wide range of CPU power and, in the form of PCM extension slots, a high flexibility for I/O. PowerPC based PMCs equipped with large user programmable FPGAs are an ideal platform for compact, high performance DAQ input channels. The BpNet system facilitates the integration of complex multi-CPU systems while increasing the portability of the system to new hardware platforms. This architecture allows to cover both I/O bound and CPU bound applications with the same family of processors.

References

- 1 Creative Electronic Systems, "RIO3 8064 Data Sheet", Geneva, Jan 2000.
- 2 Creative Electronic Systems, "RIO3 4065 Data Sheet", Geneva, Jan 2000.
- 3 M. Weymann, L. Vivolo, F. H. Worm, "Modular Function Units for efficient bridging of high speed networks and control busses", ICALEPCS'99, Trieste, October 1997.
- 4 M. Weymann, "A PMC Based Computing Core", SYSCOMMS'98, CERN, Geneva, March 1998.
- 5 M. Weymann, "Linking PCI based processor platforms", CHEP'97, Berlin, April 1997.