Rev 10/2004

HTF128

Introduction to the NEW PICMG 2.18 Serial RapidIO CompactPCI Specification

Technology Focus

Hybricon Corporation • 12 Willow Road • Ayer, Massachusetts 01432
Tel: (978) 772-5422 • Fax: (978) 772-2963
Web: www.hybricon.com • E-Mail: info@hybricon.com

Executive Summary

This paper will provide an introduction and overview of the newly released PICMG 2.18 Serial RapidIO CompactPCI specification. This is the latest in the CompactPCI family of specifications; CompactPCI is based on the PCI SIG PCI Local Bus specification, but using a more rugged IEEE 1101.10/1101.11 mechanical standards, and is used for rugged, reliable systems in applications for Telecom, Military/Aerospace, Industrial markets, among others.
The PICMG 2.18 specification was approved and released by the PICMG Executive Board in June 2004. This new specification provides, for the first time, a standardized way to implement a Serial RapidIO Switch Fabric in a standard backplane without the use of cables, and using standardized backplane connector pinouts. As the fourth Switch Fabric to be implemented on CompactPCI (this was preceded by PICMG 2.16, PICMG 2.17 and PICMG 2.20), this is the latest example of the trend for system architectures to move towards serial point-to-point interconnections and Switch Fabric topologies to supplement, and eventually replace, parallel buses such as PCI in the long term.

CompactPCI background

The CompactPCI specification defines two types of cards—3U and 6U card sizes. Many applications use a 6U card size because it provides extra space for more components, extra interconnects such as H.110, and extra rear I/O pins. Each card slot has 5 backplane connectors.

Existing CompactPCI systems suffer from a number of shortcomings:

1. Existing CompactPCI backplane interconnect alternatives (PCI, H.110) support limited bandwidth.
2. The PCI bus requires active bridge cards to extend beyond 8 slots.
3. The PCI bus is a single point of failure in a High Availability system.
4. Implementing CompactPCI CPU redundancy and fail-over for High Availability systems is very complex, and there is not yet a standardized way to do this in CompactPCI.

These limitations have prompted PICMG (PCI Industrial Computer Manufacturing Group) to form the PICMG 2.18 subcommittee to investigate a better method of transferring large amounts of data between cards, with an emphasis on real time applications that require low latency. This contrasts with PICMG 2.16, which has an emphasis on packet traffic, and PICMG 2.17, which leverages PCI and H.110 transparency.

Introduction to the Serial RapidIO CompactPCI Backplane

PICMG 2.18 has the following high-level features:

- Scalable system performance: Up to 380 Gigabits/chassis (1 chassis).
- Scalable reliability: 1 or 2 fully independent fabrics (HA capable).
- Scalable cost through support of various topologies including:
  a. Single fabric using Basic Node Slots and single Fabric Slot; this can coexist with PICMG 2.16 Ethernet and the H.110 bus.
  b. Redundant dual fabric using Basic Node Slots and dual Fabric Slots; this can coexist with PICMG 2.16 Ethernet and the H.110 bus.
  c. High performance mesh topologies using Multi Segment Node Slots; this can coexist with the H.110 bus.
  d. High performance mesh topologies using Fabric Native Node Slots; this can coexist with PICMG 2.16 Ethernet Fabrics.
- "Fine" granularity P2P architecture, single point of failure (SPF) = 1 slot.

- "Virtual backplane" capable allowing for multi-chassis architectures.
- Concise pinout requires only 32 connections per Basic Node Slot, supporting 2 redundant Serial RapidIO Links per Basic Node Slot.
- Inherently hot swappable and HA due to serial switch fabric based architecture.

Key Aspects of PICMG 2.18 include:

1. Point-to-Point interconnections across the backplane.
2. Leverages emerging Serial RapidIO capabilities on semiconductor devices.
3. Supports low latency real time applications.
4. Utilizes a Switch Fabric architecture.

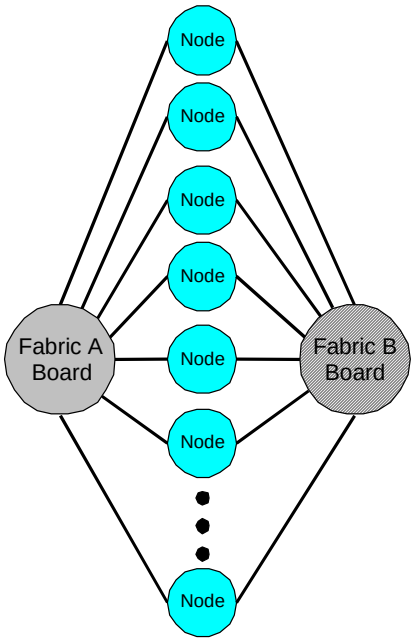
Switch Fabric Architecture

PICMG 2.18 defines a Switch Fabric Architecture, as contrasted with a Bus Architecture (e.g. PCI). A Switch Fabric is an interconnected network of switching devices. The fabric contains many input and output ports, and the switch fabric transports the data from the inputs to the outputs. This means that in order to travel from one slot to another, data passes through an active switching device.

Four new types of Slots/Cards are defined:

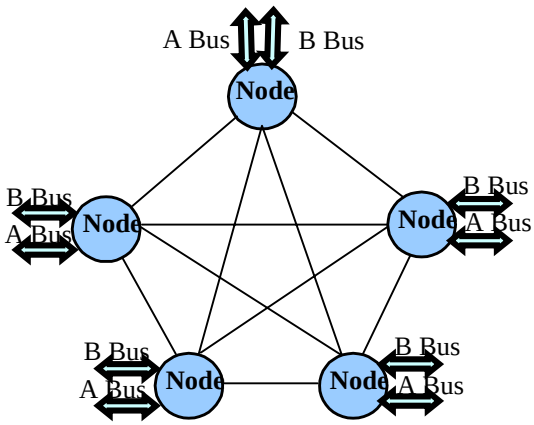
- Fabric Slots/Cards
- Basic Node Slots/Cards
- Multi Segment Node Cards
- Fabric Native Node Cards

In a system with a centralized PICMG 2.18 fabric topology, these active switching devices are located on the Fabric Card(s); in this type of PICMG 2.18 system, the other cards that connect to the Fabric Card(s) are called Basic Node Cards. The Switch Fabric architecture supports connections from every Node Card to centralized Fabric Card(s). The availability of 2 Fabric Slots/Cards supports redundant connections for high availability applications. Refer to the following figure:



Dual Centralized Fabric Topology

In a system with a distributed PICMG 2.18 mesh topology, these active switching devices are located on Multi Segment Node Cards or Fabric Native Node Cards. Refer to the following figure:



Mesh Topology

The impact on available I/O pins is minimal.

- Fabric Slots/Cards utilize pins on P1/J1 – P5/J5.
- Basic Node Slots/Cards utilize pins on P3/J3
- Multi Segment Node Cards utilize pins on P3/J3
- Fabric Native Node Cards utilize pins on P3/J3 and P4/J4

Redundancy for High Availability applications

High Availability systems strive to eliminate single points of failure. The PICMG 2.18 specification was designed to carry significant amounts of traffic, and so reliability and availability is a key issue. Each Basic Node Card can connect to two Fabric Cards, with one Fabric Card used as a redundant system element. If one Fabric Card fails, traffic can be routed through the other one without the system going down. In case a Node Card fails, PICMG 2.18 supports the full Hot-Swap specification so that the faulty card can be removed and another put in its place without rebooting the system.

Point-to-Point Architecture

To maximize bandwidth, a point-to-point architecture is used, utilizing active switching elements to route data between many processing cards, called Node Cards. These active switching elements are located on special cards, called Fabric Cards, which are located in special backplane slots, called Fabric Slots. The PICMG card types are:

- ✓ A central interface card to perform switching of Serial RapidIO signals to and from Node Slots on the backplane as well as incoming and outgoing traffic (Fabric Card(s)). Note that the system may have zero, one or two Fabric Cards.
- ✓ A distributed set of processing cards (Basic Node Cards). These are used in centralized fabric topologies with one or two Fabric switch cards and can coexist with PICMG 2.16 Ethernet and the H.110 bus.
- ✓ Multi Segment Node Cards that support 4 Serial RapidIO Links per Card. These can be used in centralized fabric topologies with one or two Fabric switch cards or in mesh topologies without requiring a central Fabric switch card, and can coexist with the H.110 bus.
- ✓ Fabric Native Node Cards that support 4 Serial RapidIO Links per Card. These can be used in centralized fabric topologies with one or two Fabric switch cards or in mesh topologies without requiring a central Fabric switch card, and can coexist with PICMG 2.16 Ethernet.

The PICMG 2.18 specification was designed for a 21-slot chassis. A PICMG 2.18 chassis can support a maximum of two Fabric Cards and 19 Node Cards.

There is a full-duplex connection between each Node Card and each Fabric Card (four pairs for transmit, and four pairs for receive, as well as grounds.)

Each Basic Node Card can connect to both Fabric Cards. The dual Fabric Cards can be used as redundancy or as additional resources.

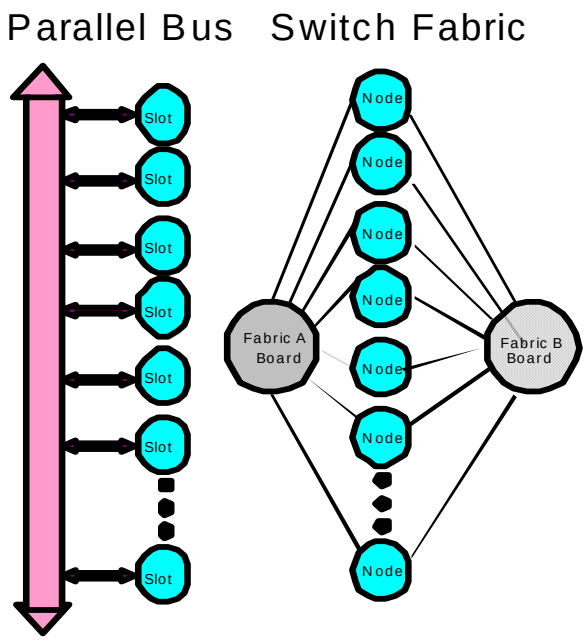
The specification also provides for various mesh topologies utilizing Multi Segment Node Cards or Fabric Native Node Cards.

The backplane for PICMG 2.18 is designed to isolate differential pairs to avoid cross-talk. This is important because signals are routed parallel to each other within the CompactPCI connector, which can lead to severe cross-talk noise. PICMG 2.18 eliminates this problem by dedicating entire rows on the connector to provide grounded electrical isolation between the Serial RapidIO Tx and Rx pairs, as well as between the Serial RapidIO pairs and the PCI/H.110 or rear I/O signals.

Each Basic Node Card can support up to a 10 gigabit full duplex connection to each Fabric Card. Thus the total bandwidth for system with a centralized fabric topology is 380 Gb/s. Mesh topology systems can support double this total bandwidth per slot.

Switch Fabrics in General

There is no doubt that industry is moving toward switch fabric architectures to supplement, and eventually replace, traditional buses in many applications. Switch Fabrics allow multiple simultaneous data flows, each at full speed, rather than sharing the bandwidth of a single bus. They also minimize the number of interconnect signals required, and eliminate bus loading skew problems that limit the performance of parallel buses. The following table contrasts Switch Fabrics with VME and CompactPCI, and the figure contrasts bus structures with fabrics.



A number of other industry specifications that utilize Switch Fabrics for interconnecting boards and systems are being developed. These new interconnect architectures are serial, point-to-point architectures. They include:

- ✓ **InfiniBand**—Designed for ultra-high bandwidth for servers, with data rates of 5 Gb/s and up. This will migrate from servers into other systems, but will likely require different backplane connector standards in order to support the very high speed signaling.
- ✓ **PCI Express**—A chip-to-chip interconnect. This is a technology to watch because of its strong industry backing; PCI Express has been selected by the PCI SIG as the successor to PCI.
- ✓ **RapidIO**—There are two versions of RapidIO. The initial version, Parallel RapidIO, is a multiple pair chip-to-chip interconnect with a separate clock; it is aimed at DSPs and CPUs, but there is also support from FPGA vendors. The second version, Serial RapidIO, with an embedded clock, is used in this specification. RapidIO is a packet-based system with low latency.
- ✓ **StarFabric**—StarFabric is a low-cost switch fabric that acts as an interconnect for existing buses, including PCI and H.110. It allows very large PCI systems to be constructed with no impact to existing PCI software, and is theoretically capable of supporting redundant system controllers (with appropriate software). It can also be used for carrying a mix of traffic types including Voice and Video with Quality of Service (QOS) provisions.
- ✓ **Hypertransport**—A multiple pair chip-to-chip interconnect with broad vendor support that is less likely find its way into backplane interconnects, but it is likely to be used with mezzanine boards. It is aimed at CPUs, but there is also support from FPGA vendors.
- ✓ **CSIX**—A specialized TDM interconnect for support of ATM packets; it is initially a parallel interconnect. A serial version of the CSIX protocol has used for the PICMG 2.20 Serial Mesh Backplane Specification.

All of these switch fabrics have their own niche. However, the PICMG 2.18 backplane has been designed especially for real time low latency applications. It can also happily coexist with PICMG 2.16. These two factors make it very likely that combinations of PICMG 2.16 and PICMG 2.18 will play a key part in future systems using CompactPCI.

Why you should talk to Hybricon

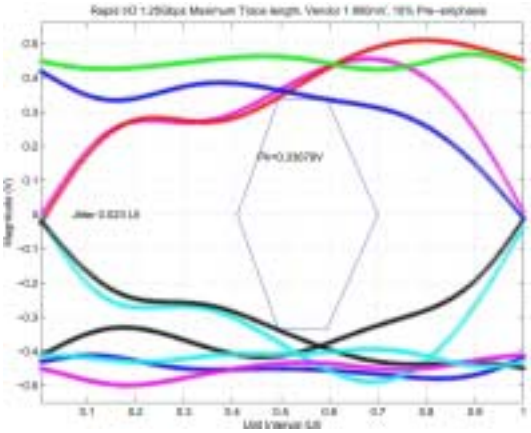
Hybricon is a PICMG Executive member that is very active in many PICMG subcommittees and is very well known and respected within the PICMG community. Our involvement in writing the specifications means that Hybricon is well versed in all approved PICMG specifications as well as draft specifications.

As one of the Sponsors of the PICMG 2.18 subcommittee, Hybricon is committed to the success of PICMG 2.18, and is intimately familiar with the specification.

In fact, as the company chosen by the RapidIO Trade Association to perform all of the signal integrity simulation work for PICMG 2.18 (as well as for PICMG in support of PICMG 2.16 and PICMG 2.17), **Hybricon is the definitive expert in designing PICMG 2.18 backplanes**. This is new technology; why trust your critical design to anyone else?

Hybricon offers standard PICMG 2.18 compliant backplanes and chassis. We anticipate that many PICMG 2.18 applications will require custom solutions, and Hybricon's world-class engineering team stands ready to craft custom solutions that meet today's aggressive time-to-market requirements.

An example of the simulation results from Hybricon's PICMG 2.18 signal integrity studies is shown in the following figure.



Hybricon's CompactPCI products cover a wide range of backplanes, power planes, card cages, rack mount chassis and desktop enclosures. While we do offer a broad range of standard products, Hybricon specializes in custom solutions. We offer a broad range of engineering services including sophisticated Signal Integrity analysis for custom buses or multi-gigabit serial links using 2D/3D field solvers and circuit simulators as well as sophisticated thermal/air flow analysis using CFD analysis tools.

For more information on Hybricon's Serial RapidIO CompactPCI Development Kit, see:
http://www.hybricon.com/products/enclosures/picmg218.html

For more information on Hybricon's Serial RapidIO CompactPCI Backplanes, see:
http://www.hybricon.com/products/backplanes/picmg218.html

For more information on Hybricon products, please e-mail us at info@hybricon.com, visit us on the Web at www.hybricon.com, or call Sales at (978) 772-5422.