



HTF120

Introduction to the NEW PICMG 2.17
StarFabric Compact PCI
Specification

Technology Focus

Executive Summary

This paper will provide an introduction and overview of the newly released PICMG 2.17 StarFabric 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 2.17 specification was approved and released by the PICMG Executive Board in May, 2002. This new specification provides, for the first time, a standardized way to implement a StarFabric Switch Fabric in a standard backplane without the use of cables, and using standardized backplane connector pinouts. The PCI transparency that is built into the StarFabric hardware allows users to leverage available software drivers for PCI hardware, and get applications up and running immediately. This provides a clean, efficient way to implement very high bandwidth non-blocking simultaneous interconnects for an arbitrarily large number of PCI slots. As the second Switch Fabric to be implemented on CompactPCI (this was preceded by PICMG 2.16), this is at the leading edge of a new era, as system architectures 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. Systems used in telecommunications usually use a 6U card 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 2.17 subcommittee to investigate a better method of transferring large amounts of data between cards, with an emphasis on PCI and TDM/H.110 traffic. This contrasts with PICMG 2.16, which has an emphasis on packet traffic.

Introduction to the StarFabric CompactPCI Backplane

PICMG 2.17 has the following high-level features:

- Scalable system performance: Up to 190 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 StarFabric Links per Basic Node Slot.
- Inherently hot swappable and HA due to serial switch fabric based architecture.

Key Aspects of PICMG 2.17 include:

1. Point-to-Point interconnections across the backplane.
2. Utilizes existing PCI drivers and application software as-is to speed Time To Market.
3. Supports Quality of Service including support for TDM/H.110 traffic with additional software support.
4. Utilizes a Switch Fabric architecture.

Time To Market

PCI driver and PCI-centric application software is available off-the-shelf. This allows system vendors to improve Time-To-Market by re-using available software drivers and application software.

StarFabric technology will be incorporated into the PCI Express Advanced Switching specification, selected by the PCI SIG as the successor to PCI. This means that companies that are interested in deploying PCI Express products in the future can implement with StarFabric now, and have a migration path to PCI Express Advanced Switching products in the future.

Switch Fabric Architecture

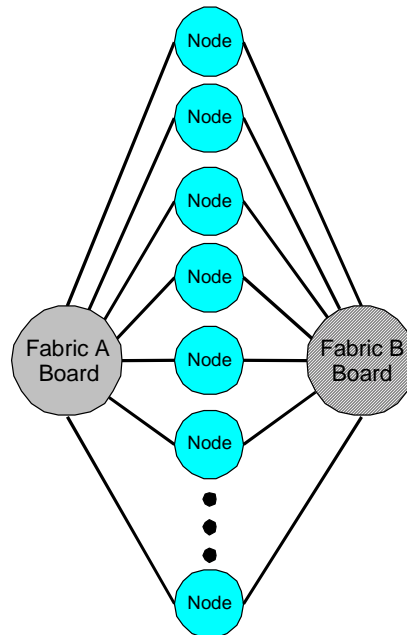
PICMG 2.17 defines a Switch Fabric Architecture, as contrasted with a Bus Architecture (e.g. PCI). A Switch Fabric is an interconnected network of switching devices. The fabric contains many input

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:

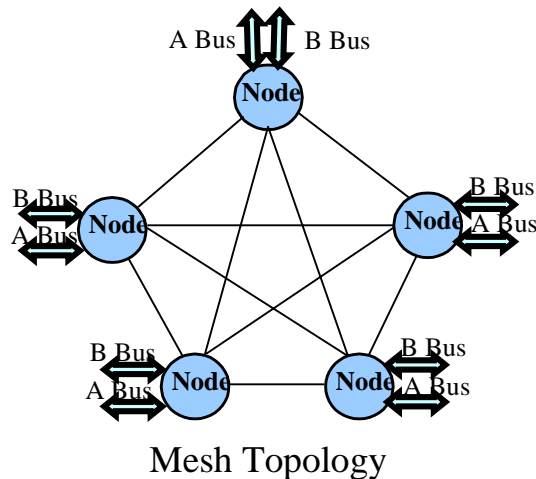
- Fabric Slots/Cards
- Basic Node Slots/Cards
- Multi Segment Node Cards
- Fabric Native Node Cards

In a system with a centralized PICMG 2.17 fabric topology, these active switching devices are located on the Fabric Card(s); in this type of PICMG 2.17 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 2.17 mesh topology, these active switching devices are located on Multi Segment Node Cards or Fabric Native Node Cards. Refer to the following figure:



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 2.17 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 route through the other one without the system going down. In case a Node Card fails, PICMG 2.17 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, that are located in special backplane slots, called Fabric Slots. The PICMG card types are:

- ✓ A central interface card to perform switching of StarFabric 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 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 StarFabric 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 StarFabric 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 2.17 specification was designed for a 21-slot chassis. A PICMG 2.17 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.

Legacy PCI software transparency

With available StarFabric to PCI bridges, PICMG 2.17 provides a convenient vehicle for extending the bandwidth, slot count, and availability of CompactPCI systems, preserving investments in legacy software and hardware. StarFabric achieves this by being totally software transparent for PCI bus operations. As such, it is a good platform for evolution into the high performance Switch Fabric paradigm, while preserving legacy software and CompactPCI hardware.

With available StarFabric to H.110 bridges, PICMG 2.17 provides a convenient vehicle for extending legacy TDM voice systems, increasing the number of voice channels that can be conveniently supported in a single chassis, and supports a mix of different types of traffic on the fabric while preserving Quality of Service.

The backplane for PICMG 2.17 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 2.17 eliminates this problem by dedicating entire rows on the connector to provide grounded electrical isolation between the StarFabric Tx and Rx pairs, as well as between the StarFabric pairs and the PCI/H.110 or rear I/O signals.

StarFabric can interconnect existing buses, including PCI and H.110. It theoretically 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.

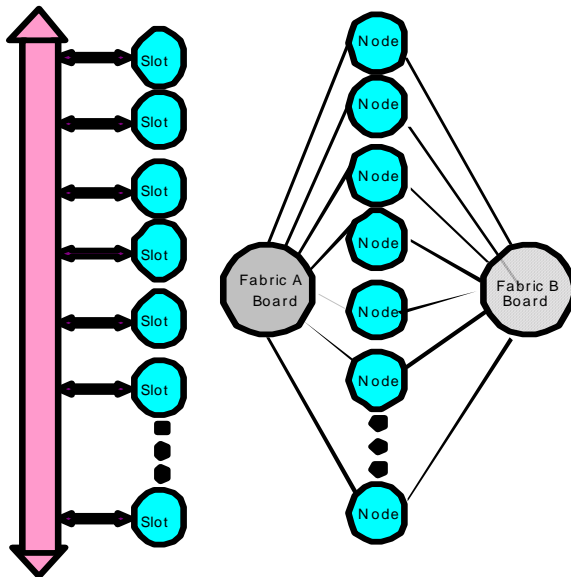
Each Node Card can support up to a 5 gigabits full duplex connection to each Fabric Card. Thus the total bandwidth for the system is 190 Gb/s.

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.

| Parameter | VME, CompactPCI | Switch Fabric |
|-------------|--|--|
| Topology | Bussed bi - directional multi-drop | Point to-point |
| Types | <ul style="list-style-type: none"> • VME – asynchronous • CompactPCI – synchronous | <ul style="list-style-type: none"> • Centralized – connect from each Node Slot to one or more dedicated Fabric Slots • Mesh – connect from each Slot to every other Slot |
| Signal Type | Single ended signals with TTL-like levels | Differential signals typically LVDS/PECL |
| Data Width | 32 – 64 bit address/data + control | 1 – 16 bits |
| Clocking | <ul style="list-style-type: none"> • VME – bussed • CompactPCI – point-to-point | <ul style="list-style-type: none"> • Point-to-point connection, where clock is typically embedded in serial bit stream with 8b/10b encoding • Clock is occasionally a separate point-to-point connection, one per N data connections |
| Data rate | <ul style="list-style-type: none"> • VME – 10 ~ 320 Mbytes/sec • CompactPCI – 132 ~ 528 Mbytes/sec | <ul style="list-style-type: none"> • Ethernet – 20 ~ 2000 Mbits/sec per link • Infiniband – 5000 Mbits/sec per link • StarFabric – 5000 Mbits/sec per link |

Parallel Bus Switch Fabric



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, now under development. 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. PCI Express will incorporate StarFabric technology as part of the PCI Express Advanced Switching specification. This means that companies that are interested in deploying PCI Express products in the future can implement with StarFabric now, and have a migration path to PCI Express Advanced Switching products in the future. For more information, see

the press release at:

http://www.starfabric.org/pr_20020812.htm.

- ✓ **RapidIO**—There are two versions of RapidIO. The initial version 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 specification for a second version, Serial RapidIO, with an embedded clock, has been recently released. RapidIO is a packet-based system with low latency. Serial RapidIO may find its way into backplane interconnects.
- ✓ **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 been proposed for the new PICMG 2.20 Serial Mesh Backplane effort.

All of these switch fabrics have their own niche. However, the PICMG 2.17 backplane has been designed especially for the CompactPCI chassis, and is designed to provide an evolutionary path from PCI into high performance switch fabrics. It can also happily coexist with PICMG 2.16. These two factors make it very likely that combinations of PICMG 2.16 and PICMG 2.17 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.

HYBRICON CORPORATION • 12 WILLOW ROAD • AYER, MASSACHUSETTS 01432

TEL: (978) 772-5422 • FAX: (978) 772-2963

WEB: [HTTP://WWW.HYBRICON.COM](http://www.hybricon.com) • EMAIL: [INFO@HYBRICON.COM](mailto:info@hybricon.com)

Bob Sullivan, Hybricon's VP of Technology, sits on the Board of Directors of the StarFabric Trade Association.

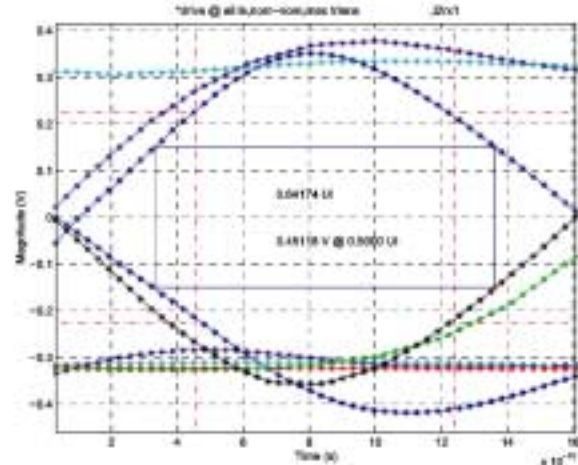
As one of the Sponsors of the PICMG 2.17 subcommittee, Hybricon is committed to the success of PICMG 2.17, and is intimately familiar with the specification.

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

Hybricon offers standard PICMG 2.17 compliant backplanes and chassis. We anticipate that many PICMG 2.17 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.

There is no such thing as a PICMG 2.17 compliant product (not even a "draft" compliant product) until the specification is approved, despite what some companies' marketing departments would have you believe. Understanding the risks of designing to a draft specification, Hybricon is happy to work with early adopters even before the specification is finalized.

An example of the simulation results from Hybricon's PICMG 2.17 work 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 StarFabric CompactPCI Development Kit, see:

<http://www.hybricon.com/coolslot/picmg217.html>

For more information on Hybricon's StarFabric CompactPCI Backplanes, see:

<http://www.hybricon.com/cpci/picmg217.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.