



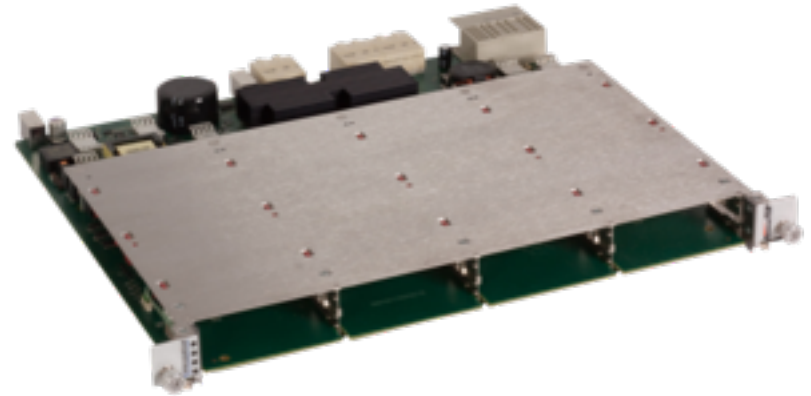
# Using Intel Processors with RapidIO: Server and High Performance Computing with ATCA

Nov 2011

Devashish Paul



- Why RapidIO
- Bringing Intel Processors into RapidIO Networks
- ATCA with RapidIO for Server and High Performance Computing
- RapidIO and Ethernet
- IDT RapidIO Gen2 Ecosystem





## **RapidIO on board**

as the single, simple interconnect among all board components



## **RapidIO On the backplane**

- Future proof
- High throughput
- Low deterministic latency
- Guaranteed packet delivery
- Prioritized traffic



## **RapidIO for fault tolerant Systems**

- Flexible sparing strategies
- Continued system operation in the event of single faults
- Rapid detection of faults
- Flexible response to faults



## **Protect your SW investment**

- S-RIO logical layer remains the same across different physical layer
- RapidIO scales per port
- Saves system total power

- Disruptive Architecture, that changes the overall economics of deployment in all multi-processor application
- Scalable solutions for board, backplane and inter chassis communication vs PCIe or Ethernet
  - Extend overall system solution by aggregating chassis
- Lowest overall system power with S-RIO
- Superior **end to end packet latency**, throughput and fault tolerance
- Flexibility and scalability to evolve system configuration in field
- Option to use PCIe enabled processors while leveraging all key attributes of RapidIO
- Incumbent interconnect in all top 10 wireless OEM 3G and 4G base station designs



**IDT RapidIO has over 150 design wins with 400% year over year revenue growth**

## Silicon Partners with RapidIO

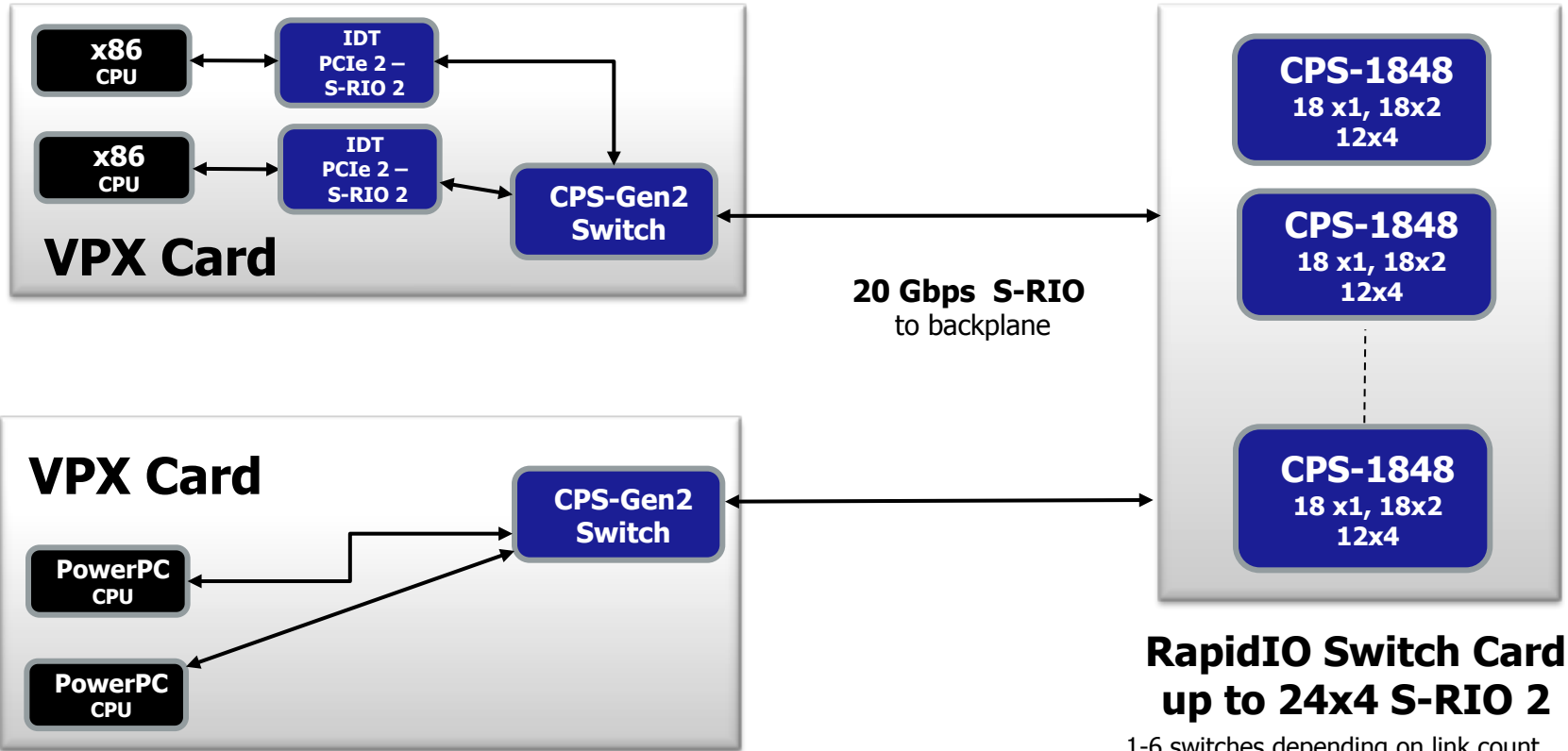


- PCIe enabled processors are leading the market in performance
- This performance is key in: Military, High Performance Computing, Server, Imaging, Wireless and other embedded applications
- The same applications need the performance of RapidIO Interconnect for:
  - Peer to peer networks with scalability
  - Lowest system power with protocol terminated in Hardware
  - Lowest end to end packet latency



**OEMs want the best of processor performance  
and RapidIO Peer to Peer Interconnect  
For Embedded Networking High Performance Computing**

All RapidIO PCIe Payload Card  
Payload Card with RapidIO Bridge



**RapidIO Switch Card  
up to 24x4 S-RIO 2**

1-6 switches depending on link count  
and non-blocking bandwidth needs

- Connect x86 CPUs to Open VPX systems
- Design heterogeneous systems
- Design RapidIO based backplanes
- Leverage peer-to-peer and messaging capability of RapidIO using x86 processors

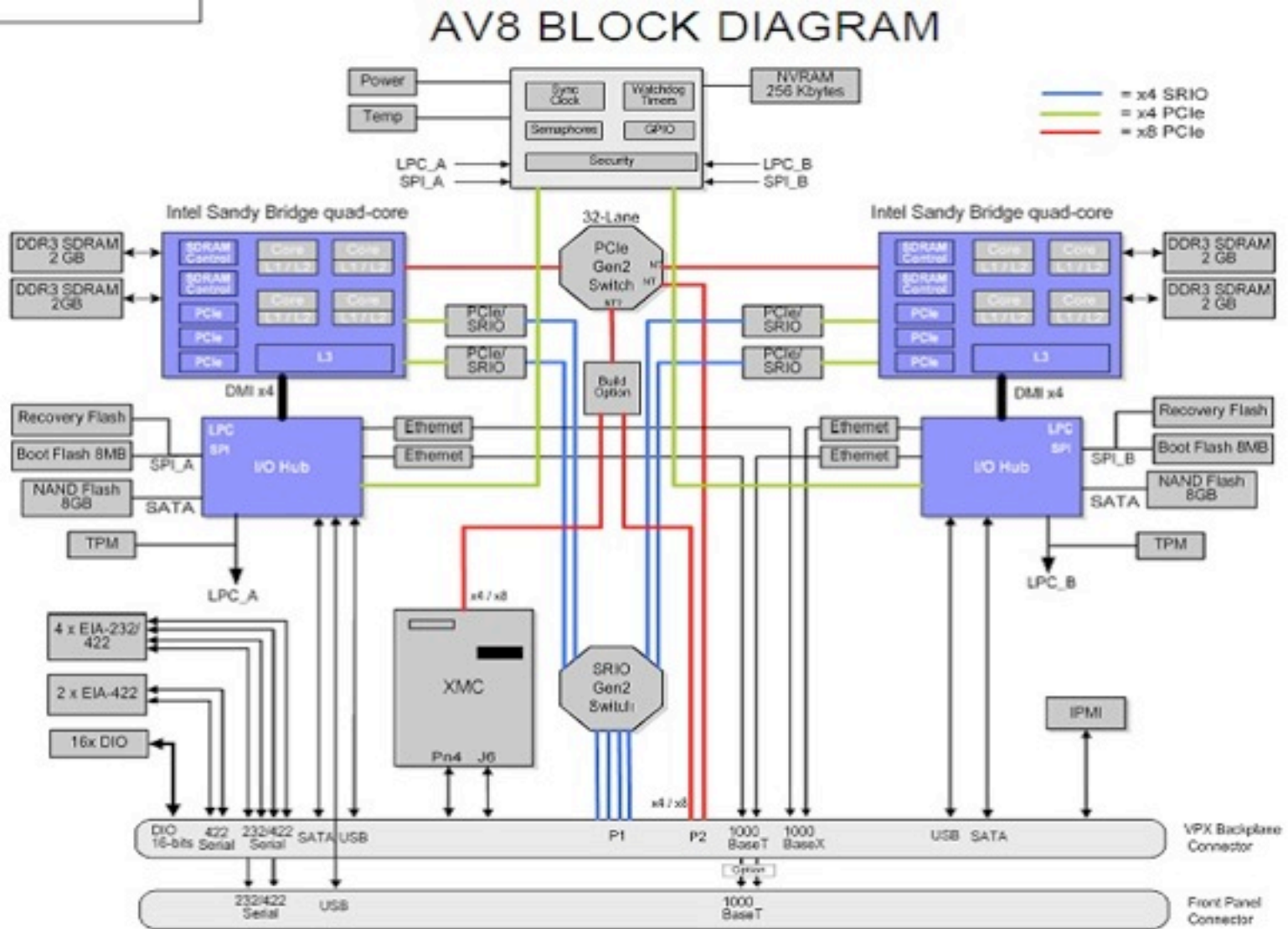
**RapidIO system performance  
is superior to 10GbE**

- CHAMP-AV8 6U OpenVPX™ DSP Engine Features two quad-core Intel® Core™ i7-2715QE processors and **IDT Gen2 PCIe to S-RIO protocol conversion**
- Interfaces Intel Processors to RapidIO Gen2 backplane
- Ultra Low latency and system power



Curtiss-Wright Controls Introduces CHAMP-AV8, Its first 6U OpenVPX™ Multiprocessing DSP Engine based on the 2nd Generation Intel® Core™ i7 Processors

**Scalable low latency modular solution  
with Intel I7 and RapidIO**



## VPX6-6902

- Supporting a centralized switch architecture in both star and dual-star topologies
- S-RIO switch fabric supports both Gen-1 (1.25, 2.5, 3.125 GBaud) and Gen-2 SRIO (2.5, 5, 6.25 GBaud) data rates.
- Each of it's 28x S-RIO ports operates in x4 lane mode, offering up to 20Gbps throughput per port.
- Provides switching infrastructure for RapidIO based VPX systems



**6U VPX Form Factor  
Highest Performance Switch Card  
In Embedded Industry**

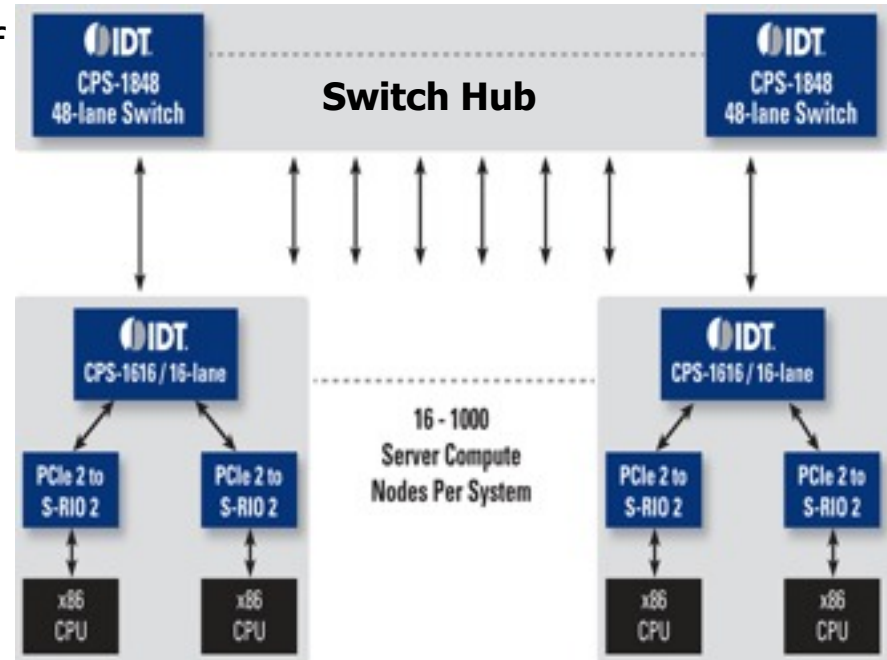


# Retargeting Mil/Aero Successes into





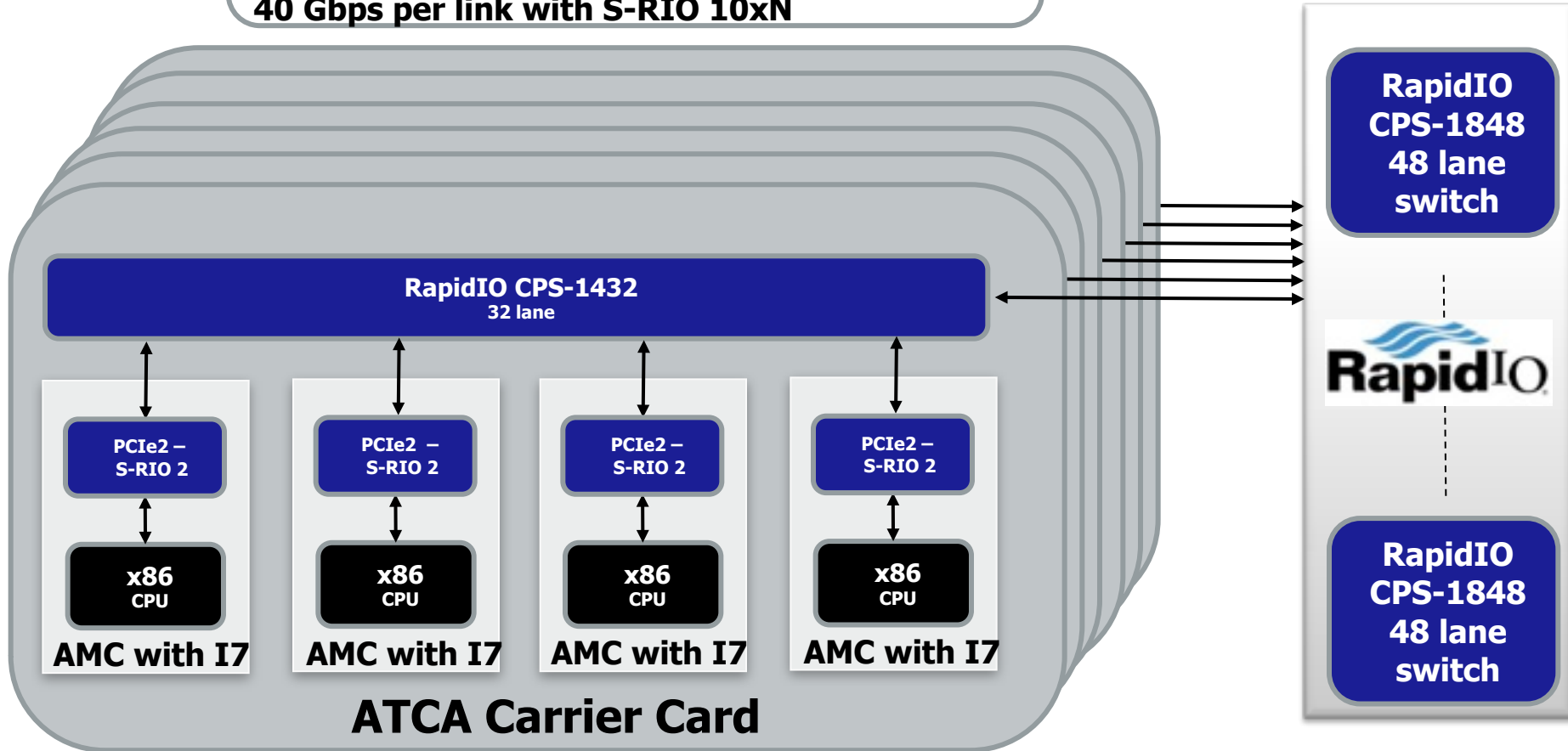
- Disruptive Architecture, that changes the overall economics of deployment
  - Reduces cabling
  - More Modular Systems
- **Superior end to end packet latency, throughput and fault tolerance**
- Remove high cost Ethernet switches from top of rack and costly Infiniband cabling
- **Scalable** Solutions for server and inter chassis communication
  - Scale within a given system with a varying number of compute nodes
  - Extend overall system solution by aggregating chassis
- Support Green Initiatives, with overall **reduced energy footprint** in datacenter with S-RIO
- Flexibility to evolve system configuration in field
- Use PCIe enabled processors while leveraging all key attributes of RapidIO that have made it successful in wireless and embedded markets

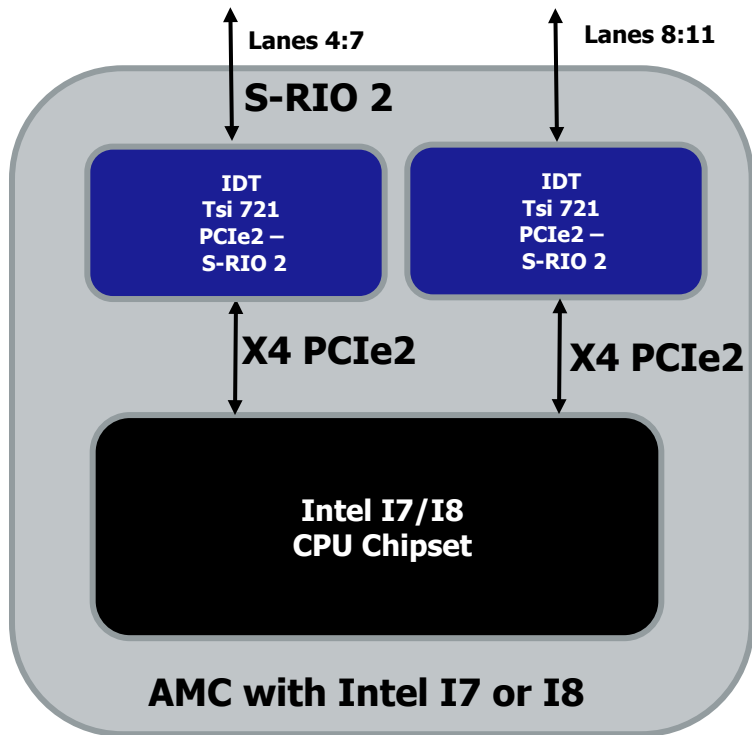


**IDT PCIe to S-RIO bridge with low latency, power consumption and form factor compared to Infiniband and Ethernet NIC options**

# IDT RapidIO in High Performance Computing/Server with ATCA

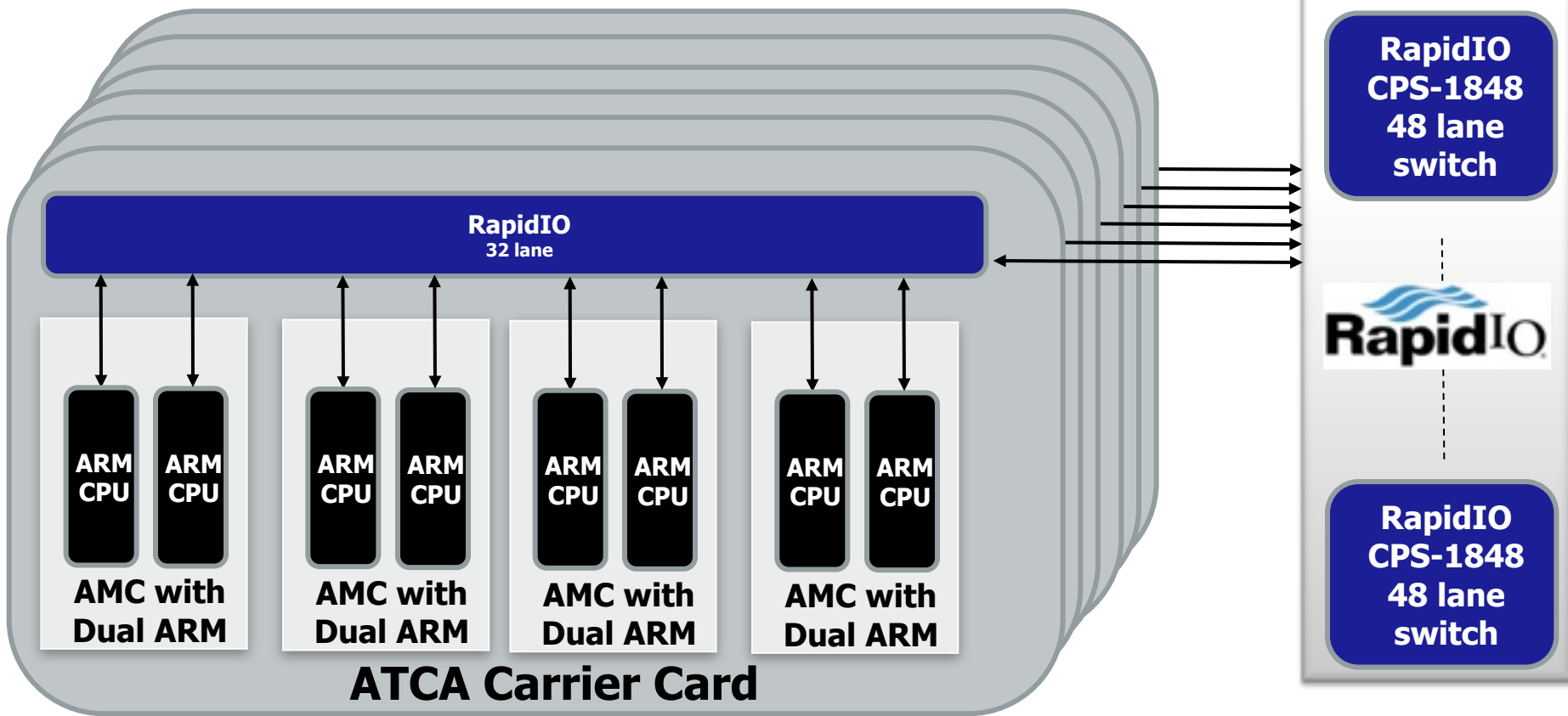
Multiple AMC Cards per ATCA Carrier  
Multiple ATCA Carrier Cards per chassis  
Multiple Chassis within a network  
20 Gbps per link now with S-RIO2  
40 Gbps per link with S-RIO 10xN





- Use Multiple AMC Cards per ATCA Carrier
- IDT PCIe2 to S-RIO 2 bridge to connect to S-RIO carrier card or in MicroTCA Chassis
- Low latency and footprint compared to Infiniband/Ethernet NIC
- Option to support 8 lanes of PCIe with two bridge devices (small form factor 13x13)
- Supports Fat Pipes for MicroTCA chassis or ATCA carrier cards
- In development with multiple vendors

Multiple AMC Cards per ATCA Carrier  
Multiple ATCA Carrier Cards per chassis  
Multiple Chassis within a network  
**Ultra Low Power with ARM Processors**  
**Low Latency with no NIC given native S-RIO**





# RapidIO Gen2 / Ethernet in ATCA

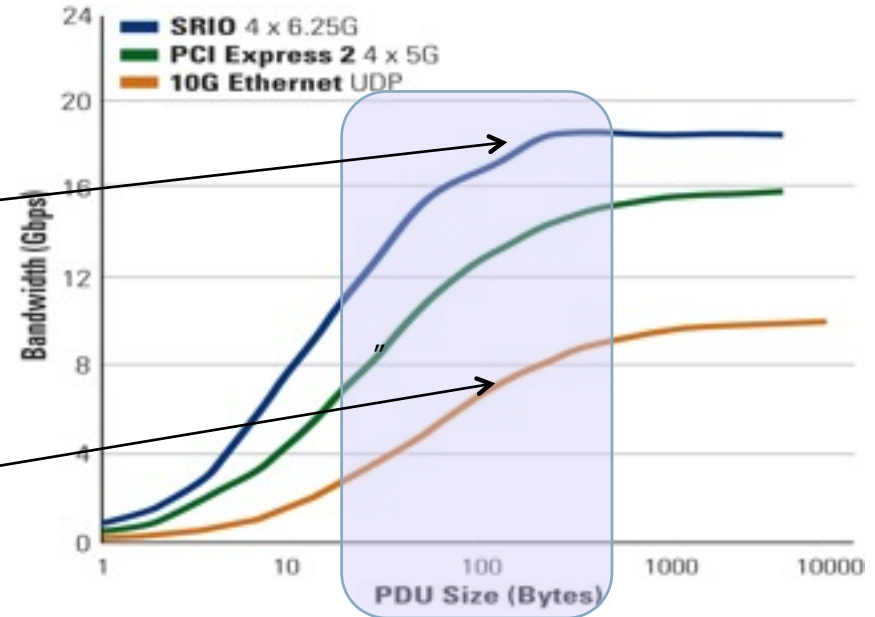


## • RapidIO Gen2 vs 10 Gigabit Ethernet

- RapidIO best peer to peer multi processor performance
- Lowest System Power
- **Guaranteed delivery mechanisms in RapidIO for mission critical peer to peer transactions**
- Payload rate 19 Gbps for S-RIO2, 8 Gbps for 10 GigE (256 byte packets)
- Better per-port economics than 10 GigE switches
- Best cut through latency 100 ns, 10 GigE cut through latency of multiple hundreds of ns

RapidIO 2  
19.6 Gbps  
actual data  
rate, 256 byte  
packet

10 GbE  
~8 Gbps,  
256 byte  
packet



### Bandwidth and Per Port Economics

System Requirement	10G Ethernet	RapidIO 2
Switch per-port performance raw data	10 G	20 G
End to end packet	>10 ms	~1-2 us
Messaging performance	Poor	Excellent
Volume pricing \$ per 10G	> \$10	< \$4.00
Hardware bridging to PCIe Express	No	Yes
Overall system power	High	Lowest

- **Packet Size and Efficiency**

- Largest S-RIO packet size is 276 bytes with up to 20 bytes header. Can achieve 95% protocol efficiency with small header

- **Supports mixed traffic flows**

- Control and Data Plane using Messaging, Logical I/O read/write etc.

- **Efficient Flow Control**

- RapidIO incorporates efficient flow control mechanism
- Manages short, to long-term congestion in a system
  - Link Level Flow control
  - VOQ Back-pressure
  - Virtual Channels and Bandwidth control
  - Transmitter and Receiver-based flow control

- **Guaranteed Delivery of Packet**

- Packets are not dropped even under congestions or transmission errors

- **Flow-control and error recovery mechanism**

- Supports Control symbols that can be embedded within packets to reduce latency

- **RapidIO guarantees packet delivery on each link**

- For example in 100 m fibre connection @ 10G, recovery can be accomplished with the exchange of 3 CSs within 2.5 usec. For chip-to-chip, the exchange can be completed in less than 300 nsec.

	<b>Ethernet</b>	<b>RapidIO</b>
<b>Lossless, in order delivery</b>	<b>No</b>	<b>Yes</b>
<b>Latency</b>	<b>HIGH</b>	<b>LOWEST</b>
<b>Quality of Service</b>	<b>More components, \$\$\$</b>	<b>Integrated,</b>
<b>Flow Control</b>	<b>Drops Packets</b>	<b>Comprehensive</b>
<b>Switch Sizes</b>	<b>Application mismatch, features vary</b>	<b>Sizes match applications.</b>

## Legacy

**RapidIO 1.0**

Parallel -LVDS  
Serial - XAUI

- Multi-processor
- DSP Cards
- CompactPCI
- Backplanes
- ATCA, VME
- VPX, VSX
- Military
- Medical
- 2.5G Wireless
- WiMax Modem
- SDRadio

## Today

**RapidIO 1.2, 1.3, 2.3**

Serial – XAUI, CES  
1.25, 2.5, 3.125GBaud  
5, 6.25GBaud  
1X, 2X, 4X, 8X, 16X

-10-20GBaud 4X  
backplanes  
-1.25GxN to 6.25GxN

**Key Applications RapidIO 1.2 and beyond**

- DSP and Processor Farms
- Wireless 3G, WiMax, 4G and future 5G
- Video servers, IPTV, HDTV, Media Gateways
- microTCA, AMC, PMC
- VME, VSX, VPX systems
- Storage/ Server Systems
- High Performance Computing

## 2012

**RapidIO 10xN**

Serial –10G KR  
10GBaud

-10GxN  
-40GBaud to  
160GBaud  
Backplanes

## Future

**RapidIO 25xN**

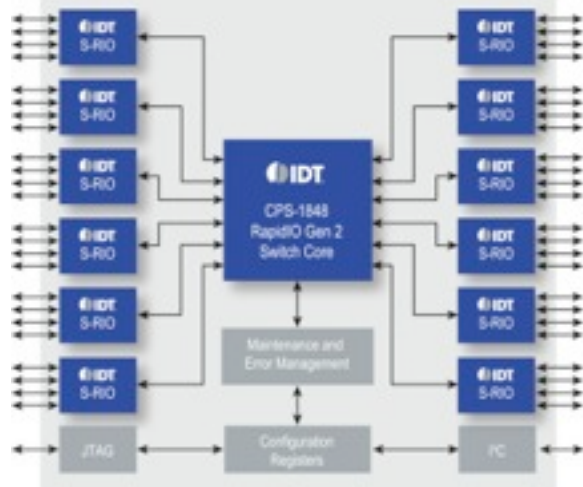
IEEE –25G  
25GBaud

-25GxN  
-25GBaud to  
400GBaud  
Backplanes



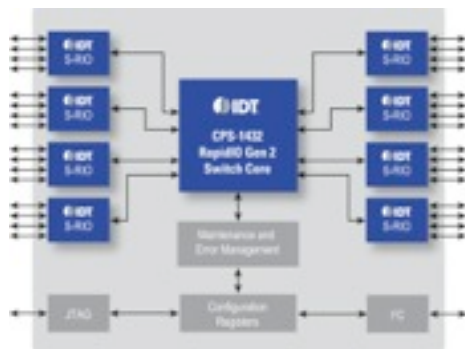
# Silicon and AMC Infrastructure





**CPS-1848**  
**12x20 Gbps**  
**18x 10 Gbps**  
**18x 5 Gbps**  
**29x29 FCBGA**

- Designed to S-RIO v1.3 and 2.1
- Up to 48 lanes –12x4, 18x2, 18x1
- **Up to Full duplex 240 Gbps non blocking bandwidth**
- Supports all RapidIO speeds: 1.25, 2.5, 3.125, 5, 6.25 Gbaud
- **Cut through latency 100 ns**
- 40% per 10 Gbps power reduction vs S-RIO 1.3
- **Several switch fabric related patents filed**



**CPS-1432**  
**8x20 Gbps**  
**14x 10 Gbps**  
**14x 5 Gbps**  
**25x25 FCBGA**

## Detailed Features

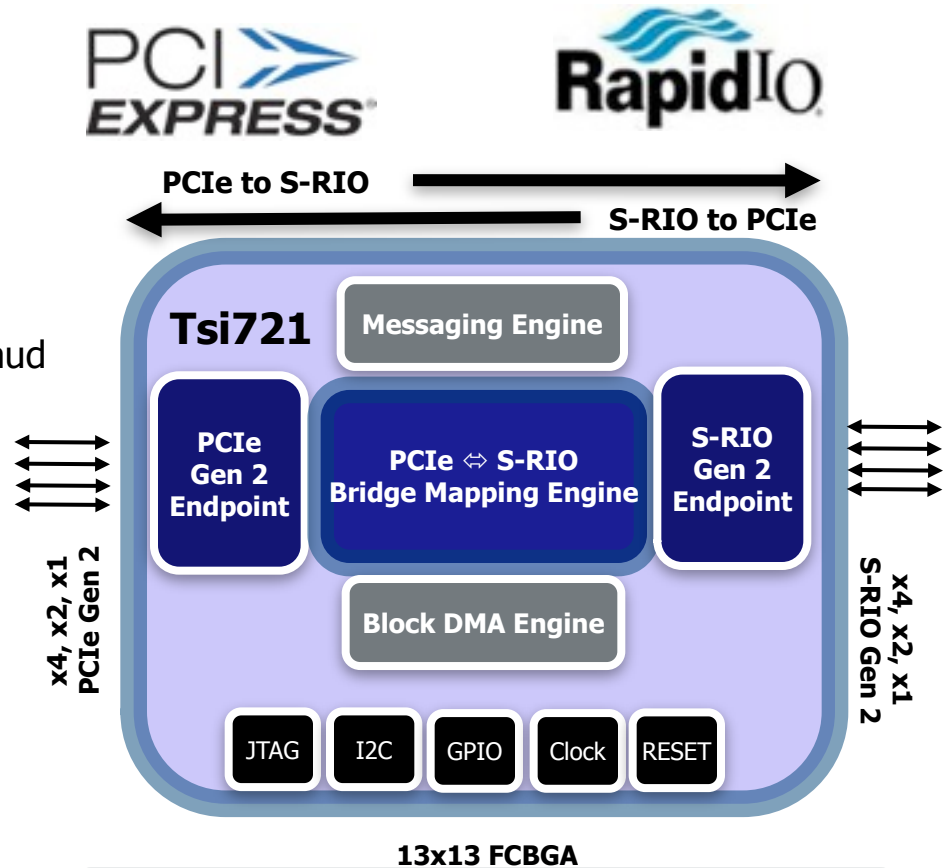
- High-Performance SerDes
  - Long reach 100 cm 2 connector with DFE support
  - Transmit Pre Emphasis and receive Equalization
- Typical 300 mW per 10 Gbps of data
- Dynamic ingress and egress buffers management:
- 40 multicast groups per port
- Supports cut-through and store & forward
- Error management extension support
- Error Log: sequence of events in time
- packet mirror, trace, filter support
- Receiver and transmitter based flow control
- Per port reset mode robust support for hot swap
- Multicast event control symbol generation input pin



**CPS-1616**  
**4x20 Gbps**  
**8x 10 Gbps**  
**16x 5 Gbps**  
**21x21 FCBGA**

## FEATURES

- Gen 1 and Gen 2 support
  - PCIe v2.1
  - S-RIO v2.1
- PCIe to S-RIO bridging
  - Non-transparent for transaction mapping
  - 8 DMA and messaging engines
- Single port, x1/x2 or x4 @ 1.25, 2.5, 3.125, 5 Gbaud
  - S-RIO 1.3 and 2.1 compliant
  - PCI-express 1.1 and 2.1 compliant (End-Point)
  - Can buffer up to 32 S-RIO max size packets
  - Full line rate throughput for 64 byte and > packets
- Low Power ~ 3-4W typical
- Power down unused lanes, when used in x1 or x2
- Lane swap and polarity inversion support
- Reach support: 60 cm over 2 connectors
- S-RIO and PCIe endpoint compatible clocking options 100 MHz, 125 MHz 156.25 MHz
- Forward Bridge (must have microprocessor on PCIe side of the network)
- JTAG 1149.1 and 1149.6
- 13x13mm FCBGA package
- Commercial and industrial variants



**13x13 FCBGA**

**Connect PCIe Processors  
To S-RIO Networks  
For Superior Performance  
Over 10 GbE and Infiniband**



## Tsi721 PCIe2 to S-RIO Eval Board Connectors

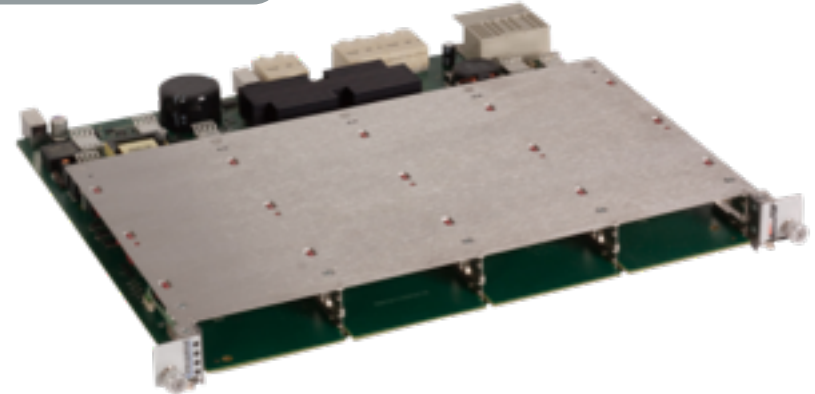
- 1 AMC Connector: with three 4x S-RIO links, AMC.0 and AMC.4 Specification (NO support on IPMC and JTAG)
- 2 SFP+ Connectors: 1 ports with 1x S-RIO link, INF-8341 Specification
- 2 InfiniBand/CX4 Connectors: 1 port with 4x S-RIO link or 2 ports with 2x S-RIO links, SFF-8470 Specification
- 1 SMA Array: 1 port with 4x S-RIO links



## SRDP2 (available from [www.silicontkx.com](http://www.silicontkx.com))

- 3 AMC B+ Connectors: 2 with two 4x S-RIO links, 1 with three 4x S-RIO links, AMC.0 and AMC.4 Specification (NO support on IPMC and JTAG)
- 2 SFP+ Connectors: 1 ports with 1x S-RIO link, INF-8341 Specification
- 1 QSFP Connectors: 1 port with 4x S-RIO link, SFF-8438i Specification
- 2 InfiniBand/CX4 Connectors: 1 port with 4x S-RIO link or 2 ports with 2x S-RIO links, SFF-8470 Specification
- 1 SMA Array: 1 port with 4x S-RIO links

## Prodrive Ecosystem Solutions with RapidIO



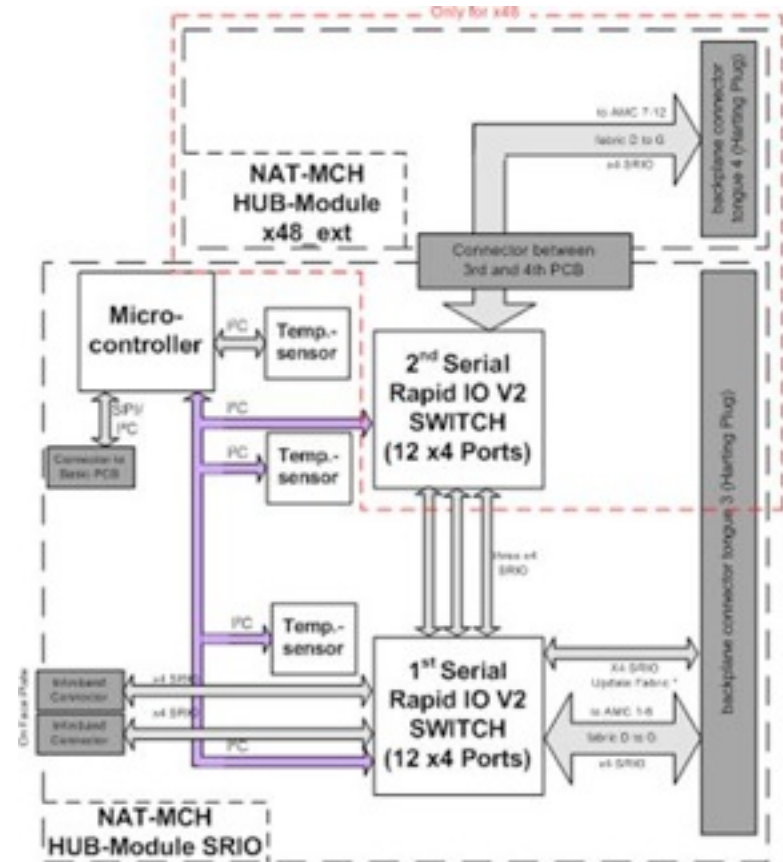
Quad AMC carrier blade (Switch)  
based on S-RIO 1.3 and 2.1



20 Gbps per PORT  
S-RIO switch board

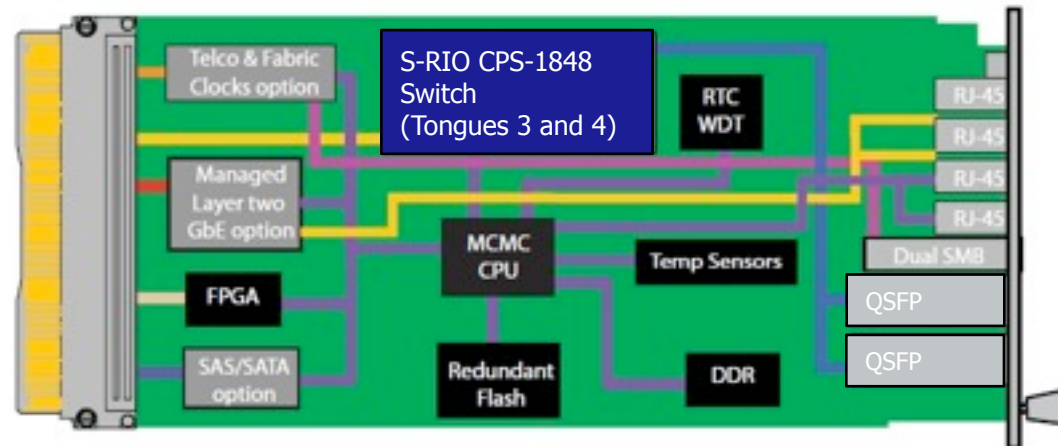
## Provides Switching for MicroTCA System with 12x20 Gbps S-RIO

- 12x4 RapidIO ports @ 20 Gbps
- Enables customers to develop entire signal processing systems with RapidIO in the backplane
- 2x IDT CPS-1848 S-RIO Switch
- Front Panel Connectivity for chassis to chassis system level expansion
- Backward compatible with numerous S-RIO 1.3 AMCs
- AMC.4 connector compliant

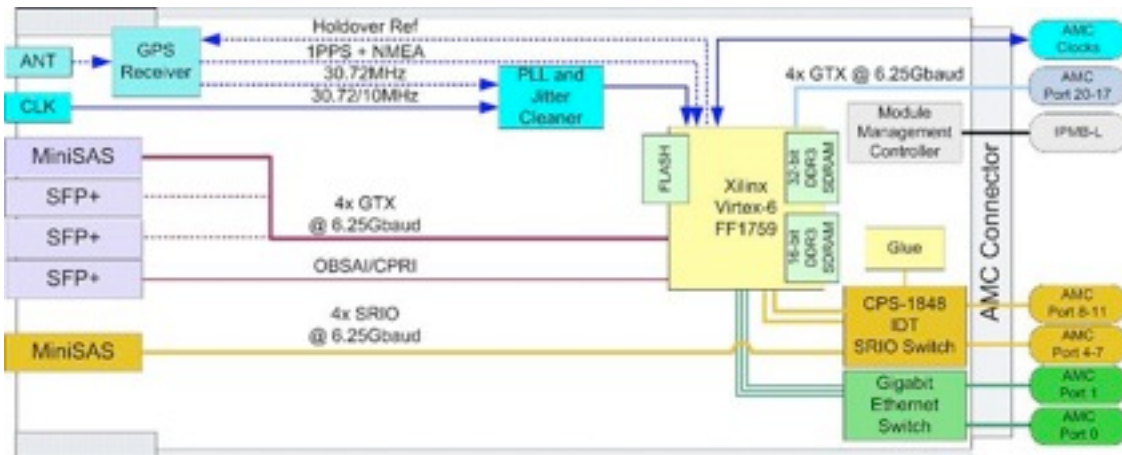


## Feature Rich MicroTCA Carrier Hub (MCH) for Chassis Management utilizing RapidIO Gen 2 Performance and Features

- Supports up to 6.25Gbps RapidIO connectivity between MCHs
- Software provided for both carrier and shelf management
- IDT CPS-1848 Gen2 S-RIO Switch
- Two 4x Front Panel S-RIO Connectivity (20Gbps each) for  $\mu$ TCA chassis to chassis expansion
- Backward compatible with existing S-RIO 1.3 AMCs

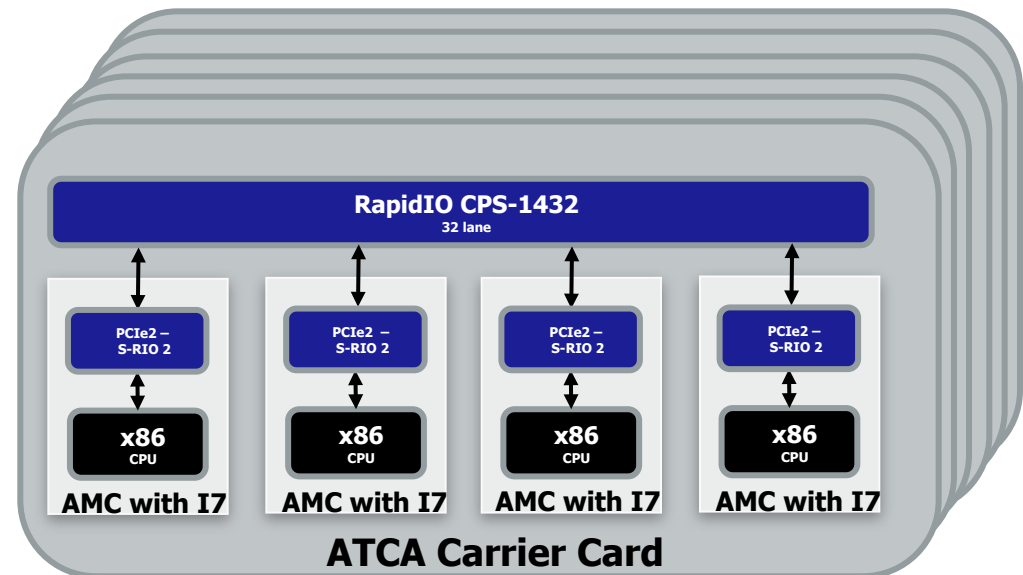
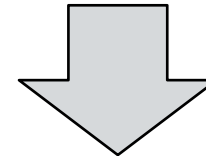


- Complete MAC-PHY Signal Processing Solution for wireless
  - Xilinx Virtex-6 FPGA (LX240T-2 FPGA standard build - up to LX550T-2 FPGA possible)
  - IDT CPS-1848 Serial RapidIO switch
    - SRIO V2.1 at up to 20Gbps per port with multiple x4 ports to backplane and front panel
  - Dual mini-SAS, to FPGA and SRIO - flexible, high-speed cabled connectivity
  - Three front panel SFP+ optical interfaces configurable as CPRI, OBSAI, GigE, S-RIO
  - Upcoming Expansion for DSP daughter card





- RapidIO + PCIe to S-RIO bridging + Intel CPUs deployed in Mil Aero already
- Ideal for servers and high performance computing
- ATCA/MicroTCA based architectures happening NOW
- Low latency, low power, high scalability





# Backup



Ethernet

RapidIO ↔ Ethernet  
Encapsulation

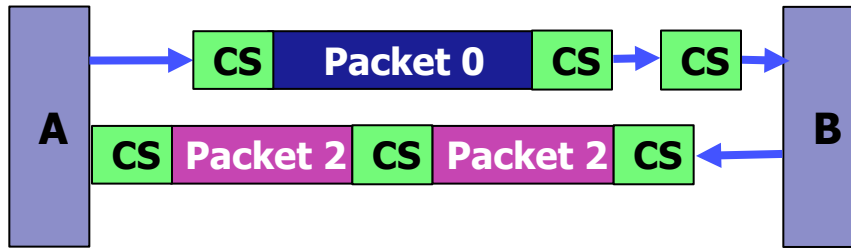
Native RapidIO  
Messaging



- Map IP addresses to/from RapidIO Device IDs
- Map VLAN/MPLS priority to/from RapidIO priority
- Could implement TCP/IP stack for small systems
- Straight forward implementation using Type 9
- Enabled by IDT IP!

Nothing To Send   
 Control Symbol  
 4 or 8 bytes

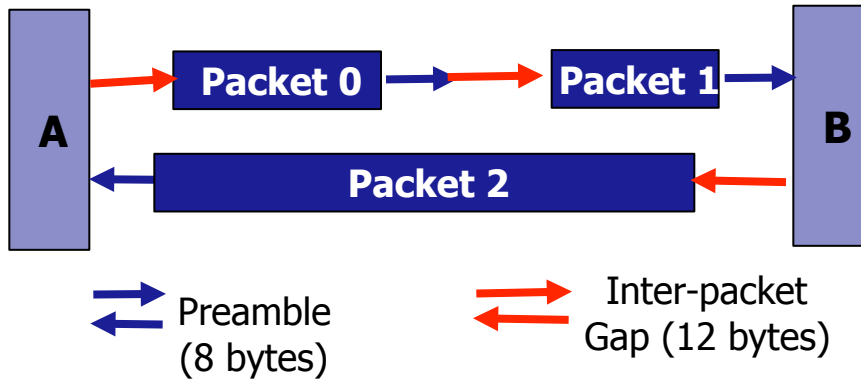
RapidIO  
Physical  
Layer



RapidIO uses control symbols to guarantee packet delivery.

Fabric ordering rules guarantee in order delivery.

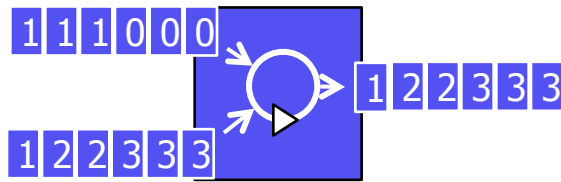
Ethernet  
Physical  
Layer



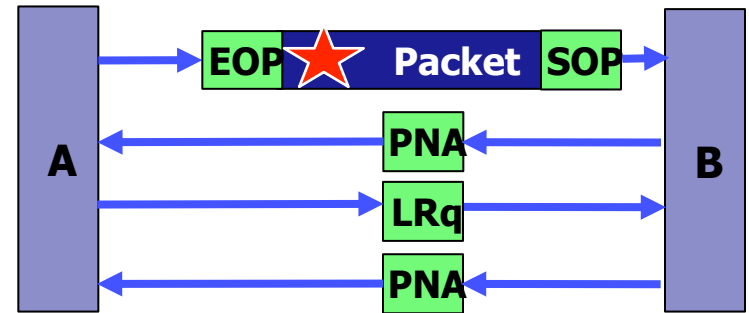
Ethernet acknowledges packets with other packets.

Ethernet is expected to drop packets under congestion and with transmission errors.

## RapidIO

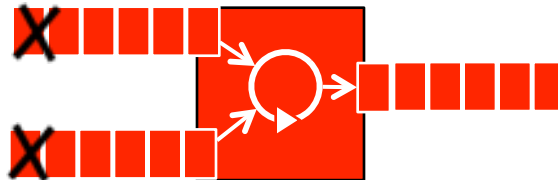


- High priority packets sent first under congestion
- Lossless delivery guaranteed

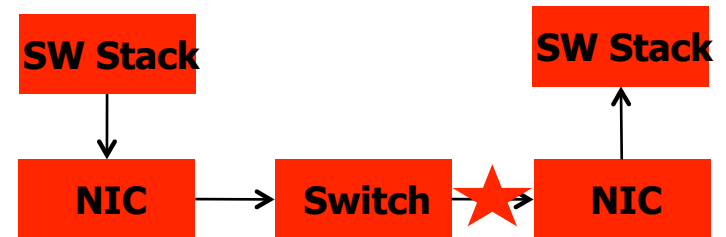


- Error recovery in < 300 nsec

## Ethernet



- Packets discarded due to congestion
- Transmission errors cause packet discard



- Recovery from packet discard requires timeout in TX SW Stack/TOE
- Timeout must be set high to account for congestion in the switch/network (milliseconds to seconds)

## RapidIO

8 Byte Minimum  
276 Byte Max

Packet Size

## Ethernet

64 byte Minimum  
9216 byte Max

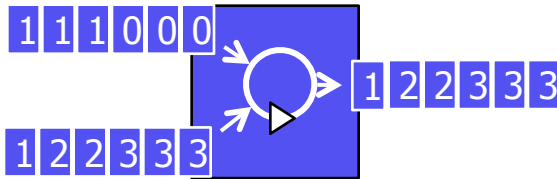


- Small HOL impact per packet
- Flow control control symbols embedded within packets



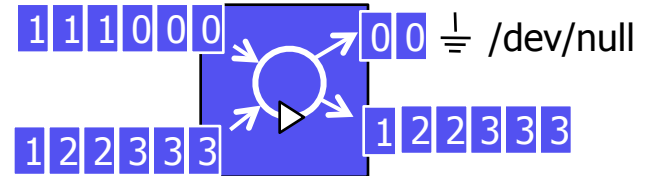
- Large HOL impact per packet
- Flow control packets delayed by larger packets

Head of Line Blocking Impact



- High priority packets sent first
- Guaranteed Delivery

Latency Guarantees Under Congestion



- High "priority" packets sent first
- Packets discarded to clear congestion

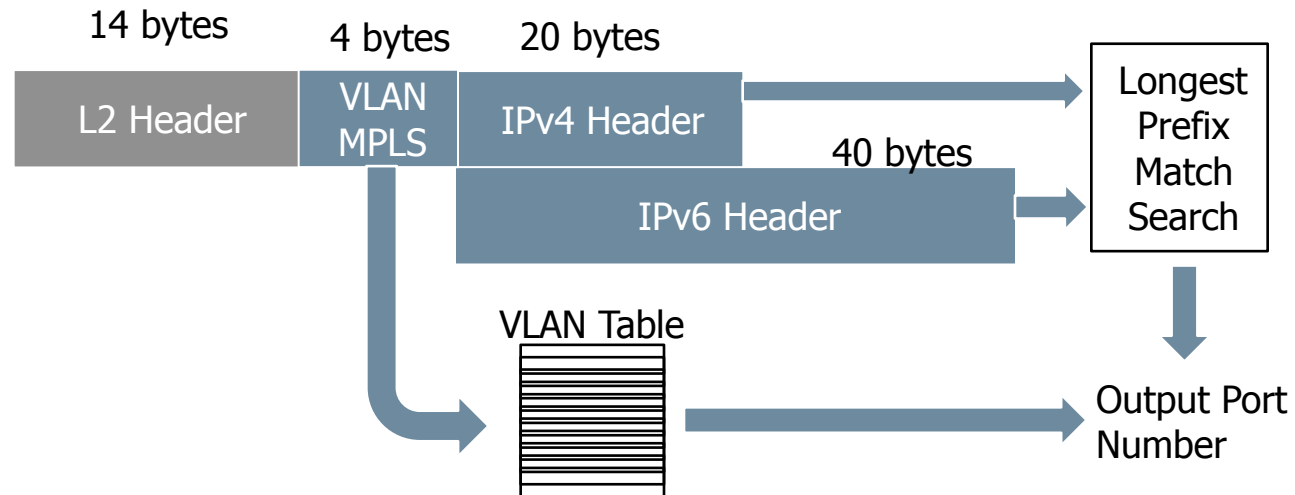
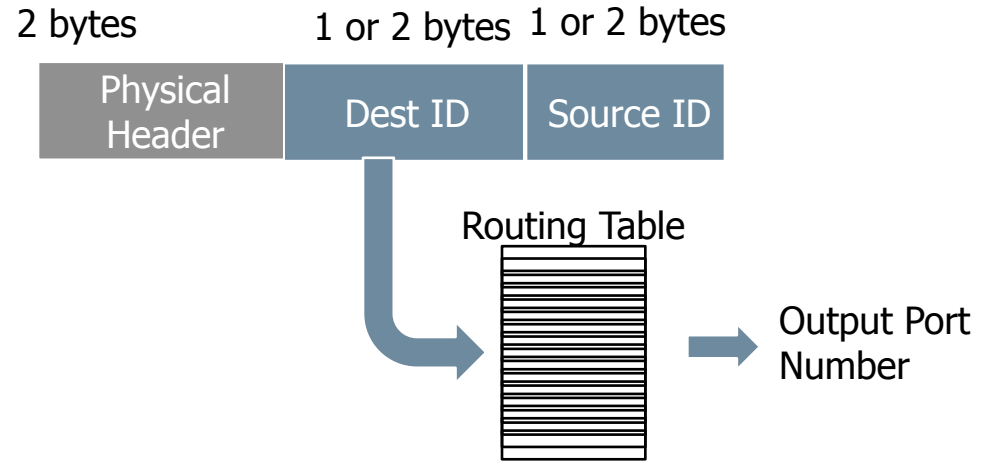
- RapidIO Virtual Channels (VC)
- VC0 – Latency Guarantee
- VC1-8 – Throughput Guarantee



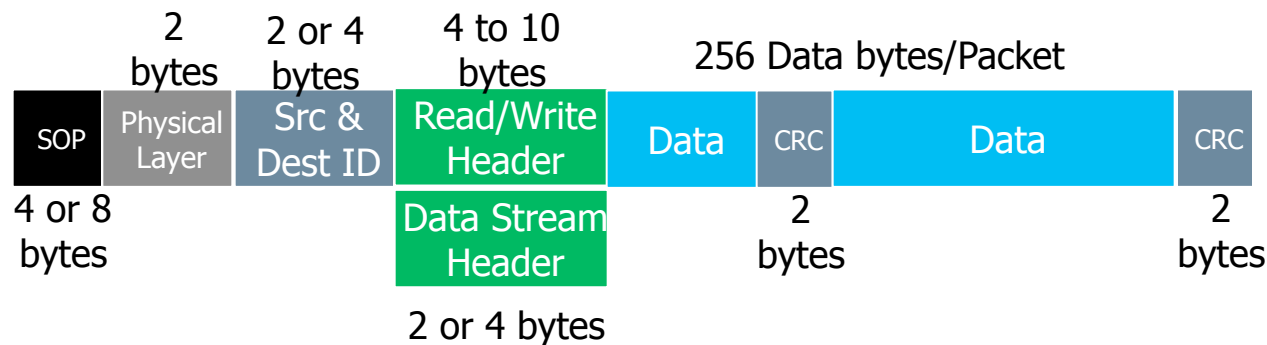
Throughput Guarantees

- TCP protocol slow to adapt, ineffective in DCE
- VPN complex scheduling
- Deep packet inspection & separate hardware may be required!

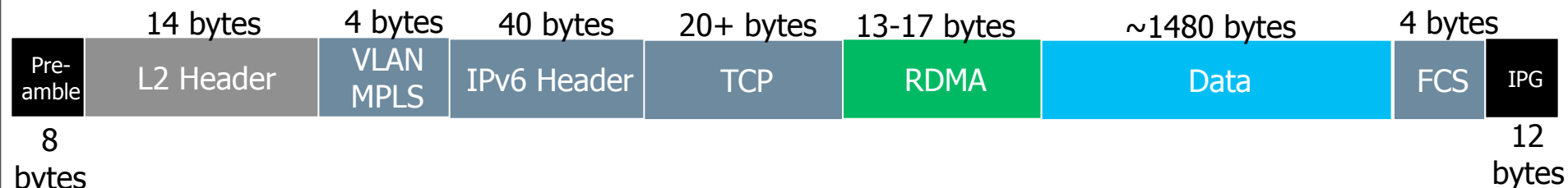
- RapidIO maps Dest ID to Output Port using an indexing operation.
- Routing can begin after the first 4 bytes of the packet have been received, minimizing latency
- Ethernet routing requires packet parsing after L2 Header
- IP routing uses expensive Longest Prefix Match search for 4 (IPv4) or 8 (IPv6) byte addresses, which increases latency and requires additional power i.e external TCAM
- VLAN/MPLS routing uses indexing operation, but also requires pushing/popping VLAN/ MPLS tags from the packet, which requires more latency.



## RapidIO



## Ethernet

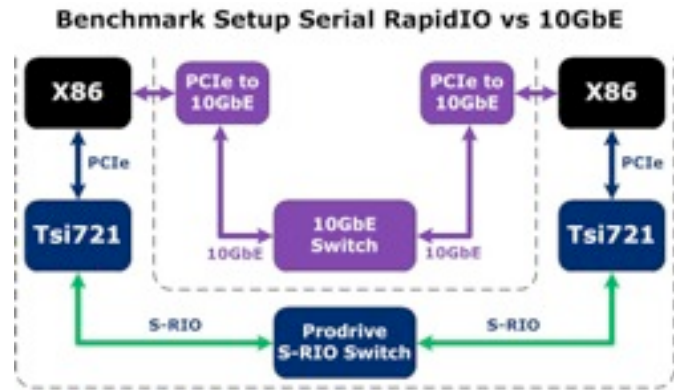


Transfer	256 B Message (Bytes)	4K RDMA
RapidIO	276	4372 (16 packets)
Ethernet	378	4501 (3 packets)

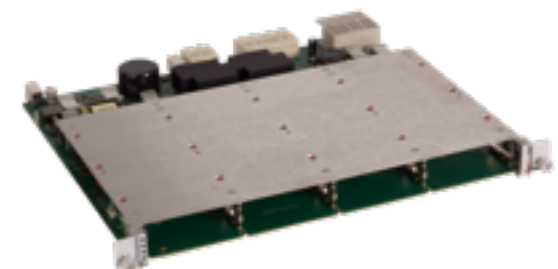
## IDT + Prodrive Technology Benchmark Exercise

### Attribute

- Throughput
  - Maximum messaging rate: different message sizes
  - Differentiated messaging throughput
- Latency and jitter performance
  - Basic latency: different message sizes
  - Latency under error conditions: demonstrate throughput / error recovery
  - Latency variation (jitter) in standard conditions and under error conditions
- Quality of Service (QoS)
  - Latency variation for each priority message
- Power optimization
  - demonstrate power efficiency with lower speed links



Setup 1: Serial RapidIO 1.3  
Setup 2: 10GbE



Quad AMC carrier blade (Switch) based on S-RIO 1.3 and 2.1



20 Gbps per PORT S-RIO switch board