

Comparing Ethernet and RapidIO®

Greg Shippen
System Architect
Freescale Semiconductor



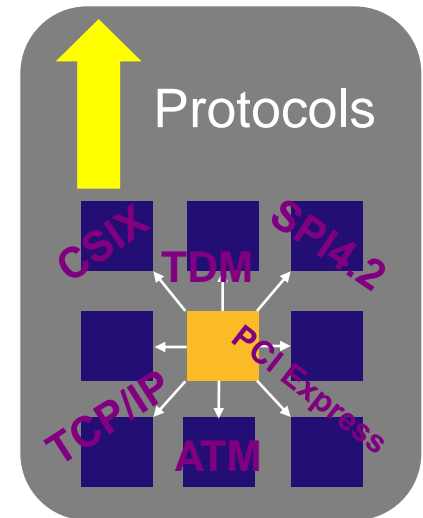
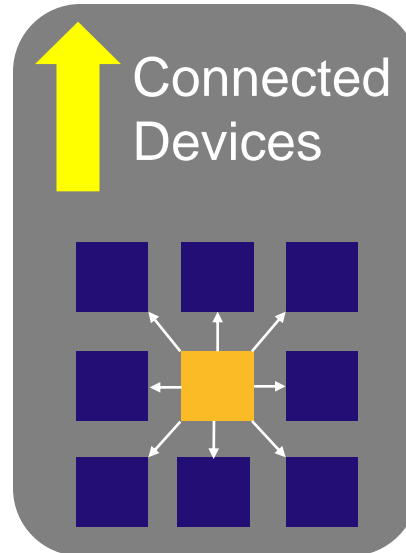
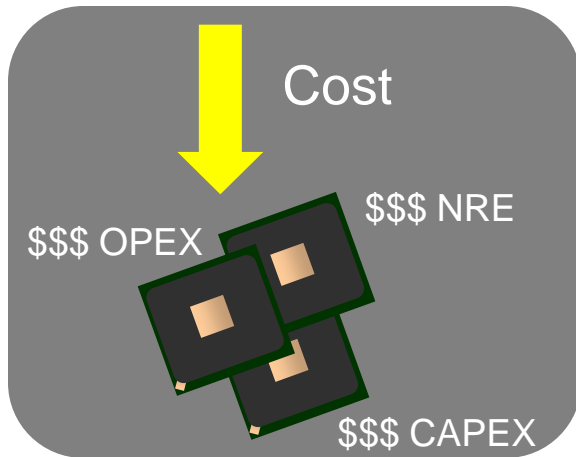
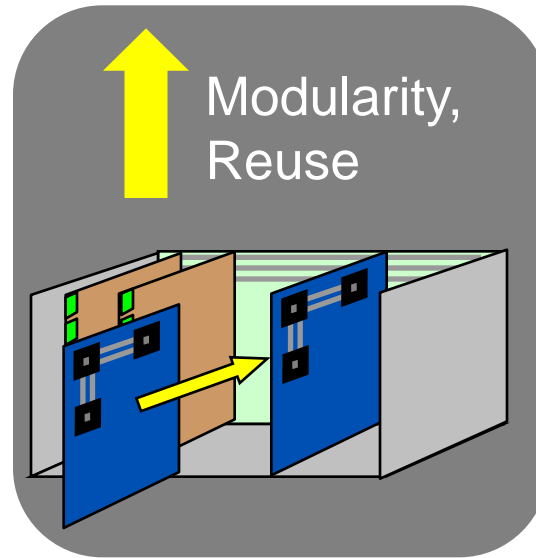
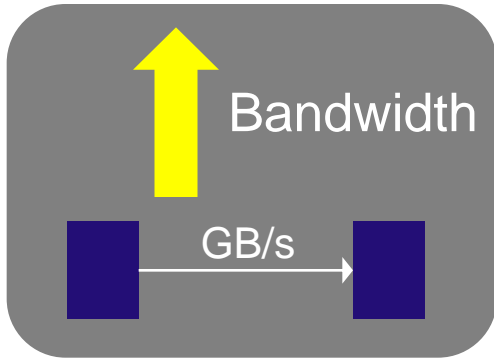
The Embedded Fabric Choice

Agenda



- Interconnect Trends
- Technical Overview
- Comparison
- Applications
- Conclusion

Market Trends



Interconnect Trends



- **1st Generation Point-to-Point**

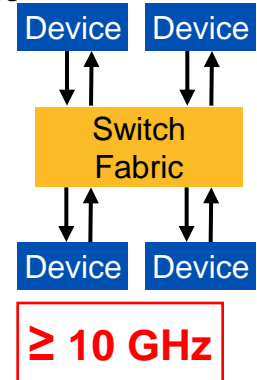
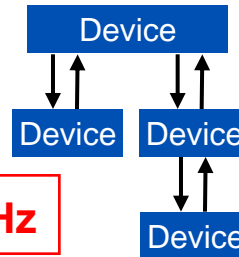
- Packet switched
- PHY: Source-sync differential
- Lower pin count

Example: HT/P-RapidIO

≤ 3 GHz

- **2nd Generation Point-to-Point**

- Packet switched
- PHY: SERDES differential
- Lowest pin count

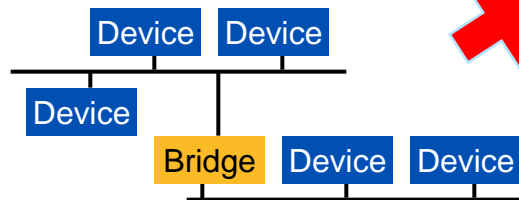


≥ 10 GHz

Ex: PCIe, Serial RapidIO, SATA, SAS

- **Hierarchical Bus**

- Bridged Hierarchy
- Broadcast
- PHY: Single-ended

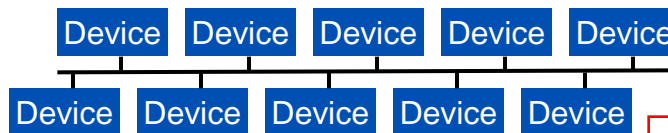


Example: PCI/PCI-X/SCSI

≤ 133MHz

- **Shared Bus**

- Single segment
- Broadcast
- PHY: Single-ended
- Highest pin count

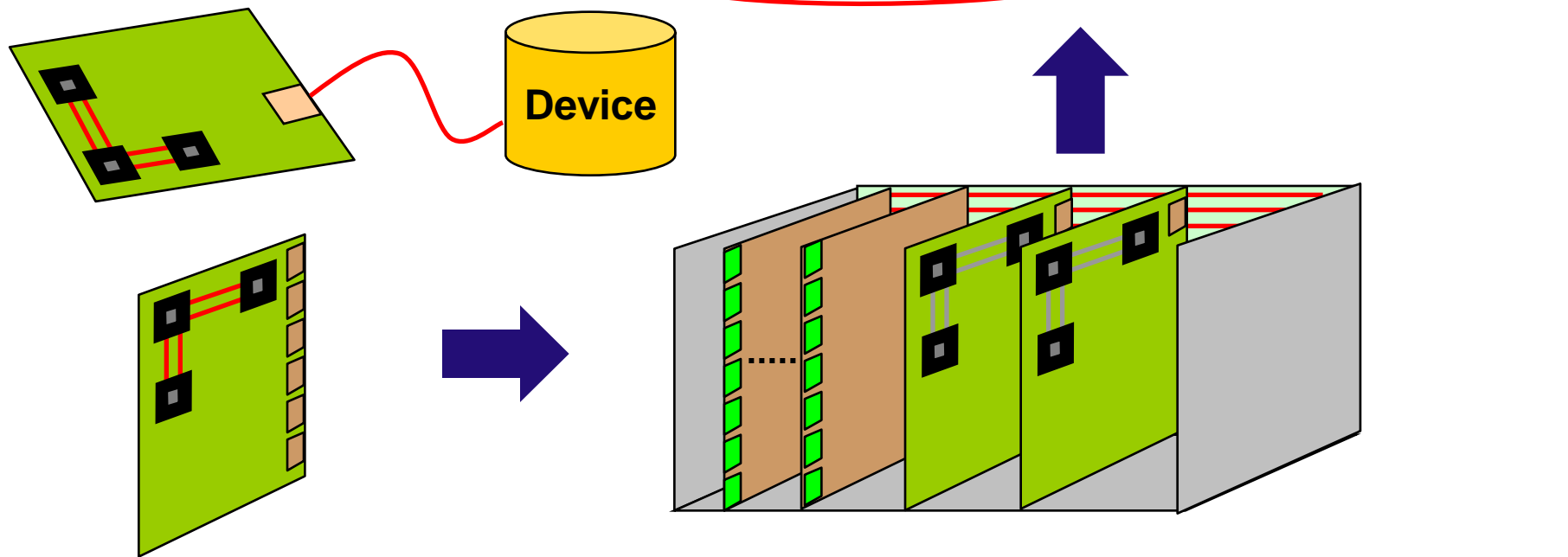


Example: VME

≤ 66MHz

Interconnect Roles

- Chip-to-chip
- Board-to-Device
- Board-to-board
- Chassis-to-chassis



Interconnect Clarifying Terminology

Serial Interconnect

- Link Protocol
- Point-to-point, no System Addressing



Lightweight

Switched Interconnect

- Serial Protocol
- System Addressing



I/O Load & Store



Ethernet

Datagram (Packets)

Switched Fabric

- IO *and* Packets Protocols
- Robust Link Layer
- Traffic Managed
- CPU offload



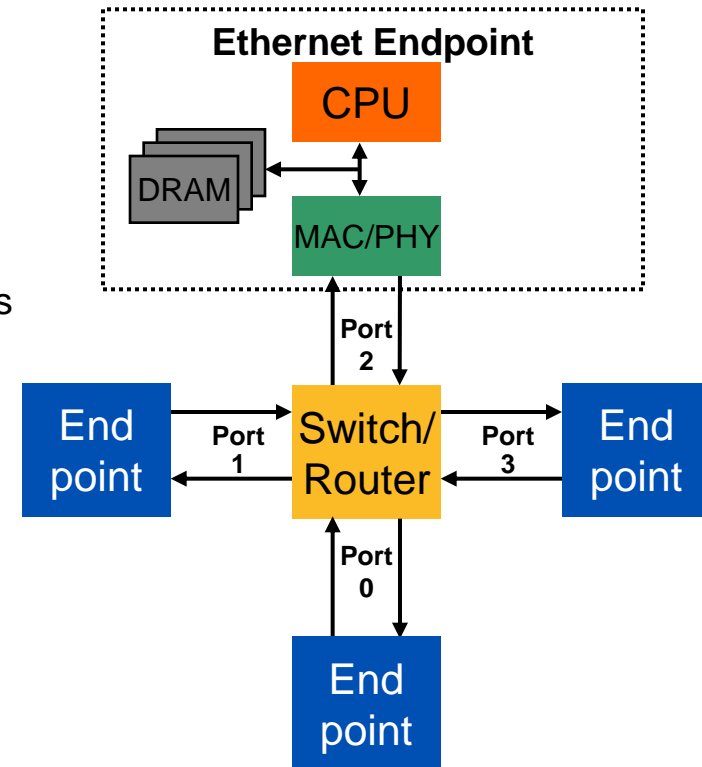
Agenda



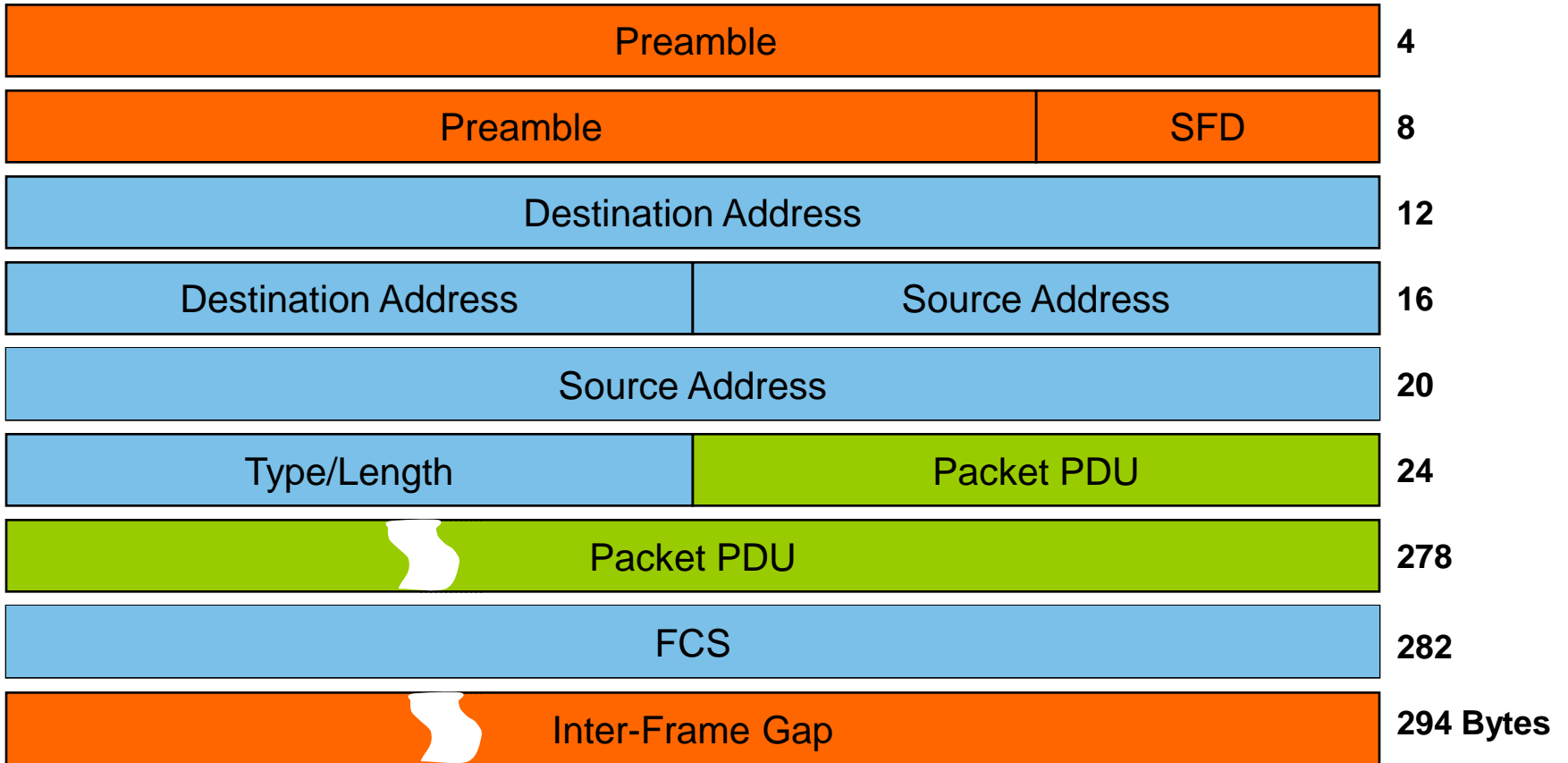
- Interconnect Trends
- Technical Overview
- Comparison
- Applications
- Conclusion

Ethernet Overview

- WAN scale interconnect
 - Box-to-box, board-to-board, chip-to-chip, backplane
 - Connect thousands to millions of endpoints
 - Physical layer defined for LAN-scale interconnection
 - Closet to computer
 - 100+ m distance
- Target market
 - Initially used in aggregation settings
 - High performance switches, routers and LAN backbones
 - Now WAN to workstations, PCs and laptops
- Gigabit Ethernet ubiquitous now
- Gigabit Specification standard completed in 1998
 - Gigabit Copper (1000Base-T) in 1999
 - No 10G Copper PHY defined yet
 - Various MAC-to-PHY Standards defined
- Extensible Layered Specification
- Point-to-point packetized architecture
 - Variable packet size
 - High header overhead
 - 46-1500 byte packet L2 PDU
 - Up to 9000 byte jumbo frames



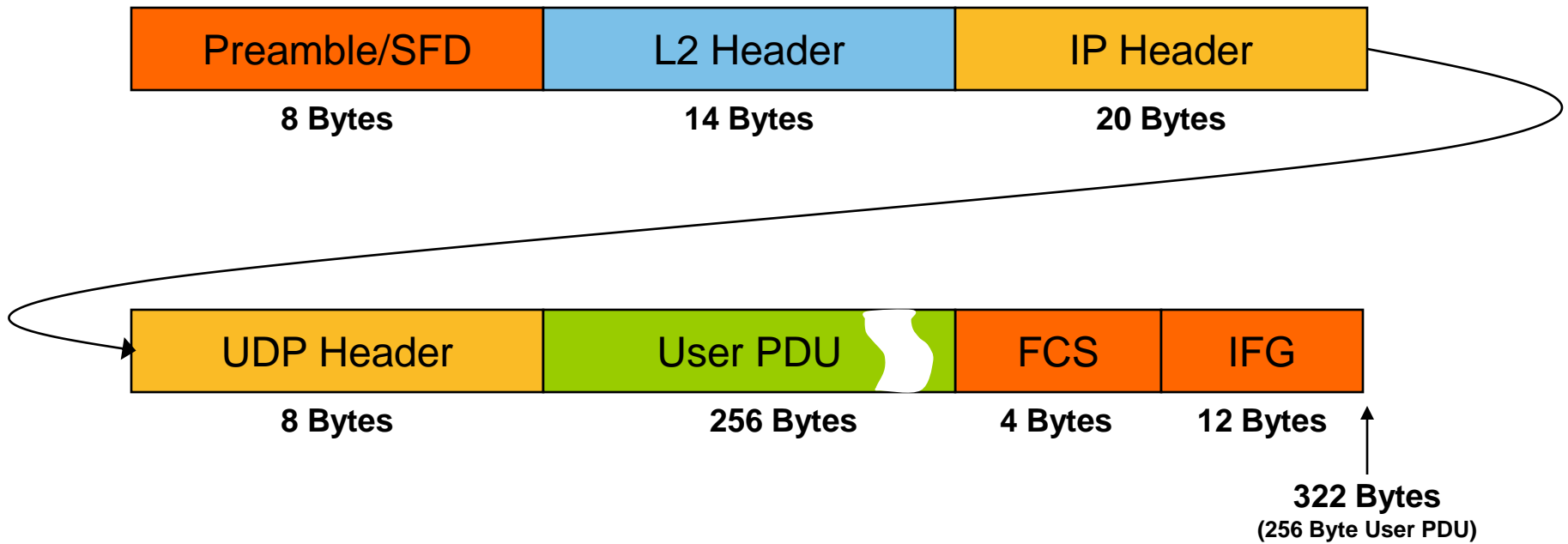
Ethernet Layer 2



Layer 2 Packet Type: 1500 Byte Max Packet PDU

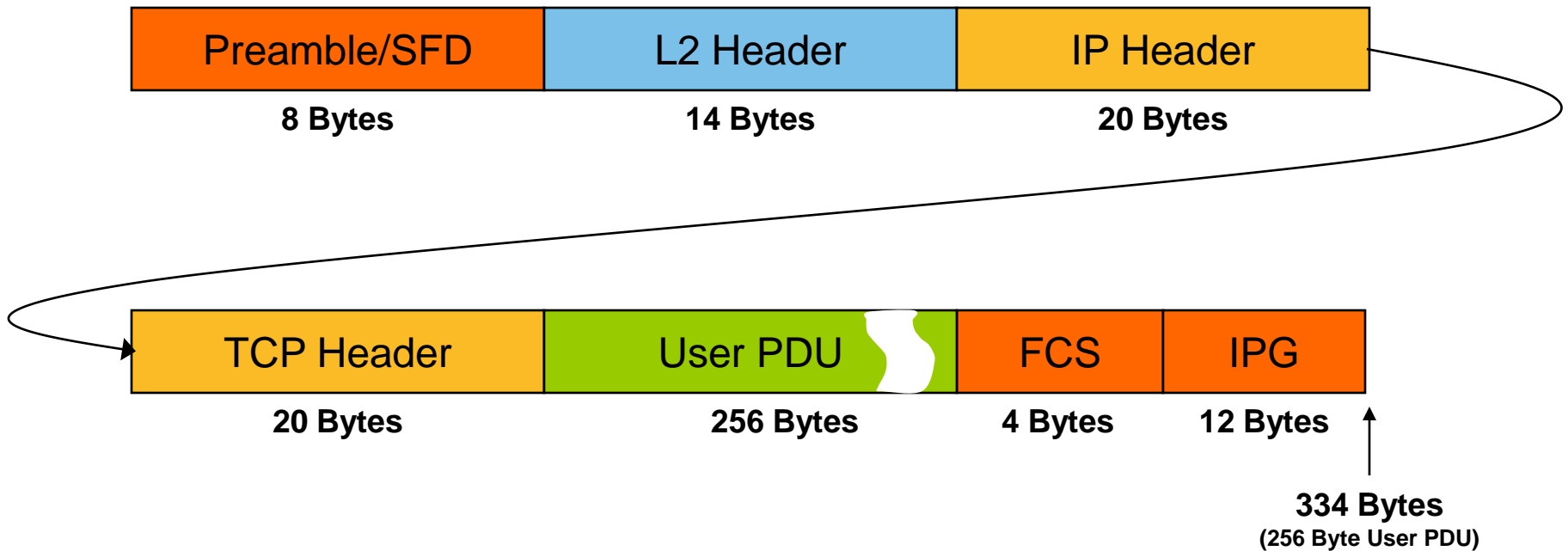
**Total = 294 Bytes
(256 Byte PDU)**

Ethernet + UDP



UDP Packet Type: 1472 byte User PDU

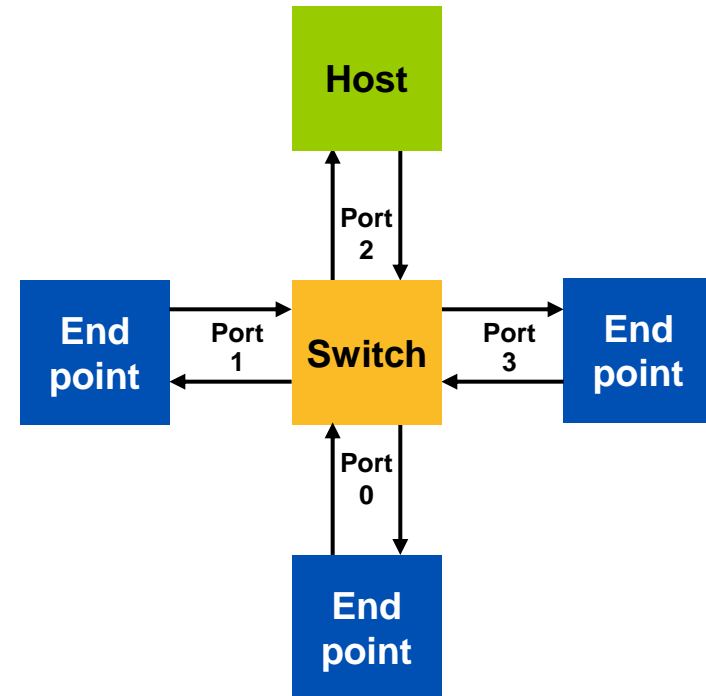
Ethernet + TCP/IP



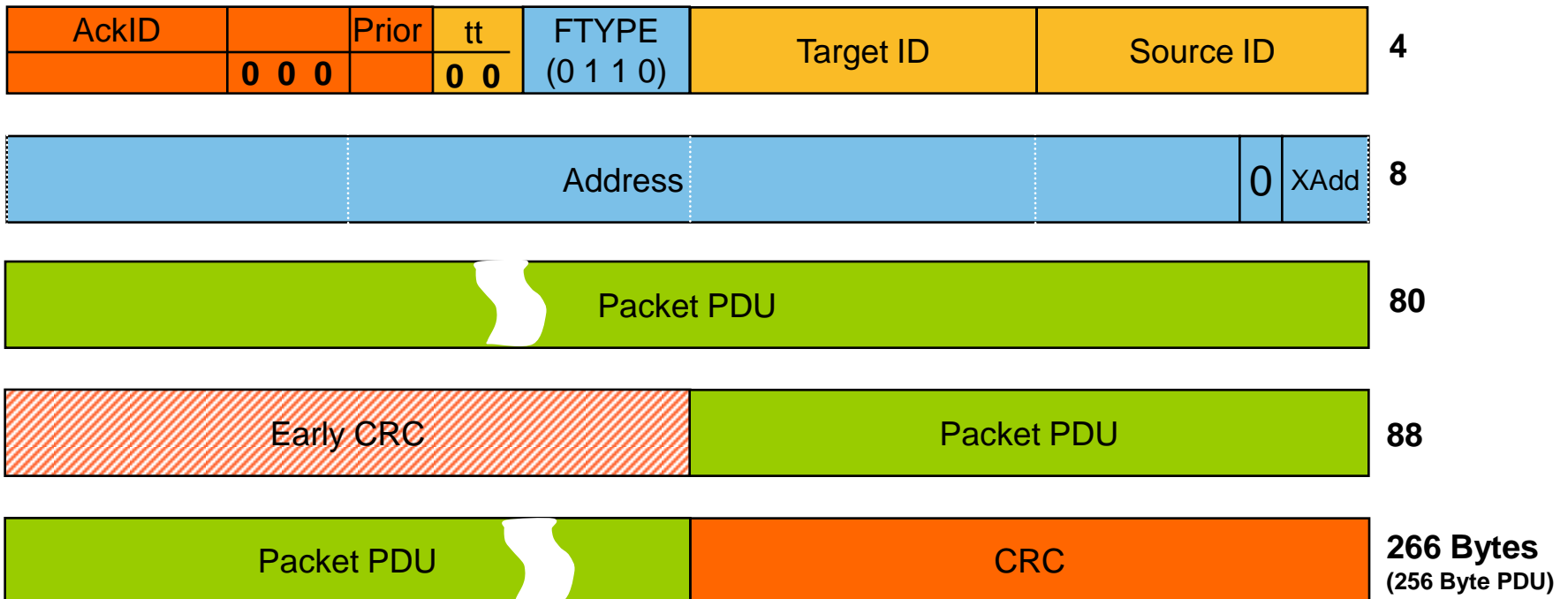
TCP/IP Packet Type: 1460 Byte Max User PDU

RapidIO Overview

- Chassis scale interconnect
 - Chip-to-chip, Board-to-board via connector or cabling
 - Physical layer defined for backplane interconnection
 - ~80-100 cm + 2 connectors (Serial)
- Target market
 - Embedded systems
 - Compute, defense, networking & telecom line cards
 - CPU I/O, Line-card aggregation, backplane
 - Serial PHY allowed expansion to data plane
 - Flow control, encapsulation, streams
- Initially a processor interconnect
 - First spec a Motorola & Mercury collaboration
- First revision standard completed in 1999
 - Rolled out with processors, bridges and switches
 - Parallel 8-bit RapidIO @ 500 MHz applied clock
 - 5-6G PHY, 2, 8 and 16x lanes + Virtual Channels nearing completion
- Extensible Architecture
 - Layered architecture
- Point-to-point packetized architecture
 - Low overhead
 - Variable packet size
 - Maximum 256 byte PDU
 - SAR support for 4 K-byte messages

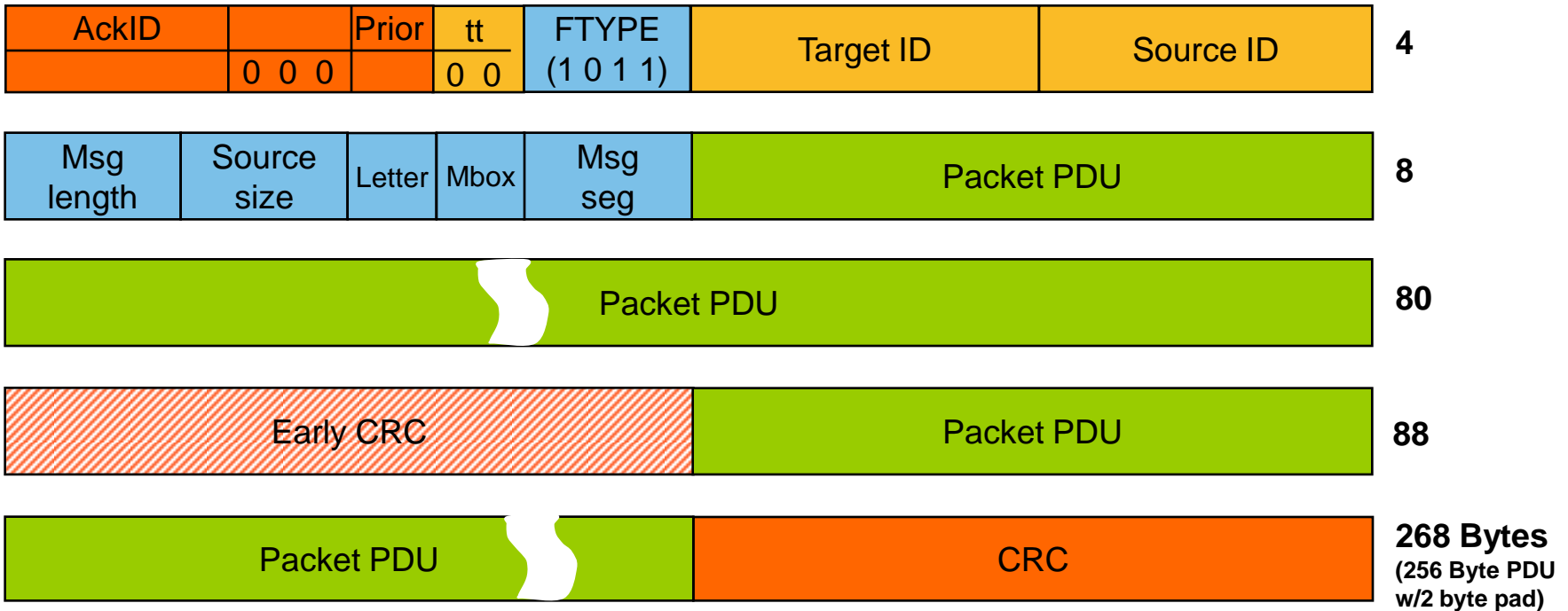


RapidIO Packet Format: SWRITE



SWRITE Packet Type: 256 Byte Max Packet PDU

RapidIO Packet Format: Message



Message Packet Type: 256 Byte Max Packet PDU, 4K w/SAR

Agenda



- Interconnect Trends
- Technical Overview
- Comparison
- Applications
- Conclusion

Logical Layer Comparison



	Ethernet		RapidIO
	Layer 2	Layer 3+	Logical Layer
Payload Size (Bytes)	46-1500 (802.3) 46-9192 ^a (Jumbo)	26-1460 (802.3) 26-9172 ^b (Jumbo)	1-256
Memory-mapped R/W	No	RDMA	Read/Write Atomics Configuration
Write w/Response	No	No	Yes
Address Size	N/A	64-bits (RDMA)	34, 50, 66-bits
Messaging Support	No	TCP	4KB user payloads, Doorbells
Datagram Support	Yes	Yes	64KB user payloads ^c
Globally Shared Memory	No	No	Yes
Deadlock Avoidance	N/A	L3+ must address	Pervasive HW Support

^aLargest common jumbo frame size

^bAssumes 20-byte IP Header

^cDataplane Extensions feature

Transport Layer Comparison



	Ethernet		RapidIO
	Layer 2	Layer 3+	Logical Layer
Topologies	Any	Any	Any
Delivery Service	Best Effort	Guaranteed (TCP, SCTP, others)	Guaranteed Best Effort ^a
Routing	MAC Address	IP Address	Device ID
Maximum Endpoints	2 ⁴⁸	2 ³² (IPv4) 2 ¹²⁸ (IPv6)	2 ⁸ (Small) 2 ¹⁶ (Large)
Header Fields which change link-to-link	None	TTL, MAC Addr, FCS	AckID Hop Count, CRC (Maint. Only)
Redundant Link Support	Yes	Yes	Yes

^aData Streaming Spec

Physical Layer Comparison

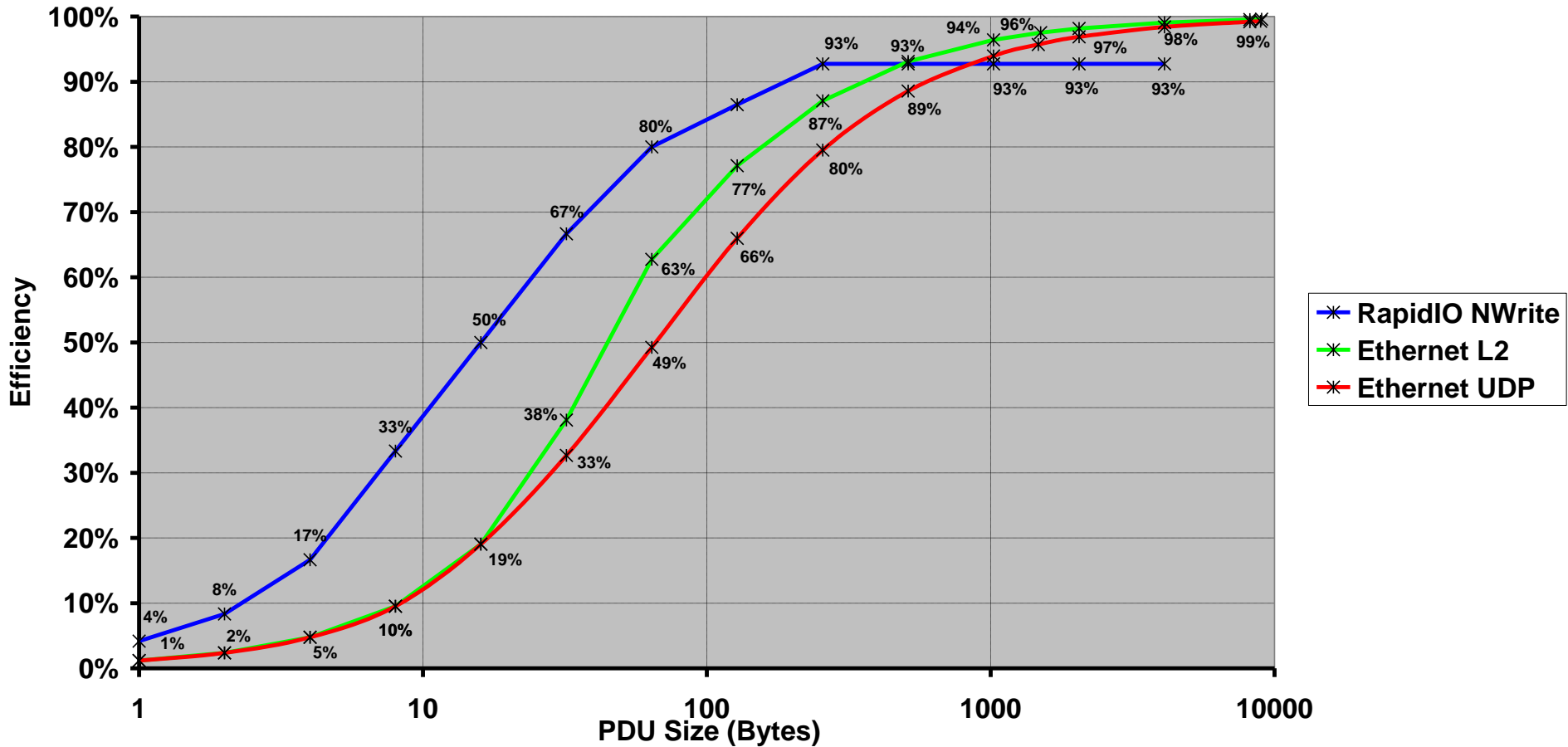


	Gigabit Ethernet		RapidIO	
	1000Base-T	SERDES ^a	LP-LVDS	LP-Serial
Channel	100m Cat5 Cable	~50cm FR4	~50-80cm FR4 + 2 connectors	~80-100cm FR4 + 2 Connectors
Data rate (Per pair)	1 Gbps	1 Gbps	500-2000 Mbps	1, 2, 2.5 Gbps
Symbol rate (Per pair)	125 Mbaud	1.25 Gbaud	250-1000 Mbaud	1.25, 2.5, 3.125 Gbaud
Encoding	8b → 4 quinary symbols	8b → 10b symbols	DDR	8b → 10b symbols
Signaling	5 Layer PAM Code (Multilevel Signaling)	NRZ	NRZ	NRZ
Signal Pairs (Per Direction)	4 ^b	1	10, 19	1, 2, 4, 8, 16
Electricals	Custom	XAUI	LVDS	XAUI (w/Long & Short Reach)
Clocking	Embedded	Embedded	Source Sync.	Embedded

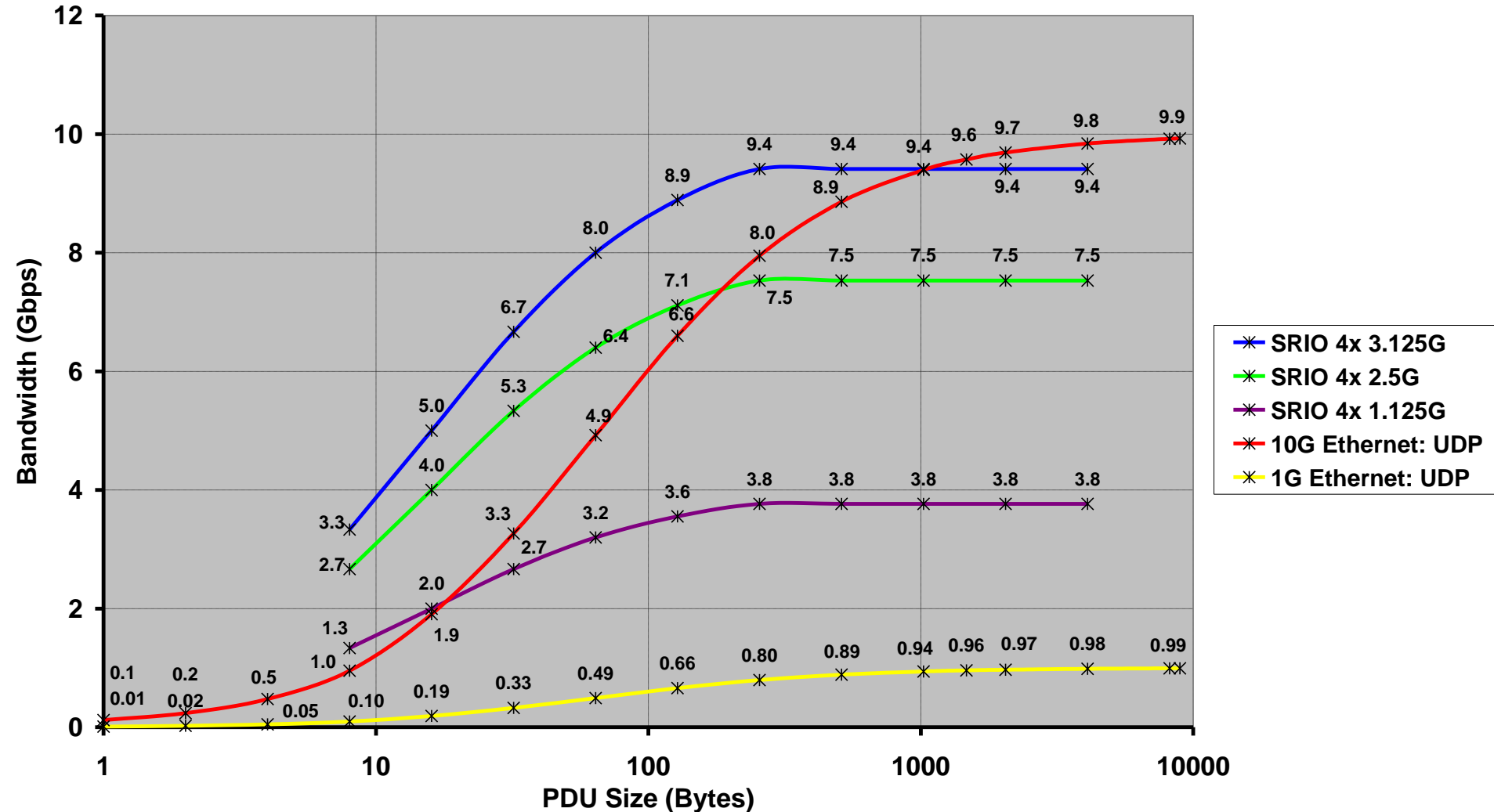
^aDefacto standard. Some using 1000Base-CX electricals over backplanes

^bEach pair carries both Tx and Rx

Protocol Efficiency



Effective Bandwidth



Quality-of-Service (QoS) Dependencies

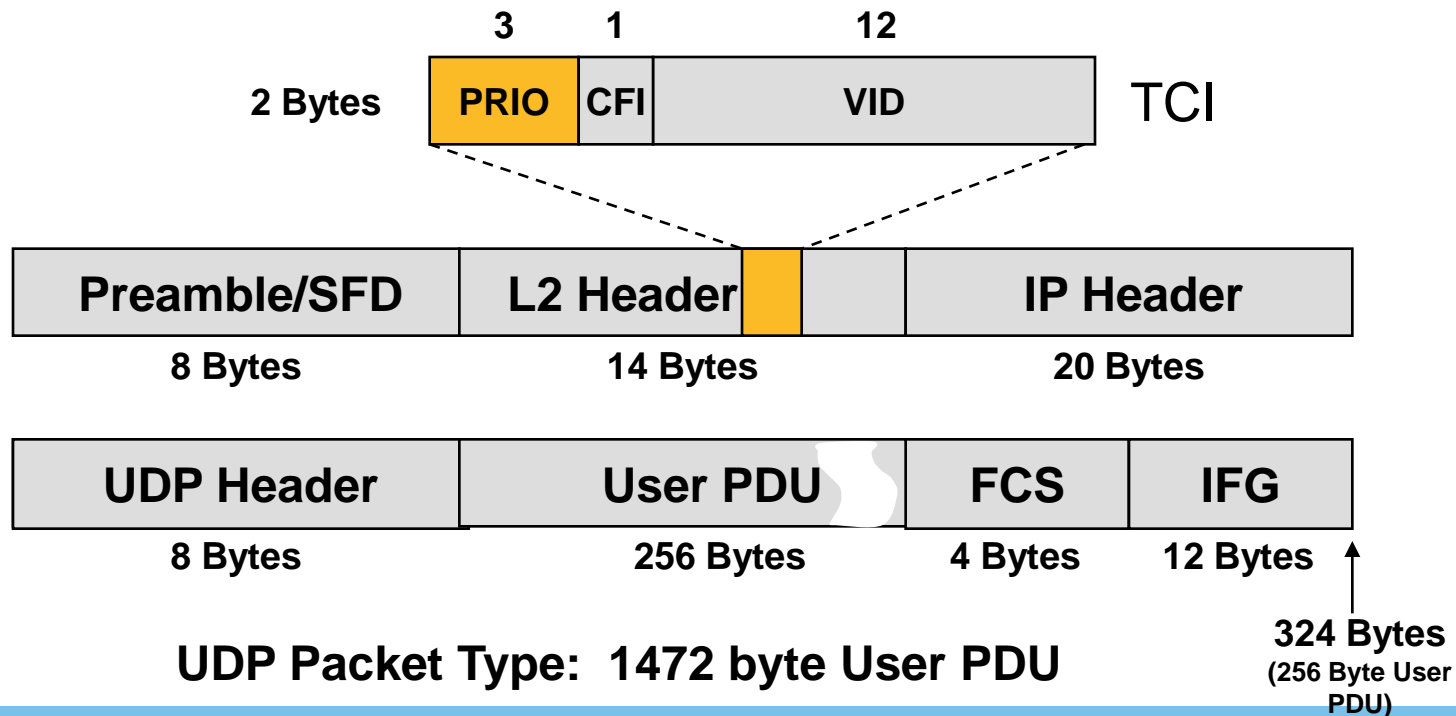
QoS depends on proper hooks across the interconnect fabric

- Hierarchical Flow Control
 - Addresses short, medium and long-term congestion events
 - Link and end-to-end
- Ability to define many streams of traffic
 - Often defined as a logical sequence of transactions between two endpoints
- Ability to differentiate classes of traffic among streams
- Ability to reserve and allocate bandwidth to streams and classes



QoS Comparison: Ethernet

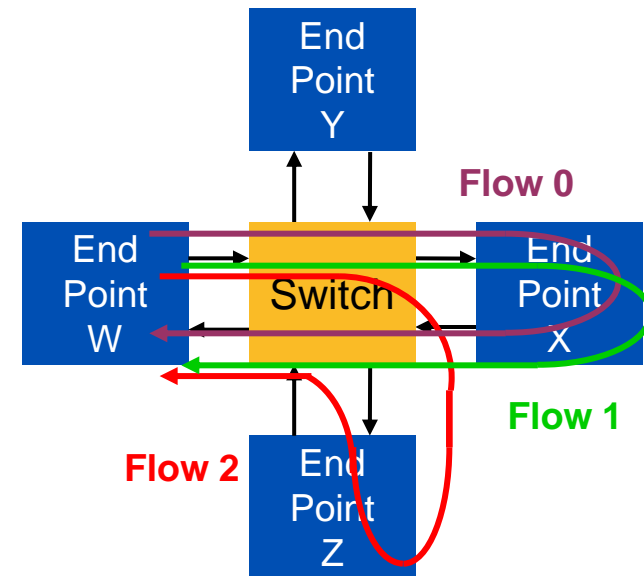
- No universal QoS standard
- Some Layer 2+ switches support Priority Tagging (802.1d/q)
 - Eight classes
- Increasing number of routers support MPLS at L3



QoS Comparison: RapidIO



- All implementations must support 3 prioritized flows
 - No ordering between flows
 - Allows shared buffer pool across flows
- Switches required to provide some improved service
 - Extent of improvement is implementation dependant
- Dataplane Extensions adds carrier-grade QoS
 - Support for thousands of flows, hundreds of traffic classes
 - End-to-end traffic management



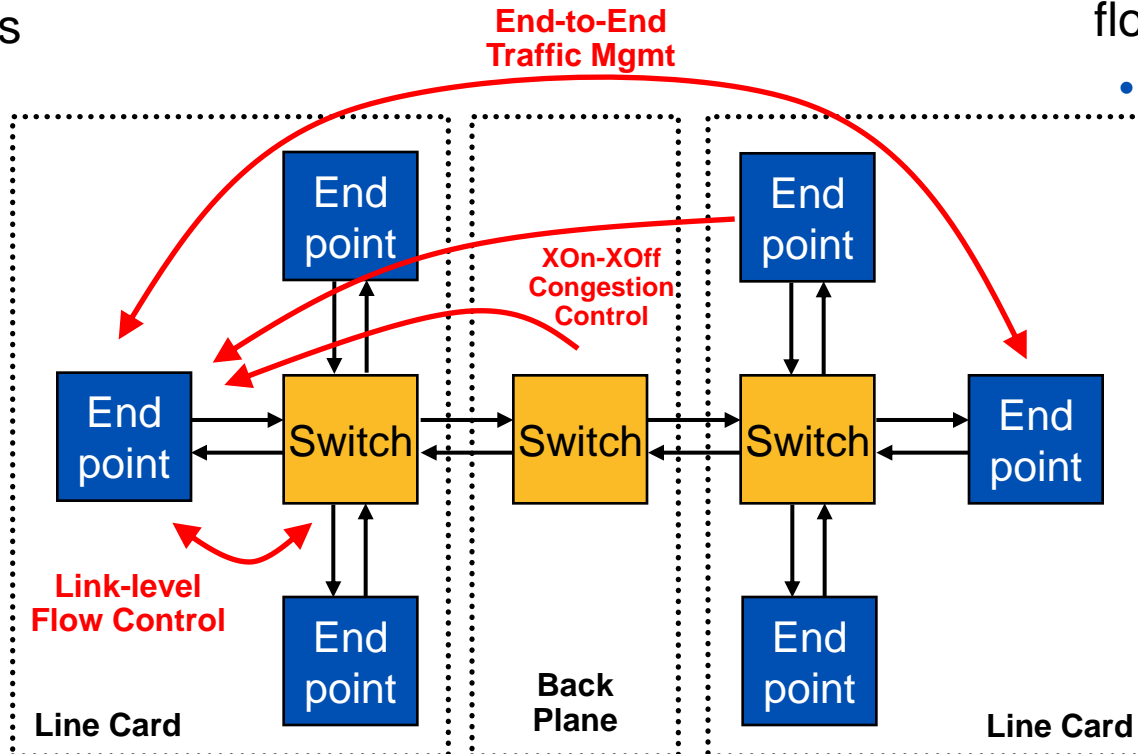
Flow Control Comparison

Ethernet

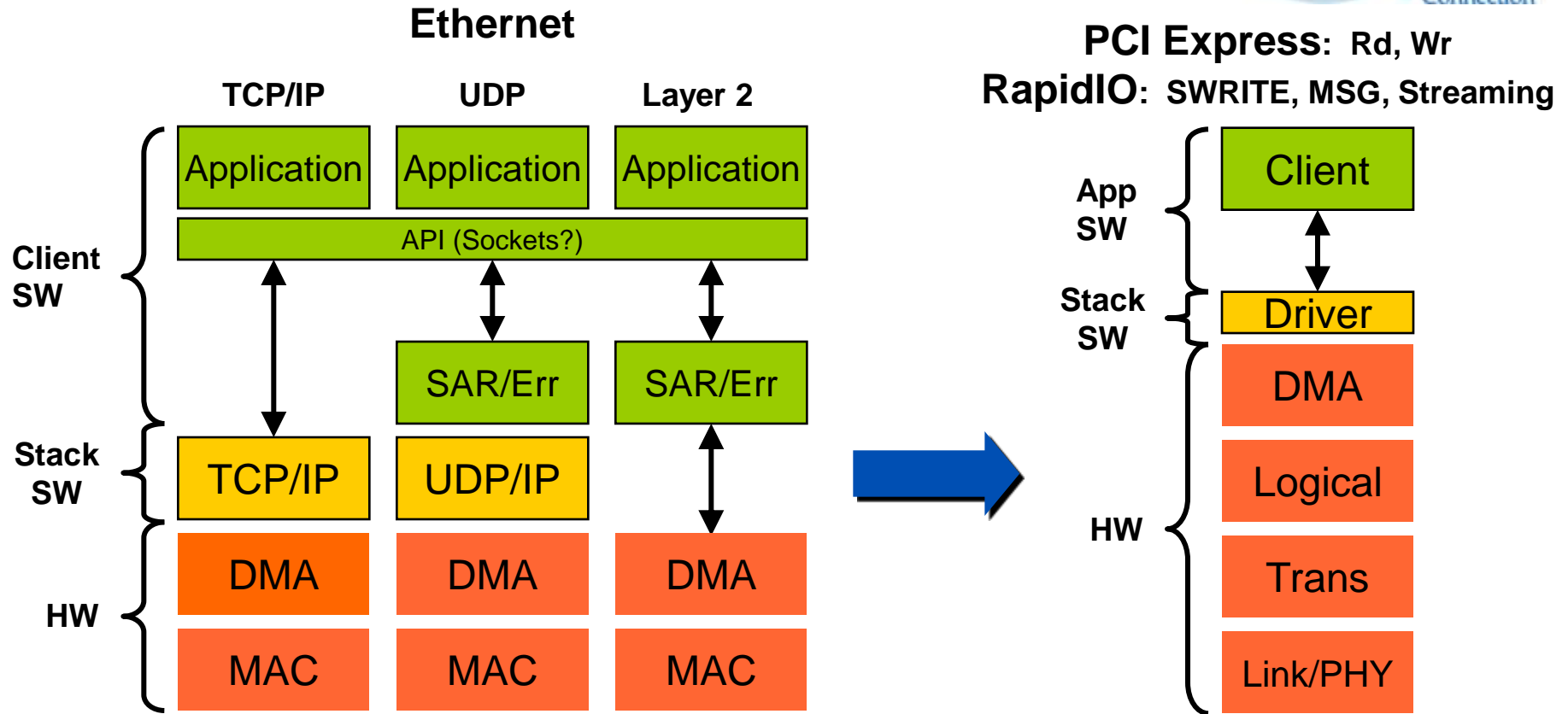
- Link-to-link flow control
 - PAUSE frames
- L3+ end-to-end flow control
 - ECN, TCP windowing, others

RapidIO

- Link-to-link flow control
- Congestion control
 - XON, XOFF
- Fine-grained end-to-end flow control
 - Data Streaming Logical Layer



Software/Hardware Interface



- High-bandwidth interconnects require low CPU overhead usage model
 - Hardware support for logical, transport and link layer
 - Low overhead DMA with QoS support

Ethernet Performance



- Microsecond+ fall through latencies (~100us?)
 - Not just the hardware, data has to traverse the SW stack
- High CPU overhead
 - Rule of thumb appears to be borne out in data for TCP/IP SW overhead
 - 1 Hz of CPU per bit of throughput (per direction)
 - Wire speed achievable with GHz class processors
 - Some CPU will be left but how much depends on
 - Protocol being terminated
 - Offload features of GigE interfaces
 - Too often advanced off-load features cannot be leveraged
 - OS & SW stack support issues
- UDP or MAC/Layer 2 solutions sometimes also use proprietary higher layer protocols
 - Can defeat the value of off-the-shelf “standards-based” solution
- Error correction at endpoint stacks introduce latency jitter and determinism issues
- Works well for application bandwidth < ~300Mbps
 - Lack of flow control problematic for systems that can't significantly overprovision

- Latency
 - Sub-microsecond switch latencies
- End-to-end latency
 - Lower latency than Ethernet since latency does not include a SW stack
- Architecture
 - RapidIO switches straightforward and orthogonal in architecture
 - Strict peer-to-peer
 - Packet headers architected to reduce logic
 - No need to recalculate CRC

- RapidIO and Ethernet with modest TCP/IP offload have similar underlying silicon costs
 - Aggressive TCP/IP Offload engine larger than a RapidIO endpoints
 - GigE Copper PHY is very large (~20mm² in 130nm)
- Leveraging Ethernet volume economics not always a reality
 - L2+ Ethernet switches suitable for aggregation and backplanes are not high volume
 - 16-24 ports, QoS features and SERDES PHYs for backplane
 - 12-16 ports, QoS features for aggregation
 - Terminating TCP/IP demands significant processor overhead
 - Dedicate processor or reduce performance and/or application features

Agenda



- Interconnect Trends
- Technical Overview
- Comparison
- Applications
- Conclusion

Summary



- Ethernet widely used in low bandwidth embedded applications
 - Undisputed standard for wide area networks (WAN)
 - 10G Ethernet role in backplane is yet to be determined
 - Broad endpoint silicon and software support
 - Flexible and adaptable software protocol stack
 - High overhead & latency
 - Significant cost jump for bandwidth above 1Gbps
 - No standard Off-load or backplane SERDES PHY
- RapidIO has captured DSP and line card aggregation role
 - Looking to expand role onto the backplane
 - Low overhead protocol supporting both control and data plane
 - Superior Quality-of-Service
 - Variety of PHY speeds
 - Cost competitive against 1G and 10G Ethernet
 - Growing ecosystem

Application Fit By Bandwidth



Application	Ethernet	RapidIO
Interconnect Bandwidth < 300 Mbps		
Control Plane (Low latency, Reliable)	Stack Latency	Good Fit
Data Plane (QoS, Streams, HA)	Good Fit	Good Fit
300 Mbps < Bandwidth < 1 Gbps		
Control Plane	Bonding, 10Ge	Good Fit
Data Plane	Bonding, 10Ge	Good Fit
1 Gbps < Bandwidth < 10 Gbps		
Control Plane	Bonding, 10Ge	Good Fit
Data Plane	Bonding, 10Ge	Good Fit
Bandwidth > 10 Gbps		
Control Plane	10Ge Bonding	Marginal
Data Plane	10Ge Bonding	Marginal

Good Fit

Marginal

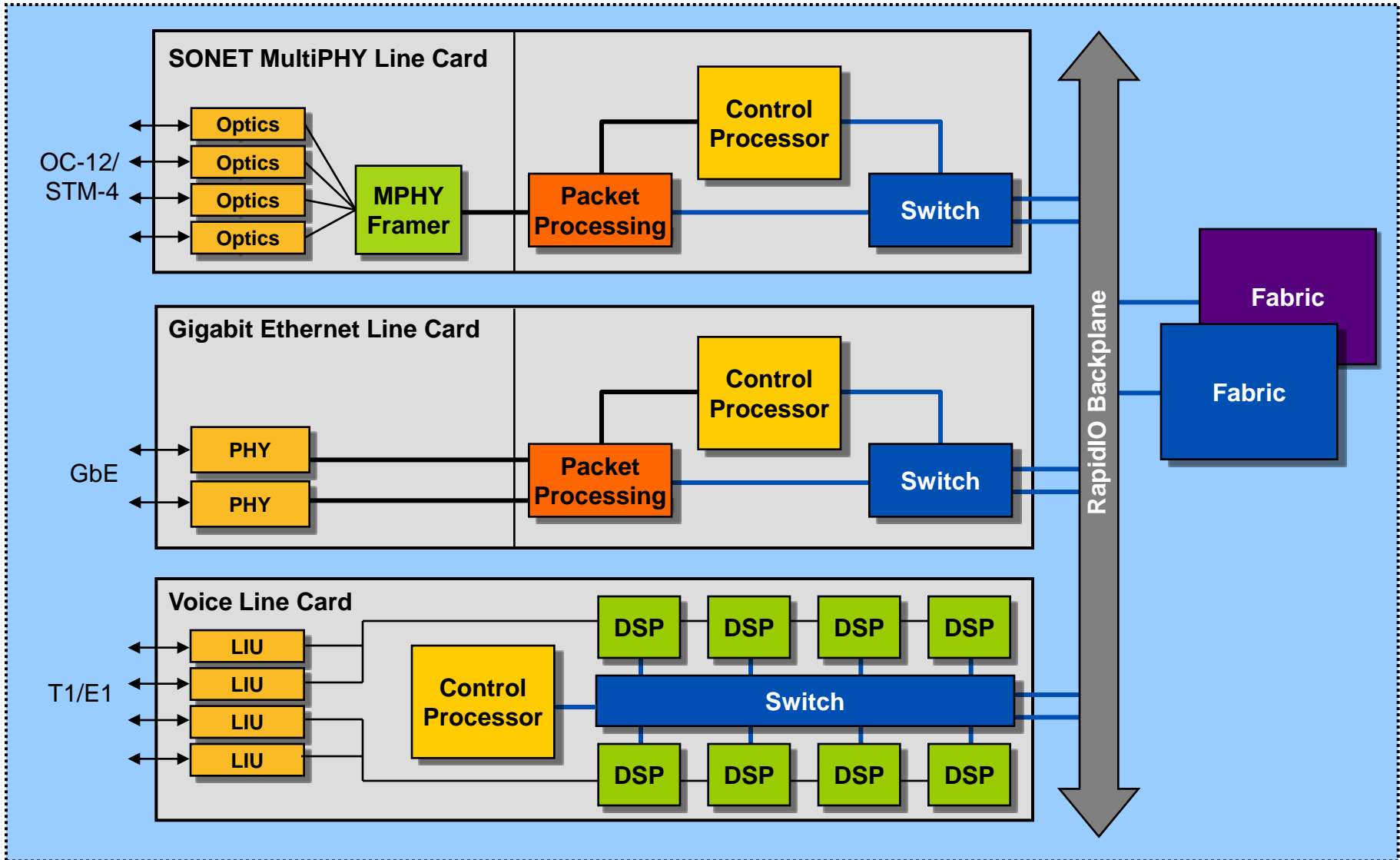
Poor Fit

Application Fit By Attribute

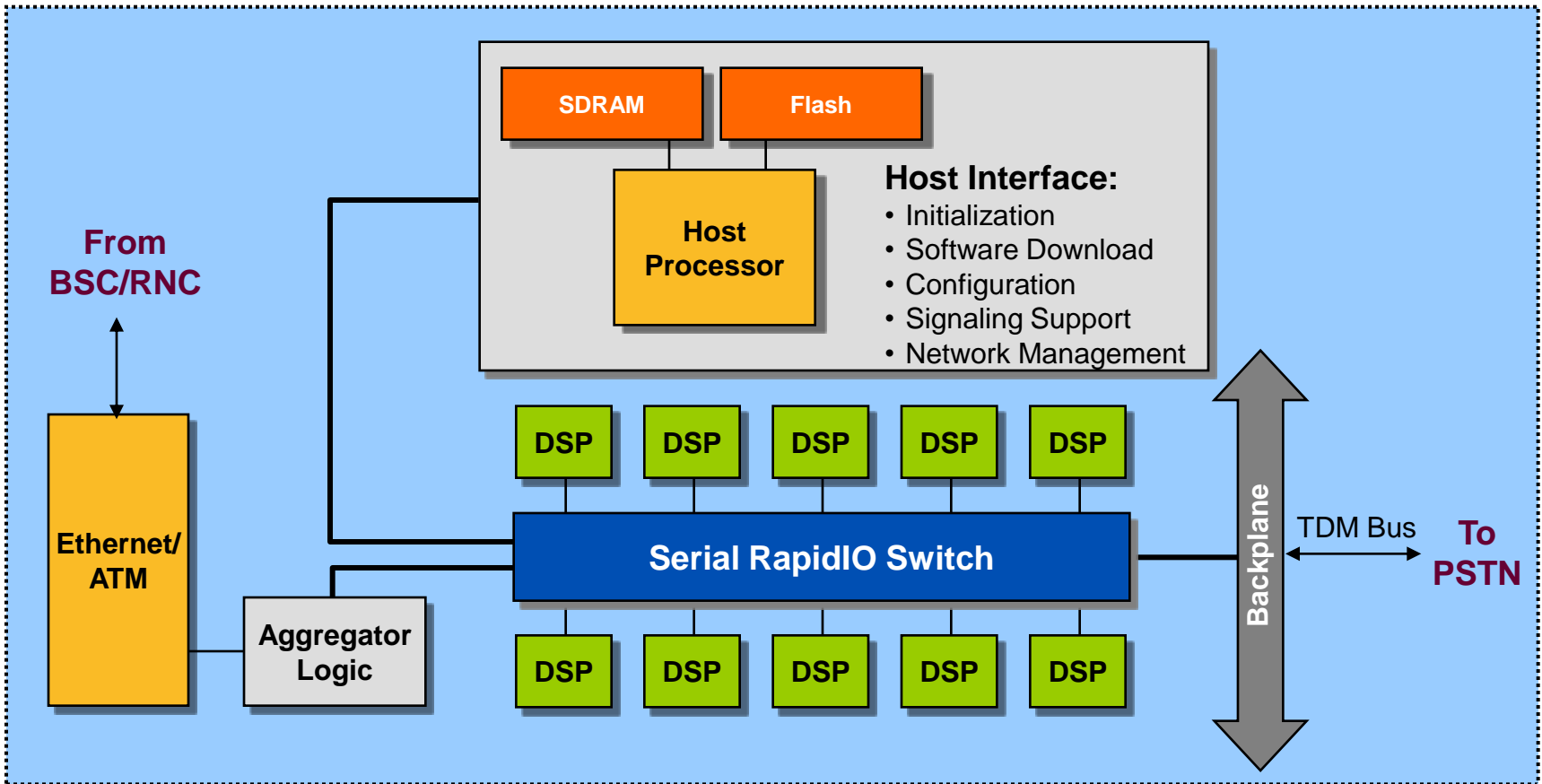
System Requirement	Ethernet	RapidIO
High backplane bandwidth (3–10 Gbps+)	Requires CPUs	Good Fit
Low latency	Stack Traversal	Good Fit
Low CPU overhead	SW Stack	Good Fit
High Availability	Good Fit	Good Fit
High QoS (low latency jitter, many streams, flow control)	Flow Control	Good Fit
Distributed computing	High Latency	Good Fit
5+ endpoints in a network	Good Fit	Good Fit

Good Fit
Marginal
Poor Fit

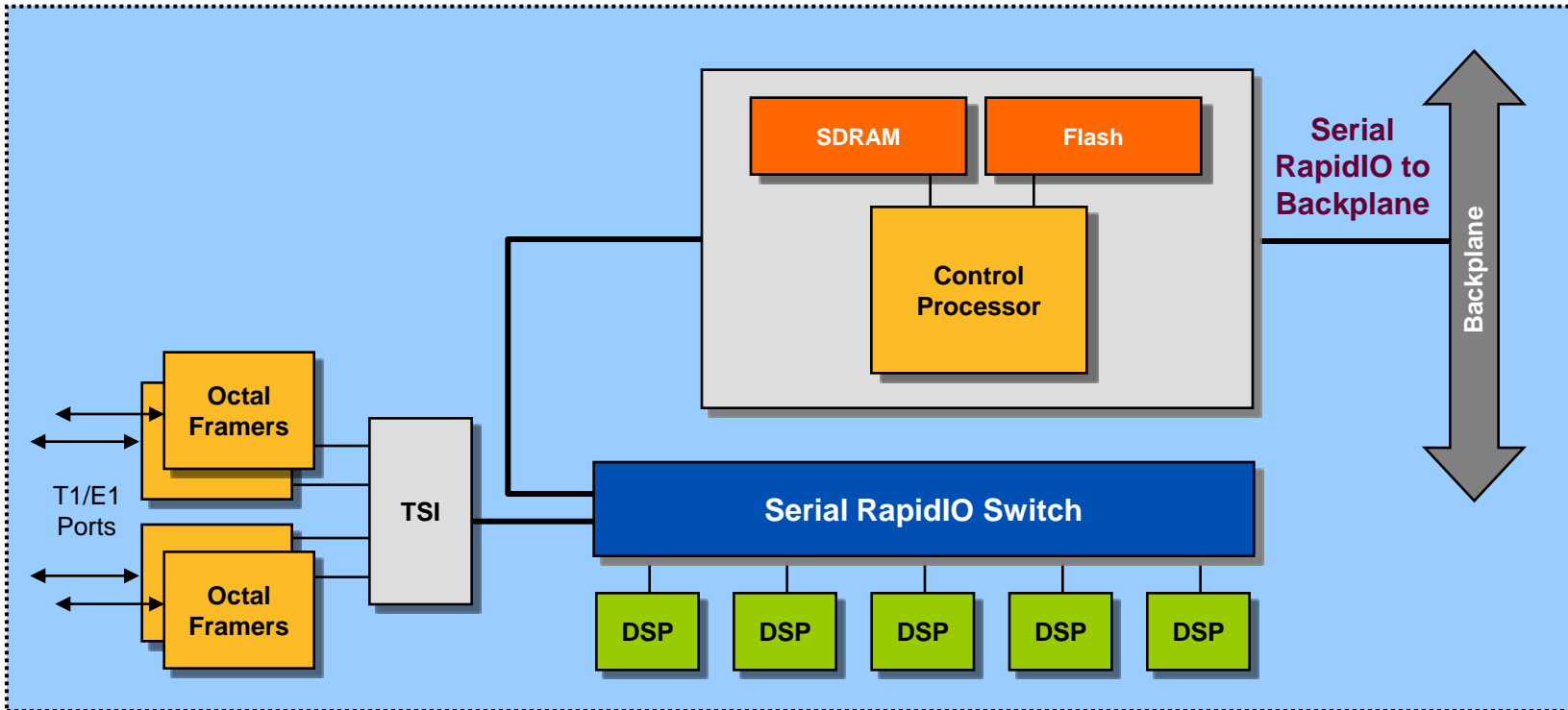
Multiservice Switch



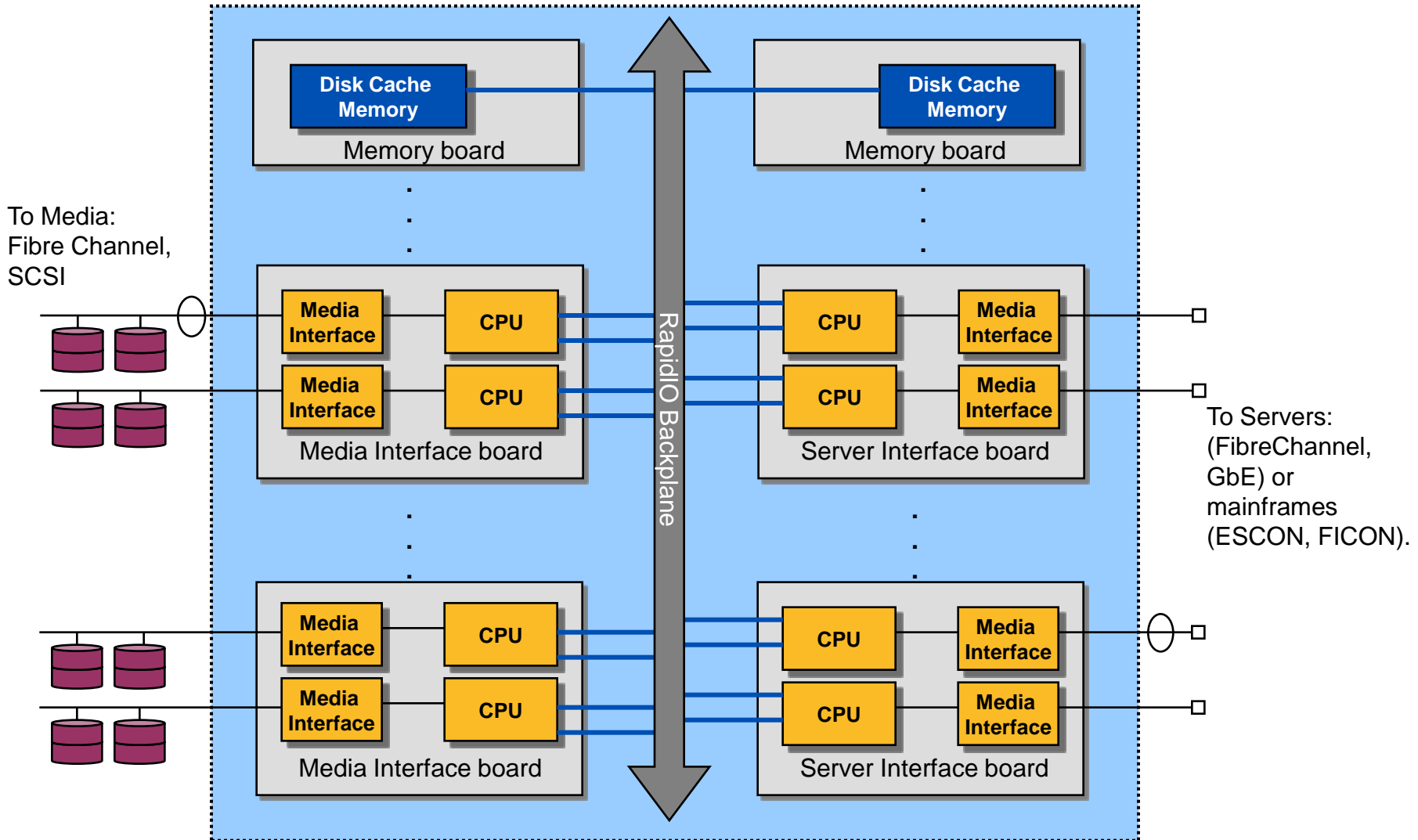
Mobile Switching Center



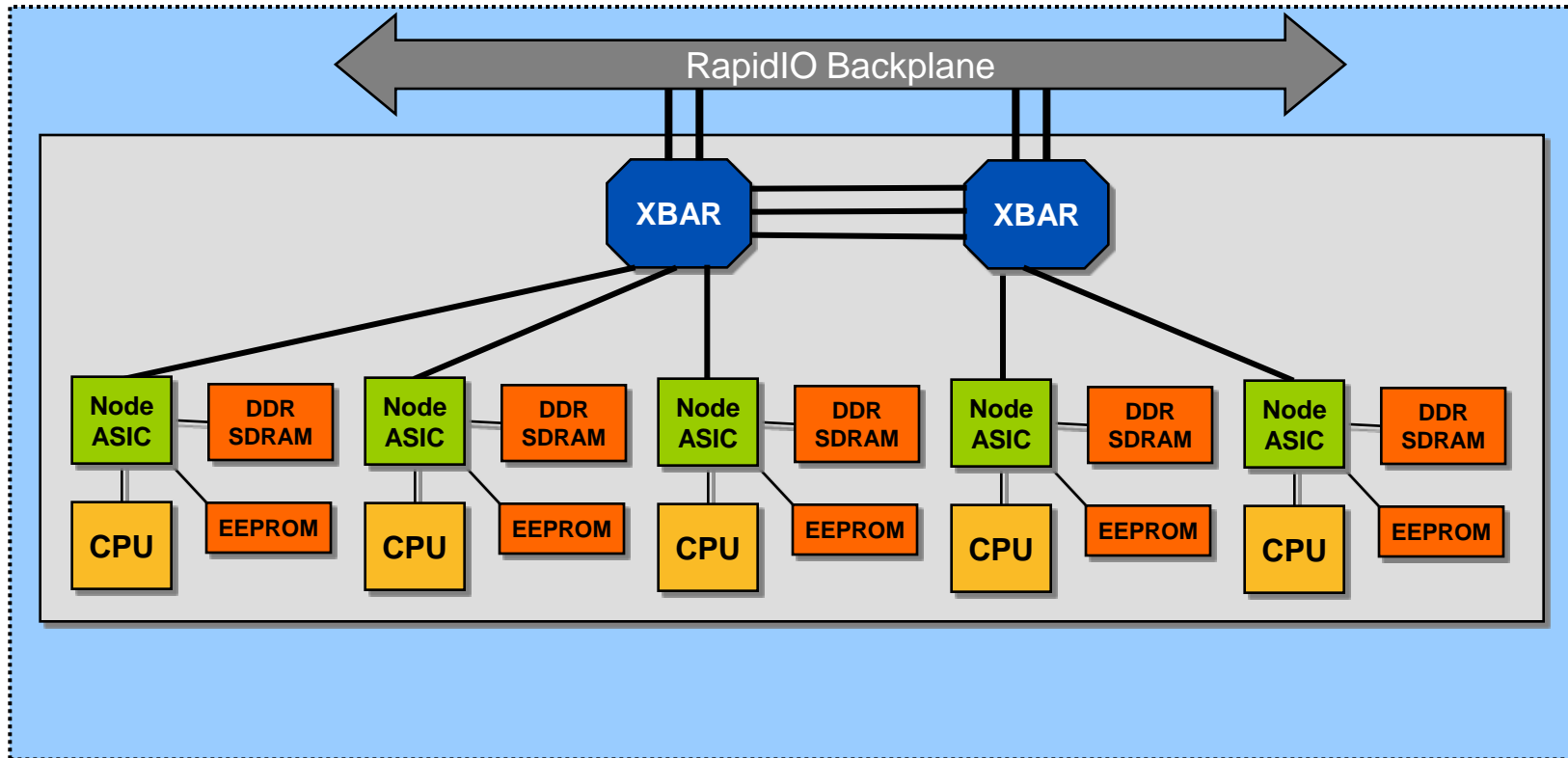
DSLAM – Voice Gateway Module



Enterprise Storage Switch



Signal and Image Processing



Conclusion



- Gigabit Ethernet will serve a limited role as a system interconnect
 - Works in low performance settings where significant over provisioning is possible
- RapidIO will expand its existing role as the standard system fabric
 - Available now
 - Efficient protocol supporting both control and data plane
 - Variety of PHY speeds
 - Cost competitive underlying economics