RapidIO.org Update

rickoco@rapidio.org

RapidIO®

RapidIO® – the unified fabric for Performance Critical Computing

© 2015 RapidIO.org

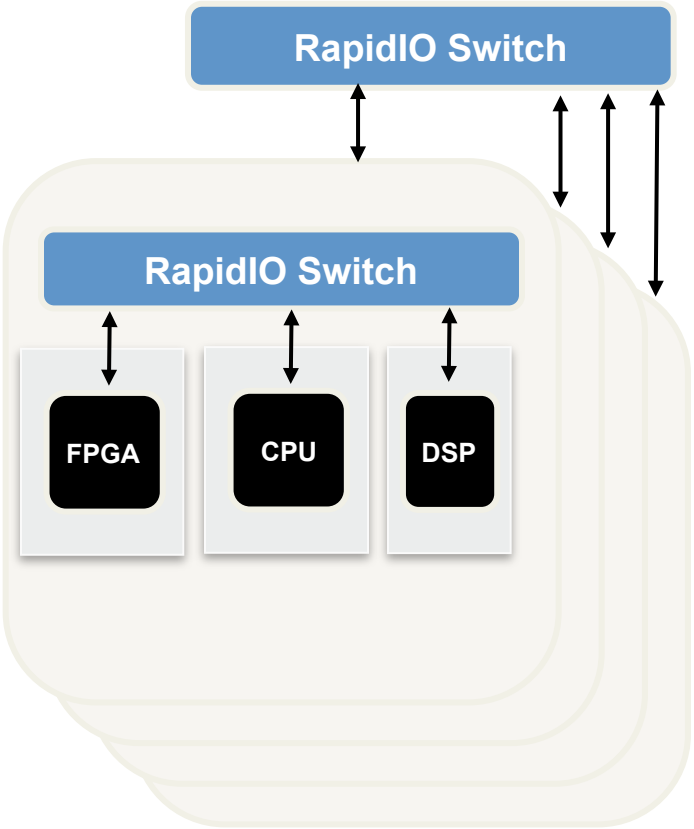
Outline

- RapidIO Overview
  - Benefits
  - Interconnect Comparison
  - Ecosystem
  - System Challenges
- RapidIO Markets
  - Data Center & HPC
  - Communications Infrastructure
  - Industrial Automation
  - Military & Aerospace
- RapidIO.org Task Groups
  - ARM 64-bit Coherent Scale Out
  - SW Task Group
  - Verification Task Group
  - 100G+ Phy Task Group
- Summary

© 2015 RapidIO.org

RapidIO Benefits

- Proven technology > 10 years of market deployment
- Supported by major CPU, DSP, NPU, FPGA & system vendors
- 10-160 Gbps/port Specification (10xN) released Q4 2013
- 10xN 3.1 Spec released Q3 2014 with increased fault tolerance for wireless & space applications
- ARM 64-bit Coherent Scale Out
- 25-400 (25xN) Gbps/port spec in development
- Hardware termination at PHY layer
- Lowest Latency Interconnect ~ 100 ns
- Inherently scales to 10,000's of nodes

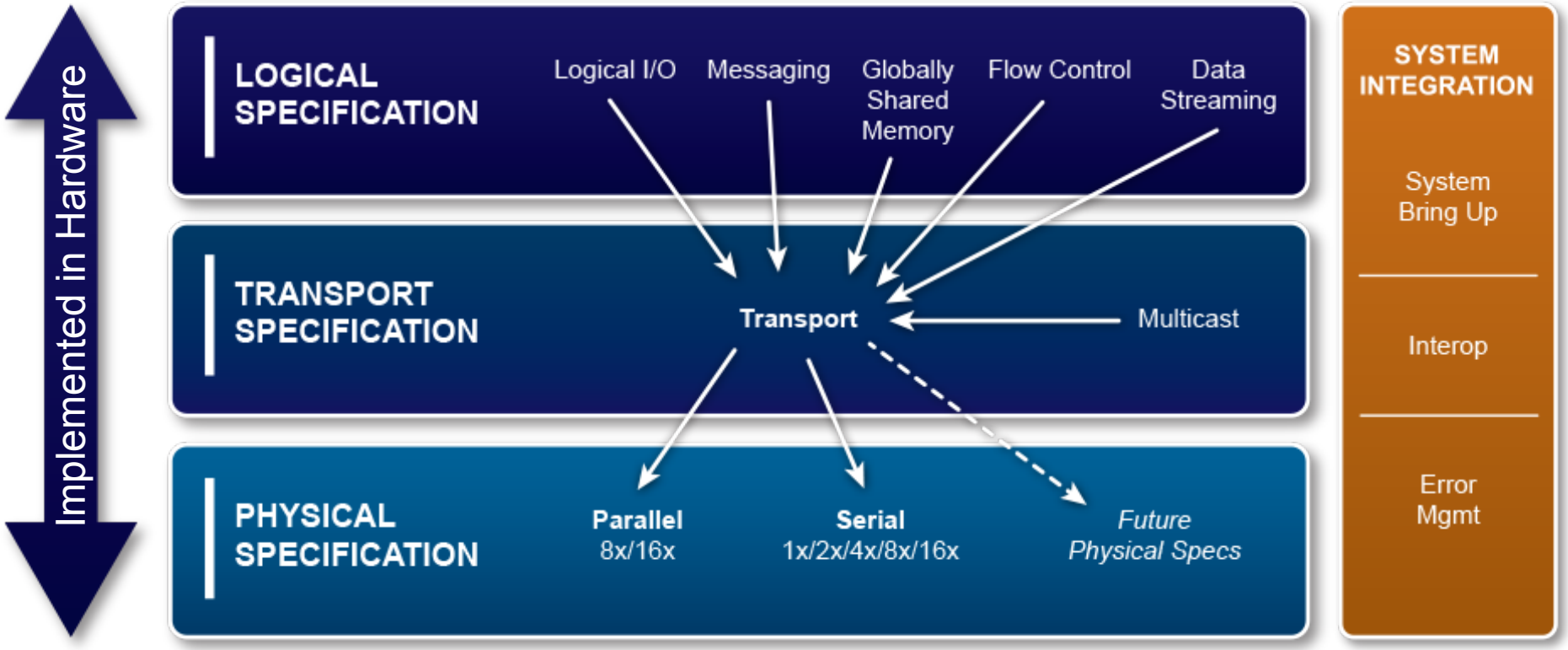


• Over 100 million 10-20 Gbps ports shipped worldwide
• 100% 4G/LTE interconnect market share
• 60% Global 3G & 100% China 3G interconnect market share

© 2015 RapidIO.org

RapidIO Specification Hierarchy

Hardware Terminated Protocol Stack – no CPU overhead



© 2015 RapidIO.org

Interconnect 'Check In'

Interconnect Requirements	RapidIO	Infiniband	Ethernet	PCIe	The Meaning of ●
Low Latency	●	●	✕	●	switch silicon: ~100 nSec memory to memory : < 1 uSec
Scalability	●	●	●	✕	support any topology, 1000's of nodes, true peer-to-peer
Integrated HW Termination	●	✕	✕	●	integrated into SoCs and guaranteed, in order delivery without software overhead
Power Efficient	●	✕	✕	●	3 layers terminated in hardware, Integrated into SoC's
Fault Tolerant	●	●	●	✕	supports hot swap and fault tolerance
Deterministic	●	●	✕	●	guaranteed, in order delivery with deterministic flow control
Top Line Bandwidth	●	●	●	✕	supports > 8 Gbps/lane

© 2015 RapidIO.org

RapidIO Strong and Growing Ecosystem

© 2015 RapidIO.org

System Challenges

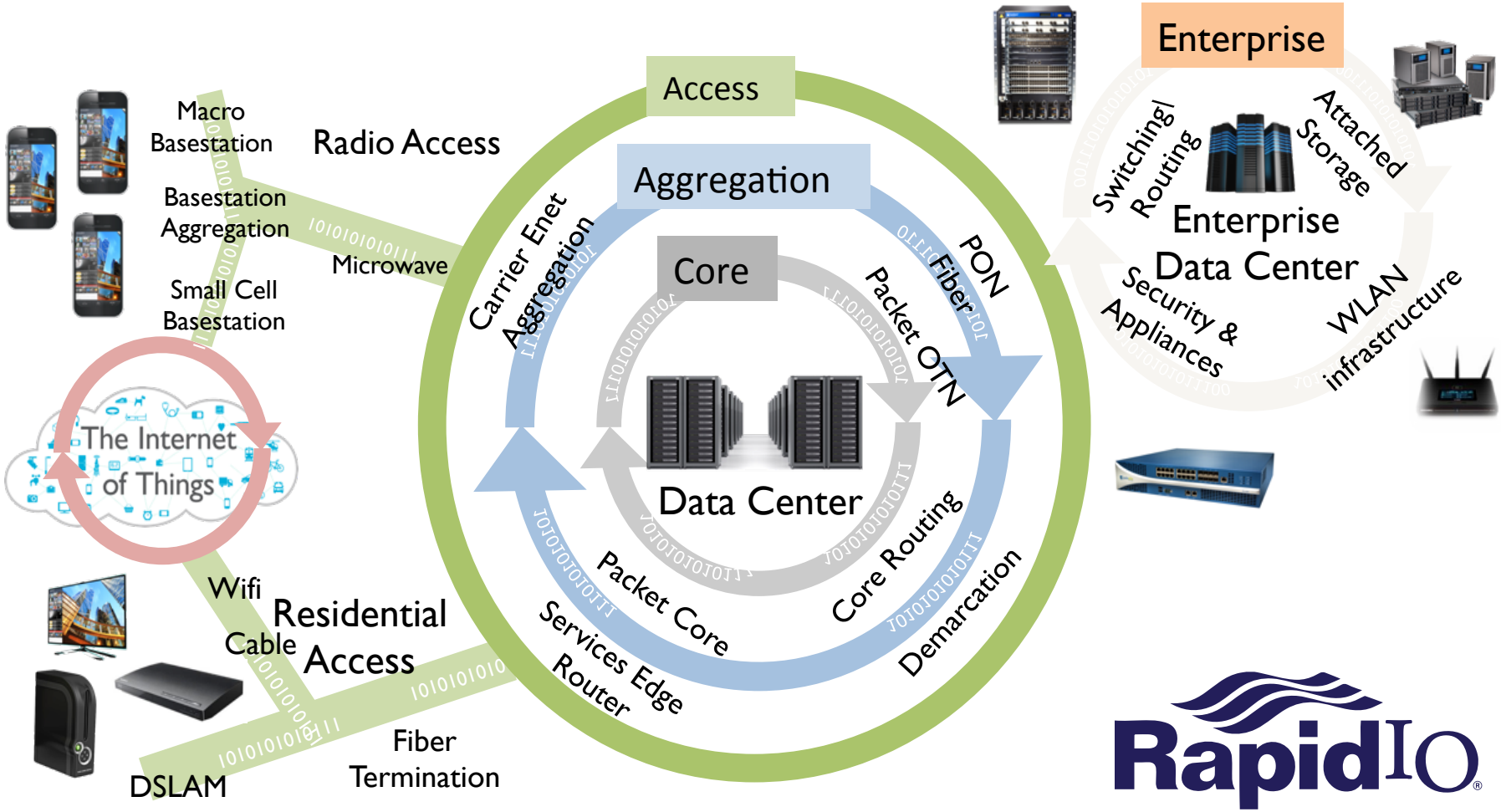
Networking and Computing Infrastructure

RapidIO®

RapidIO® – the unified fabric for Performance Critical Computing

© 2015 RapidIO.org

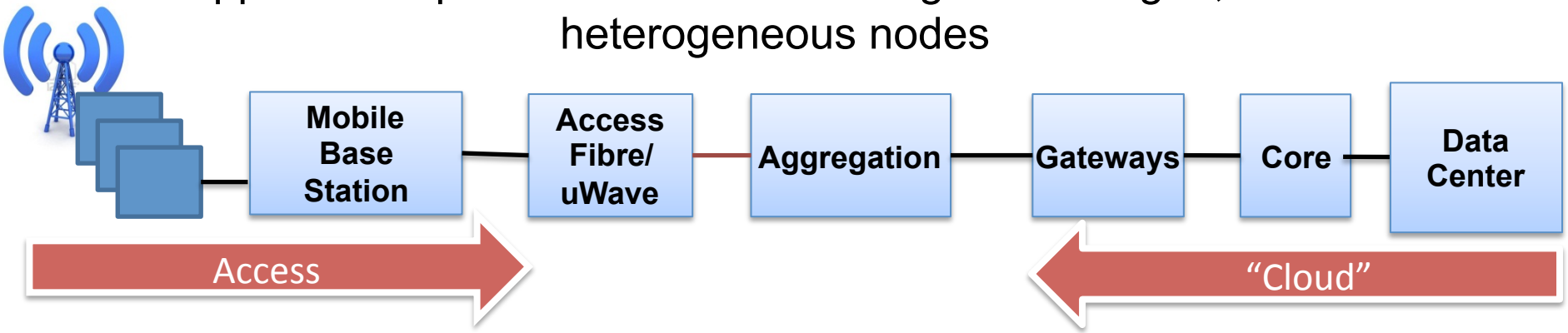
Networking & Computing Infrastructure View



© 2015 RapidIO.org

Hybrid Cloud - Networking and Computing functions Colliding

Application specific boxes transitioning to converged, scaled heterogeneous nodes



Many power constrained form factors

- Small cell: ~5W SoC
- Macro cell: ~25W SoC

Data center power budget

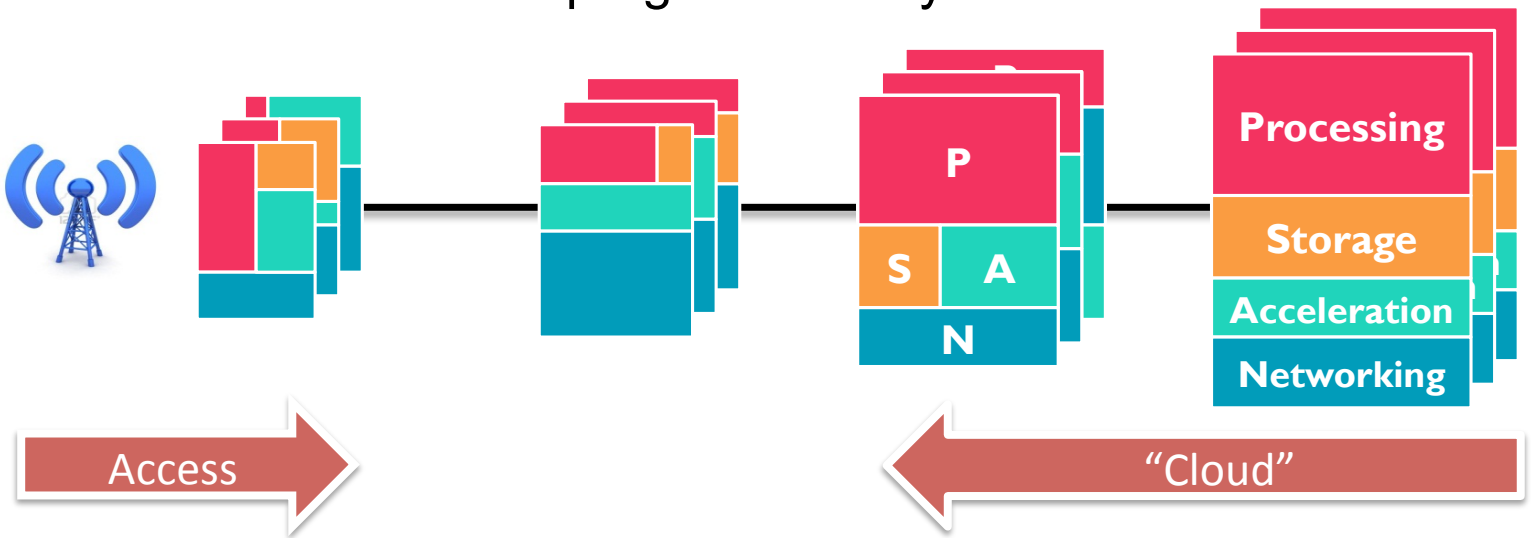
- ATCA blade: ~200W

Move to general purpose processors
Virtualisation
Separation of operational plane (hardware) and control plane
Consolidation (different boxes/subsystems within a box)
Flexibility for evolving workloads

© 2015 RapidIO.org

Distribution of Flexible Intelligence End-to-End

Pool of scalable heterogeneous nodes with increased visibility and programmability



Access to distributed intelligence for flexibility of workloads, traffic, services
Heterogeneous platforms to meet diversity of physical requirements, workload requirements

© 2015 RapidIO.org

RapidIO Markets

© 2015 RapidIO.org

Data Center & HPC

- System flexibility is at the core of RapidIO unified fabrics
- Can support up to 64K processors (or end points), each with their own complete address space & supporting peer-to-peer transactions
- Server class ARM 64-bit Coherent Scale Out specification
- Data Center Compute and Networking (DCCN) platform has been released
  - Targeted at Server, Data Center & Supercomputing applications + showcases strength of the RapidIO ecosystem with multi-vendor collaboration

© 2015 RapidIO.org

HP Moonshot Proliant m800

- 2D Torus RapidIO unified fabric
- up to 45 m800 cartridges capable of providing 5Gbs per lane connections in each direction to its north, south, east and west neighbors
- highest density DSP solution in an industry standard infrastructure in the market today:
  - 1,440 C66x DSP cores
  - 720 ARM A15 cores
  - up to 11.5TB of storage in a single Moonshot chassis
  - all connected via a 5Gbs per lane RapidIO unified fabric

© 2015 RapidIO.org

Communications Infrastructure

- Dominant market share in global deployment of cellular infrastructure 3G, 4G & LTE networks
- Over 100 million RapidIO ports shipped into wireless base stations to date
- RapidIO unified fabric clusters general purpose, digital signal, FPGA & communication processors together in a tightly coupled system with low latency and high reliability

© 2015 RapidIO.org

Industrial Automation

- System topology flexibility and scalability
- Lossless fabric with guaranteed forward progress & deterministic data delivery
- Low-latency for real-time applications & low footprint enabling high density designs
- Mainstream Linux supported functions:
  - Hot-swap capable, Dynamic configuration, Real-time diagnostics with fast localization of failures, System discovery & enumeration, OpenMPI ready

RapidIO Fabrics enable guaranteed data delivery and system topology flexibility for tightly controlled Industrial Automation applications

© 2015 RapidIO.org

Military & Aerospace

- RapidIO Fabrics deliver key performance metrics such as:
  - best in class switch cut through latency of ~100ns
  - lossless data flow
  - small footprint and lowest overall power consumption
- RapidIO.org Community works closely with the VITA Standards Organization
  - to specify use of RapidIO Fabrics within the OpenVPX (VITA 65) backplane standard & the SpaceVPX (VITA 78) backplane standard

© 2015 RapidIO.org

RapidIO.org Tasks Groups

© 2015 RapidIO.org

ARM 64-bit Coherent Scale Out Task Group Charter

- The ARM 64-bit Coherent Scale Out over RapidIO Task Group shall be responsible for developing a specification for multi SoC / core coherent scale out of ARM 64-bit cores with the following functionality:
  - coherent scale out of a few 10s to 100s cores & 10s of sockets
  - ARM AMBA® protocol mapping to RapidIO protocols
    - AMBA 4 AXI4/ACE mapping to RapidIO protocols
    - AMBA 5 CHI mapping to RapidIO protocols
  - Migration path from AXI4/ACE to CHI and future ARM protocols
  - support for GPU/DSP floating point heterogeneous systems
  - HW hooks and definition to support RDMA, MPI, secure boot, authentication, SDN, Open Flow, Open Data Plane, etc
  - Other functionality as necessary to for performance critical computing - support Data Center, HPC and Networking Infrastructure system development and deployment

© 2015 RapidIO.org

ARM 64-bit Coherent Scale Out Task Group Contributors

© 2015 RapidIO.org

Coherent Scale Out Use Cases

	Coherency Class	Number of Devices	Coherency Domains	Coherency Granule	Distances	Latency characteristics	Interconnet PHY	Interconnect Architecture	Examples Verticals
1	Homogeneous	2	1	cache line	~inch	highly optimized	parallel, electrical	p2p	HPC
2	Homogeneous	2-8	1-4	cache line	single board	real-time limits	serial, electrical or optical	p2p or switched	HPC
3	Homogeneous	4-16	1-8	cache line	two boards	real-time limits	serial, electrical or optical	switched	HPC
4	Heterogeneous	2-4	1		single board	real-time limits	serial, electrical	p2p or switched	HPC
5	Heterogeneous	2-8 Processor SoCs	1-8	cache line	single or dual board	real-time limits	serial, electrical	p2p or switched	embedded compute
6	Heterogeneous	2-8 mixed component	1-8	cache line	single or dual board	real-time limits	serial, electrical	switched	Graphics, BTS, GPU/DSP with OpenCL
7	Heterogeneous	2-8 mixed component	1-8	cache line	single board	soft	serial, electrical	switched	SDN/NFV, router
8	Heterogeneous	4-100	many	cache line	multi-board or shelf	best effort	serial, electrical or optical	switched	switch/ router
9	Large System	10s-100s	many	tbd	distributed system	best effort	serial	switch	Storage, Cloud, fail-over
10	Large System 2	10s-100s	many	tbd	distributed system	best effort	serial	switch	HPC

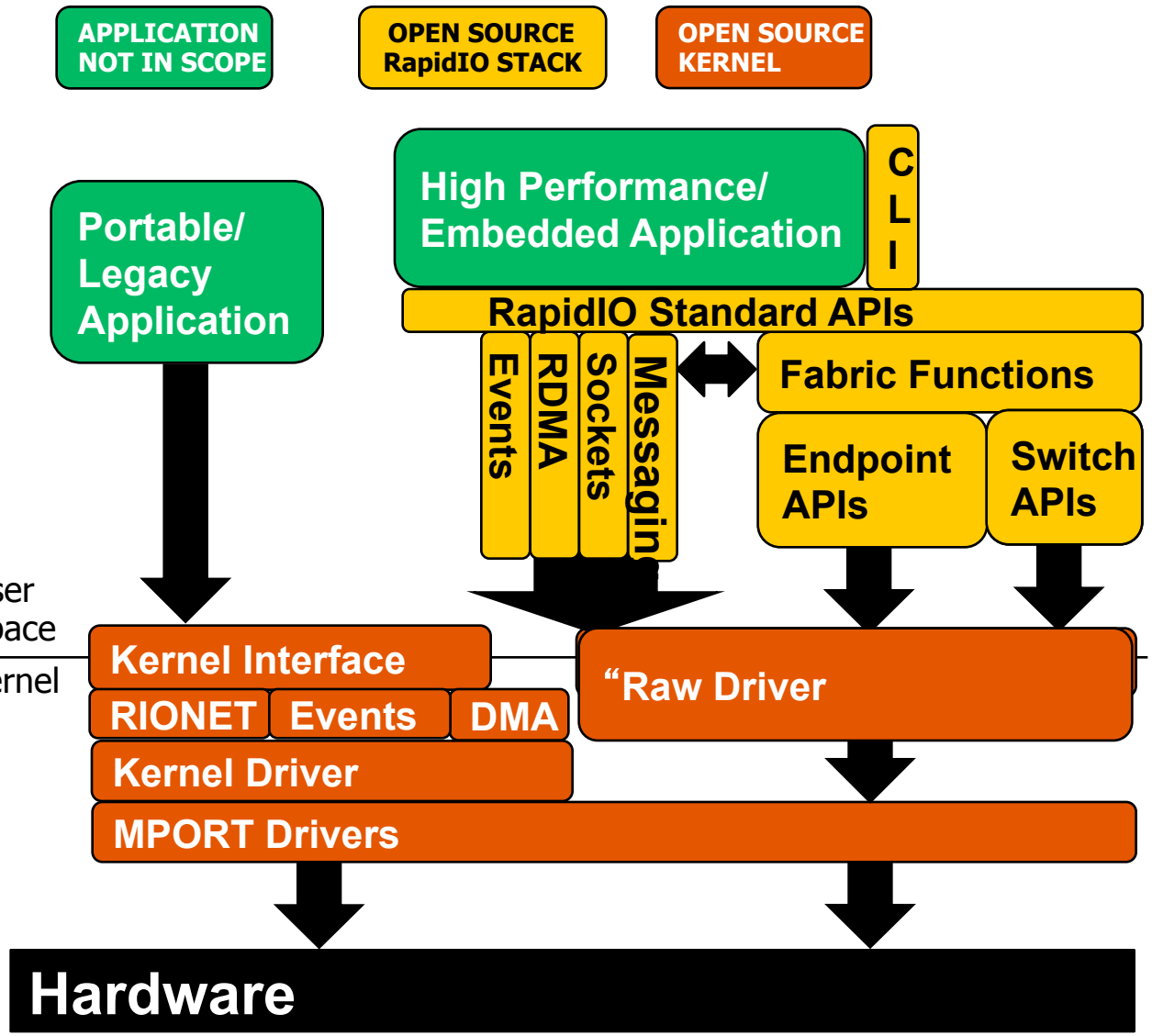
© 2015 RapidIO.org

SW Task Group

- Kernel enhancements
  - "raw driver" interfaces which support all endpoints
  - Fault tolerance/hot swap support
  - Customizable/optional enumeration/discovery
- Fabric management stack
  - Uses "Raw Driver" interfaces
  - Portable across endpoints and operating systems
  - Command Line Interpreter for basic configurability
- Interoperable, standardized data path
  - RDMA, Sockets, and Channelized Messaging
  - Reference implementation uses "Raw driver"
  - User Mode Driver for maximum performance

© 2015 RapidIO.org

Open Source Stack



CLI - Open source command line interpreter for fabric management

RapidIO Standard APIs
RapidIO standard interface definitions and behavior

Fabric Functions
Implementation of RapidIO Standard APIs Fabric Management

Endpoint APIs - Universal programming model for endpoint functions

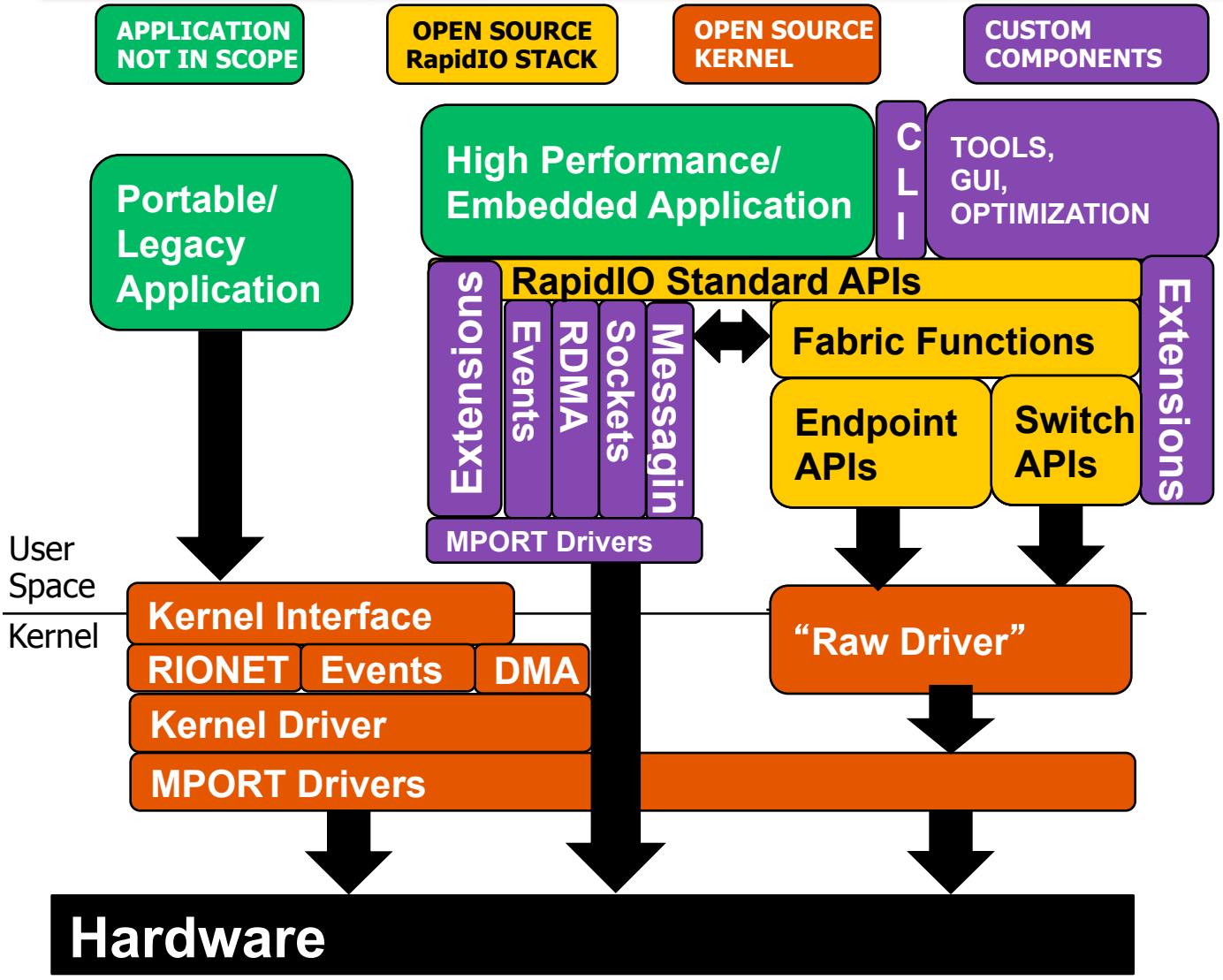
Switch APIs - Universal programming model for switch functions.

Messaging/Sockets/RDMA /Events - Implementation of RapidIO Standard API's Data Path

"Raw Driver" - Standardized hardware functions.

© 2015 RapidIO.org

Customization, Optimization



TOOLS, GUI, OPTIMIZATION
Value adders for debug/ monitoring, system visualization, data interpretation, and topology specific functions

MPORT Drivers
Drivers optimized for and aware of executing hardware operations in user mode.

Extensions
Additional data path and/or fabric management services provided to applications

© 2015 RapidIO.org

Verification Task Group

- Define features to be added to the BFM
  - 3.1 extensions including the NGSIS Space Profile
  - ARM 64-bit coherency support
- Define licensing model for BFM
  - Open source?
  - Eliminate membership requirement for library?
- Define interfaces to support customization / integration with AXI BFMs
- Develop baseline verification environment

© 2015 RapidIO.org

100G+ PHY Charter

- The 100G+ PHY Task Group charter is to develop higher performance PHY specifications for the RapidIO protocol
  1. Determine and define 25G or greater PHY leveraging industry standard specification(s)
  2. Evaluate and confirm PHY definition will improve or maintain existing PHY characteristics
     a) Improve power efficiency
     b) Maintain asymmetric link operation
     c) Maintain low latency characteristics
     d) Maintain fault tolerant characteristics
     e) Maintain reliability characteristics assuming a bit error rate of less than 10-12
  3. Specify a low latency PHY suitable for dual- and quad-socket cache coherency applications
  4. Outline and prioritize additions/improvements to the PHY feature set
  5. Other additions and improvements to the PHY feature set as requested

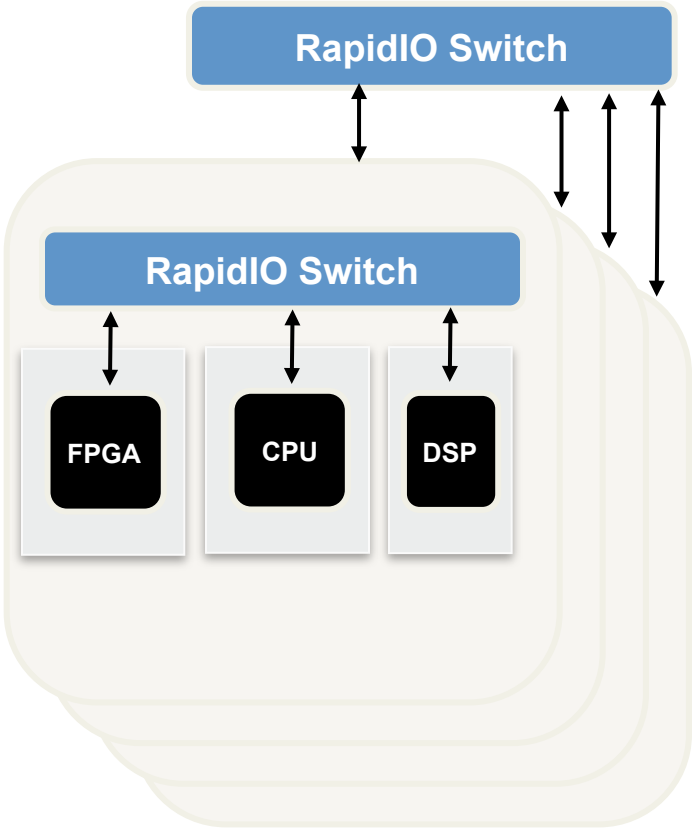
© 2015 RapidIO.org

RapidIO.org Summary

© 2015 RapidIO.org

RapidIO Summary

- Proven technology > 10 years of market deployment
- Supported by major CPU, DSP, NPU, FPGA & system vendors
- 10-160 Gbps/port Specification (10xN) released Q4 2013
- 10xN 3.1 Spec released Q3 2014 with increased fault tolerance for wireless & space applications
- ARM 64-bit Coherent Scale Out
- 25-400 (25xN) Gbps/port spec in development
- Hardware termination at PHY layer
- Lowest Latency Interconnect ~ 100 ns
- Inherently scales to 10,000's of nodes



• Over 100 million 10-20 Gbps ports shipped worldwide
• 100% 4G/LTE interconnect market share
• 60% Global 3G & 100% China 3G interconnect market share

© 2015 RapidIO.org