RapidIO® Technology Overview and Applications

Technical Discussion of the RapidIO Interconnect and System Design Examples

RapidIO
The Embedded Fabric Choice

© Copyright 2005 RapidIO® Trade Association
Revision 03, May 2005
Embedded Systems Driving Standard Interconnect Requirements

**Development Needs**
- Higher performance
- Lower system costs (NRE, CAPEX, OPEX)
- Modularity – reuse across platforms
- Common components
- Distributed processing
- Standardized connectivity

**Standard Interconnect Requirements**
- High-performance
- Multiple hosts – distributed processing
- Direct peer-to-peer communications
- Multiple heterogeneous operating systems
- DMA and message passing
- Support complex topologies
  - Discovery mechanism
  - Redundant paths – fail over
- Multicast
- High Reliability
  - Loss less
  - Automatic retraining and device synchronization
  - System level error management
- Time of day synchronization
- Quality of Service (QoS)
RapidIO Technology Genesis

• Started as embedded interconnect in 1997
  – Conceived as processor interconnect
  – Motorola & Mercury collaborated on initial specification
• RapidIO design goals:
  – Define a light-weight protocol
  – Limit software impact
  – Focus on ‘inside-the-box’ communications
• First rev standard completed 1999
  – 2001 first silicon: FPGAs, processors, bridges and switches
• Expanded to support carrier-grade fabric requirements
RapidIO Technology System Value

• Scalable, modular architecture
  – Layered architecture with common transport layer
  – Chip-to-chip, board-to-board, backplane

• High-speed connectivity
  – Physical layer defined for backplane interconnection
    • ~80-100 cm + 2 connectors (Serial)
  – Up to 10Gbps bandwidth today

• Robust feature set
  – Carrier-grade reliability
  – Traffic management
  – Multi-protocol/convergence
RapidIO Architecture

Logical Specifications
- I/O System
- Message Passing
- Global Shared Memory
- Flow Control
- Data Streaming

Transport Specifications
- Common Transport

Physical Specifications
- 8,16 Parallel
- 1x/4x Serial
- Higher Speed PHYs

Ancillary Specifications
- Interoperability
- Error Management
- System Bring up
- Multicast
Physical Layer Specifications

- **Parallel**
  - 8/16b LVDS
  - 250 MHz to 1 GHz DDR
- **Serial**
  - 1 or 4 lane XAUI AC coupled
  - 8/10b encoded
  - 1.25 / 2.5 / 3.125 Gbaud
  - Aggregated bandwidth of up to 10Gbps in a x4 sRIO link
- Point-to-point topology
- Device synchronization and training
- Every packet is acknowledged or retried (16 or 32b CRC applied)
- Explicit hardware-based error recovery
- Control symbols provide the main mechanism for physical layer control (can be embedded within packets)
- Four fixed priorities
  - Mechanism for higher and lower priority traffic
  - Order is maintained for traffic at a given priority
  - Transaction acknowledge/reply higher priority than request
- Link-level flow control
Transport Specification

- Switches operate at the transport layer
  - All logical protocols use a common transport header
  - Today’s switches will work with future logical protocols
- Device-based routing
  - 8 or 16 bit device ID
    - Simplifies classification and routing compared to Ethernet or IP
    - Any RapidIO device ID can be used as a unicast ID or a multicast group
      - Unicast packets are forwarded out a single port
      - Multicast packets are elaborated to multiple ports at the switch connected to leaf nodes
      - Can multicast using a transaction that omits a logical layer acknowledge, i.e., NWRITE, SWRITE, Data Streaming
  - Fail-over events only affect routing table entries of nearest neighbors
    - Unlike path based routing architectures
Input/Output Logical Specification

- Load / Store / DMA operations
  - Processor is not necessarily RapidIO aware
- 34, 50, 66-bit address space
- Transactions:
  - NREAD – read operation
    - Data returned is the response
  - NWRITE – write operation, no response
  - NWRITE_R – robust write with response from the target end-point
  - SWRITE – streaming write
  - ATOMIC – atomic read-modify-write
  - MAINTENANCE – system discovery, exploration, initialization, configuration and maintenance operations
Message Passing Logical Specification

- Hardware support for segmentation and reassembly of 4kB datagrams
  - Segments are automatically reordered
- Logical Layer acknowledge
  - Auto retry on timeout
- Messages
  - Organized into 4 mailboxes and 4 letters within each mailbox
    - Sending device can send 4 concurrent letters to each target mailbox
- Doorbells
  - Short 8 or 16-bit messages
Data Streaming Logical Specification

- Segmentation and reassembly
  - 64Kb PDUs
  - System-wide MTU size
- Efficient logical protocol for communications
  - Start, continuation, end segments
    - continuation has 20 bit header
- Interworking
  - Ethernet, UTOPIA, SPI-3/4, CSIX, etc
- Virtual Streams
  - Flow identification
- Traffic Management framework
  - End-to-end Flow control
  - Millions of streams
  - 256 traffic classes
  - Lossy operations
    - No logical layer acknowledge to support traffic managed data plane applications
Other Features

- Specification includes a user-defined logical protocol for custom extensions
- Packets are 256 payload bytes or less

Ancillary Specifications
  - Interoperability
    - PCI transaction mapping
    - Hardware Interoperability Platform (HIP)
    - Interoperability checklist
  - Error Management
  - System Bring-up
RapidIO Packet Format

**Logical Layer**

- Logical Layer format delineated by 4 bit Ftype field
  Listed as “TYPES” in the Specification
- Transaction field indicates the specific request or response type
- All other fields dependent on Ftype and Transaction

![Diagram of RapidIO Packet Format]

- **Ftype**
  - Size: 4 bits
  - Source TID: 8 bits
  - Device Offset Address: 32, 48, 64 bits
- **Transaction**
  - Optional Data Payload: 8 to 256 Bytes
RapidIO Packet Format TYPES

• I/O non-coherent functions
  – NREAD, NWRITE, NWRITE_R, SWRITE, ATOMIC
• Message functions
  – DOORBELL, MESSAGE
• System support functions
  – MAINTENANCE
• Flow Control
• User Defined functions
• Cache coherence functions
  – READ, READ_TO_OWN, CASTOUT, IKILL, DKILL, FLUSH, IO_READ
RapidIO Packet Format

**Transport Layer**

- RapidIO uses source based addressing
- Switches use route tables to determine destination port
- TT Field indicates size of route address
- A destination may have more than 1 target address for redundant routes

<table>
<thead>
<tr>
<th>Size</th>
<th>Source TID</th>
<th>Device Offset Address</th>
<th>Optional Data Payload</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>8</td>
<td>32, 48, 64</td>
<td>8 to 256 Bytes</td>
</tr>
</tbody>
</table>
RapidIO Packet Format

**Physical Layer**

- Physical Layer adds device to device handshake info
- Packet partitioned to simplify assembly/disassembly in controllers
RapidIO networks are built around two “Basic Blocks”
- **Endpoints** to source and sink packets
- **Switches** to pass packets between ports without interpreting them

- All devices support maintenance transactions for access to configuration registers
- Resulting applications range from
  - Wireless infrastructure to
  - Edge devices to
  - Storage equipment to
  - High-end computing to
  - Military, scientific and industrial applications
Edge Boxes and Other High-Availability Systems

• The interconnect is the backbone of a system’s fault tolerance architecture
• RapidIO technology designed for 24/7 with hardware-based reliability:
  – Redundancy – supports all sparing schemes
  – Hot-swap support
  – Fault detection
    • CRC
    • 8B/10B encoding
    • Performance and reliability monitors to detect degradation
  – Fault isolation – physical layer handshaking
  – Fault containment – table-based routing algorithm
Multiservice Switch

SONET MultiPHY Line Card
- OC-12/STM-4
- Optics
- MPHY Framer
- Host Processor
- Switch Interface

Gigabit Ethernet Line Card
- GbE
- PHY
- Switch Interface
- Packet Processing
- Host Processor

Voice Line Card
- T1/E1
- LIU
- Switch Interface
- Packet Processing
- Host Processor

Revision 03 © Copyright 2005 RapidIO® Trade Association
Wireless Infrastructure and Other Areas of Convergence

- Includes systems such as RNCs, mobile switching centers, and DSLAMs – where the increase and mix of voice, video and data continue to push compute density.
- RapidIO technology provides high-speed, low latency, multi-protocol connectivity for:
  - DSP farms, processors, ASICs and FPGAs in peer-to-peer or master/slave configurations
  - Control and data backplane
  - Baseband to RF board
  - Chip to symbol rate processor
DSLAM – Voice Gateway Module

- TSI Serial RapidIO Switch
- SDRAM
- Flash
- Control Processor
- Serial RapidIO to Backplane
- Backplane

- Octal Framers
- T1/E1 Ports

- Serial RapidIO Switch
  - DSP

Revision 03 © Copyright 2005 RapidIO® Trade Association
Storage Systems and Other Compute Intensive Environments

- RAID arrays and other mass storage devices serving as repositories for imaging on-demand, high-definition video streaming, news groups, clustering environments, financial applications and more
- RapidIO technology offers the throughput and reliability required for these applications
Enterprise Storage Switch

To Media:
- Fibre Channel
- SCSI

To Servers:
- (FibreChannel, GbE)
- mainframes (ESCON, FICON)
Signal and Image Processing

Diagram showing a network of nodes connected by XBARs. Each node contains an ASIC, CPU, DDR SDRAM, and EEPROM. The network is connected to a backplane.
Standards-Based Solutions
E.G., AdvancedTCA® Platforms

- Leveraged for development and deployment in a wide range of application areas including routers, wireless infrastructure, soft switches, media gateways, optical transport systems
- RapidIO interconnect is the best option for chip, card and backplane connectivity in ATCA platforms
  - High-speed (up to 3.125Gbps/lane)
  - No software overhead
  - Hardware-based reliability
  - Built-in traffic management
- PICMG 3.5 RapidIO Standard

ATCA Fabric Requirements:
- Bandwidth
  - 10 Gbps per blade slot
  - 4 channels at 3.125 Gbaud using 8b/10b encoding = 10 Gbps
- Cost-effective switch silicon
  - Integrated SERDES (1600 mV maximum)
  - Large enough for a multi-chassis configuration
    - 20-24 ports (16 slots + fabric extension)
  - Around 300 Gbps switch throughput
- Fabric Characteristics
  - Congestion management
  - Quality of Service (QoS)
    - Low overhead, jitter
  - Error reporting
- Mesh, Star topology options
**AdvancedTCA® Platform**

**PICMG 3.5 RapidIO Standard**

**AdvancedTCA Mezzanine Cards**
- RapidIO connecting the DSP farm AMC to the ATCA card
- RapidIO connecting the control AMC to the ATCA card
- RapidIO switch card connecting the AMCs to the RapidIO backplane

**PICMG 3.0 ATCA Platform**

**Real-World Deployment**
Design with RapidIO® Today

- Scalable
  - Chip-to-chip, card-to-card, chassis-to-chassis
  - Extensible architecture
- High Speed
  - Up to 10Gbps bandwidth today (next-generation RapidIO PHY, to be completed in 2005, will provide up to 40Gbps)
  - Efficient protocol
- Robust
  - High reliability
  - Traffic management