RapidIO Whitepaper

A Comparison of Specification Changes in the Recent Versions of RapidIO and PCI Express

By Geoff Roddick,
Technical Consultant at RapidIO Trade Association

Introduction

There are many interconnect technologies to connect components in a system and an embedded designer is faced with an array of standards and technologies to choose from. This paper explores the latest changes in two prominent options, in the hopes of helping make the selection of the most appropriate interconnect technology.

The changes in each specification and how they compare is shown in the table below. Following the table is a discussion of each of these changes.

Legend

X	New feature in specification revision
x	Comparable feature
√	Comparable feature (not exact match)
	No comparable feature

Change	RapidIO		PCI Express			Discussion Section
	1.3 to **2.0**	2.0 to **2.1**	1.x to **2.0**	2.0 to **2.1**	2.0 to **3.0**	
Addition of 2.5 GBaud	x	x	**X**	x	x	1.1 Lane Rates and Widths
Addition of 4.0 GBaud	x	x			**X**	1.1 Lane Rates and Widths
Addition of 5.0 GBaud	**X**	x				1.1 Lane Rates and Widths
Addition of 6.25 GBaud	**X**	x				1.1 Lane Rates and Widths
Addition of support for 2x, 8x and 16x ports	**X**	x	x	x	x	1.1 Lane Rates and Widths
Software control of link speed, reporting of speed and width changes, and control of loopback.	x	x	**X**	x	x	1.2 Software Control of Links
Dynamic link speed optimization	√	√	**X**	x	x	1.3 Baud Rate Discovery

Change	RapidIO		PCI Express			Discussion Section
	1.3 to **2.0**	2.0 to **2.1**	1.x to **2.0**	2.0 to **2.1**	2.0 to **3.0**	
Baud rate discovery	**X**	x	√	√	√	1.3 Baud Rate Discovery
Compliance speed			**X**	x	x	1.4 Compliance Speed Testing
Testability enhancements		**X**	**X**	x	x	1.5 Testability Enhancements
Pre-emphasis	**X**	x	√	√	√	1.6 Emphasis
De-emphasis	√	√	√	√	**X**	1.6 Emphasis
Non-operational lanes	x	x	**X**	x	x	1.7 Lane Failure
Interconnect latency	√	√		**X**	x	1.8 Ordering Attribute
Multicast	√	√		**X**	x	1.9 Multicast
Support for Virtual Channels	**X**	x				1.10 Data Plane
Reliable Traffic and Continuous Traffic	**X**	x				1.10 Data Plane
Virtual Output Queuing Backpressure	**X**	x				1.10 Data Plane
Flow Control for Streaming Datagrams	**X**	x				1.10 Data Plane

1 Discussion of Changes

1.1 Lane Rates and Widths

The lane rate and number of lanes on an interconnect technology directly affects the data throughput an end application can expect. In the latest revision of specifications, both RapidIO and PCI Express have increased the lane rates. However RapidIO is faster at 6.25 GBaud than PCI Express' 4 GBaud.

Both PCI Express and RapidIO leverage the XAUI SerDes technology, developed for Ethernet, for the lower speeds. For 4.0 GBaud and higher, both have based their standards on work done by the OIF and have added scrambling to support these higher speeds.

While PCI Express has always supported support lane widths of 1x, 2x, 4x, 8x and 16x, the new additions to the Serial RapidIO specification have been updated to match this.

The combination of widths and speeds now allows designers a huge amount of flexibility to tune their application to the required speeds. With the current specifications, RapidIO leads when it comes to bandwidth per lane.

1.2 Software Control of Links

PCI Express has added features surrounding the software control of link speeds and widths that did not exist in the previous specifications. RapidIO has had these features since the 1.2 version of the specification. The speed signaling controls allow software to determine whether a device can operate at a specific signaling rate and width. This can then be used to change the signaling rate and width to reduce power consumption when the full bandwidth is not required.

1.3 Baud Rate Discovery

PCI Express has added a speed negotiation and rate fall-back feature, which is an improvement over previous specifications. With speed negotiation, two devices powered up at different speeds will both automatically fall back to 2 GBaud.

RapidIO has added the optional feature of baud rate discovery. This feature is very similar to the PCI Express speed negotiation and rate fall-back feature, but it is different in two manners. Firstly, this feature is optional in RapidIO. Secondly, the fall back rate is the lowest rate supported by the two devices, which may be higher than the set 2 GBaud specified by PCI Express. If the received baud rate is less than the baud rate of the idle sequence being transmitted by the port, the port reduces the baud rate to the next lowest baud rate that it supports and will continue to step the baud rate down until the rates match. This allows for the highest possible link speed between the two devices.

Besides potentially improving line speeds these features are also effective in error recovery. If a device sees errors and enters into recovery, it will try to achieve lock at the current speed. If the link is damaged and it cannot achieve symbol lock, it will then fall back to a lower data rate, which may allow for a symbol lock and a more stable link.

PCI Express, being a consumer technology addressing the add-in card use model, common in PCs, must be ‘plug and play’, hence PCI Express must to support automatic lane width negotiation and, since the advent of PCI Express 2.0, lane speed negotiation. This forces the complexity of the configuration into dedicated hardware; while RapidIO does not force the implementation as embedded systems often ship with a fixed hardware

configuration, which reduced the need for full software and hardware speed-width negotiation.

With the current specifications, RapidIO is more versatile for this feature as it allows for both automatically negotiation and/or under software control – which means that the software can be written to detect and abort training in circumstances where failures occur. Without a 'failure notification', it's possible for a PCI Express link to fail to retrain and not send out notification of this failure.

1.4 Compliance Speed Testing

PCI Express has modified and added a new method for compliance testing. This change allows for flexibility in performing compliance testing at multiple speeds. The two methods are an inband method using a compliance load board and the Control and Status Register (CSR) method. The inband method has been modified slightly from the previous revision of the specification, where as the CSR method is using new registers defined in the 2.0 revision of the specification.

Since RapidIO does not specify compliance testing, such a feature does not exist in its specification.

1.5 Testability Enhancements

PCI Express has further added to its compliance testing features by adding receiver compliance, transmitter margining and loopback enhancements. These features allow for a cost effective way to test the transmitter and receiver. The receiver compliance is a new mode in polling compliance where the receiver sends out modified compliance patterns including a receiver error count. This is intended to margin the receiver. The transmitter margining has added multiple transmit voltage levels. The loopback enhancements are related to speed changes and the ability to loop back without achieving lock.

Since RapidIO does not specify compliance testing, such a feature does not exist in its specification.

RapidIO 2.1 has added the ability to enable ‘disabled’ lanes for testing purposes. This allows for further control over the connectivity between two link partners. Software can enable the inactive lanes on both link partners and perform its testing using implementation specific SerDes control registers and errors count in registers to determine the electrical quality of the link in each direction.

1.6 Emphasis

De-emphasis and pre-emphasis are methods to improve the overall signal-to-noise ratio on a link.

RapidIO had added pre-emphasis as a feature for the higher speeds (5.0 GBaud and above for the 2.1 specification and 3.5 Gbaud and above in the 2.0 specification). A port must provide status information about itself to the connected port to control the transmit emphasis settings. The emphasis control can be performed automatically in hardware, or under software control. The receiver can therefore change the transmitter settings, and then vary it’s own equalization to get the best possible eye

PCI Express has made de-emphasis mandatory it was optional.

These additions to both specifications provide for more robust links. The RapidIO addition provides more flexibility as it can be done automatically by the hardware, or by software.

1.7 Lane Failure

In early revisions of the PCI Express specification, the full link would become non-operational if a single lane became non-operational. In the newer specifications, PCI Express links continue to function even when one or more of the lanes become non-operational. The automatic link width negotiation feature in conjunction with auto-lane reversal enables continued system operation in the presence of a failure of a single lane within a link. The link can be reset and will then automatically negotiate to employ the usable half of the link.

RapidIO has had this feature since the first revision of the specification. In RapidIO, if a lane fails, the link falls back to 1x mode, either on lane 0 or on the redundant lane.

The lane failure feature of PCI Express, which allows the connection to continue using more than one lane, is better than the current RapidIO solution of falling back to a single lane as it will continue operation with a higher bandwidth link.

1.8 Ordering Attribute

PCI Express has added a new ordering attribute, which devices may optionally support, to provide enhanced performance for certain types of workloads and traffic patterns. The new ordering attribute relaxes ordering requirements between unrelated traffic by comparing the Requester/Completer IDs of the associated TLPs. This provides opportunities for independent request streams to bypass another congested stream, yielding performance improvement.

RapidIO has had this feature since the first revision of the specification. In RapidIO, a destinationID and sourceID pair are seen as a stream and other pairs are ordered independently.

1.9 Multicast

PCI Express has added multicast functionality by means of an Extended Capability structure. Multicast allows a single posted request TLP sent from a single source to be

distributed to multiple recipients. Multicast TLPs are identified based on their address matching the multicast window in a root complex, switch or endpoint.

RapidIO has had this feature since the second revision of the specification. In RapidIO, the destinationID is used to select whether a packet is multicast or not. Within RapidIO, any type of packet can be multicast.

1.10 Data Plane

The 2.0 revision of the RapidIO specification has added significant data plane enhancements because PCI Express is mainly focused on the PC and server markets it does not have these features. The additions to the RapidIO specification surrounding the data plane are listed below. These features add further support for RapidIO as a communication fabric providing encapsulation and traffic management features.

1.10.1 Support for Virtual Channels

Virtual Channels allow the physical channel to be subdivided into nine independently managed sub-channels called Virtual Channels (VCs). There is no ordering guaranteed between VCs and each has its own link layer flow control and buffers.
The nine VCs allow for the reserving of bandwidth and Quality of Service (QoS) on a sub-channel granularity, thus enabling different traffic management for individual VCs.

PCI Express has defined a virtual channel concept which is conceptually equivalent to priorities in RapidIO.

1.10.2 Reliable Traffic and Continuous Traffic

The 2.0 revision of the RapidIO specification added the concept of reliable traffic (RT – lossless) and continuous traffic (CT – lossy under load). Continuous traffic is time-sensitive data with no retransmissions of data under congested states. Lossless traffic is guaranteed and is retransmitted until successful.

1.10.3 Virtual Output Queuing Backpressure

Virtual Output Queuing Backpressure (VOQ) is a method that provides physical layer status messages to communicate the congestion status of downstream devices' ports. This greatly reduces head-of-line (HOL) blocking, which increases fabric performance as congestion cannot spread. Data that would normally be caught behind traffic that is competing for congested resources can be placed at the head of the queue and make forward progress.

1.10.4 Flow Control for Streaming Datagrams

The Data Streaming Logical Layer Flow Arbitration in the 2.0 RapidIO specification extends the existing congestion management to allow endpoints to manage and arbitrate for resources through the use of Xon and Xoff messages to the transmitting device.