

# IBTA Update

**Lloyd Dickman, QLogic**  
InfiniBand Trade Association

**March 24, 2009**



**INFINIBAND<sup>SM</sup>**  
TRADE ASSOCIATION

# Strong Industry Participation

## Steering Committee



Amphenol Interconnect Products  
Avago Technologies  
Bay Microsystems  
Brocade Communications Systems, Inc.  
C+M Holdings  
Cinch Connectors  
Emcore Corporation  
FCI USA, Inc.  
Finisar Corporation  
Flextronics  
Foxconn  
Fujikura America, Inc.  
Fujitsu Components America, Inc.  
Fujitsu Limited  
Hitachi, Ltd.  
JAE Electronics, Inc.  
Lamprey Networks, Inc.  
Lawrence Livermore National Laboratory  
LEONI High Speed Cables  
LSI Logic  
Luxtera, Inc.  
Meritec  
Molex Inc.  
NEC Corporation  
Network Appliance Inc.  
Obsidian Research Corporation  
Panduit Corporation  
Reflex Photonics  
Silicon Graphics, Inc.  
Tyco Electronics Corporation  
Volex  
W.L. Gore & Associates, Inc.  
Xsigo Systems  
Zarlink Semiconductor

# IBTA Structure

## □ Steering Committee

- IBM – Chet Mehta, Co-chair
- Intel – Jim Pappas, Co-chair
- Mellanox – Marc Sultzbaugh
- QLogic – Lloyd Dickman
- Sun – David Brean
- Voltaire – Yaron Haviv

## □ Administrator

- Jessica Zuver, VTM

## □ Marketing Working Group

- Brian Sparks, Mellanox and Kevin Judd, QLogic
- Active involvement of Owen Media - Samantha Chernak

## □ Technical Working Group

- Bill Magro, Intel and Diego Crupnicoff, Mellanox
  - Applications Working Group
  - ElectroMechanical Working Group – Alan Benner, IBM
  - Link Working Group – Nir Arad, Mellanox and John Russo, QLogic
  - Management Working Group – Jimmy Hill, IBM and Eitan Zahavi, Mellanox
  - Software Working Group – Tom Sand, IBM

## □ Compliance and Interoperability Working Group

- Tuan Phamdo, Intel and Lenny Dingle, IBM
- Active involvement of Lamprey Networks and UNH/IOL – Rupert Dance

# Marketing Working Group

□ March 2008 - IDC InfiniBand Market Forecast Update

□ April 2008 – Taneja Group InfiniBand Storage

□ April 2008 – Brochure

□ September 2008 Technology Forum



## MARKET ANALYSIS Worldwide InfiniBand 2007-2011 Forecast Update

Stephen L. Josselyn Lucinda Borwick

### IDC OPINION

There have been some changes in the demand for InfiniBand (IB) since our previous forecast, which was published in May 2007. The compound annual growth rate (CAGR) for host channel adapter (HCA) revenue is now set at 20%, with revenue reaching \$279.7 million in 2011. The CAGR for InfiniBand switch revenue is 47.2%, with revenue growing to \$556.3 million in 2011. We still believe the main drivers for InfiniBand adoption are increased use of high-performance computing (HPC), scale-out database environments, and shared and virtualized I/O and increasing demand from financial applications with HPC-like characteristics. Key elements driving changes in this forecast include the following:

- The underlying server market forecast has not been changed significantly from the version that was the basis of our previous forecast. However, there have been some changes in terms of the rate of its adoption and the rate of changeover from single data rate (SDR) to double data rate (DDR) solutions. The move to DDR has proceeded slightly faster than we anticipated in our previous market assessment, and our forecast has been adjusted accordingly.
- The demand for increased bandwidth continues to grow, and the requirement to meet these needs without increasing the complexity or cost is still a major factor. The establishment of InfiniBand as a standards-based architecture is a key for future customer adoption.
- The need to understand customer deployment models for the use of interconnect technology remains an important part of InfiniBand supplier product development in the future. The ability for InfiniBand to coexist with Ethernet and Fibre Channel in the currently installed LAN and SAN infrastructures, respectively, goes a long way toward solving those customer concerns. Development continues for both Ethernet and InfiniBand solutions, and the availability of cost-effective and high-performance Ethernet/Fibre Channel gateway products continues to be a focus, with a method for employing the solution that workload requirements.

File Information: March 2008, IDC #1171, Volume 1, Title: Host Channel Adapter, Technology Markets, Market Analysis



## Technology Brief

**InfiniBand in the Enterprise:  
Next Generation Data Center Computing**  
April 2008



There is a flurry of activity afoot in the industry today around the demands of the next generation data center. Without a doubt, the circumstances are challenging; space and power are at a premium and only getting more so; there are thousands of different demands for I/O even in small data centers, with new demands popping up at frequencies that require every IT manager to have scenarios of prediction when designing the infrastructure; and the pace of change and reconfiguration around every server and storage array in the enterprise is phenomenal. In the midst of this, the typical storage manager finds they are managing a broader range of connectivity and protocols than ever before – spanning Ethernet, Fibre Channel, IP, SCSI, NFS, and more. It is little wonder that all corners of the industry are grasping after new solutions and coming up with new standards for better I/O management with less complexity.

In the middle of this storm of activity, stands InfiniBand. Long looked to as the leader for high performance computing, InfiniBand is gaining increasing adoption in general purpose enterprise computing. Why? A few reasons include: cost effectiveness, simplicity, flexibility and most of all performance. At first glance, those characteristics might be easily dismissed as a set of buzzwords, but their combination in one I/O platform in fact brings the next generation of data center connectivity to bear today. Moreover, InfiniBand is leaping and bounding beyond the competition in each of those capabilities. In this Technology Brief, we'll look at the challenges surrounding enterprise I/O management, dive into key InfiniBand capabilities, and take a look at how InfiniBand is being adopted by enterprise today.

### The Data Center Built of Sticks

Much like the story of the three little pigs, many current generation data centers are figuratively built of sticks. Those sticks are the I/O fabrics that connect systems together. Today's common fabrics are not up to the task of supporting an ever-changing and evolving infrastructure that is being bombarded by a deluge of requests from business users for more compute power and storage than ever before. In the face of these

requests, the data center requires more ability to scale I/O bandwidth, manage I/O allocation, and consolidate I/O than ever before.

IT managers are searching for ways to obtain these capabilities in a data center fabric made out of heterogeneous strands of connectivity. The result? These managers must add more switches, more cables, more adapters, more protocols, and more network administration and control devices in



### ADVANTAGES OF INFINIBAND

**Powerful multi-core CPUs, server virtualization, blade architectures, increasing storage demands, and the need to reduce power usage all place increasing demands on I/O solutions. InfiniBand meets these demands with:**

**Superior performance.** InfiniBand is the only shipping solution that supports 20Gb/s host connectivity and 50Gb/s switch-to-switch links.

**Low latency.** InfiniBand's ultra-low latency, with measured delays of 1µs end-to-end, greatly accelerates many data center and high-performance computing (HPC) applications.

**High efficiency.** InfiniBand provides direct support of advanced, remote transport protocols such as Remote Direct Memory Access (RDMA) to enhance the efficiency of customer workload processing.

**Cost effectiveness.** InfiniBand Host Channel Adapters (HCAs) and switches are very competitively priced and create a compelling price performance advantage over alternative technologies.

**Fabric consolidation and low energy usage.** With less than 5W per 20Gb/s port, InfiniBand can consolidate networking, clustering, and storage data over a single fabric, which significantly lowers the overall power, real estate, and management overhead required for servers and storage. Enhanced Quality of Service (QoS) capabilities support storing and managing multiple workloads and traffic classes.

**Reliable, stable connections.** InfiniBand is perfectly suited to meet the mission-critical needs of today's enterprise by enabling fully redundant and lossless (0 fabric), with automatic path failover and link layer re-routing abilities to meet the highest levels of availability.

**Data integrity.** Mission-critical applications require that data has not been corrupted as it travels across the fabric. InfiniBand enables the highest levels of data integrity by performing cyclic redundancy checks (CRC) at each fabric hop and end-to-end across the fabric to ensure the data is correctly transmitted.

**Rich, growing ecosystem.** InfiniBand is at the center of an ecosystem that includes open-source software distribution from the OpenFabrics Alliance, innovative and cost-effective scaling, and long-lived solutions that reach outside the data center and across the globe.

These advantages combine to make InfiniBand the industry's most powerful interconnect.



### INFINIBAND MARKET HOOKS

According to IDC research, high-end HPC, scale-out database environments, virtualized I/O, and increasing demand applications with HPC-like character will continue to drive the next wave of

### Worldwide Revenue for Host Channel



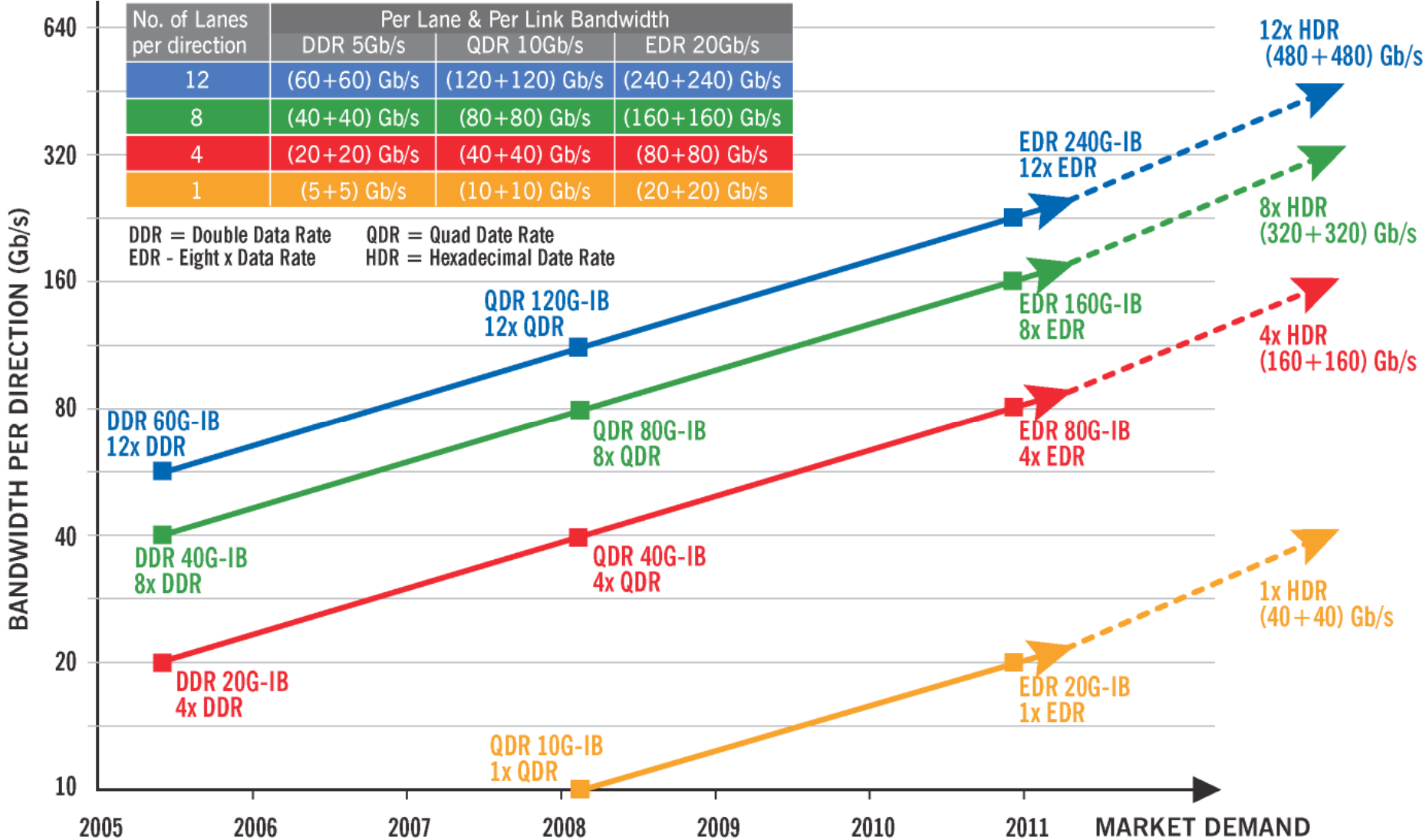
### Worldwide Revenue for InfiniBand Switches



Source: IDC, "Worldwide InfiniBand for HCA", March 2008

Copyright © 2008 TANEJA Group, Inc. All Rights Reserved. 1 of 4

# InfiniBand Link Speed Roadmap



# Technical Working Group

## □ Future HPC Enhancements

- Observers have noted that IBA extensions will be needed to play in future ultra-high tiers of HPC
- Candidate issues
  - Limits on node count (LID address space)
  - Limitations on routing (destination routing)
  - Scalability of SM/SA & connection management
  - Memory requirements of connections @ scale
  - Signaling rates
  - Point-to-point nature vs. collective requirements of apps
  - RDMA vs. MPI architecture limits communication/computation overlap

# EWG – Electro-mechanical Working Group

- QDR electrical specs
- QSFP specs (4x connector)
- CXP (12x connector)
- EDR (coming attractions)



# LWG – Link Working Group

## □ IB Routers

- Published Router strawman and distributed to other working groups
- Reached consensus with other workgroups on requirements and high level approach to router definition
- High level approach/status presented at OFA Conference

## □ XRC

- XRC Annex version 1.0

## □ IBTA Specification Maintenance

- Processed errata process

**Supplement to  
InfiniBand™ Architecture  
Specification Volume  
1.2.1**

**Annex A14:  
Extended Reli-  
able Connected  
(XRC) Transport  
Service**

March 2, 2009  
Revision 1.0

Copyright © 2003-2009, by InfiniBand™ Trade Association.  
All rights reserved.

This document contains information proprietary to the InfiniBand™ Trade Association. Use or disclosure without written permission by an officer of the InfiniBand™ Trade Association is prohibited.



# SWG – Software Working Group

- XRC
- IB Routers
- Limited Partition Membership for Multicast

# New Specification Errata Process

- Accelerated Errata Process
  - Bypass need to revise spec, multiple approval levels, have member review
  - Quarterly, or if necessary, more urgent
  
- Errata Classification
  - Each WG categorizes WG APPROVED errata into two classes: material and immaterial
  - Material errata are reviewed by TWG/SC whereas immaterial ones are not
  - Examples
    - Immaterial ones are typos in the spec or other cosmetic changes
    - Material ones are typos in the spec and other behavioral changes

# MgtWG – Management Working Group

- Approved 50 comments vs. 1.2.1 spec (currently 13 open)
- Hierarchy Information Annex
  - Standardize device hierarchy information retrieval and encoding
  - Annex now ready for cross group review
- Vendor Specific Mads Table
- Add error conditions for
  - LinkWidthSpeedPairsTable
  - VendorSpecificMadsTable
- SM Performance Optimization
- Partial P\_Key support by MCG
- Performance Monitoring Enhancements

*Supplement to  
InfiniBand™ Architecture  
Specification Volume  
1.2.1*

**Annex A15:  
Hierarchy Infor-  
mation**

March 22, 2009  
Revision 1.0

Copyright © 2003-2009, by InfiniBand™ Trade Association.  
All rights reserved.  
This document contains Information proprietary to the InfiniBand™ Trade Association. Use or disclosure without  
written permission by an officer of the InfiniBand™ Trade Association is prohibited.

# Hierarchy Example

Figure 1 Equipment Rack

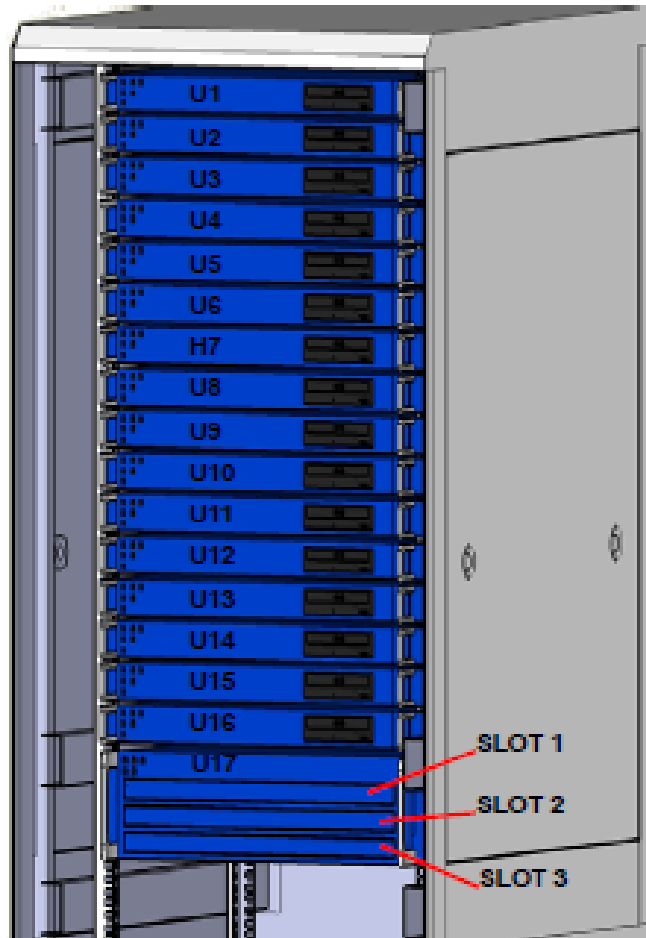
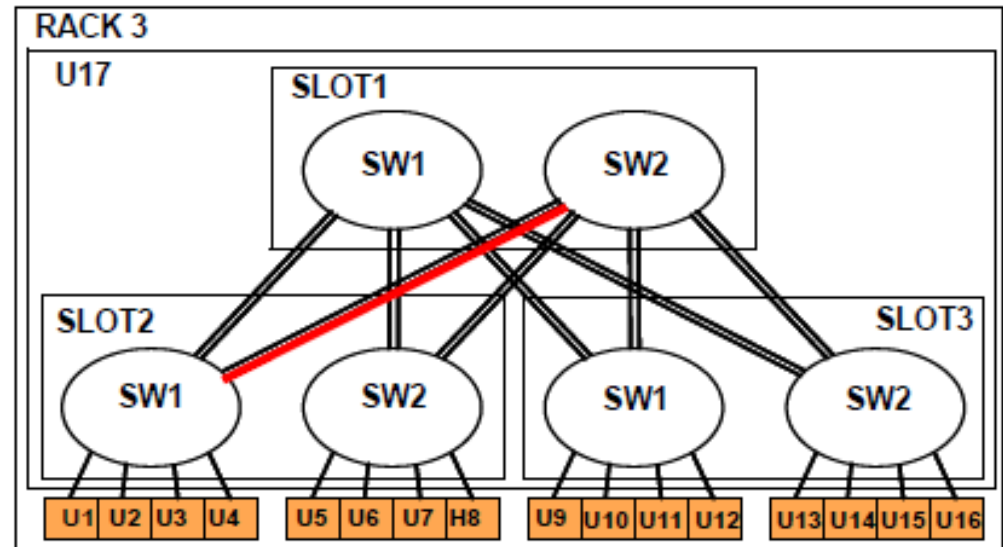


Figure 3 Physical hierarchy view



# Compliance and Interoperability Working Group

## Integrators' List

### Integrators Lists

- Devices
- Cables

### Compliance Plugfests

- #13: March 2008
- #14: September 2008

18 vendors

16 devices

- 10 HCAs, 3 switches, 2 Range Extenders, 1 SRP Target

114 cables tested

IBTA Integrators' List  
September 2008 Plugfest

#### HCAs (Host Channel Adapter)

(Device Listing indicates Core Compliance)

Manufacturer	Product Description	Model	HW	SW	Core Compliance	Additional Capabilities			
						SM/S.A	IPv6	SRP	SRP Target
Mellanox Technologies LTD	InfiniHost™ HCA card, dual-port 10Gb/s, PCI-X host bus, 128MB memory	MHETC-1TC Cragger SDR	A3	3.50	✓	✓	✓	✓	✓
	InfiniHost™ E8 HCA card, dual-port 10Gb/s, PCIe1.2 x8 host bus, 128MB memory	MHGA8-1TC Lon-Cat SDR	A2	4.8.2	✓	✓	✓	✓	✓
	InfiniHost™ E8 HCA card, dual-port 10Gb/s, PCIe1.2 x8 host bus, MemFree	MHGA8-1TC Lon-MN SDR	A2	5.3.0	✓	✓	✓	✓	✓
	InfiniHost™ E8 HCA card, single-port 10Gb/s, PCIe1.2 x8 host bus, MemFree	MHES18-1TC Chester SDR	A0	1.2.0	✓	✓	✓	✓	✓
	InfiniHost™ E8 HCA card, single-port 10Gb/s, PCIe1.2 x8 host bus, MemFree	MHES14-1TC Tiger SDR	A2	1.2.0	✓	✓	✓	✓	✓
	InfiniHost™ E8 HCA card, dual-port 20Gb/s, PCIe1.2 x8 host bus, 128MB memory	MHGA28-1TC Lon-Cat SDR	A3	4.8.2	✓	✓	✓	✓	✓
	InfiniHost™ E8 HCA card, dual-port 20Gb/s, PCIe1.2 x8 host bus, MemFree	MHGA28-1TC Lon-MN SDR	A3	5.3.0	✓	✓	✓	✓	✓
QLogic	SDR FCx8 HCA	QLB-7240		2.2	✓	✓	✓	✓	✓
	SDR FCx8 HCA	QLB-7280		2.2	✓	✓	✓	✓	✓

#### Storage Systems

(Device Listing indicates Core Compliance)

Manufacturer	Product Description	Model	HW	SW	Core Compliance	Additional Capabilities			
						SM/S.A	IPv6	SRP	SRP Target
LSI Logic	Duplex-B Storage System	7501-B	2.0	07.32.15.00	✓	✓	✓	✓	✓

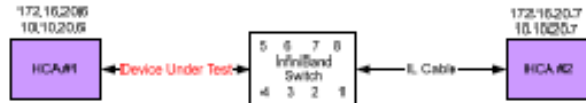
# Upcoming IBTA Plugfest

- April 20-24, 2009
- Hosted by Lamprey Networks and University of New Hampshire Interoperability Lab (UNH-IOL)
- Validation of new QDR Active Cable testing process
  - Applies to QDR Active Cable Interface, for cables which incorporate nonlinear limiting amplifiers – and includes active optical cables, active copper cables, half-active copper cables, and transceivers.
- Proposal for Updated QDR Specifications Based on WDP
  - Prior methodology based on specification of a good “eye” is insufficient at higher signaling rates. Waveform Dispersion Penalty is a measure of the deterministic penalty of a waveform with a reference equalizing receiver.
  - If this testing process proves itself at the April Plugfest, it will be incorporated into the InfiniBand Vol. 2 specification at a later date.

# Cable Interoperability Testing

## Topology Diagrams

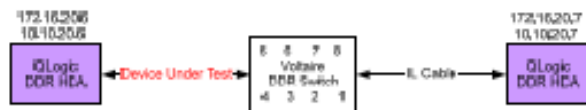
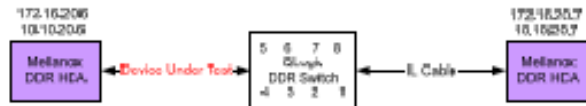
### Channel #1



### Channel #2



## Example Configurations



- All Cables must pass the Interoperability Tests defined by the CIWG in addition to the IBTA Compliance tests.

- Addresses the situation in which a cable may be compliant with the current IBTA Spec and yet fail to interoperate when combined with a worse case transmitter and receiver.

# Integrators' List Logo Program

- Branding logos are available to all vendors with InfiniBand products that have passed rigorous compliance testing
- Vendors can affix the new logo to cables and/or devices, and use it in press releases and advertising materials
- Benefits:
  - End users can easily identify the IB products that are qualified for use in high performance computing and enterprise environments that have passed compliance testing
  - Manufacturers of IB equipment will be given additional recognition for achieving Integrators' List status



# IBTA Integrators List Logos

## Device Logo



## Cable Logos 4X

## Cable Logos 12X



# Sample Logo Usage



# Summary

- The IBTA is committed to furthering the InfiniBand specification and related compliance process.
- InfiniBand community is vibrant and will continue to innovate.
- The IBTA roadmap and analyst reports project InfiniBand's continuing momentum.

**Thank You**



**INFINIBAND<sup>SM</sup>**

TRADE ASSOCIATION MEMBER