Recent Topics in the IBTA…
…and a Look Ahead

Bill Magro, IBTA Technical WG Co-Chair
Intel Corporation
InfiniBand Trade Association (IBTA)

- Global member organization dedicated to developing, maintaining and furthering the InfiniBand specification
  - Develops architecture specification
    - RDMA (Remote Direct Memory Access) software architecture
    - InfiniBand, up to 100Gb/s per port
    - RDMA over Converged Ethernet (RoCE)
  - Supports compliance and interoperability testing of commercial products
  - Markets and promotes InfiniBand and RoCE
    - Online, marketing and public relations engagements
    - IBTA-sponsored technical events and resources

Steering committee members

© InfiniBand Trade Association
IBTA Governance

- Steering Committee
- Technical Working Group
  - Marketing Working Group
- ElectroMechanical WG
  - Link WG
  - Software WG
  - Compliance & Interop WG
Industry Trends Shape InfiniBand’s Evolution

• Ever-growing bandwidth demands
• Cloud computing & extreme-scale data centers
• New non-volatile memory technologies
InfiniBand Technical Updates

• Volume 2, Release 1.3.1
  – Enhanced EDR and FDR functionality
    • Enables improved EDR cable management through additions to connector memory map
    • Additional FEC option supporting lower latency
    • Enables SM to optimize signal integrity with lowest power through enhanced CDR management
  – Improved interoperability
    • Electrical specification updates for EDR interoperability
    • Specification corrections for FDR interoperability and test methodologies
    • Test methodology improvements for EDR Limiting Active Cables
InfiniBand Technical Updates

• Virtualization Annex to Volume 1, Release 1.3
  – Brought about by increasing use of virtualized workloads from data center and cloud environments that leverage InfiniBand
  – Extends support for multiple virtualized endpoints within InfiniBand hardware
  – Simplifies management of virtual machines and improves scalability
NETWORK VIRTUALIZATION
New Annex Defines InfiniBand Network Virtualization

- Introduces concept of virtualized endpoints
- Provides more efficient view of logical endpoints
- Improves scalability of subnet management
INFINIBAND ARCHITECTURE
LINK SPEED ROADMAP
InfiniBand Roadmap

<table>
<thead>
<tr>
<th>4x Link Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDR</td>
</tr>
<tr>
<td>56 Gb/s</td>
</tr>
</tbody>
</table>

Link Bandwidth per direction, Gb/s

©2015 InfiniBand® Trade Association
InfiniBand EWG

• Volume 2 of the InfiniBand specification defines
  – Link speeds – SDR/DDR/QDR/FDR/EDR and coming speeds HDR, NDR,…
  – Link Initialization – Port-Port bringup and negotiation of capabilities (speed, coding, FEC, equalization tuning,…)
  – Coding and modulation: Data format and forward error correction coding
  – Analog Electrical Interface Specifications – Jitter, Amplitude, Crosstalk, Signal/Noise Ratio,…
  – Transceiver/Module Packaging – QSFP/SFP/CXP specs

• Key Current Work:
  – HDR and NDR interface specifications
    • New methods to spec passive cables for HCA↔Switch links
    • Tighter specs as speeds get higher
Overall objectives:
– Improve link speed & latency
– Preserve signal integrity, connectors, power efficiency, electrical/optical options
– Align with other industry efforts

Technical Details:
– Electrical interfaces: from 2-level (NRZ) to 4-level coding (PAM-4)
  • Improve signal integrity at higher speeds
– Improving methods for specifying cables
  • "Channel Operating Margin" model simplifies testing & improves yield
– Strengthen support for Active Optical Cables
  • Benefits of optical transmission without hassle of optical connectors
– Preserving EDR methods for Forward Error Correction
  • Adaptable FEC enables tradeoff of latency vs. signal-to-noise ratio
– Preserving connector/module definition (QSFP28, etc.)
  • Minimize impact to system packaging
Compliance and Interop Program

• Bi-annual plugfests held at UNH-IOL
  – Results featured in the InfiniBand Integrators’ List and RoCE Interoperability List
  – Supports InfiniBand and RoCE deployment planning

• Most recent plugfest (Oct’17)
  – RoCE interoperability testing
    • Number of vendors testing RoCE products at IBTA Plugfests continues to grow
    • 10, 40, 50 (NEW) and 100 GbE (NEW) solutions tested
  – InfiniBand interoperability testing
    • Similar group of vendors
    • EDR solutions tested

• 31st IBTA plugfest coming in April
PERSISTENT MEMORY & RDMA
New Non-Volatile Memory

• Several innovative technologies emerging
• Attractive characteristics
  – Performance approaching that of DRAM
  – Capacity approaching that of NAND
  – High durability
• Storage and memory applications
Possible RDMA Enhancements

• Define ordering of memory visibility and persistence
  • Separate rules today for ordering on fabric vs. in the server
• Provide means to trigger persistence (e.g., FLUSH)
• Provide confirmation and notification of persistence
  • To sender, avoiding need for software-based reply from target
  • To target, a la RDMA Write with immediate data
• Define atomicity characteristics of persistent RDMA Writes
Summary

• Industry trends inform InfiniBand’s evolution
• Recent updates define new link speeds and network virtualization
• Enhanced EWG & CIWG efforts improve interoperability
• Work underway to define 200Gb/s link speed, HDR
• Emerging non-volatile memory technologies pose new opportunities for InfiniBand