



Education

Fabric Consolidation with InfiniBand

Dror Goldenberg, Mellanox Technologies

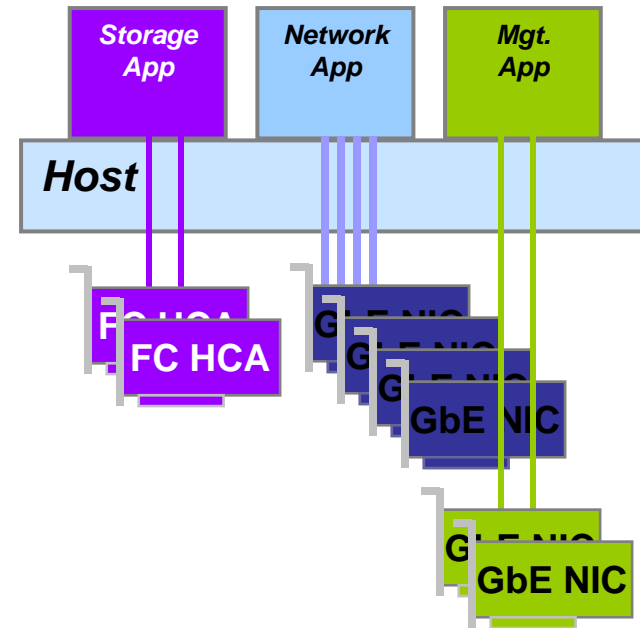
- The material contained in this tutorial is copyrighted by the SNIA.
 - Member companies and individuals may use this material in presentations and literature under the following conditions:
 - ◆ Any slide or slides used must be reproduced without modification
 - ◆ The SNIA must be acknowledged as source of any material used in the body of any document containing material from these presentations.
 - This presentation is a project of the SNIA Education Committee.
 - Neither the Author nor the Presenter is an attorney and nothing in this presentation is intended to be nor should be construed as legal advice or opinion. If you need legal advice or legal opinion please contact an attorney.
 - The information presented herein represents the Author's personal opinion and current understanding of the issues involved. The Author, the Presenter, and the SNIA do not assume any responsibility or liability for damages arising out of any reliance on or use of this information.
- NO WARRANTIES, EXPRESS OR IMPLIED. USE AT YOUR OWN RISK.**

- In the era of exploding datacenter requirements for networking and storage services, and with the increasing power, space, and budget concerns over the infrastructure, fabric consolidation becomes inevitable. InfiniBand was designed from day one for fabric consolidation. With 120Gb/s links and with ultra low-latency characteristics, InfiniBand provides a well provisioned foundation for consolidation of networking and storage. Additional features such as QoS, partitioning, virtual lanes, lossless fabric, and congestion management facilitate true consolidation of fabrics along with connectivity of InfiniBand islands to Ethernet and Fibre Channel clouds through gateways. This session highlights the features for fabric consolidation and the various protocols that run over InfiniBand with emphasis on storage protocols.

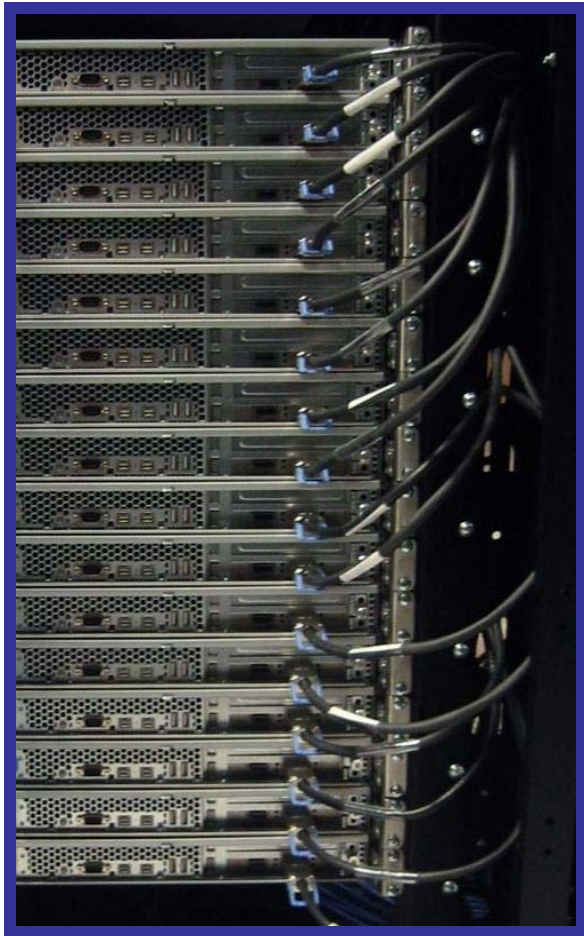
- Learning objectives
 - ◆ Understand the InfiniBand architecture and feature set.
 - ◆ Understand the benefits of InfiniBand for fabric consolidation.
 - ◆ Understand the standard InfiniBand storage protocols.

- Motivation and General Overview
- Protocol Stack Layers
- Storage Protocols over InfiniBand
- Benefits

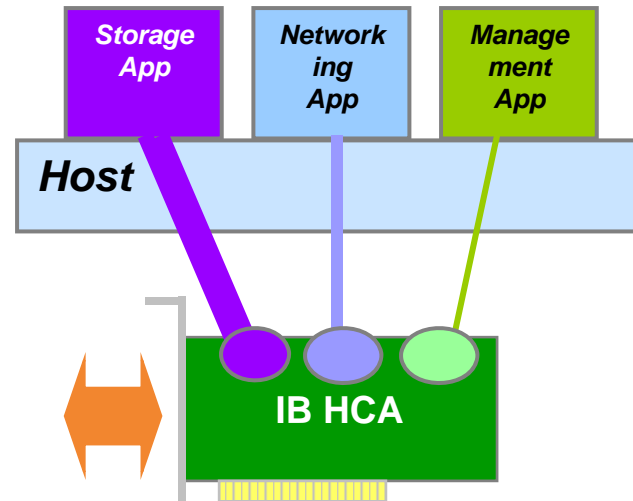
Motivation for Fabric Consolidation



- ◆ Slower I/O drives more NICs, HBAs, wires
- ◆ Different service needs drive different fabrics
- ◆ No flexibility
- ◆ More ports, fabrics to manage
- ◆ More power
- ◆ More space
- ◆ ...Higher TCO



One
Wire

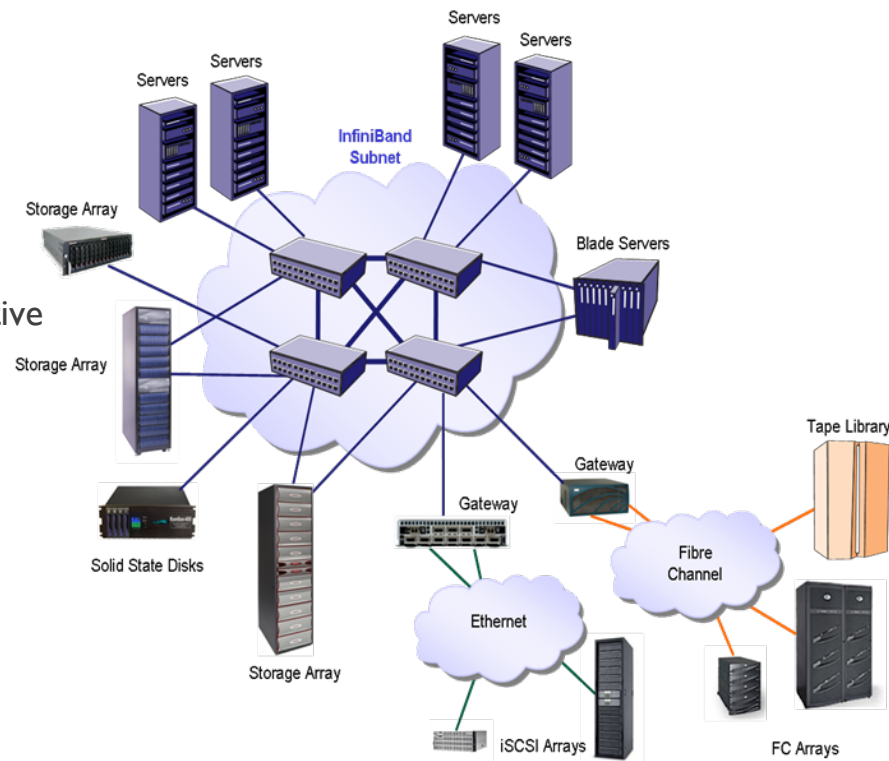


- High bandwidth pipe for capacity provisioning
- Dedicated I/O channels enable convergence
 - ◆ For Networking, Storage, Management traffic
 - ◆ QoS – across different traffic types
 - ◆ Partitions and isolation provided
- Flexibility
 - ◆ Soft servers and fabric repurposing

Why InfiniBand ?



- ◆ The InfiniBand standard - strong alternative to
 - ◆ Fibre Channel (SAN)
 - ◆ Ethernet (NAS)
- ◆ Superior performance
 - ◆ 40Gb/s host/target ports
 - ◆ 120Gb/s switch to switch
 - ◆ Sub 1µs end to end latencies
 - ◆ Aggressive roadmap
- ◆ Unified fabric for the Data Center
 - ◆ Storage, networking and clustering over a single wire
 - ◆ Scalable to 1000s nodes
- ◆ Cost Effective
 - ◆ Compelling price/performance advantage over alternative technologies
- ◆ Low power Consumption – Green IT
 - ◆ Less than 0.15W per Gb/s
- ◆ Mission Critical
 - ◆ Highly reliable fabric
 - ◆ Multi-pathing
 - ◆ Automatic failover
 - ◆ Highest level of data integrity



Fabric Technologies Comparison

Features	Fibre Channel 8G FC	Ethernet 10GigE	InfiniBand 4X QDR
Line Rate (GBaud)	8.5	10.3125	40
Unidirectional Throughput (GBytes/s)	800	1,250	4,000*
Reliable Service	Practically no	No	Yes
Fabric Consolidation	Practically no	FCoE coming soon ...	Yes
Copper Distance	15m	10GBASE-CX4 15m 10GBASE-T 100m	Passive 7m Active 15m
Optical Distance†	100m	10GBASE-SR 300m 10GBASE-LR 220m	100-300m

* Theoretical, 3.2 GB/s measured due to server I/O limitations

† Data center oriented media

Physical Layer

- Width (1X, 4X, 8X, 12X) including auto-negotiation
- Speed (SDR/DDR/QDR/EDR) including auto-negotiation
 - ◆ 4X QDR HCAs are currently shipping
- Power management
- Connector
 - ◆ Board: MicroGiGaCN*
 - ◆ Pluggable: QSFP
- 8/10 encoding
 - ◆ Maintain DC Balance
 - ◆ Limited run length of 0's or 1's
- Control symbols (Kxx.x)
 - ◆ Lane de-skew
 - ◆ Auto negotiation / training
 - ◆ Clock tolerance
 - ◆ Framing

Link Speed (10⁹ bit/sec)

Lane Speed →	SDR (2.5GHz)	DDR (5GHz)	QDR (10GHz)	EDR (20GHz)
Link Width ↓				
1X	2.5	5	10	20
4X	10	20	40	80
8X	20	40	80	160
12X	30	60	120	240

* MicroGiGaCN is a trademark of Fujitsu Components Limited

Physical Layer – Cont'd

Copper Cables*:

Width	Speed	Connector	Reach	Type / Power ‡
4X	SDR/ DDR/ QDR	Micro-GiGaCN/ QSFP	20m/ 10m/ 7m	Passive
4X	DDR	Micro-GiGaCN	15-25m	Active 0.5-1.5W
4X	QDR	QSFP†	12-15	Active 1-2W
12X	SDR/ DDR	24pin Micro-GiGaCN	20m/ 10m	Passive

4X MicroGiGaCN



4X QSFP



12X MicroGiGaCN



Fiber Optics*:

Width	Speed	Connector	Reach	Type / Power ‡	Fiber Media
4X	SDR/ DDR	Micro-GiGaCN / QSFP	300m/ 150m	Media Converter 0.8-1W	12 strand MPO fiber
4X	DDR/ QDR	Micro-GiGaCN / QSFP†	100m/ 300m	Optical Cable 1-3W	12 strand attached

4X MicroGiGaCN
Media Converter



4X QSFP
Media Converter



4X MicroGiGaCN
Optical Cable



* Currently deployed
 † Sampling
 ‡ Per End

➤ Addressing and Switching

- ◆ Local Identifier (LID) addressing
- ◆ Unicast LID - 48K addresses
- ◆ Multicast LID – up to 16K addresses
- ◆ Efficient linear lookup
- ◆ Cut through switching supported
- ◆ Multi-pathing support through LMC

➤ Independent Virtual Lanes

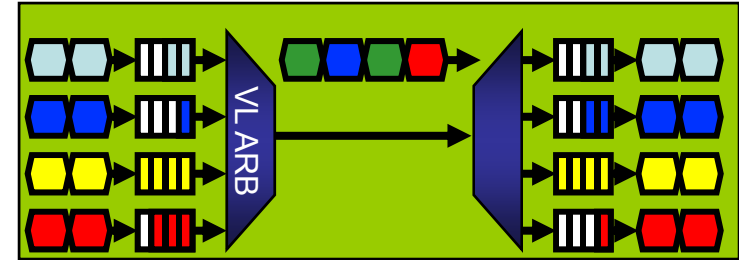
- ◆ Flow control (lossless fabric)
- ◆ Service level
- ◆ VL arbitration for QoS

➤ Congestion control

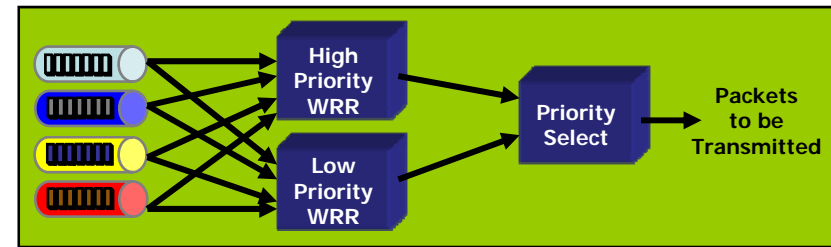
- ◆ Forward / Backward Explicit Congestion Notification (FECN/BECN)

➤ Data Integrity

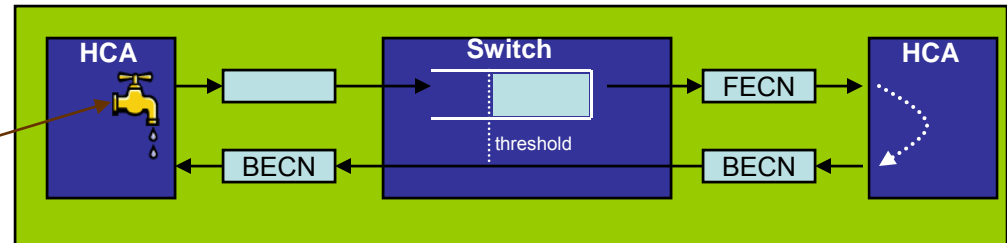
- ◆ Invariant CRC
- ◆ Variant CRC



Independent Virtual Lanes (VLs)



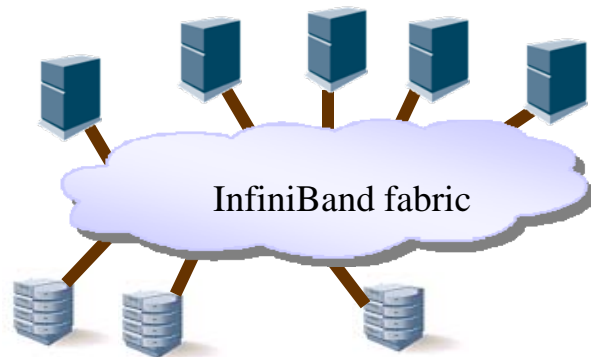
H/L Weighted Round Robin (WRR) VL Arbitration



Efficient FECN/BECN Based Congestion Control

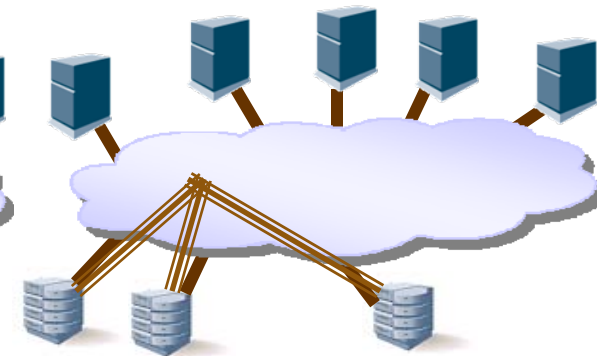
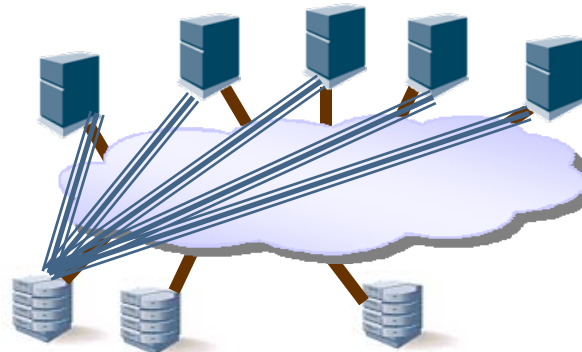
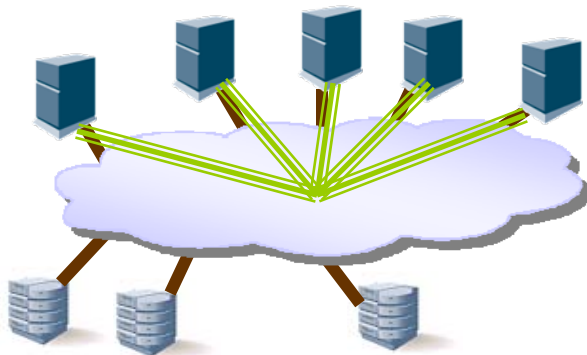
Fabric Consolidation – VLs and Scheduling Example

Physical:



➤ VLs and scheduling can be dynamically configured and adjusted to match application performance requirements

Logical:



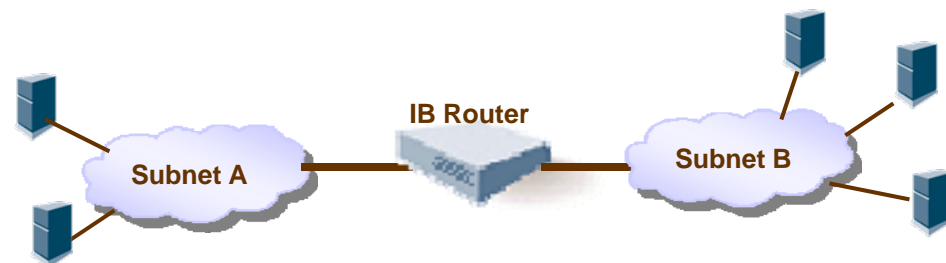
==== Low Latency VL
For Clustering

==== Mainstream Storage VL
Day - at least 40% BW
Night - at least 20% BW

==== Backup VL
Day - at least 20% BW
Night - at least 60% BW

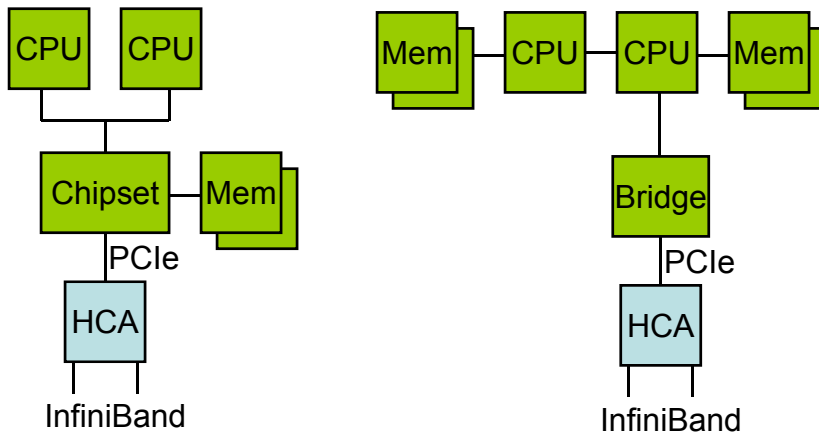
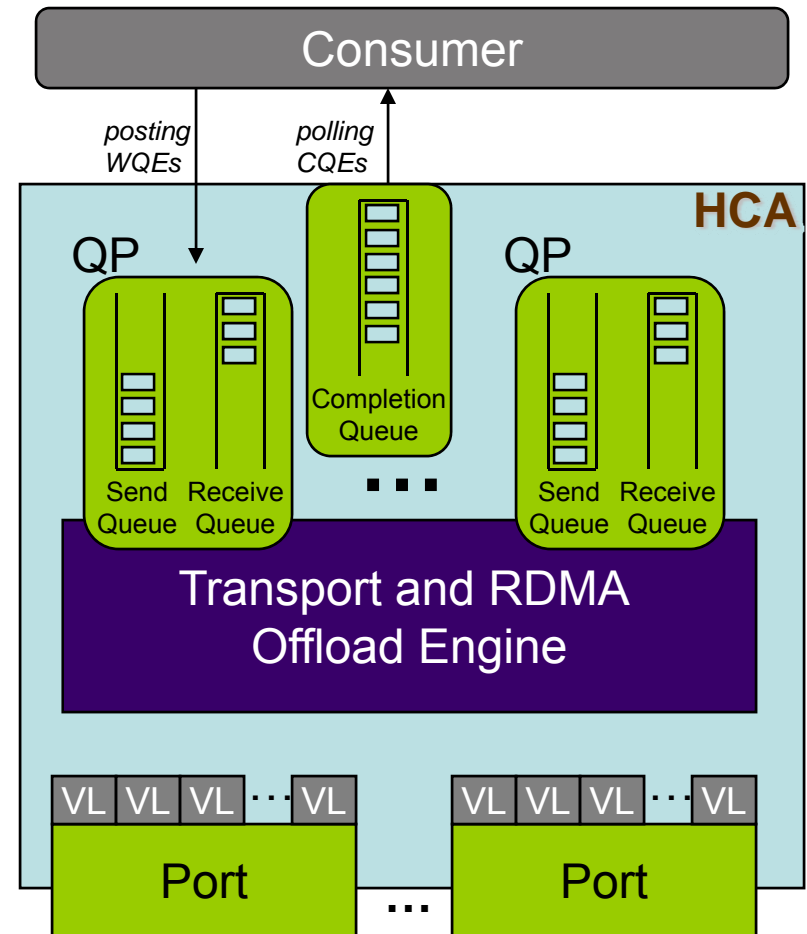
Network Layer

- Global Identifier (GID) addressing
 - ◆ Based on IPv6 addressing scheme
 - ◆ $GID = \{64 \text{ bit GID prefix}, 64 \text{ bit GUID}\}$
 - › GUID = Global Unique Identifier (64 bit EUI-64)
 - › GUID 0 – assigned by the manufacturer
 - › GUID 1..(N-1) – assigned by the Subnet Manager
- Optional for local subnet access
- Used for multicast distribution within end nodes
- Enables routing between IB subnets
 - ◆ Definition underway in IBTA
 - ◆ Will leverage IPv6 routing algorithms

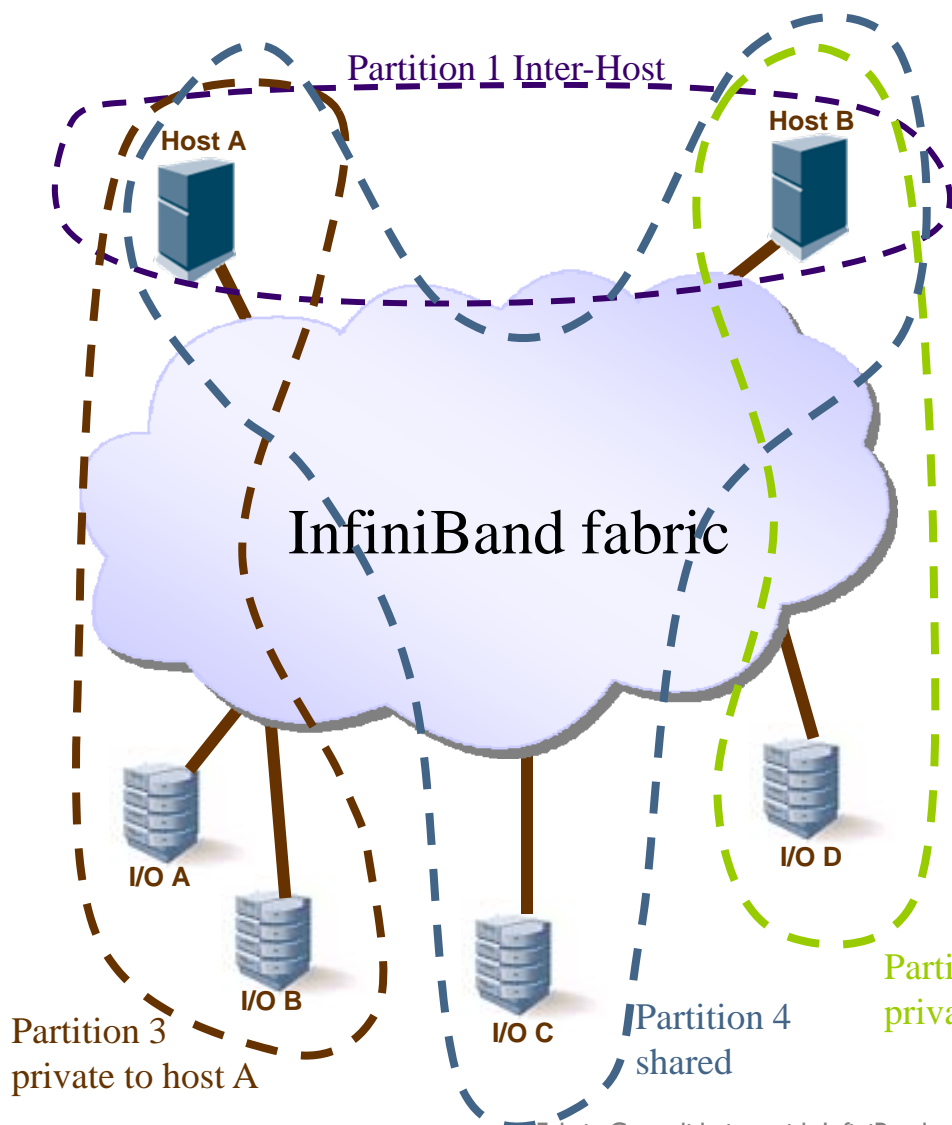


Transport – Host Channel Adapter (HCA) Model

- Asynchronous interface
 - ◆ Consumer posts work requests
 - ◆ HCA processes
 - ◆ Consumer polls completions
- I/O channel exposed to the application
- Transport services
 - ◆ Reliable / Unreliable
 - ◆ Connected / Datagram
 - ◆ Send/Receive, RDMA, Atomic operations
- Offloading
 - ◆ Transport executed by HCA
 - ◆ Kernel bypass
 - ◆ RDMA



Partitions



- Logically divide the fabric into isolated domains
- Partial and full membership per partition
- Partition filtering at switches

- Similar to
 - ◆ 802.1Q VLANs
 - ◆ FC Virtual Fabrics (VFs)

InfiniBand Data Integrity

➤ Hop by hop

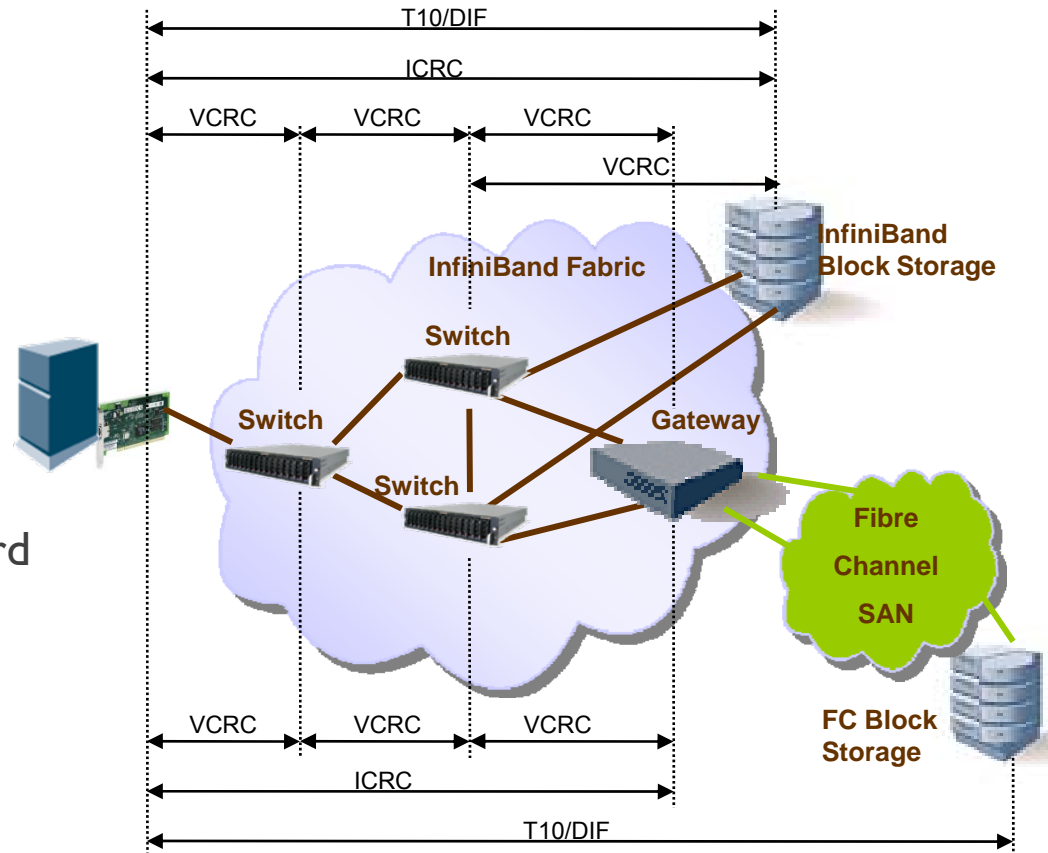
- ◆ VCRC – 16 bit CRC
- ◆ CRC16 0x100B

➤ End to end

- ◆ ICRC – 32 bit CRC
- ◆ CRC32 0x04C11DB7
- ◆ Same CRC as Ethernet

➤ Application level

- ◆ T10/DIF Logical Block Guard
 - > Per block CRC
- ◆ 16 bit CRC 0x8BB7



- Multi-port HCAs
 - ◆ Covers link failure
- Redundant fabric topologies
 - ◆ Covers link failure
- Link layer multi-pathing (LMC)
- Automatic Path Migration (APM)
- ULP High Availability
 - ◆ Application level multi-pathing (SRP/iSER)
 - ◆ Teaming/Bonding (IPoIB)
 - ◆ Covers HCA failure and link failure

Upper Layer Protocols

- ULPs connect InfiniBand to common interfaces
- Supported on mainstream operating systems

➤ Clustering

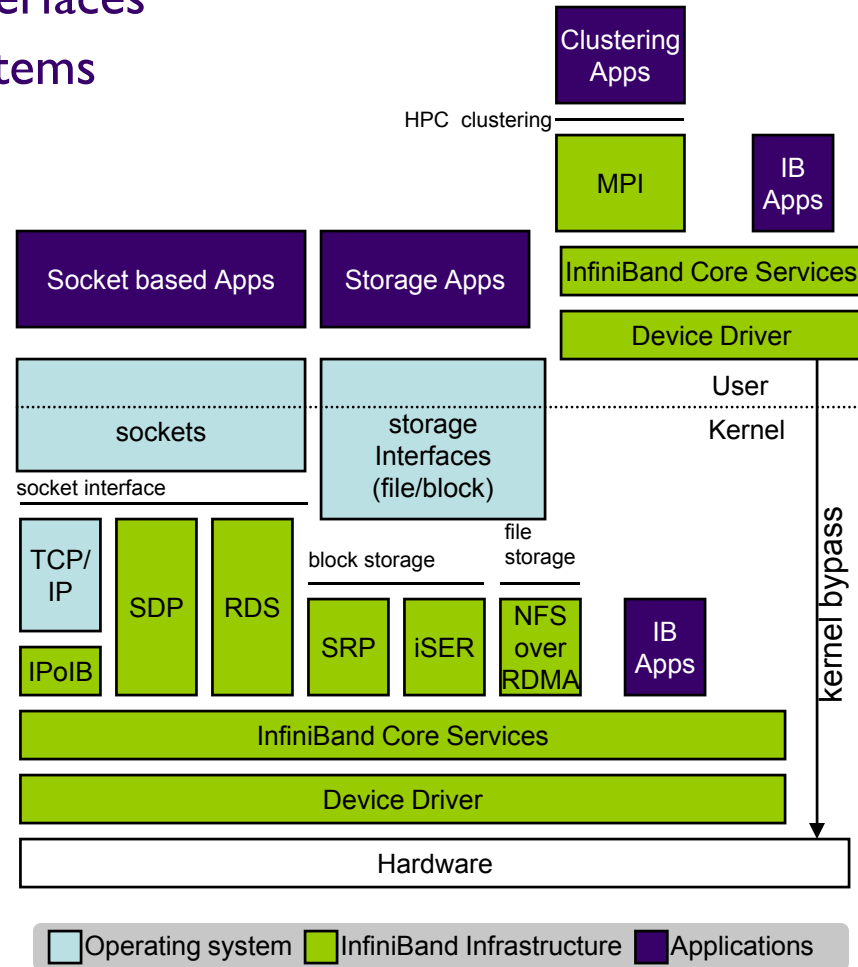
- ◆ MPI (Message Passing Interface)
- ◆ RDS (Reliable Datagram Socket)

➤ Network

- ◆ IPoIB (IP over InfiniBand)
- ◆ SDP (Socket Direct Protocol)

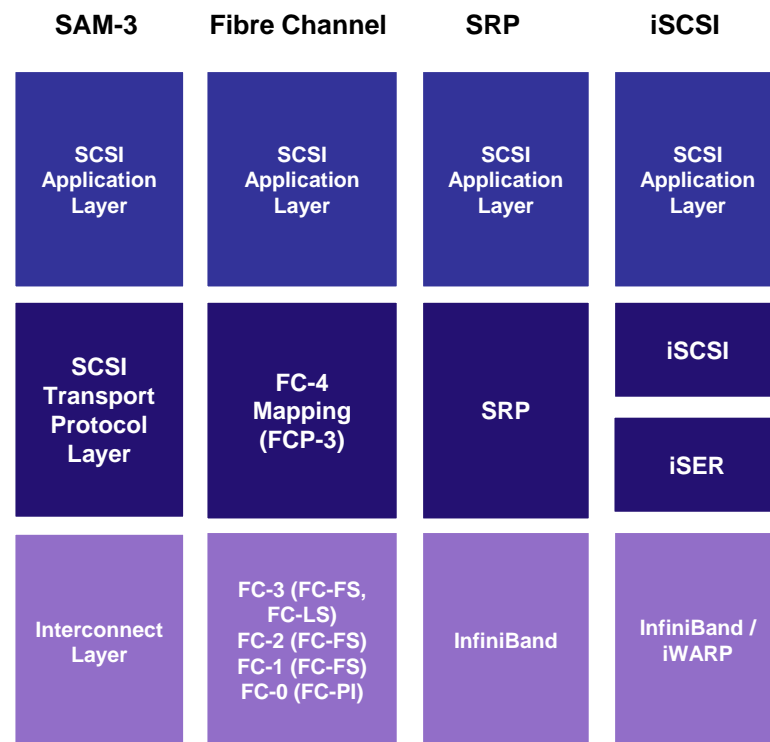
➤ Storage

- ◆ SRP (SCSI RDMA Protocol)
- ◆ iSER (iSCSI Extensions for RDMA)
- ◆ NFSoRDMA (NFS over RDMA)



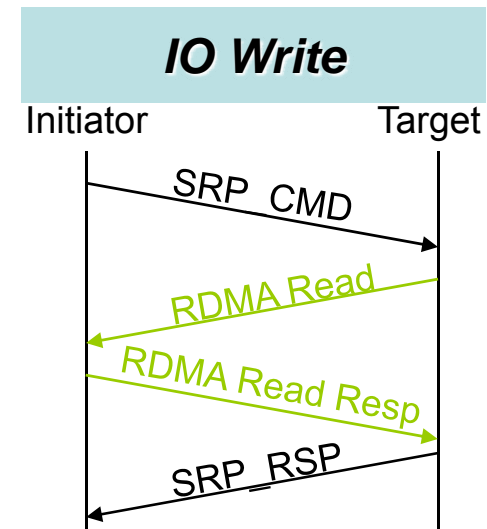
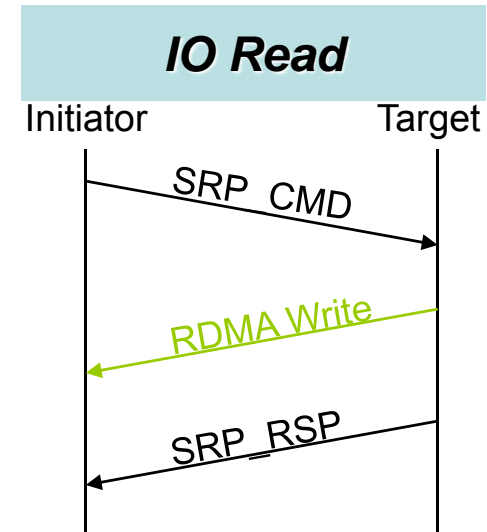
InfiniBand Block Storage Protocols

- SRP - SCSI RDMA Protocol
 - ◆ Defined by T10
- iSER – iSCSI Extensions for RDMA
 - ◆ Defined by IETF IP Storage WG
 - ◆ InfiniBand specifics (e.g. CM) defined by IBTA
 - ◆ Leverages iSCSI management infrastructure
- Protocol offload
 - ◆ Use IB Reliable Connected
 - ◆ RDMA for zero copy data transfer



SRP - Data Transfer Operations

- Send/Receive
 - ◆ Commands
 - ◆ Responses
 - ◆ Task management
- RDMA – Zero Copy Path
 - ◆ Data-In
 - ◆ Data-Out
 - ◆ Target issues the RDMA operations
- iSER uses the same principles
 - ◆ Immediate/Unsolicited data allowed through Send/Receive
- iSER and SRP are part of mainline Linux kernel

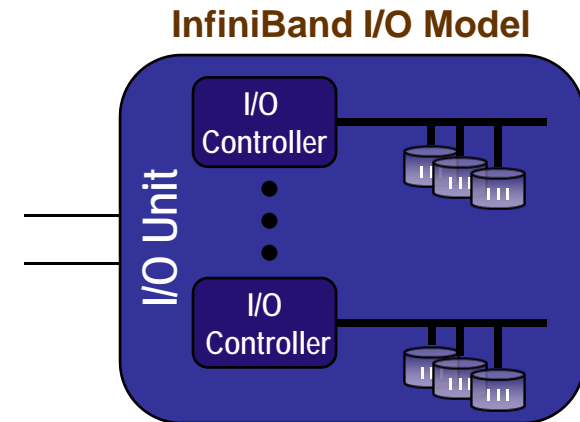


➤ SRP

- ◆ Persistent Information {Node_GUID:IOC_GUID}
- ◆ Subnet Administrator (Identify all ports with CapabilityMask.IsDM)
- ◆ Identifiers
 - › Per LUN WWN (through INQUIRY VPD)
 - › SRP Target Port ID {IdentifierExt[63:0], IOC GUID[63:0]}
 - › Service Name – SRP.T10.{PortID ASCII}
 - › Service ID – Locally assigned by the IOC/IOU

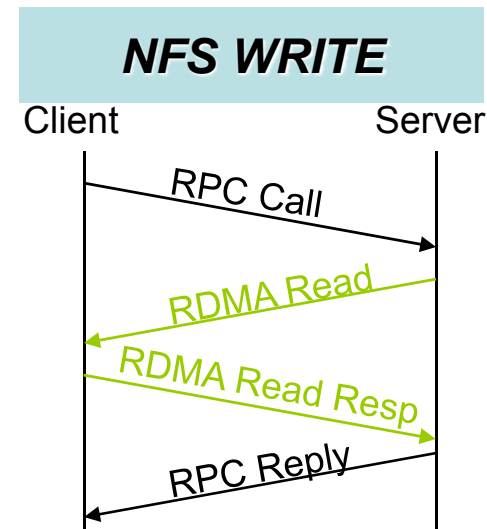
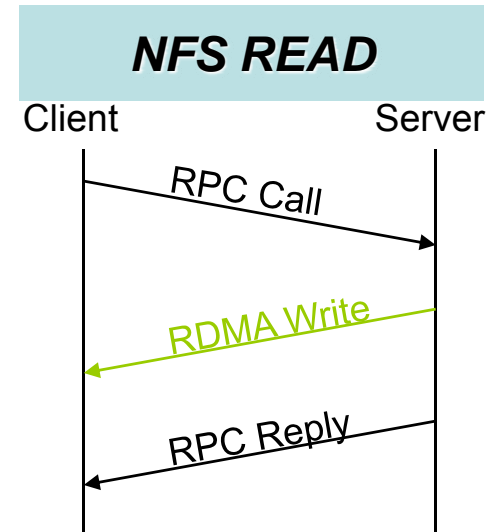
➤ iSER – uses iSCSI's (RFC 3721)

- ◆ Static Configuration {IP, port, target name}
- ◆ Send Targets {IP, port}
- ◆ SLP
- ◆ iSNS
- ◆ Target naming (RFC 3721/3980)
 - › iSCSI Qualified Names (iqn.), IEEE EUI64 (eui.), T11 Network Address Authority (naa.)



NFS over RDMA

- **Defined by IETF**
 - ◆ ONC-RPC extensions for RDMA
 - ◆ NFS mapping
- **RPC Call/Reply**
 - ◆ Send/Receive – if small
 - ◆ Via RDMA Read chunk list - if big
- **Data transfer**
 - ◆ RDMA Read/Write – described by chunk list in XDR message
 - ◆ Send – inline in XDR message
- **Uses InfiniBand Reliable Connected QP**
 - ◆ Uses IP extensions to CM
 - ◆ Connection based on {IP, port}
 - ◆ Zero copy data transfers
- **NFS over RDMA is part of mainline Linux kernel**



➤ Benefits

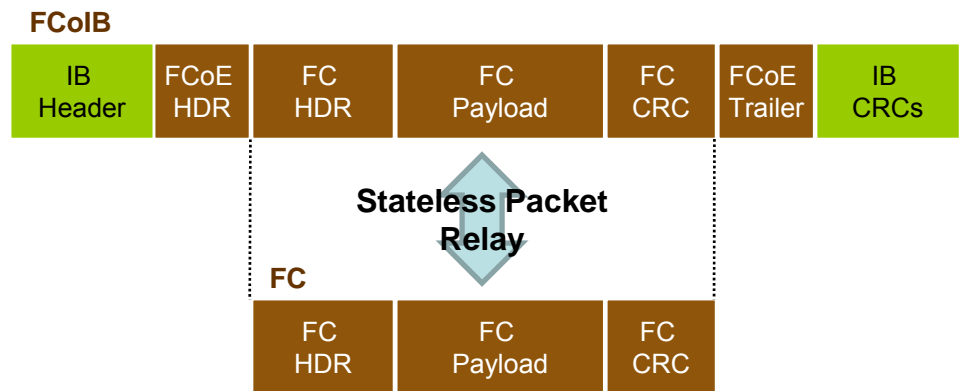
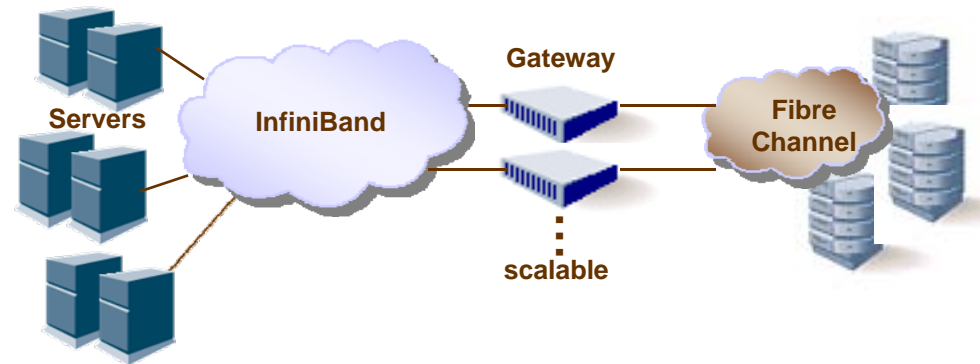
- ◆ Connectivity of InfiniBand islands to SAN
- ◆ I/O scales independently of computing
- ◆ Design based on average load across multiple servers

➤ Current Gateways

- ◆ SRP → FC
- ◆ iSER → FC
- ◆ Stateful architecture

➤ Gateways Futures

- ◆ FCoIB → FC
- ◆ FCoE sibling
- ◆ Stateless architecture
- ◆ Scalable, high performance



What Drives Efficient Clustering?

- **Ultra Low Latency**
 - ◆ Kernel bypass
 - ◆ Flexible polling vs. interrupt model
 - ◆ Cut through switching
- **High Bandwidth/Message Rate**
 - ◆ HCAs match server I/O available bandwidth
 - ◆ Efficient offloaded implementation
 - ◆ Congestion management
- **Scalability**
 - ◆ Fat tree with equal bisectional bandwidth
 - ◆ Linear routing for up to 48K LIDs
- **Overlapped Communication and Computation**
 - ◆ RDMA – saves data copy, runs in parallel

What Drives Efficient Storage Access?

➤ Full I/O offload

- ◆ Zero copy
- ◆ Interrupt avoidance (moderated per I/O interrupt)
- ◆ Offloaded segmentation and reassembly
- ◆ Transport reliability
- ◆ Lossless fabric – VL credit based flow control

➤ Fabric Consolidation

- ◆ Partitions - isolation
- ◆ VL Arbitration – QoS
- ◆ Host virtualization friendliness
- ◆ High throughput
- ◆ Performance counters

Performance Metrics

➤ InfiniBand Verbs (Native)

- ◆ Latency
 - › RDMA Write 0.9us
 - › RDMA Read 1.87us (roundtrip)
- ◆ Bandwidth
 - › 3.2GB/s (unidirectional)
 - › 6.5GB/s (bidirectional)
- ◆ Measured on IB QDR

➤ Clustering (MPI)

- ◆ Latency 0.97us
- ◆ Message rate 40M msg/sec
- ◆ Measured on IB QDR

➤ Block Storage (SRP)

- ◆ Bandwidth (1MB I/O, no RAID)
 - › I/O Read 1.7GB/s
 - › I/O Write 1.6GB/s
 - › Single DDR port
- ◆ Measured on IB DDR

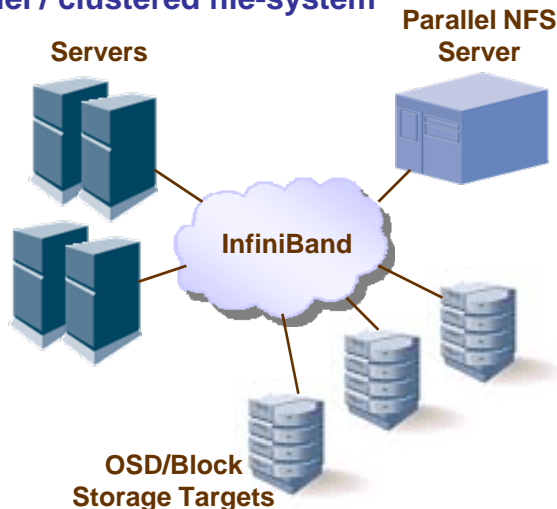
➤ File Storage (NFSoverRDMA)

- ◆ Read 1.6GB/s
- ◆ Write 0.59GB/s
- ◆ Measured on IB DDR

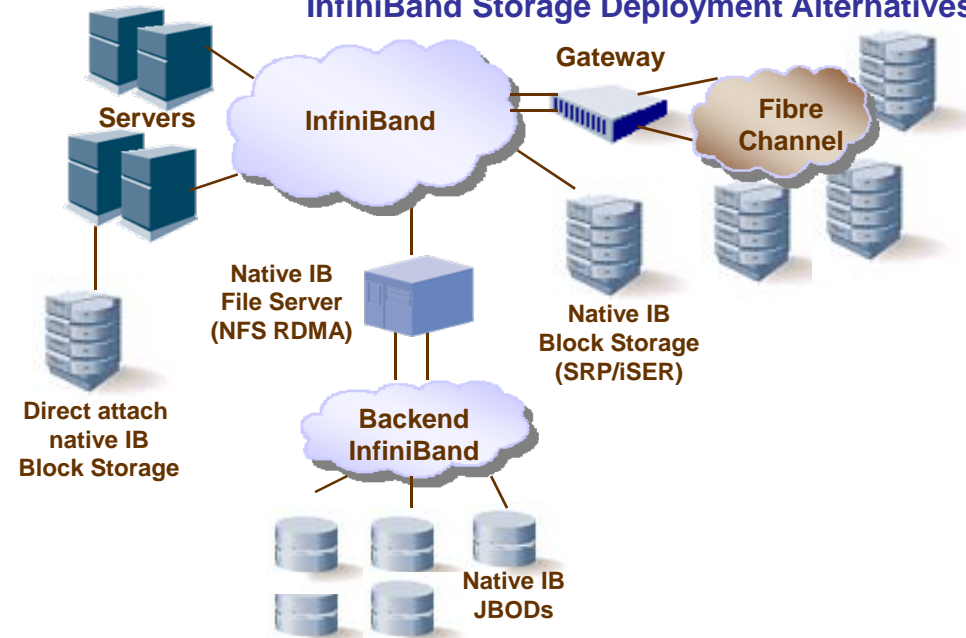
InfiniBand Storage Opportunities & Benefits

- Clustering port can connect to storage
- High Bandwidth Fabric
- Fabric consolidation (QoS, partitioning)
- Efficiency – full offload and zero copy
- Gateways
 - ◆ One wire out of the server
 - ◆ Shared remote FC ports – scalability
- Network cache

Parallel / clustered file-system



InfiniBand Storage Deployment Alternatives



- Clustered/Parallel storage, Backend fabric benefits:
 - ◆ Combined with clustering infrastructure
 - ◆ Efficient object/block transfer
 - ◆ Atomic operations
 - ◆ Ultra low latency
 - ◆ High bandwidth

- **InfiniBand I/O is a great fit for the datacenter**
 - ◆ Layered implementation
 - ◆ Price/Performance, power
 - ◆ Enables efficient SAN, Network, IPC and Management traffic

- **InfiniBand brings true fabric consolidation**
 - ◆ Gateways provide scalable connectivity to existing fabrics
 - ◆ Fabric is fully featured for consolidation (partitions, QoS, over provisioning, etc.)

- **Existing storage opportunities with InfiniBand**
 - ◆ E.g. connectivity to HPC clusters, where IB is the dominant fabric



➤ Check out SNIA Tutorials

- ◆ **Comparing Server I/O Consolidation Solutions: iSCSI, InfiniBand and FCoE**
- ◆ **Fibre Channel over Ethernet (FCoE)**
- ◆ **Ethernet Enhancements for Storage: Deploying FcoE**

- Please send any questions or comments on this presentation to SNIA: tracknetworking@snia.org

**Many thanks to the following individuals
for their contributions to this tutorial.**

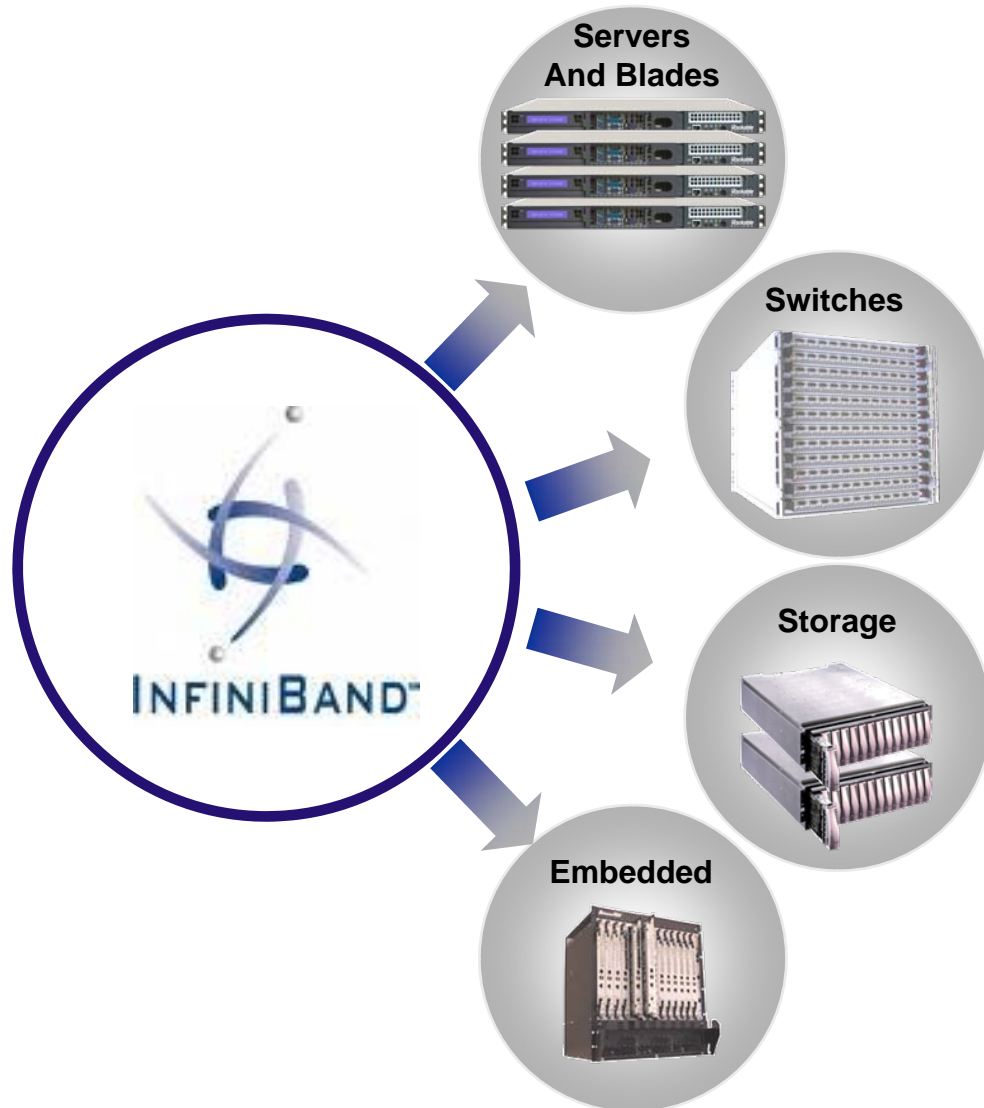
- SNIA Education Committee

**Bill Lee
Graham Smith
Joe White
Sujal Das**

**Gilad Shainer
Howard Goldstein
Ron Emerick
Walter Dey**

Backup

Interconnect: A Competitive Advantage



End-Users

Enterprise Data Centers

- Clustered Database
- eCommerce and Retail
- Financial
- Supply Chain Management
- Web Services

High-Performance Computing

- Biosciences and Geosciences
- Computer Automated Engineering
- Digital Content Creation
- Electronic Design Automation
- Government and Defense

Embedded

- Communications
- Computing and Storage Aggregation
- Industrial
- Medical
- Military

- **Data Centers**
 - ◆ Clustered database, data warehousing, shorter backups, fabric consolidation, power savings, virtualization, SOA, XTP
- **Financial**
 - ◆ Real-time risk assessment, grid computing and fabric consolidation
- **Electronic Design Automation (EDA) and Computer Automated Design (CAD)**
 - ◆ File system I/O is the bottleneck to shorter job run times
- **High Performance Computing**
 - ◆ High throughput I/O to handle expanding datasets
- **Graphics and Video Editing**
 - ◆ HD file sizes exploding, shorter backups, real-time production

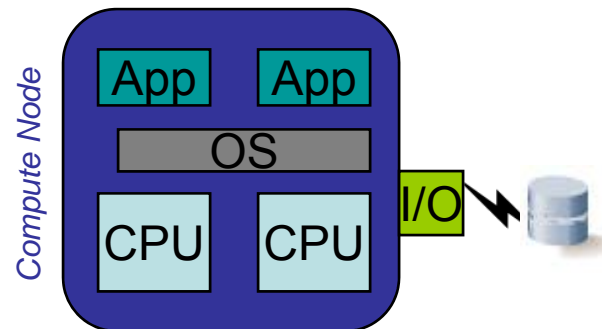
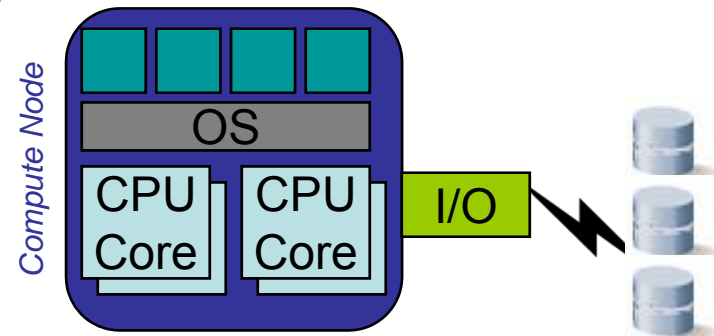
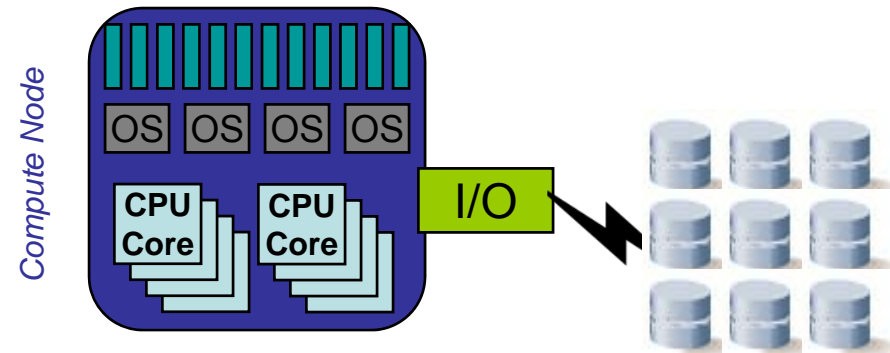
The Need for Better I/O

➤ Datacenter trends

- ◆ Multi-core CPUs
- ◆ Bladed architecture
- ◆ Fabric consolidation
- ◆ Server virtualization & consolidation
- ◆ Increasing storage demand

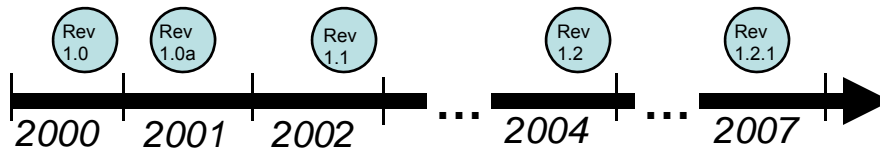
➤ Better I/O is required

- ◆ High capacity
- ◆ Efficient
 - › Low latency
 - › CPU Offload
- ◆ Scalable
- ◆ Virtualization friendly
- ◆ High availability
- ◆ Performance
- ◆ Low power
- ◆ TCO reduction

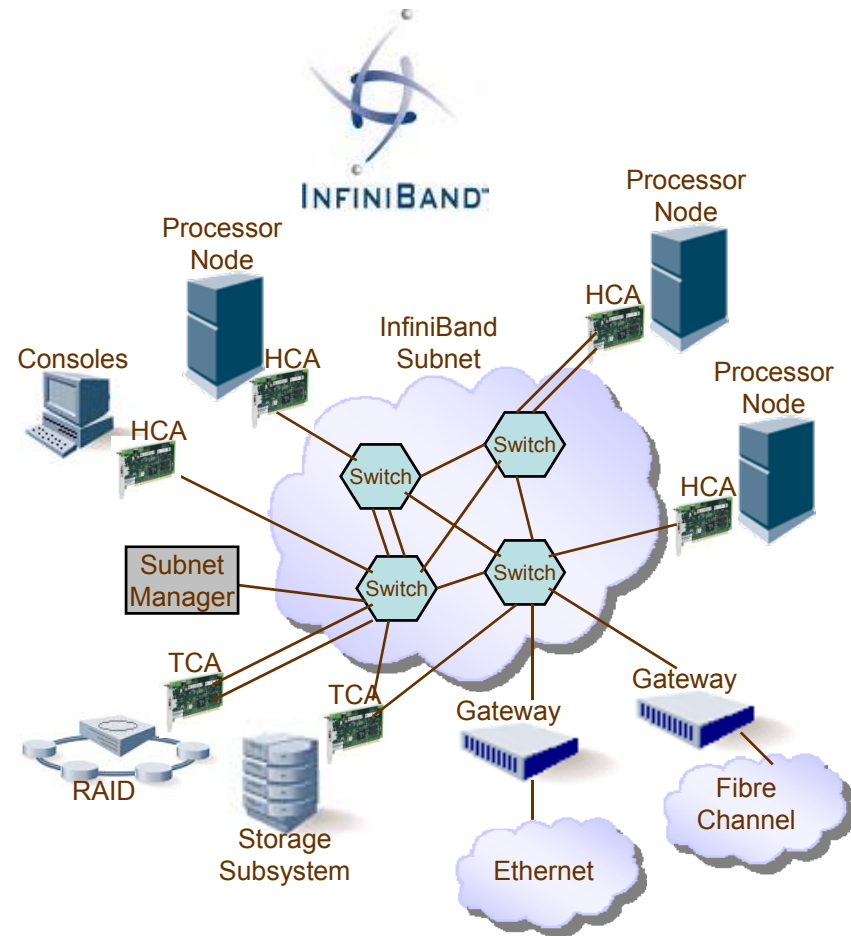


The InfiniBand Architecture

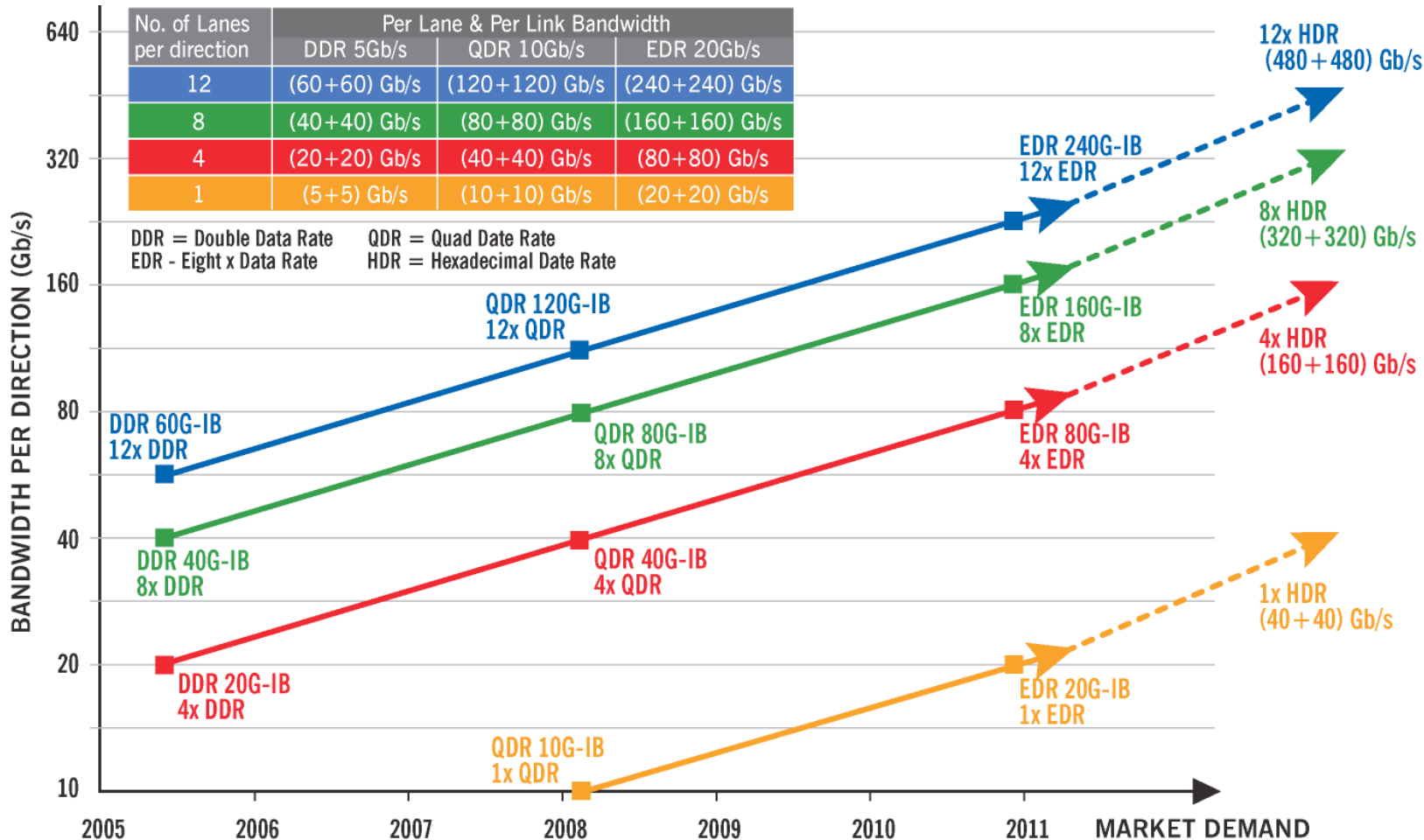
- Industry standard defined by the InfiniBand Trade Association
- Defines System Area Network architecture
 - ◆ Comprehensive specification:
 - from physical to applications



- **Architecture supports**
 - ◆ Host Channel Adapters (HCA)
 - ◆ Target Channel Adapters (TCA)
 - ◆ Switches
 - ◆ Routers
- **Facilitated HW design for**
 - ◆ Low latency / high bandwidth
 - ◆ Transport offload



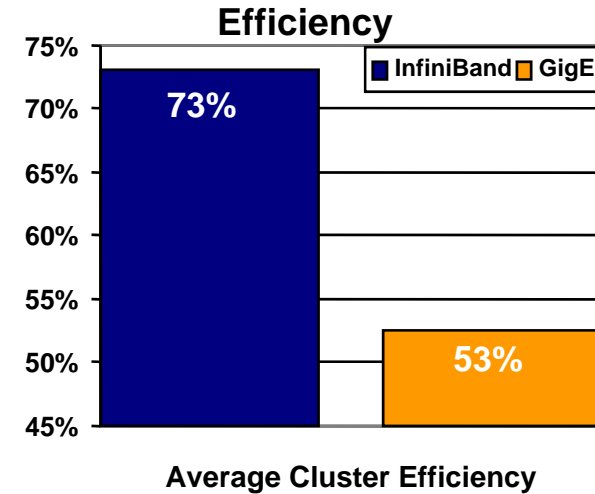
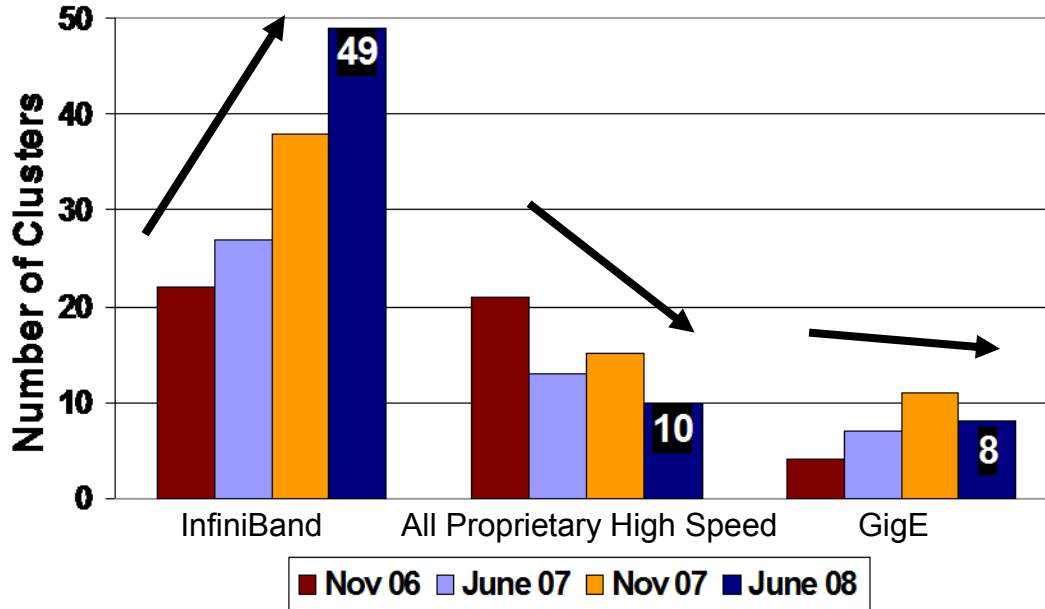
InfiniBand Wire Speed Roadmap



INFINIBANDSM
Trade Association

Interconnect Trends – Top500

Top100 Interconnect Trends

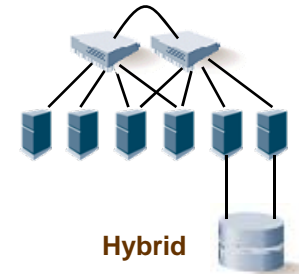
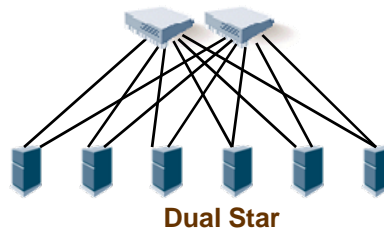
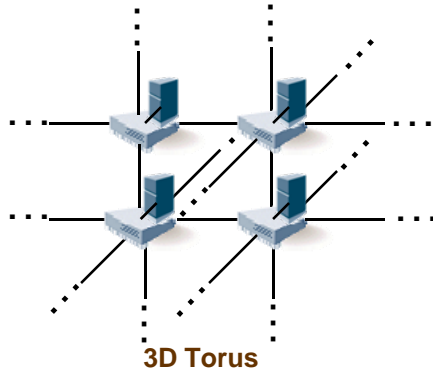
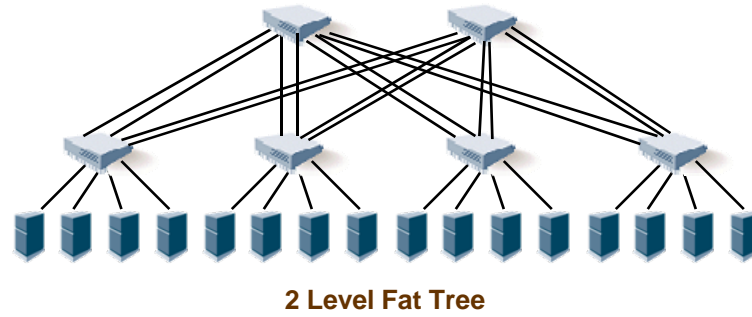
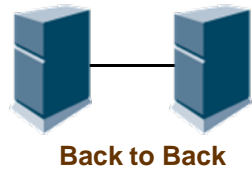


➤ InfiniBand cluster scale beyond Petaflop

Source: <http://www.top500.org/>

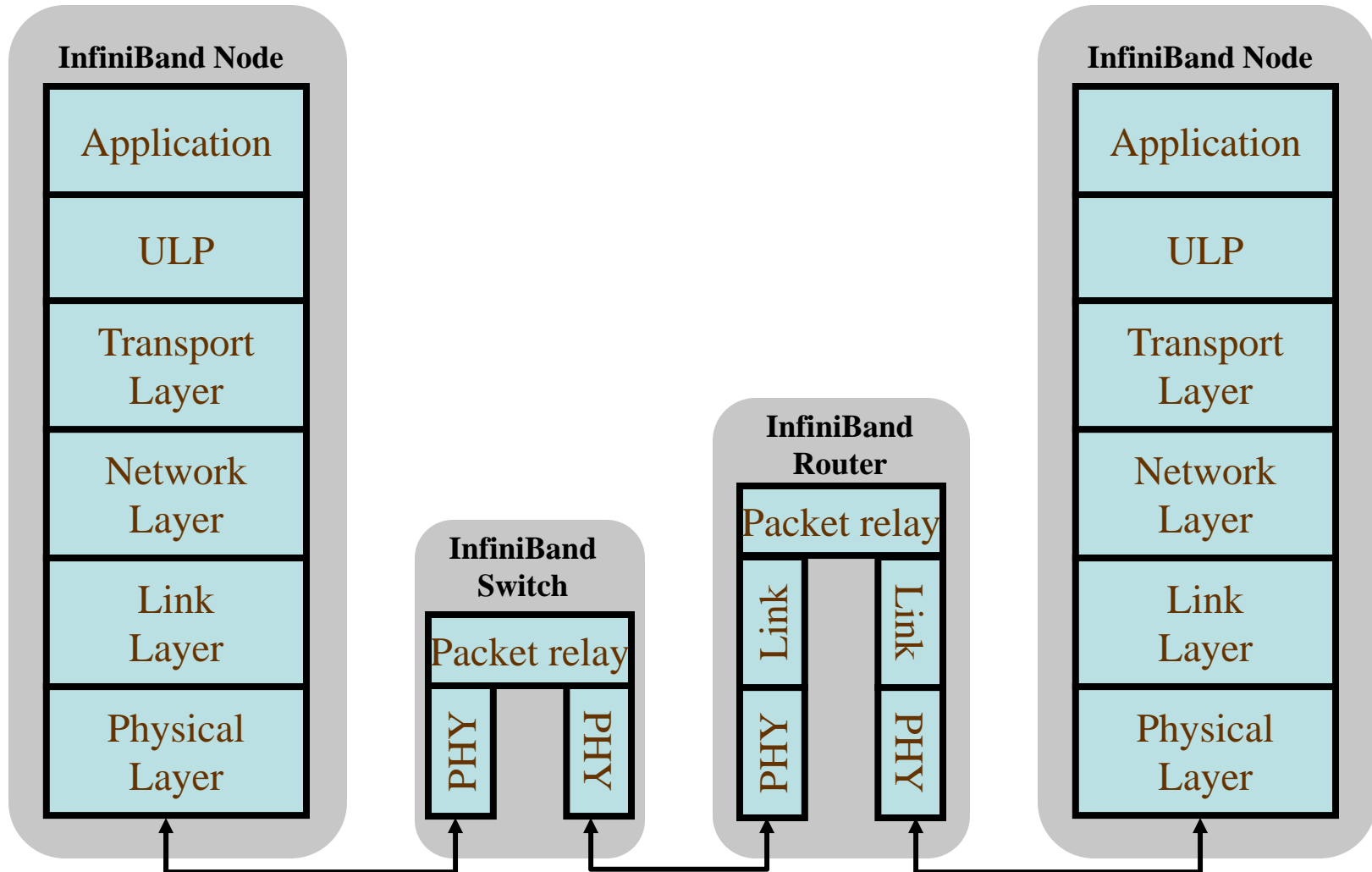
The TOP500 project was started in 1993 to provide a reliable basis for tracking and detecting trends in high-performance computing.

InfiniBand Topologies

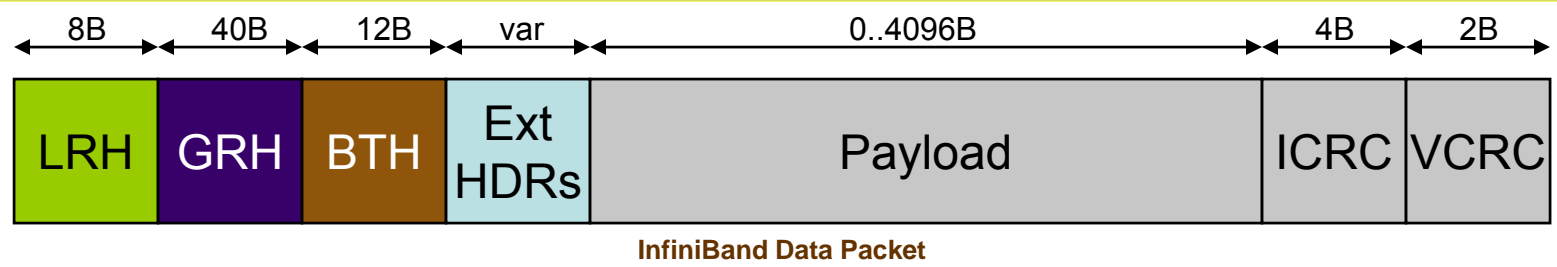


- Example topologies commonly used
- Architecture does not limit topology
- Modular switches are based on fat tree architecture

InfiniBand Protocol Layers



InfiniBand Packet Format



VL	LVer	SL	rsvd	LNH	DLID
rsvd	Len			SLID	

LRH

Opcode	Pad	TVer	Partition Key
rsvd	Destination QP		
A	rsvd	PSN	

BTH

IPVer	TClass	Flow Label	
Payload Len		Next Header	Hop Lim
SGID[127:96]			
SGID[95:64]			
SGID[63:32]			
SGID[31:0]			
DGID[127:96]			
DGID[95:64]			
DGID[63:32]			
DGID[31:0]			

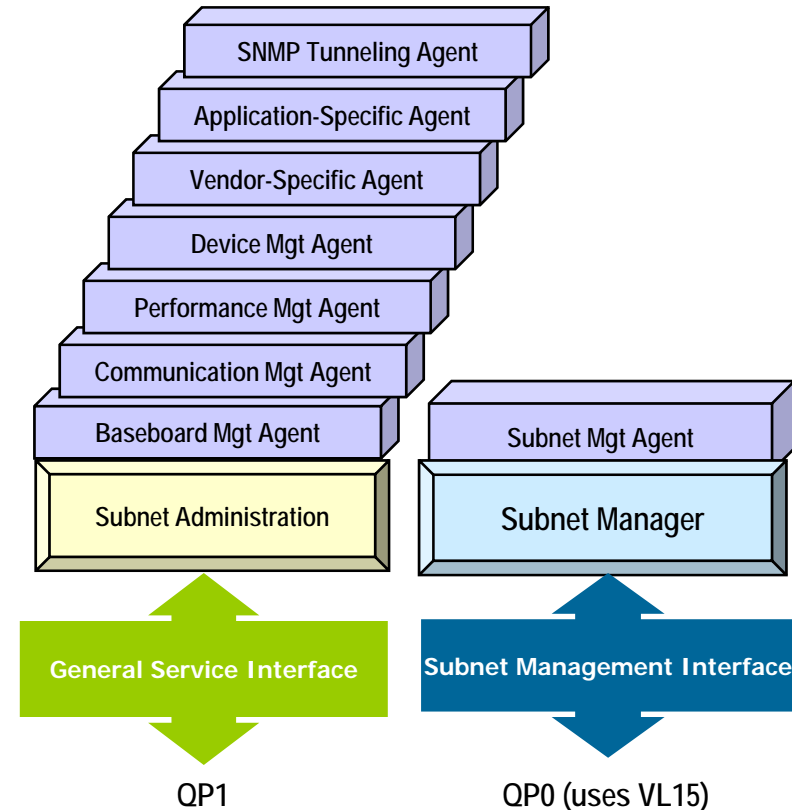
GRH (Optional)

- Extended headers:**
- Reliable Datagram ETH (4B)
 - Datagram ETH (8B)
 - RDMA ETH (16B)
 - Atomic ETH (28B)
 - ACK ETH (4B)
 - Atomic ACK ETH (8B)
 - Immediate Data ETH (4B)
 - Invalidate ETH (4B)

- Queue Pair (QP) – transport endpoint
 - ◆ Asynchronous interface
 - › Send Queue, Receive Queue, Completion Queue
 - ◆ Full transport offload
 - › Segmentation, reassembly, timers, retransmission, etc
 - ◆ Operations supported
 - › Send/Receive – messaging semantics
 - › RDMA Read/Write – enable zero copy operations
 - › Atomics – remote Compare & Swap, Fetch & Add
 - › Memory management - Bind/Fast Register/Invalidate
- Kernel bypass
 - ◆ Enables low latency and CPU offload
 - ◆ Exposure of application buffers to the network
 - › Direct data placement / zero copy
 - ◆ Enabled through QPs, Completion Queues (CQs), Protection Domains (PD), Memory Regions (MRs)
- Polling and Interrupt models supported

Management Model

- **Subnet Manager (SM)**
 - ◆ Configures/Administers fabric topology
 - ◆ Implemented at an end-node or a switch
 - ◆ Active/Passive model when more than one SM is present
 - ◆ Talks with SM Agents in nodes/switches
- **Subnet Administration**
 - ◆ Provides path records
 - ◆ QoS management
- **Communication Management**
 - ◆ Connection establishment processing

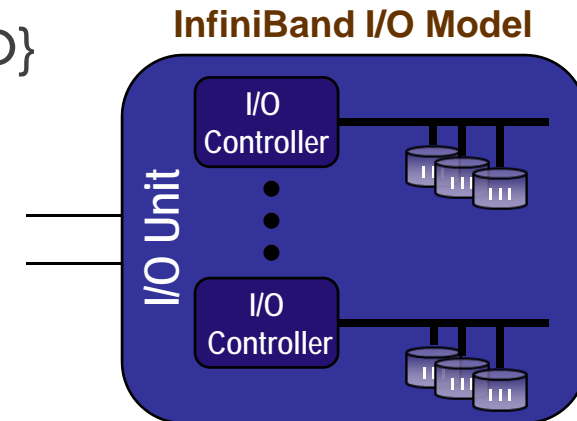


➤ Discovery methods

- ◆ Persistent Information {Node_GUID:IOC_GUID}
- ◆ Subnet Administrator (Identify all ports with CapabilityMask.IsDM)
- ◆ Configuration Manager (CFM)*
 - Locate the Device Administrator through Service Record
- ◆ Boot Manager*
- ◆ Boot Information Service*

➤ Identifiers

- ◆ Per LUN WWN (through INQUIRY VPD)
- ◆ SRP Target Port ID
{IdentifierExt[63:0], IOC GUID[63:0]}
- ◆ Service Name – SRP.TIO.{PortID ASCII}
- ◆ Service ID – Locally assigned by the IOC/IOU



- Leverages all iSCSI infrastructure
 - ◆ Using IP over InfiniBand
- Same iSCSI mechanisms for discovery (RFC 3721)
 - ◆ Static Configuration {IP, port, target name}
 - ◆ Send Targets {IP, port}
 - ◆ SLP
 - ◆ iSNS
- Same target naming (RFC 3721/3980)
 - ◆ iSCSI Qualified Names (iqn.)
 - ◆ IEEE EUI64 (eui.)
 - ◆ TII Network Address Authority (naa.)

Block Storage Data Transfer Summary

	SRP	iSER	iSCSI	FCP
Request	SRP_CMD (SEND)	SCSI-Command (SEND)	SCSI-Command	FCP_CMND
Response	SRP_RSP (SEND)	SCSI-Response (SEND)	SCSI-Response (or piggybacked on Data-In PDU)	FCP_RSP
Data-In Delivery	RDMA Write	RDMA Write	Data-In	FCP_DATA
Data-Out Delivery	RDMA Read RDMA Read Resp.	RDMA Read RDMA Read Resp.	R2T Data-Out	FCP_XFER_RDY FCP_DATA
Unsolicited Data-Out Delivery		Part of SCSI-Command (SEND) Data-Out (SEND)	Part of SCSI- Command Data-Out	FCP_DATA
Task Management	SRP_TSK_MGMT (SEND)	Task Management Function Request/ Response (SEND)	Task Management Function Request/ Response	FCP_CMND

InfiniBand software is developed under
OpenFabrics Open source Alliance

<http://www.openfabrics.org/index.html>



InfiniBand standard is developed by the
InfiniBand® Trade Association

<http://www.infinibandta.org/home>



Reference

- InfiniBand Architecture Specification Volumes 1-2 Release 1.2.1
 - ◆ www.infinibandta.org
- IP over InfiniBand
 - ◆ RFCs 4391, 4392, 4390, 4755 (www.ietf.org)
- NFS Direct Data Placement
 - ◆ <http://www.ietf.org/html.charters/nfsv4-charter.html>
- iSCSI Extensions for RDMA (iSER) Specification
 - ◆ <http://www.ietf.org/html.charters/ips-charter.html>
- SCSI RDMA Protocol (SRP), DIF
 - ◆ www.tl0.org

Glossary

- ◆ APM - Automatic Path Migration
- ◆ BECN - Backward Explicit Congestion Notification
- ◆ BTH - Base Transport Header
- ◆ CFM - Configuration Manager
- ◆ CQ - Completion Queue
- ◆ CQE - Completion Queue Element
- ◆ CRC - Cyclic Redundancy Check
- ◆ DDR - Double Data Rate
- ◆ DIF - Data Integrity Field
- ◆ FC - Fibre Channel
- ◆ FECN - Forward Explicit Congestion Notification
- ◆ GbE - Gigabit Ethernet
- ◆ GID - Global IDentifier
- ◆ GRH - Global Routing Header
- ◆ GUID - Globally Unique IDentifier
- ◆ HCA - Host Channel Adapter
- ◆ IB - InfiniBand
- ◆ IBTA - InfiniBand Trade Association
- ◆ ICRC - Invariant CRC
- ◆ IPoIB - Internet Protocol Over InfiniBand
- ◆ IPv6 - Internet Protocol Version 6
- ◆ iSER - iSCSI Extensions for RDMA
- ◆ LID - Local IDentifier
- ◆ LMC - Link Mask Control
- ◆ LRH - Local Routing Header
- ◆ LUN - Logical Unit Number
- ◆ MPI - Message Passing Interface
- ◆ MR - Memory Region
- ◆ NFSoRDMA - NFS over RDMA
- ◆ OSD - Object based Storage Device
- ◆ OS - Operating System
- ◆ PCIe - PCI Express
- ◆ PD - Protection Domain
- ◆ QDR - Quadruple Data Rate
- ◆ QoS - Quality of Service
- ◆ QP - Queue Pair
- ◆ RDMA - Remote DMA
- ◆ RDS - Reliable Datagram Socket
- ◆ RPC - Remote Procedure Call
- ◆ SAN - Storage Area Network
- ◆ SDP - Sockets Direct Protocol
- ◆ SDR - Single Data Rate
- ◆ SL - Service Level
- ◆ SM - Subnet Manager
- ◆ SRP - SCSI RDMA Protocol
- ◆ TCA - Target Channel Adapter
- ◆ ULP - Upper Layer Protocol
- ◆ VCRC - Variant CRC
- ◆ VL - Virtual Lane
- ◆ WQE - Work Queue Element
- ◆ WRR - Weighted Round Robin