Infiniband Overview

What is it and how we use it

What is Infiniband

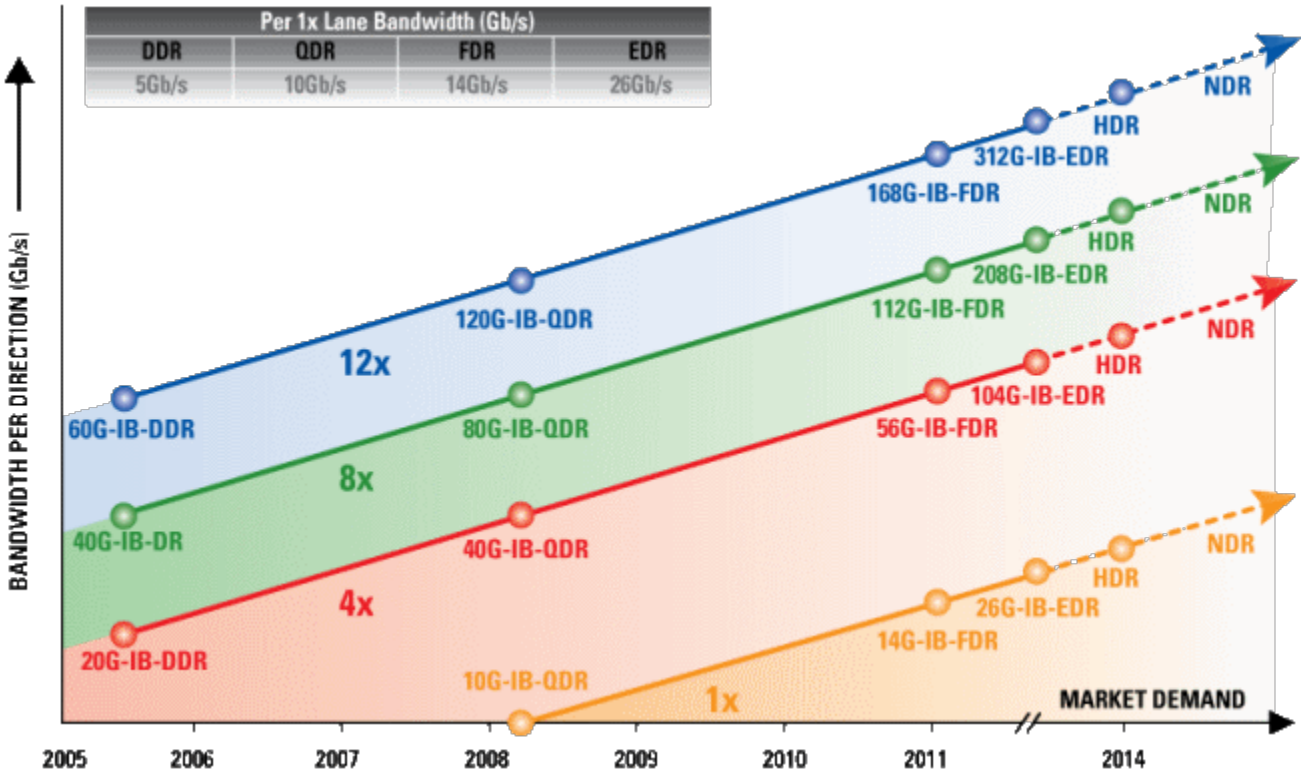
- Infiniband is a contraction of "Infinite Bandwidth"
  - can keep bundling links so there is no theoretical limit
  - Target design goal is to always be faster than the PCI bus.
- Infiniband should not be the bottleneck.
- Credit based flow control
  - data is never sent if receiver can not guarantee sufficient buffering

What is Infiniband

- Infiniband is a switched fabric network
  - low latency
  - high throughput
  - failover
- Superset of VIA (Virtual Interface Architecture)
  - Infiniband
  - RoCE (RDMA over Converged Ethernet)
  - iWarp (Internet Wide Area RDMA Protocol)

What is Infiniband

- Serial traffic is split into incoming and outgoing relative to any port
- Currently 5 data rates
  - Single Data Rate (SDR), 2.5Gbps
  - Double Data Rate (DDR), 5 Gbps
  - Quadruple Data Rate (QDR), 10 Gbps
  - Fourteen Data Rate (FDR), 14.0625 Gbps
  - Enhanced Data Rate (EDR) 25.78125 Gbps
- Links can be bonded together, 1x, 4x, 8x and 12x



HDR - High Data Rate
NDR - Next Data Rate

Infiniband Road Map (Infiniband Trade Association)

What is Infiniband

- SDR, DDR, and QDR use 8B/10B encoding
  - 10 bits carry 8 bits of data
  - data rate is 80% of signal rate
- FDR and EDR use 64B/66B encoding
  - 66 bits carry 64 bits of data

Signal Rate	**Latency**
SDR	200ns
DDR	140ns
QDR	100ns

Hardware

2 Hardware vendors

- Mellanox
  - bought Voltaire
- Intel
  - bought Qlogic Infiniband business unit

Need to standardize hardware. Mellanox and Qlogic cards work in different ways.

Hardware

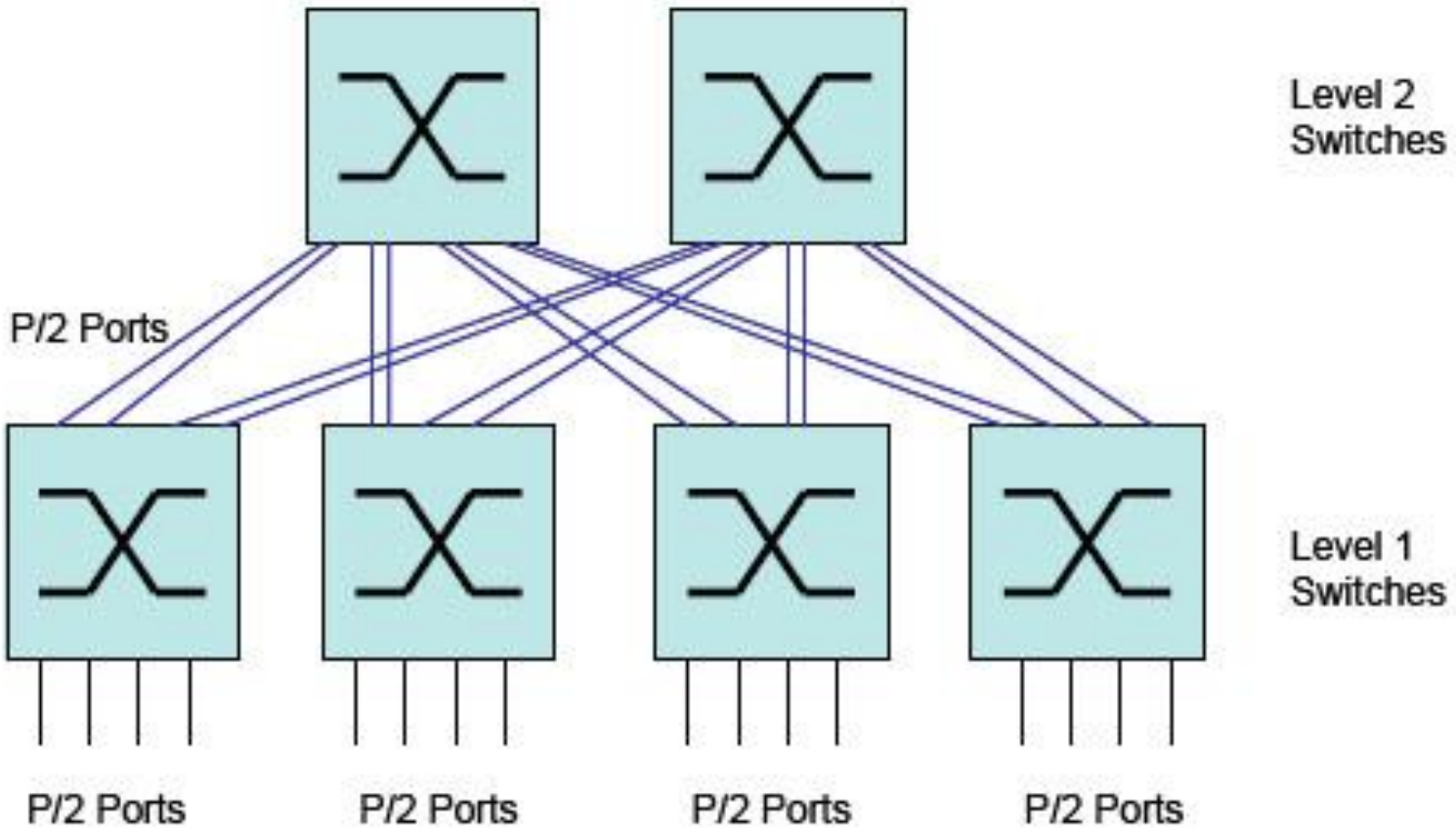
Connections

- SDR and DDR: copper CX4
- QDR and FDR: copper or fiber
  - QSFP: Quad Small Form Factor Pluggable
  - also called mini-GBIC (Gigabit Interface Converter)

Topologies

Common network topologies

- Fat tree
- Mesh
- 3D torus
- CBB (Constant Bi-sectional Bandwidth)
  - type of Fat tree
  - can be oversubscribed 2:1 to 8:1
  - oversubscription can reduce bandwidth but most applications do not fully utilize it anyway



Two level CBB; Source: Mellanox

Software

- No standard API, only a list of "verbs"
- The de-facto standard is the syntax developed by the Open Fabrics Alliance.
- Obtained via the Open Fabrics Enterprise Distribution (OFED)
  - in RHEL-5 and above
- Need to build MPI software to support queueing systems

Software

OFED stack includes

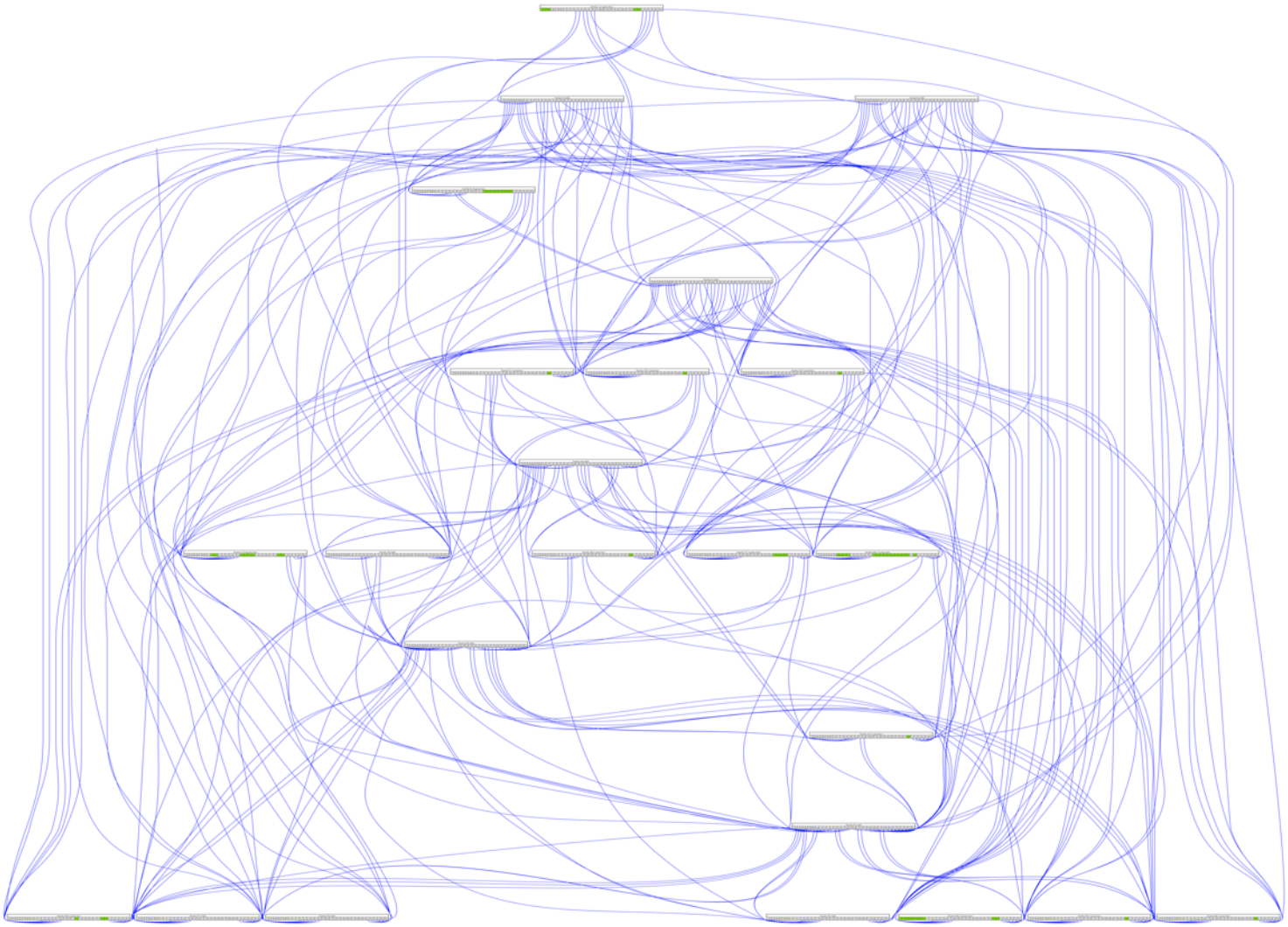
- device drivers
- performance utilities
- diagnostic utilities
- protocols (IPoIB, SDP, SRP,...)
- MPI implementations (OpenMPI, MVAPICH)
- libraries
- subnet manager

How we use Infiniband

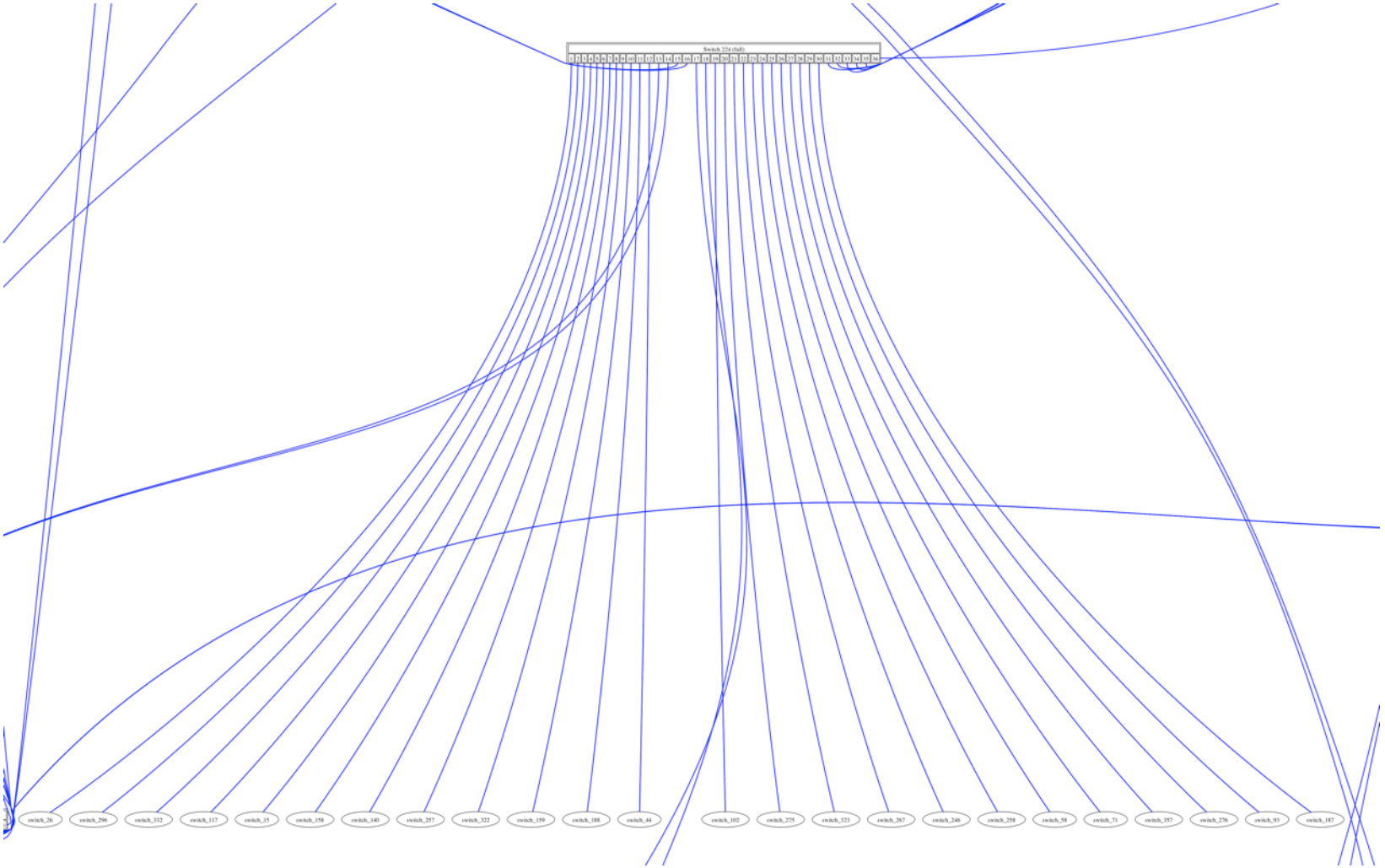
Helium cluster uses a 2:1 blocking CBB fabric consisting of (24) 36-port Voltaire switches

- 18 leaf switches
- 6 spine (root) switches

For every leaf switch, 24 ports are used to connect to server nodes and 12 ports are used to connect to the spine switches.



Switches



Leaf switch

How we use Infiniband

Servers are connected with Mellanox ConnectX QDR cards.

We use RDMA and IPoIB for communication

- RDMA for MPI and other communication for user jobs
  - standardized on OpenMPI for MPI
- IPoIB for the scratch storage file system
  - About 10Gbps bandwidth
- Mix of copper and fiber cables

How we use Infiniband

Our switches can run a subnet manager but limited in routing engine choices.

We run the OFED subnet manager on a few non-compute nodes.

We use the fat-tree routing engine, which is not available on the switch subnet manager.

Subnet manager

The main purpose of the subnet manager is to establish routes between ports.

- Routing is typically static. The subnet manager tries to balance the routes on the switches.
- A sweep is done every 10 seconds to look for new ports or ports that are no longer present. Established routes will typically remain in effect if possible.

Subnet manager

While there are several tunables the most important is the routing engine. We use fat-tree which handles the CBB topology. Other routing engines can be more general, such as min hop.

Newer engines can use virtual lanes to provide deadlock free routing

Hot spots

Since routes are static, hot spots can develop.

- The typical way to handle this is by dropping all packets at a port and starting over. The idea is that a restart will have a different distribution of packets.
- Dynamic fabric managers can take other approaches to try to redirect traffic around hot spots
  - very expensive

Metrics

It is fairly easy to diagnose the fabric at high level. There are several tools for this.

- ibnetdiscover
- ibdiagnet
- rdma_bw, rdma_lat
- ib_read_bw, ib_read_lat

There are also tools for port level measurements.

Metrics

Each IB card has a set of port counters that can be read to get performance metrics

- The problem is that the counters only go so high and then roll over.
- During periods of high (any) traffic they can roll over quickly

This makes it difficult to get accurate measurements. Efforts are being made to address this shortcoming.

Metrics

In addition to performance counters there are also error counters.

- these also have a limit but unlike performance counters do not roll over
- they (hopefully) increase at a much lower rate than performance counters

Metrics

- When gathering metrics it is important to reset the counters at the start of measuring.
- It is fairly easy to get performance of a port but difficult to get performance of the entire fabric.
- When looking at error counts, one is generally looking at the port level

Troubleshooting

Infiniband works well most of the time but ...

- cable problems are very common
- sometimes the first sign of trouble is when a job fails
- similar but different utilities can be used
  - ibcheckerrors
  - ibqueryerrors
- ibqueryerrors is newer and faster and can also filter common errors

ibqueryerrors

1. reset error counters

   ```
   ibclearerrors
   ```

2. run ibqueryerrors, suppressing XmtWait errors, which are not really errors

   ```
   /usr/sbin/ibqueryerrors -s XmtWait
   ```

3. May have to wait for errors to start showing up but serious problems will show up quickly

Errors

Not always clear cut

- Symbol errors
  - can be rather common but maybe not necessarily a major problem
  - According to the InfiniBand specification, 10E-12 BER, the maximum allowable symbol error rate is 120 errors per hour.
- LinkRecovery
  - this is a serious error but not always logged

Errors

- PortRcvRemotePhysicalErrors
  - These can be quite common when there is a bad link somewhere
  - We saw a lot of these at one time but they are typically secondary errors and difficult to track down.
  - will go away once the source error is found.
- PortXmitDiscards
  - may indicate congestion
- VL15Dropped
  - indicates a buffer problem but not usually serious

Future

Things to improve:

- performance monitoring
- subnet manager
- topology aware job scheduler
- more advanced fabric management

Links

- http://www.openfabrics.org
- http://www.mellanox.com
- http://www.intel.com/content/www/us/en/infiniband/truescale-infiniband.html

Contact:

glenn-johnson@uiowa.edu