AI/ML Backend Fabric Dynamic/Adaptive Routing Ruminations
_or_
No Man's Land (AC⚡DC 2020)

Why Traditional IP Routing Will Struggle to Do a "Really Good Job"™ in Highly Scaled-Up AI/ML Backend Fabrics

IETF 117 AC⚡DC SideMeeting

Tony P, Juniper

Unusual Disclosure

- This talk is a somewhat pretty high level cogitation on the contumacity of the world that keeps engineers employed forever
  - Another iteration of the ever more complex, contradicting sets of requirements driving IP routing and technology in general
  - Do NOT expect an "easy solution" being presented here
  - IP "ate" lots of other technologies and problems, this one does not seem to be an easy lunch

Non-Juniper

At Least Topologies Are Regular

- Dragonfly/Hypercube/… Relevant in HPC Still and in Limited Scale (Sub-)-IP-Fabric
  - Lack of Feasible "Information Summarization Vector" makes that stuff very hard to scale up economically from dynamic routing perspective
- Local Large Scale E'thing is IP Clos Pretty Much Now
  - Flashback to the "DC Routing BoF" @ IETF 100
  - Predictable outcome given it's hard to argue with linear programming solutions based on economic cost function
  - Little Sugarplum: CLOS turns out with correct routing to prevent credit loop deadlocks ;-)

Non-Juniper

But the _Real_ AI Protocol turned out to not be IP ;-)

- Infiniband in some form or fashion
  - IP Suite of protocols not particularly adept at shared distributed memory semantics
  - IP's lack of cellified, scheduled Layer-2
    - DetNet solves a mildly different problem (and not at large scale IMO)
  - IP's lack of concept of prioritized, scheduled queue under a socket
  - Sockets are at least 4x slower than IB Verb APIs
  - IP forwarding is slower than IB
    - MPLS can do on good chips sub-micro but that's not really "simple IP"

Non-Juniper

Everything Else is Scary

- Move to RDMA from Map-Reduce which drives lots of devilishly hard requirements (the wisdom of 20 years that distributed shared memory is not workable went to the scrapheap)
  - Low delay, sorry, _ultra_ low delays (we're talking microseconds e2e or fraction microsecond for switching hops)
  - Losless (since RDMA does not like complex windowing protocols)
  - Order-Preserving (to prevent complex window reassemblies again)
  - Congestion that can separate flows and push back selectively all the way to source in microseconds
  - Deep buffers can help but create all kind of different problems in turn and they cost $
  - Tons and tons of BW necessary (shortest path is very long-toothed here and good entropy will become very important)
- Possibly Multi-Topology necessary
  - separate different classes in RoCE
  - Deal with multiple involved L2 layers
- Changing sets of leaves (computes) that need a bi-dir S+I-PMSI like BIER
- Scale and even more scale
  - Insatiable demand for larger and larger clusters that has barely started
  - (Multi-Homed) HCAs multiply scale by 20-40x (remember RotH)
    - Even if HCA is not RoCE multi-homed it may be multi-homed into multiple l2 layers

Non-Juniper

Subnet Manager (Day-0 ZTP)

- Infiniband is relying on centralized architecture to do Day-0 ZTP of LIDs
  - Centralized is always nice on small
  - At massive scale centralized becomes limited especially if it does and distributes path computation results as IB architecture seems to mandate
- RoCE does not standardize any ZTP/management plane for IP

Non-Juniper

Should we all go to work for IB Forum now and do routing there ? ;-)

- Well, no, Ethernet (and IP) always won so far
  - Ethernet is equivalent of concrete, basically price of dirt, easy to pour and outsells everything even if it's necesserily not that pretty
  - IP is not easy to pour but is also price of dirt and sheer production volumes make it the 4-strokes engine of the communication world
  - IP and Ethernet are joined at the hip by now ;-)
- So, what kind of IP routing/technologies do we need for RoCE ?

Non-Juniper

Very Brave New World to Be Invented

- True Day-0 ZTP for IP Clos Fabris
- Scale all the way down to HCAs
  - HCAs are better not loaded with full topology/information
- Non-shortest path loop-free routes
  - Allows best utilization and provides rough load balancing based on load metrics
- Scale information aggregation along some predictable gradient
  - Must encompass multi-topology, DSCP and rough load on link/node per DSCP equivalent
- As reactive as possible (i.e. distributed rather than diffused computation)
- Pull IB entropy into multi path forwarding somehow
- Very fast congestion control or inband telemetry reaction
  - 802.1Qau or HPCC++ or something ;-)
- Radical Stuff like BIER based PMSI where AI/ML folding is happening on the fly in network nodes
  - Network becoming the other "AI co-processor"

Non-Juniper