Low Latency COTS GigE

Lessons Learned : IP limitations are difficult to code around. It is easier to build a cluster-specific solution.
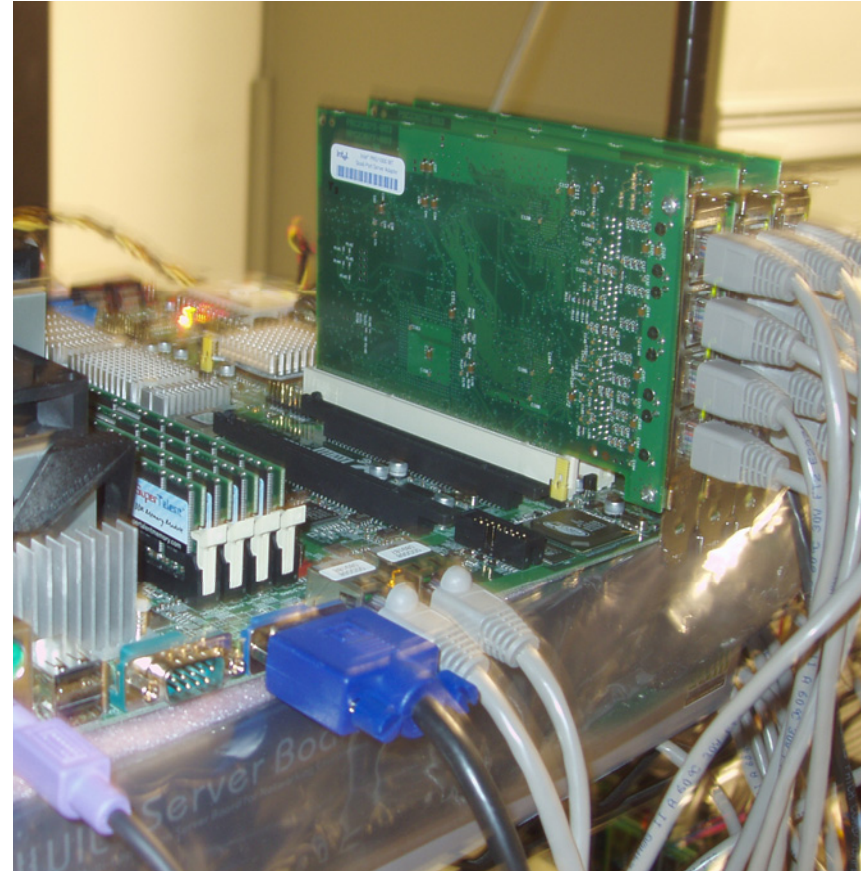
Limits; NIC (node) count, NIC Bandwidth, Bus Bandwidth

Low Latency COTS GigE

Scaling an 8 node cluster : You get incremental performance gains with NIC count after 8 NICs and reduced performance after 10. So we kept the node count at 8 and replaced our single-core CPUs with dual-cores. Additionally we went to a diskless system with a 200MB RAM disk on each client node.

Low Latency COTS GigE

GAMMA beats IP for specialized apps in Bandwidth, Latency and CPU usage.

For an 8-node cluster (7 interconnected GAMMA NICs) we observed 5.8 µsec latencies and 5% CPU overhead.

Future Plans : The Cluster will be reorganized to take advantage of the minimal amount of IP traffic we have. We can devote one NIC to it in production (primarily for NFS file transfers). Additionally, we can now scale the cluster up in traditional fashion by adding switches (as suggested by the FNN approach) to introduce single switch latencies that will no more than double our internode latencies while affording us to scale up to 64 nodes (256 PEs).