LQCD Clusters
at JLab

Chip Watson

Jie Chen, Robert Edwards

Ying Chen, Walt Akers

Jefferson Lab

Jefferson Lab SciDAC Prototype Clusters

The SciDAC project is funding a sequence of cluster prototypes which allow us to track industry developments and trends, while also deploying critical compute resources.

Myrinet + Pentium 4

128 single 2.0 GHz P4 Xeon (Summer 2002)

64 Gbytes memory

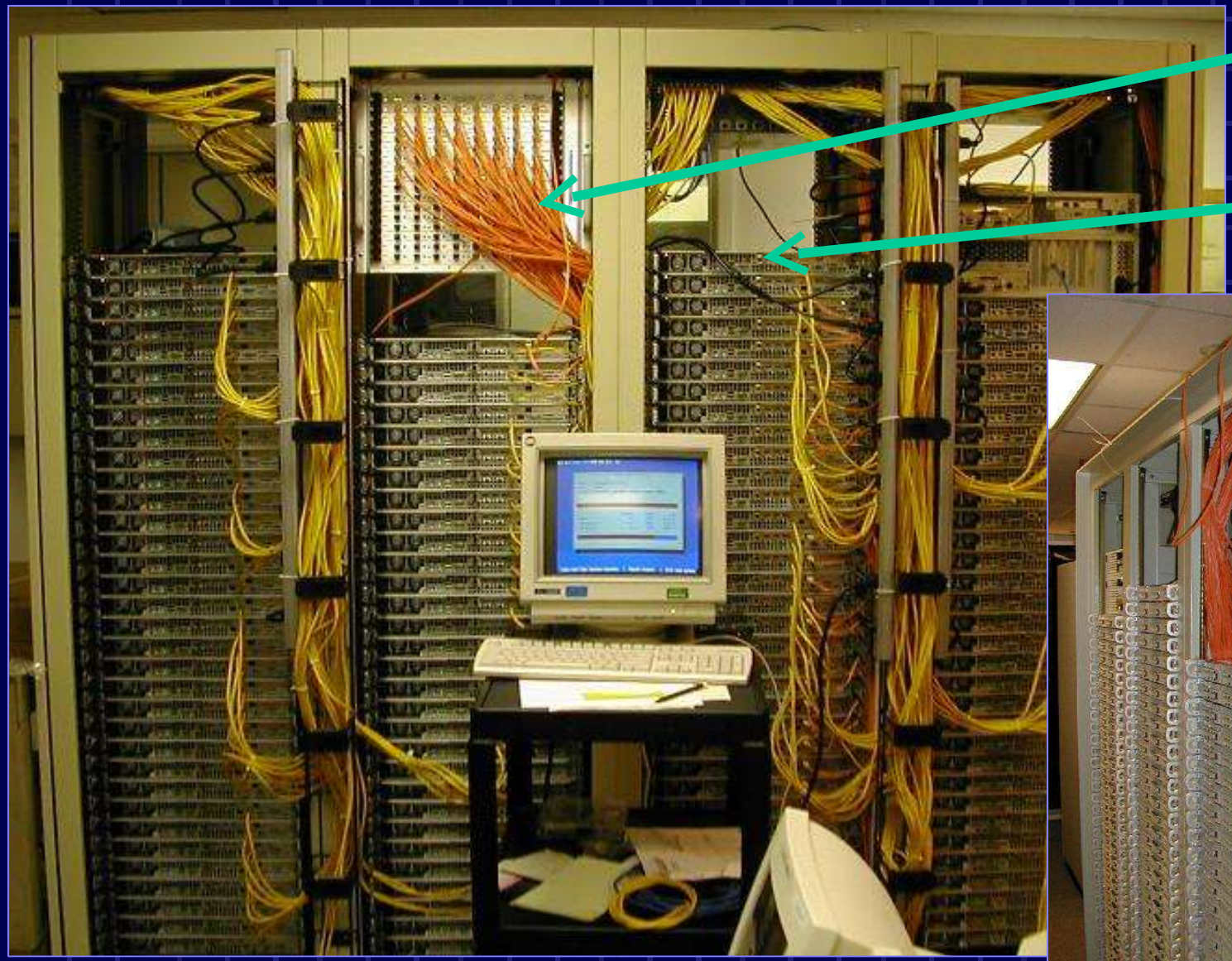
Gigabit Ethernet Mesh + Pentium 4

(An alternative cost effective cluster design now being evaluated)

256 (8x8x4) single 2.66 GHz P4 Xeon (Fall 2003)

64 Gbytes memory

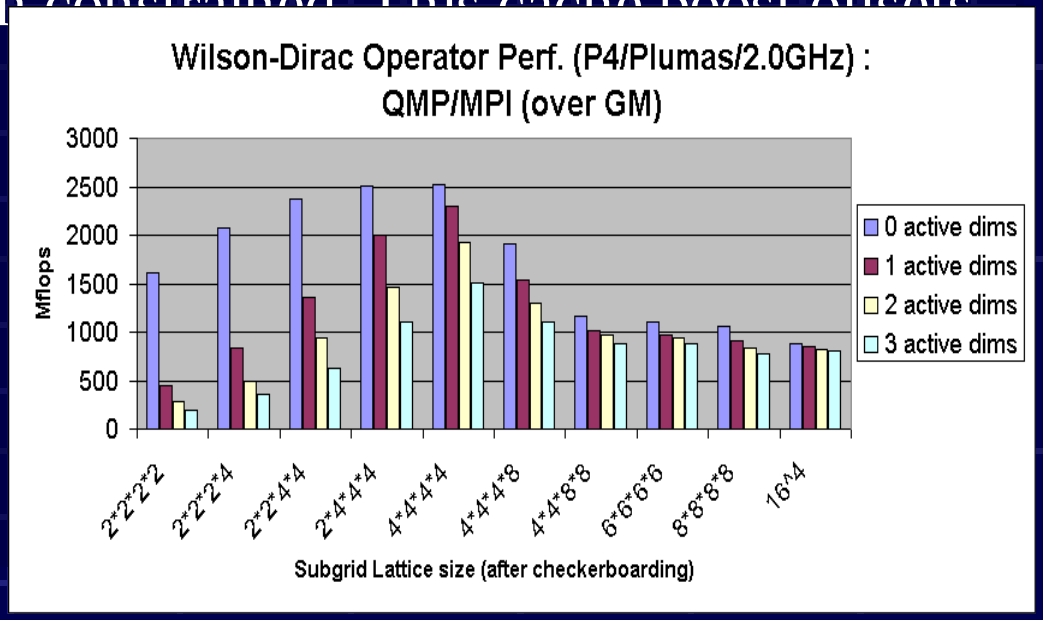
128 Node Myrinet Cluster @ JLab



2002 Myrinet Cluster Performance

Each Myrinet cluster node delivers ~600 for the DWF inverter for large problems (16^4), and sustains this performance down to 2×4^3, yielding ~75 Gflops for the cluster (rising as more SSE code is integrated into SZIN and Chroma; 150 Gflops on Wilson Dirac)

Small problem (cache resident) performance on a single processor is 3x that of memory bandwidth constrained. This cache boost offsets network overhead.

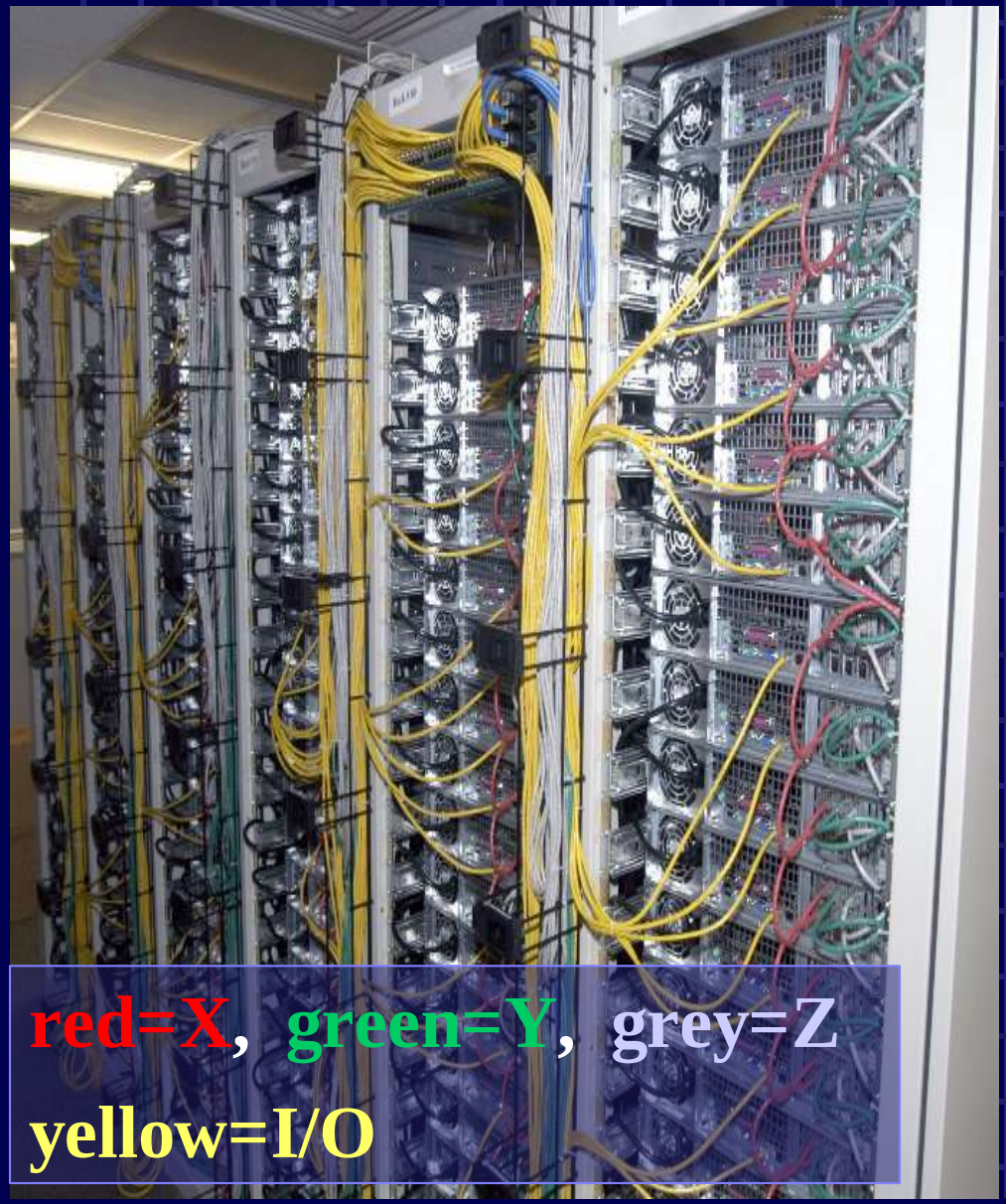


Memory bandwidth is not enough for 2nd processor.

Network characteristics:

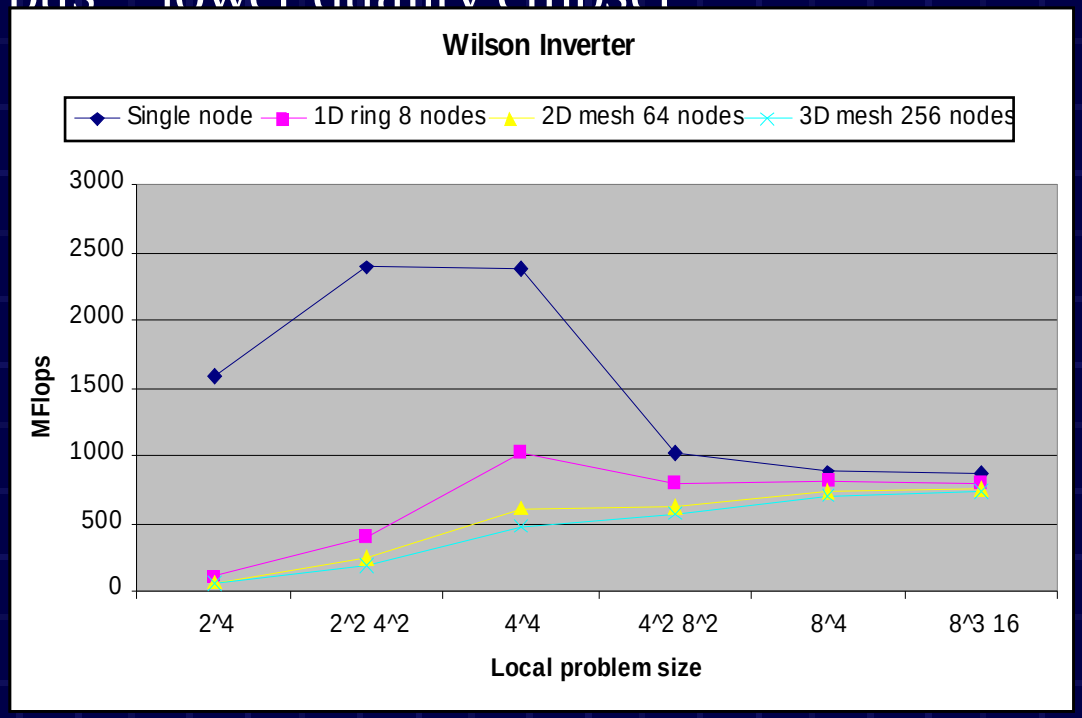
- 130 + 130 MB/s bandwidth

256 Node GigE Mesh Cluster @ JLab



2003 GigE Cluster Performance

The gigE cluster nodes deliver ~700 Mflops/node for the inverter on large problems (16^4), but degrades for small problems due to the not-yet-optimized gigE software. Nodes are barely faster than earlier cluster despite 33% faster bus – lower quality chipset implementation.

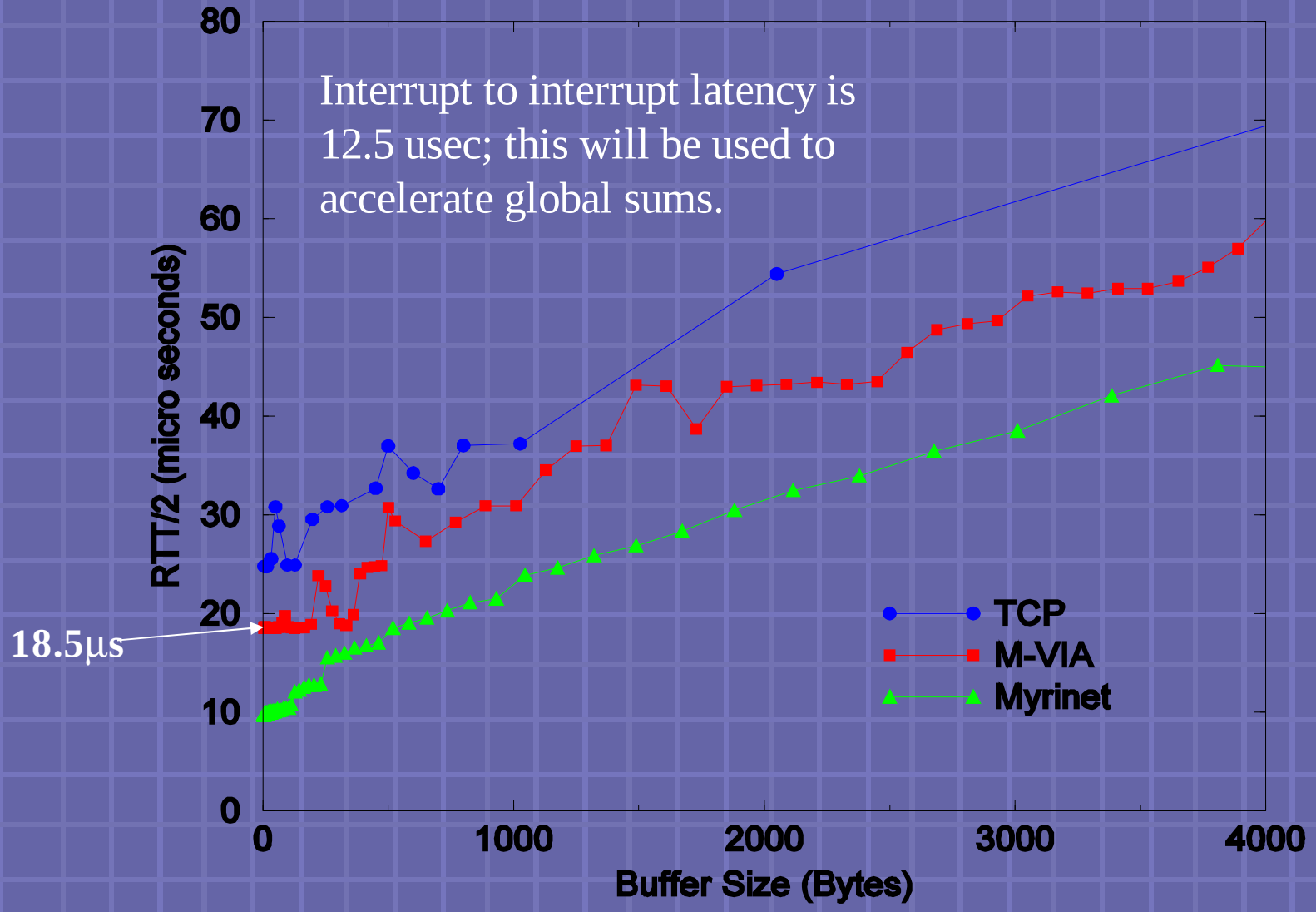


Network characteristics:

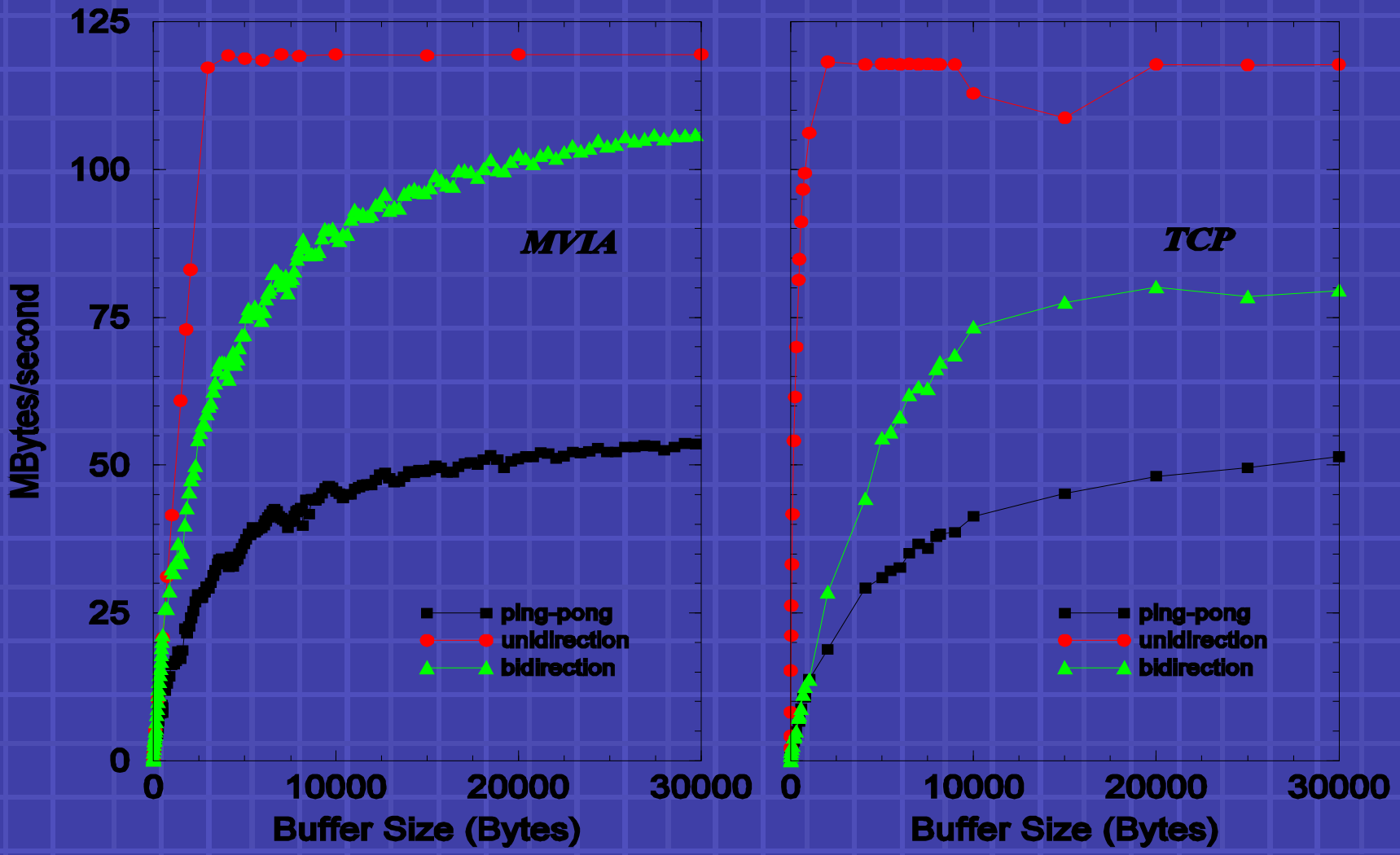
- 120 + 120 MB/s b/w 1 link
  220 + 220 MB/s b/w 3 link
- 19 usec RTT/2 latency
  (12 usec effective latency at interrupt level, planned for fast global sum)
- 13 usec software overhead

Wilson Inverter 12% faster, proportional to Streams

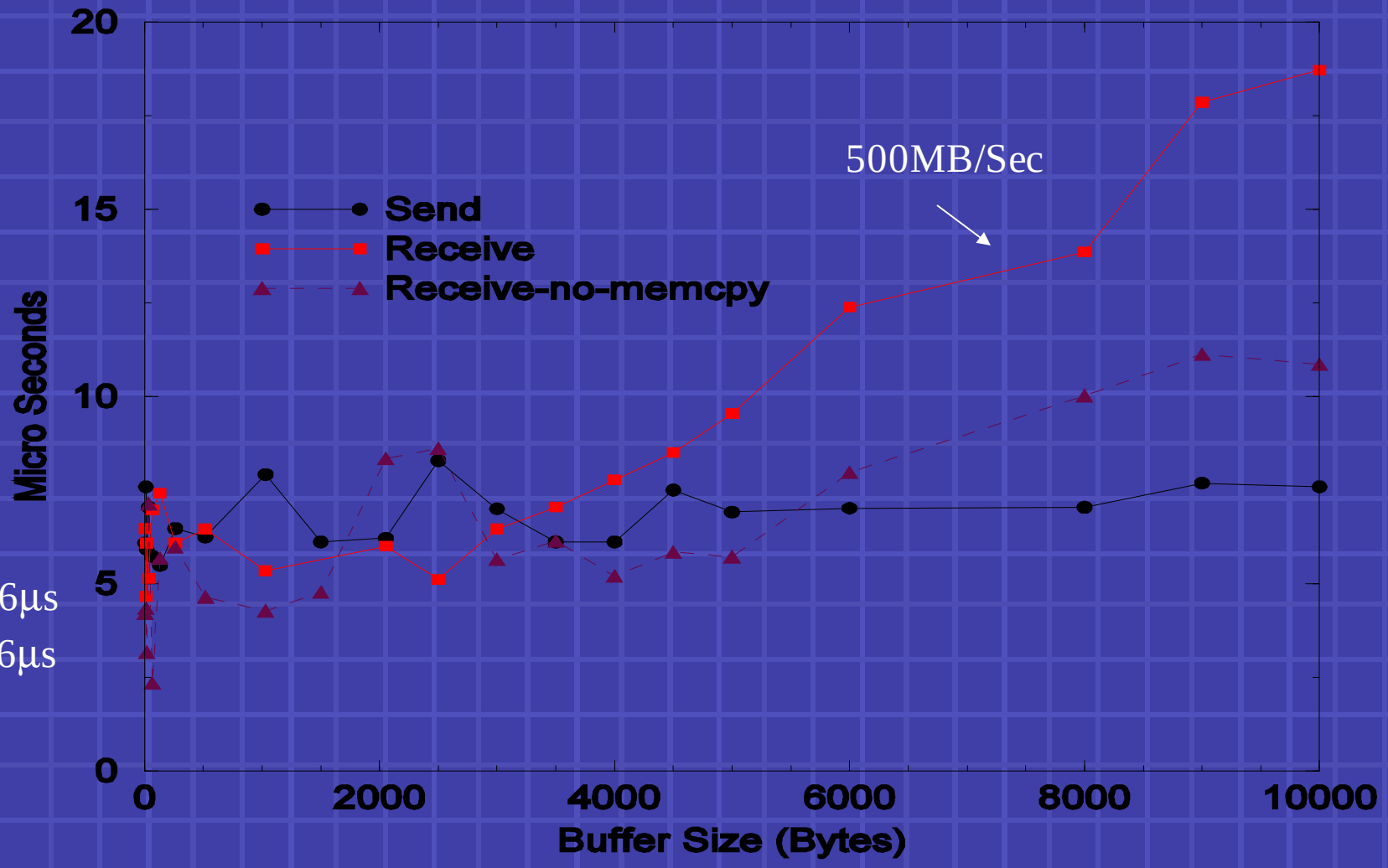
GigE Application to Application Latency



GigE point to point bandwidth



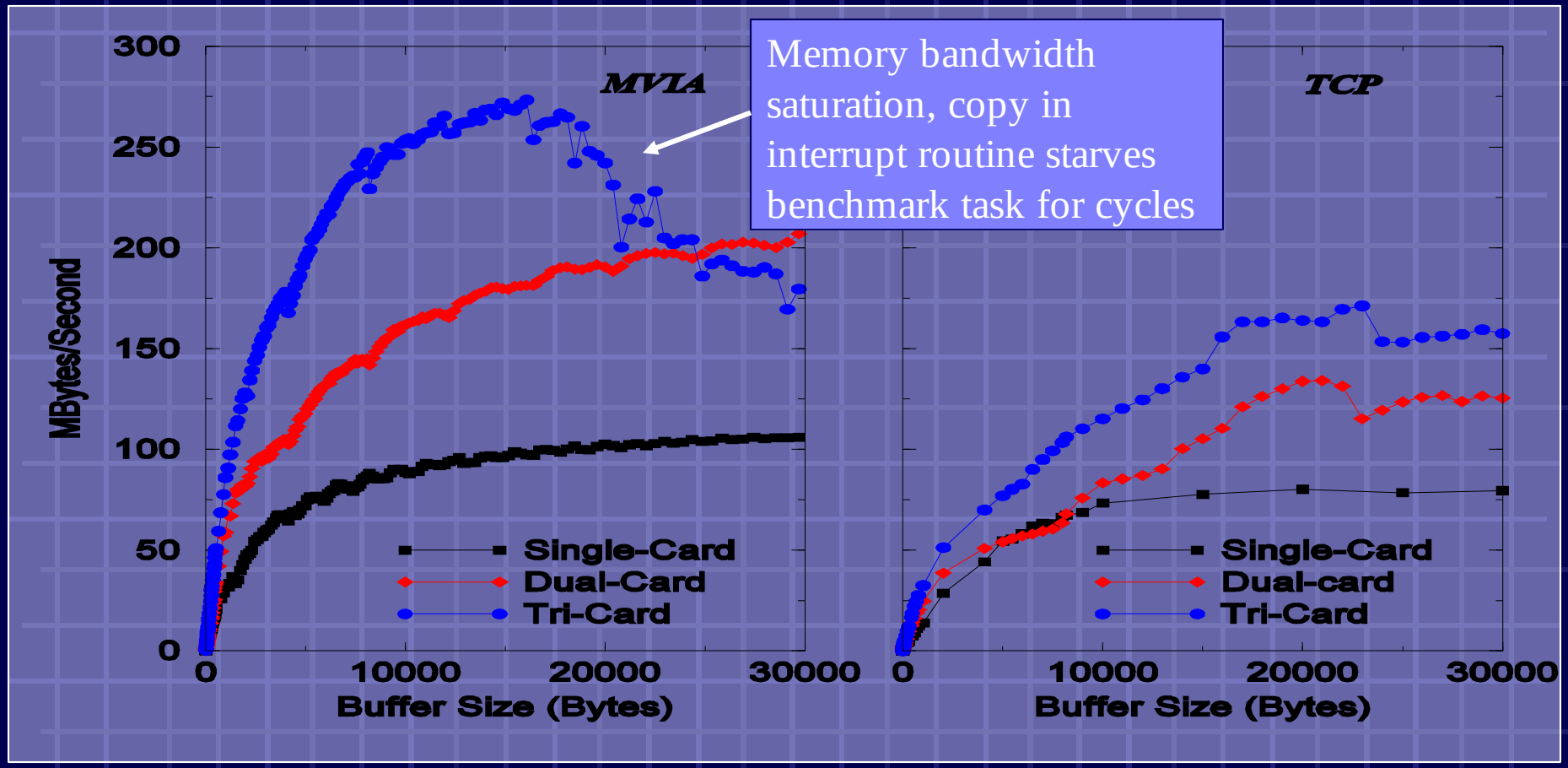
Host Overheads of M-VIA



Os~6μs

Or~6μs

GigE Aggregated Bandwidths



	Bi-directional Bandwidth	Latency	Cost
Myrinet/GM	260 MB/s (*480MB/s*)	9 µs (7)	400 + 1000
GigaE/VIA	440 MB/s	18.5 µs (12.5)	75 x 6

GigE Evaluation & Status

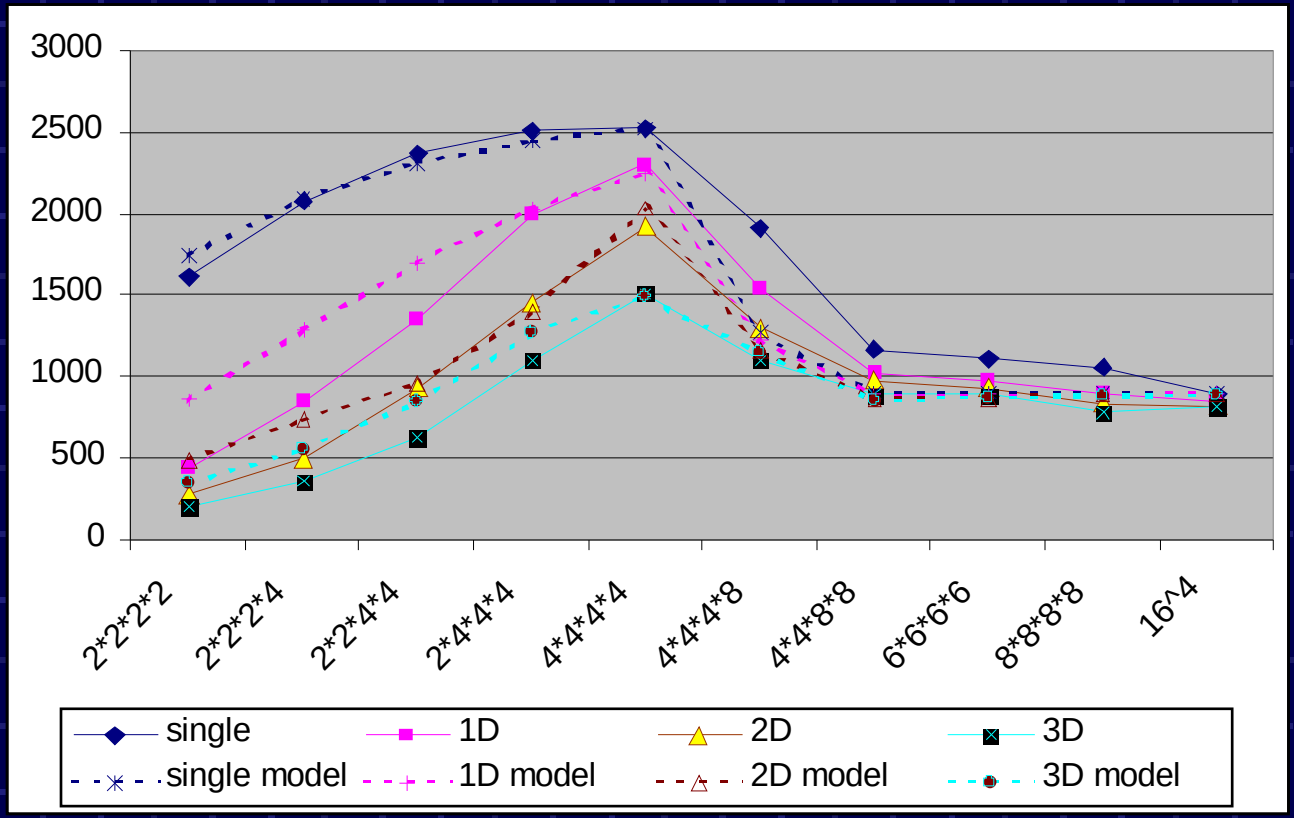
- GigE mesh reduced the system cost by 30%, allowing JLab to build a 256 node cluster instead of a 128 node cluster (largest 2^N partition) within the SciDAC + NP matching budget.
- Software development was hard!
  - The initial VIA low level code development was quick, but QMP was more lengthy (message segmentation, multi-link management)
  - Strange VIA driver bug only recently found (only occurs when running a process consuming all of physical memory)
  - One known bug in VIA finalize at job end sometimes hangs a node
  - System is perhaps only now becoming stable
- Hardware seems fairly reliable

  (assuming all or most hangs are due to the pesky VIA bug, now deceased)
  - However, IPMI is faulty, had to disable SMbus on gigE chips
  - Handful of early card & cable failures, since then modest, about 1 disk

Modeling Cluster Performance

Reminder:

We can effectively model the cluster performance using node and network characteristics.



Model curves include CPU in- and out-of-cache performance, PCI and link bandwidth, latency, etc.

Here a moderately simple model predicts cluster Wilson operator performance pretty well. Also can do inverter.

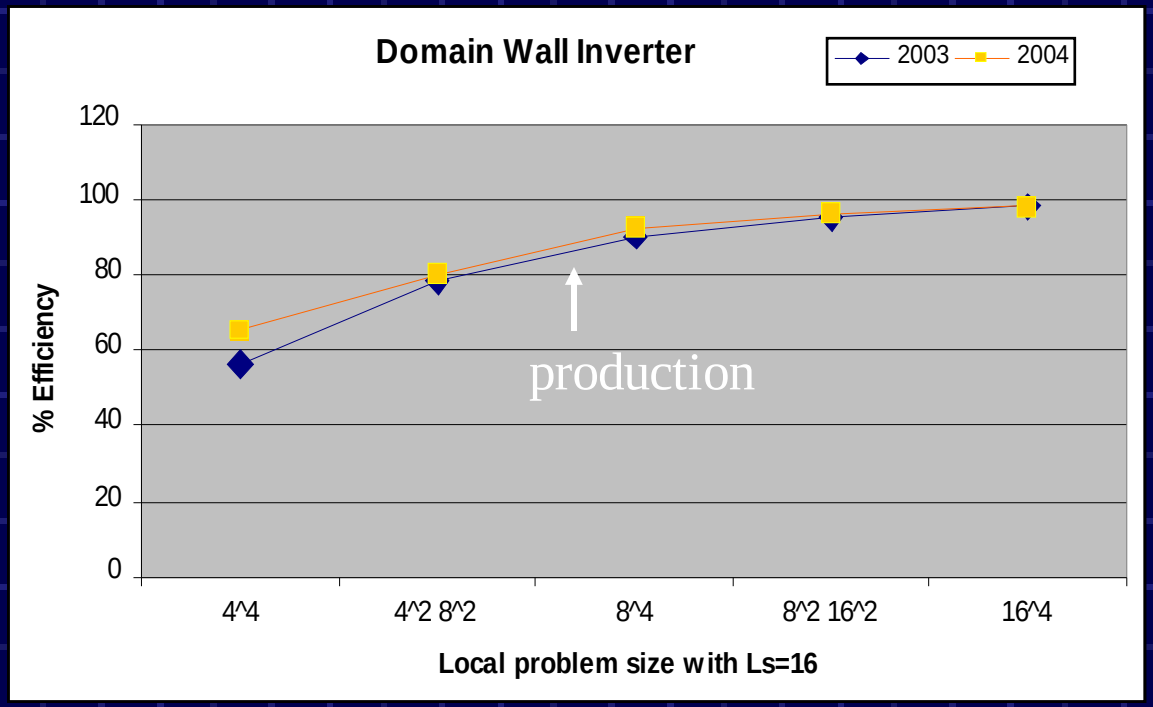
GigE Mesh Cluster Efficiency

2003:

- 533 MHz FSB
- 900 Mflops out of cache
- 256 nodes
- Production running is near 7^4 x 16, or ~85%

2004:

- 800 MHz FSB
- 1500+ Mflops out of cache
- 512 nodes
- Vectorize in 5th dimension to lower # messages



Even though the hypothetical 2004 cluster is bigger and has 66% faster nodes, the efficiency is improved because

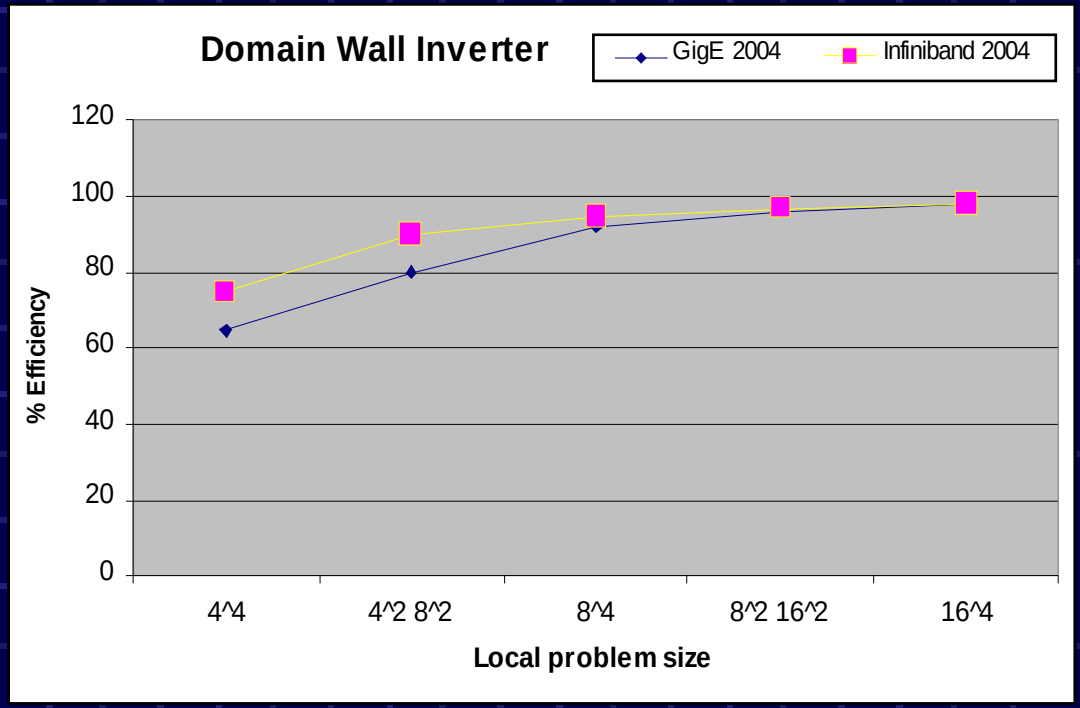
a) a faster memory bus – the copies to run the protocol become cheaper (modest effect)

b) fewer messages (better algorithm) – big effect! (dilutes software overhead)

256 Node Infiniband Cluster Efficiency

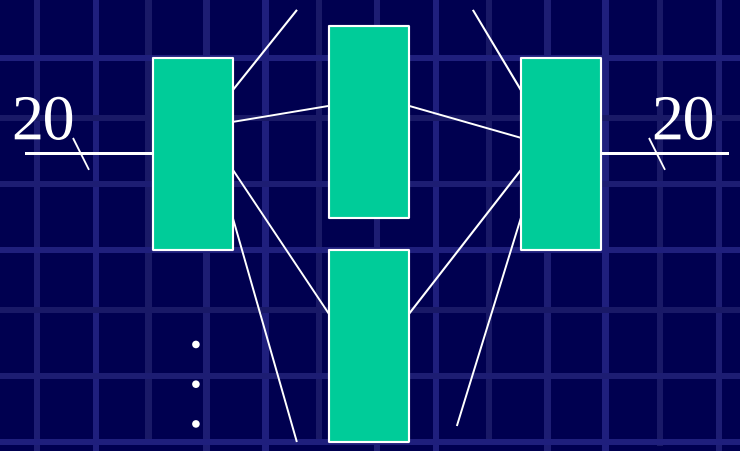
GigE vs Infiniband

- at 16^4, no difference in efficiency
- At 4^4, Infiniband does 15% better, at 30% higher cost
- Only need bandwidth of 100 +100 MB/sec for well structured code (good I/O, compute overlap)



Infiniband cost can be trimmed by not building a full fat tree; use 20:4 instead of 16:8 on edge switches

(20:4 has been assumed in cost model)



What about “SuperClusters”?

A supercluster is one big enough to run LQCD in cache. For domain wall, this would be 2x4x4x4x16 (about 1 MB single precision).

At this small volume, cluster efficiency drops to around 30% (gigE) or 40% (Infiniband), but CPU performance goes up by 3x, yielding performance BETTER than for $l6^5$ for the Infiniband cluster.

Performance for clusters thus has 2 maxima, one at large problems, one at the cache “sweet spot”.

Summer 2004 JLab Cluster

- 512 node 8x8x8 GigE Mesh
  - 500-750 Gflops @ $1.60-$1.25 / Mflops (est, problem size dependent)
  - 2U rackmount
  - On node failure, segment to plane pairs

- 384 node Infiniband (plus a few spares)
  - 420-560 Gflops @ $2.10-$1.60 / Mflops (est)
  - Pedestal instead of rackmount to save cost
  - Can run as 384 nodes if problem has a factor of 3
  - Fault tolerance since spare nodes are in the same fabric
  - More flexibility in scheduling small jobs

If these estimates are born out by firm quotes, and if operational experience on existing gigE mesh is good, then GigE is the favored solution; expectation is that FNAL will deliver the

JLab Computing Environment

1. Red Hat 9 compute and interactive nodes
2. 5 TB disk pool

   Auto migrate to silo (JSRM daemon)

   < 100 GB/day is OK; more if purchase dedicated drive

   Pin / unpin, permanent / volatile
3. PBS batch system

   Separate servers per cluster
4. Two interactive nodes per cluster
5. Some unfriendly features (to be fixed)

   2 hop ssh from offsite, scp (must go through JLab gateway)

For More Information

- Lattice QCD Web Server / Home Page:

  http://www.lqcd.org/

- The Lattice Portal at JLab

  http://lqcd.jlab.org/

- High Performance Computing at JLab

  http://www.jlab.org/hpc/