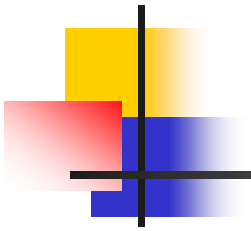
Fermilab Status

Don Holmgren

USQCD All-Hands Meeting

Fermilab

March 22-23, 2007

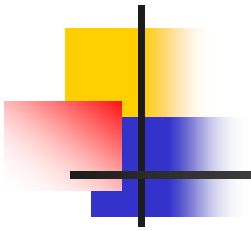


Outline

- Fermilab Status
  - Hardware
  - Statistics
  - Storage
  - Computer Security
  - User Support
- FY2008/FY2009 Procurement

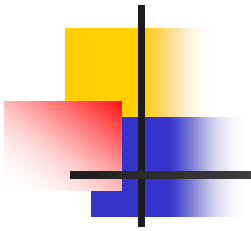
Hardware – Current Clusters

Name	CPU	Nodes	Cores	Network	DWF	Asqtad	Online
QCD	Single 2.8 GHz Pentium 4	127	127	Myrinet 2000	1400 MFlops per Node	1017 MFlops per Node	June 2004 0.15 TFlops
Pion	Single 3.2 GHz Pentium 640	518	518	Infiniband Single Data Rate	1729 MFlops per Node	1594 MFlops per Node	June 2005 / Dec 2005 0.86 TFlops
Kaon	Dual 2.0 GHz Opteron 240	600	2400	Infiniband Double Data Rate	4703 MFlops per Node	3832 MFlops per Node	Oct 2006 2.56 TFlops



Hardware

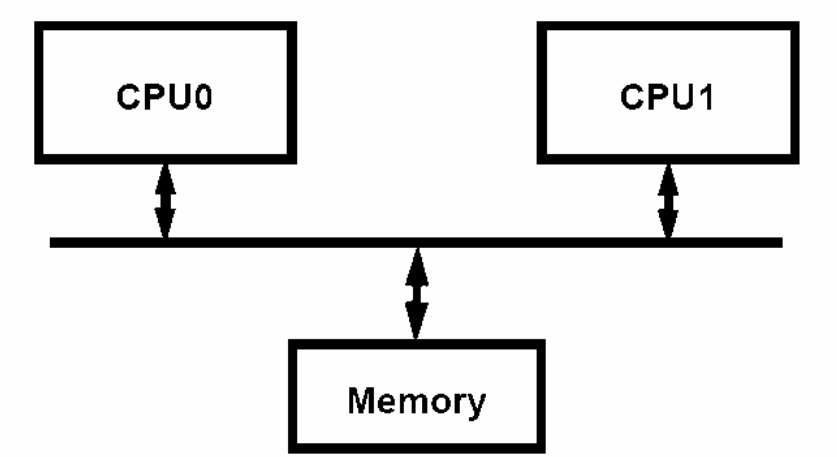
- QCD/Pion
  - Run 32-bit version of Scientific Linux 4.1, so large file support (files > 2.0 Gbytes in size) requires the usual *#define*'s
  - Access via *lqcd.fnal.gov*
- Kaon
  - Runs 64-bit version of Scientific Linux 4.2, so large file support is automatic
  - Access via *kaon1.fnal.gov*
  - Not compatible with QCD/Pion binaries
  - Will convert Pion to 64-bit after USQCD review



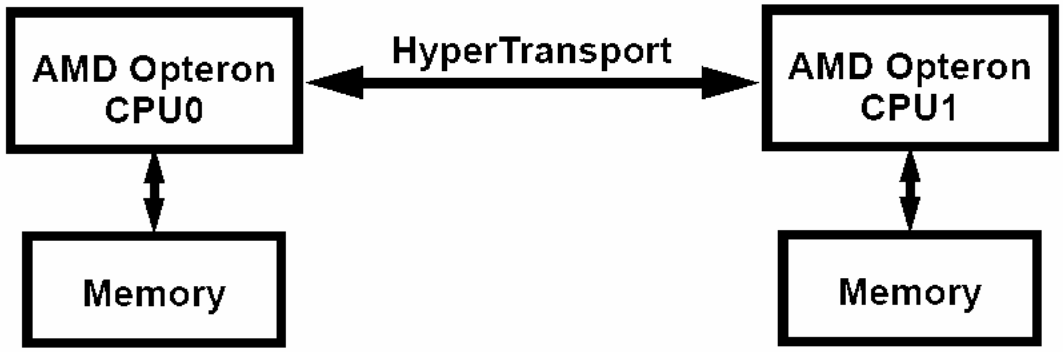
Hardware

- Kaon NUMA (non-uniform memory access) implications:
  - Kaon nodes have two Opteron processors, each with two cores
  - There is a separate memory bus for each processor
  - Access to the other processors memory bus is via hypertransport and incurs a latency penalty
  - MVAPICH and OpenMPI will automatically do the right thing – users don't have to worry
  - Non-MPI codes should use *libnuma* or be invoked via *numactl* to lock processes to cores and use local memory

Memory Architectures

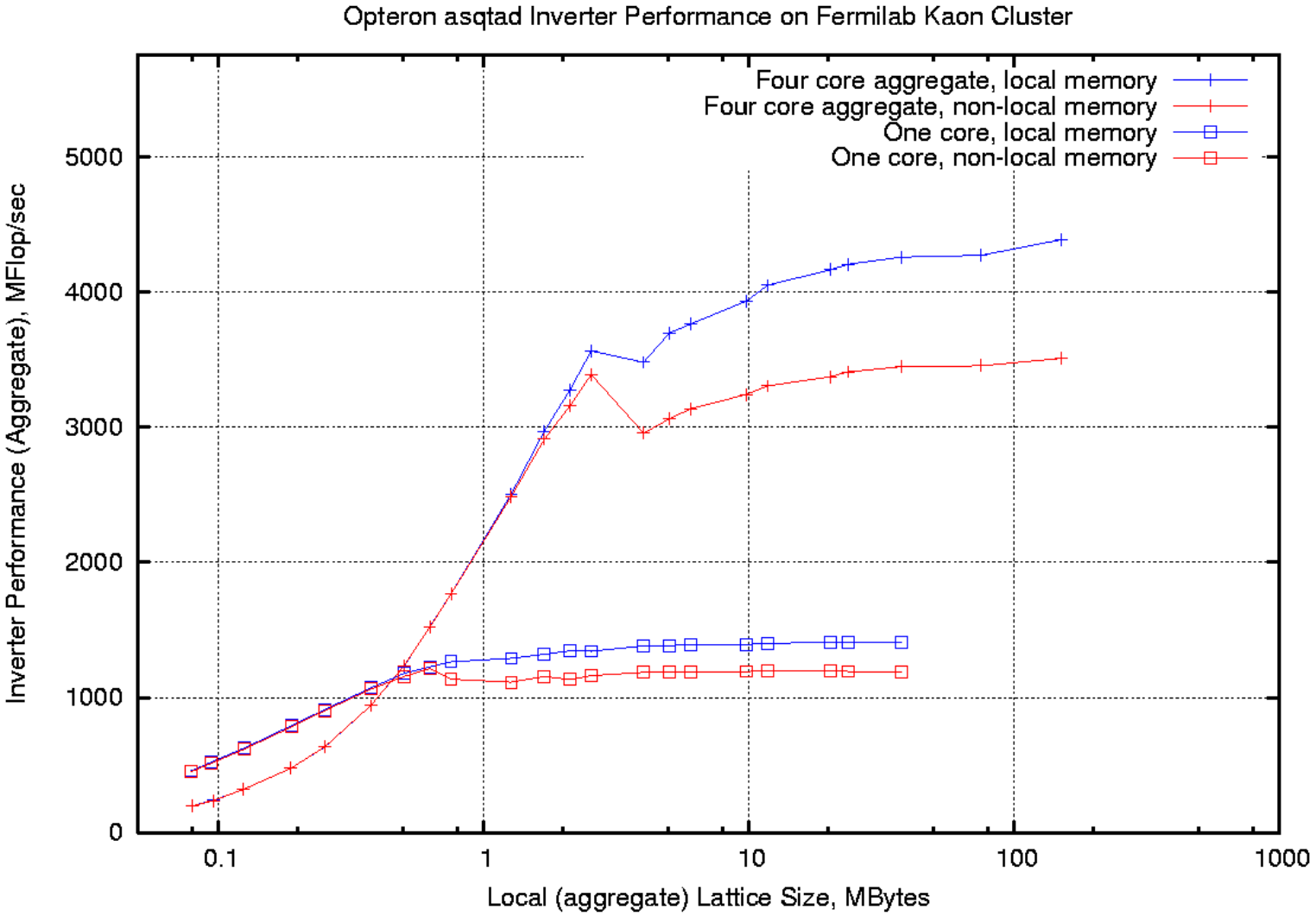


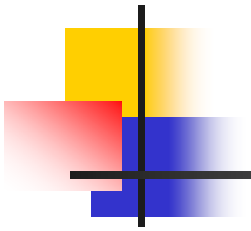
Intel Xeon SMP Architecture



AMD Opteron SMP Architecture

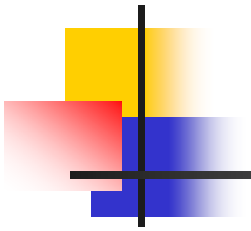
NUMA Effects





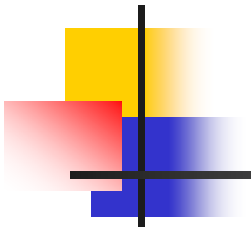
Hardware

- Kaon memory troubles:
  - In December, MILC configuration generation runs using 1024 processes (256 nodes) had high failure rates because nodes were rebooting or crashing
  - ASUS (motherboard manufacturer) suggested switching to single-ranked memory DIMMs
  - We replaced all dual-ranked DIMMs in early January
  - Since the replacements, lost node hours on these jobs have decreased from ~ 30% to less than 5%
  - Mean time to node reboot/crash on Kaon is about 18 KHrs → a 256-node, 3 hour job has about a 4% chance of failure



Hardware

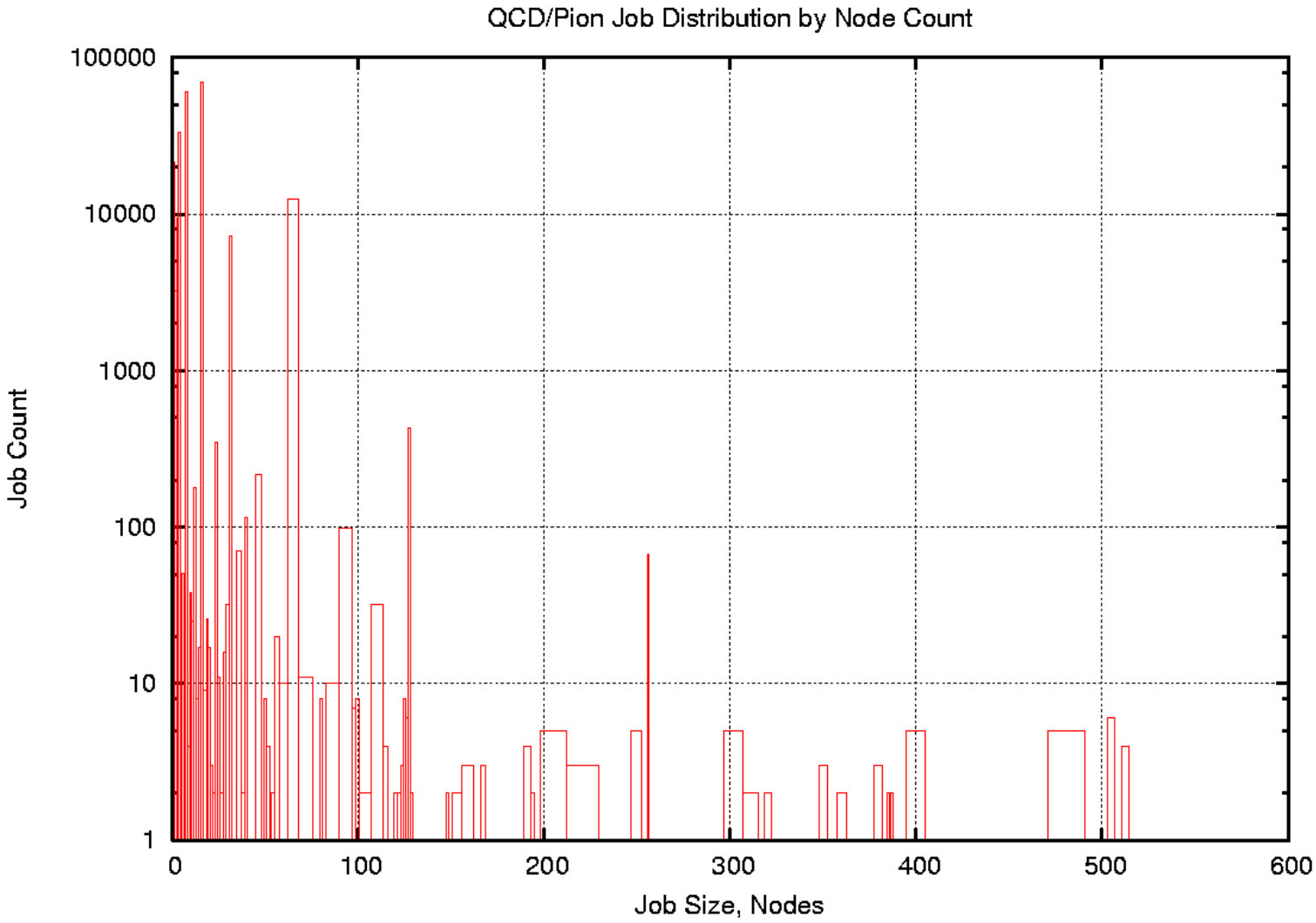
- Pion disk problems
  - Some local disks (~ 30 out of 260) on second half of Pion cluster exhibited bit error rates 100x the specification (1 in 10^13, instead of 1 in 10^15)
  - Vendor (Western Digital) confirmed bad cache memory, and replaced all disks
  - We now test all disks on all clusters monthly
- Users are urged to take advantage of CRC checks in QIO (or implement their own)
  - Observed CRC error rates on Kaon (a few a week) are likely consistent with B.E.R. of 1 in 10^15



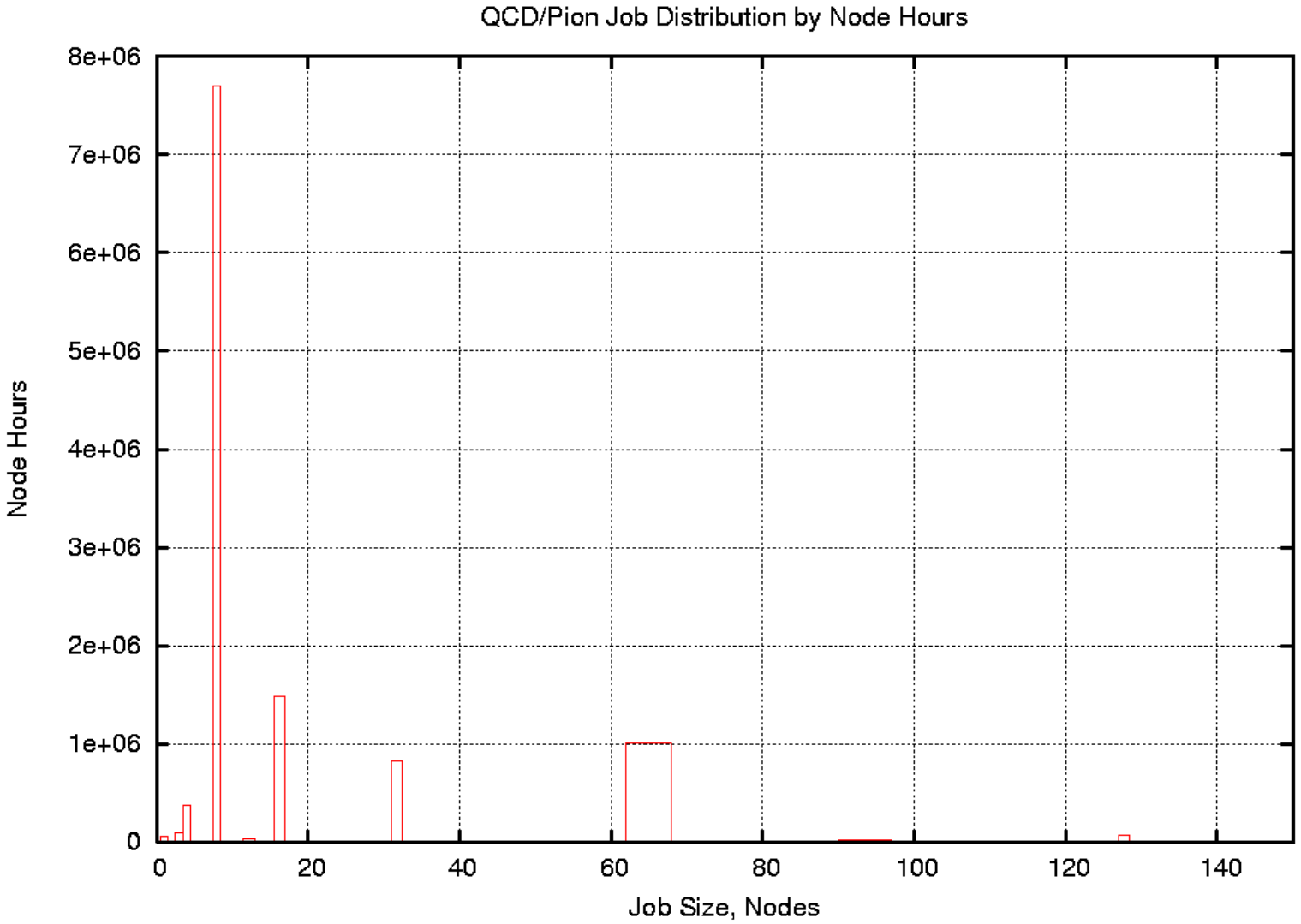
Statistics

- Since March 1, 2006:
  - Users submitting jobs:
    37 LQCD, 12 administrators or other
  - 287,708 jobs (262,838 multi-node)
  - 13.63 million node-hours
- USQCD Project deliverables (FY06 thru Feb):
  - 2.56 TFlops new capacity (3.58 TFlops total)
  - 1.47 Tflops-yrs delivered (112% of pace to goal of 3.19 Tflops-yrs)
  - 96.7% uptime (weighted by cluster capacity)

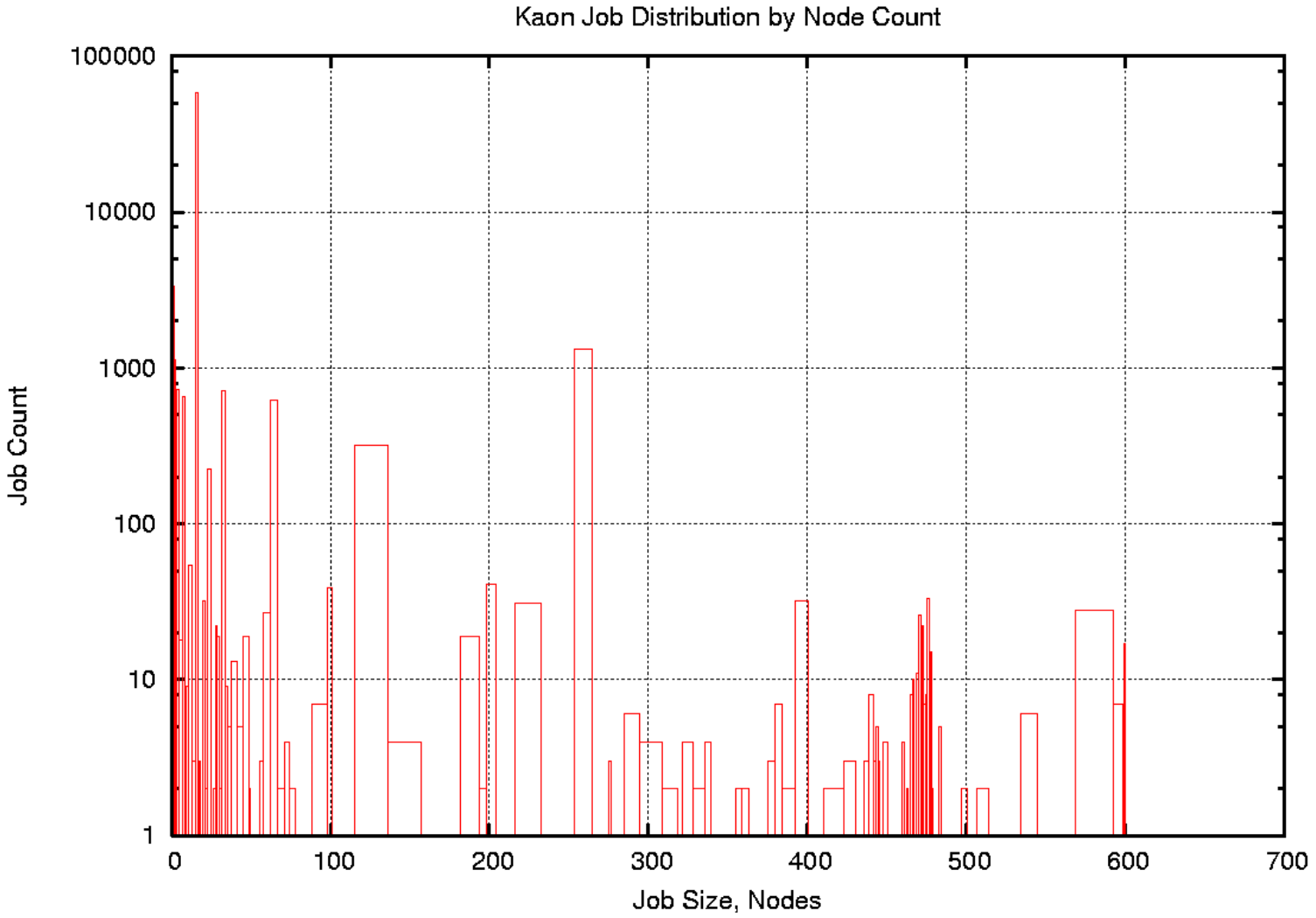
QCD/Pion Statistics



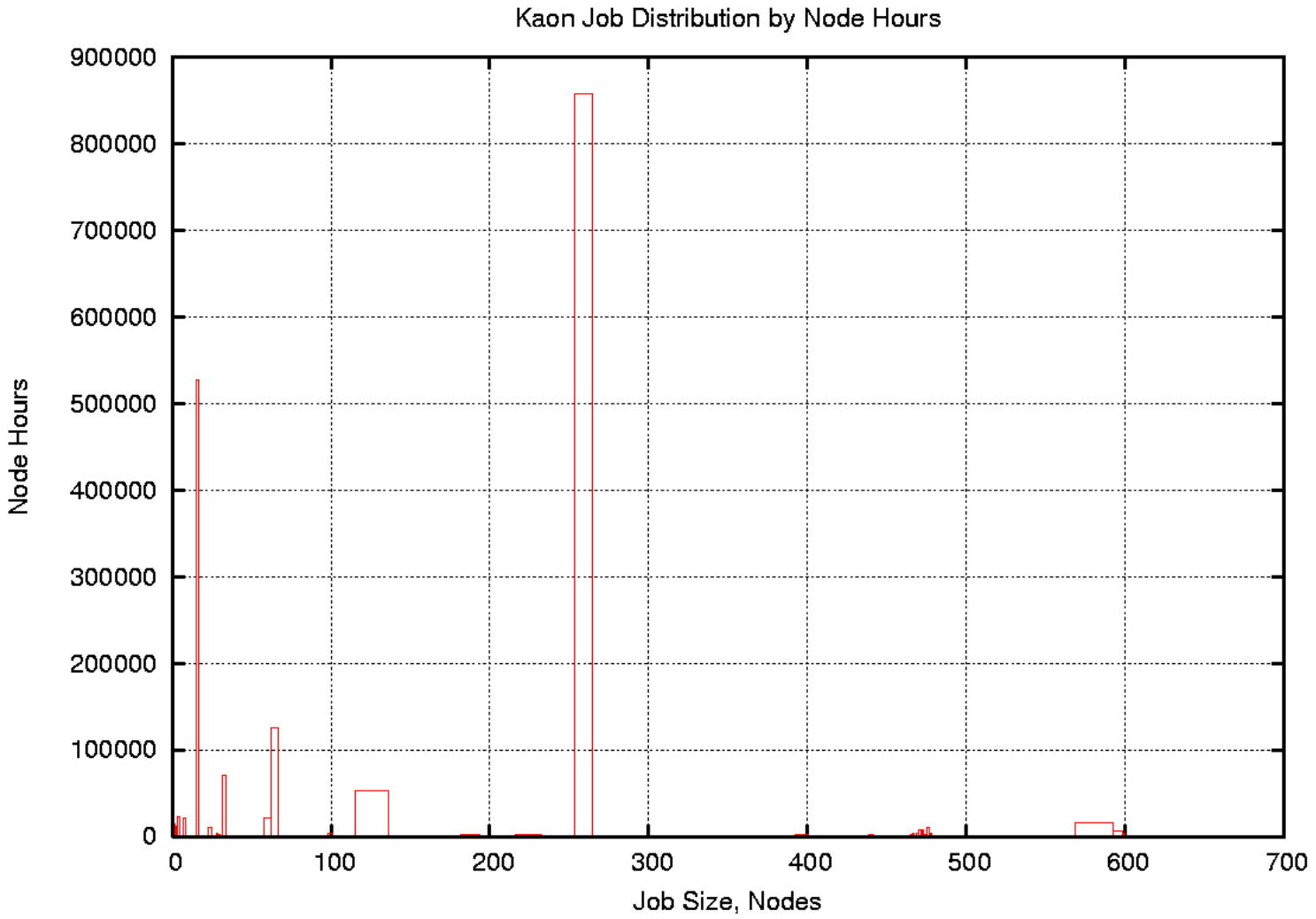
QCD/Pion Statistics



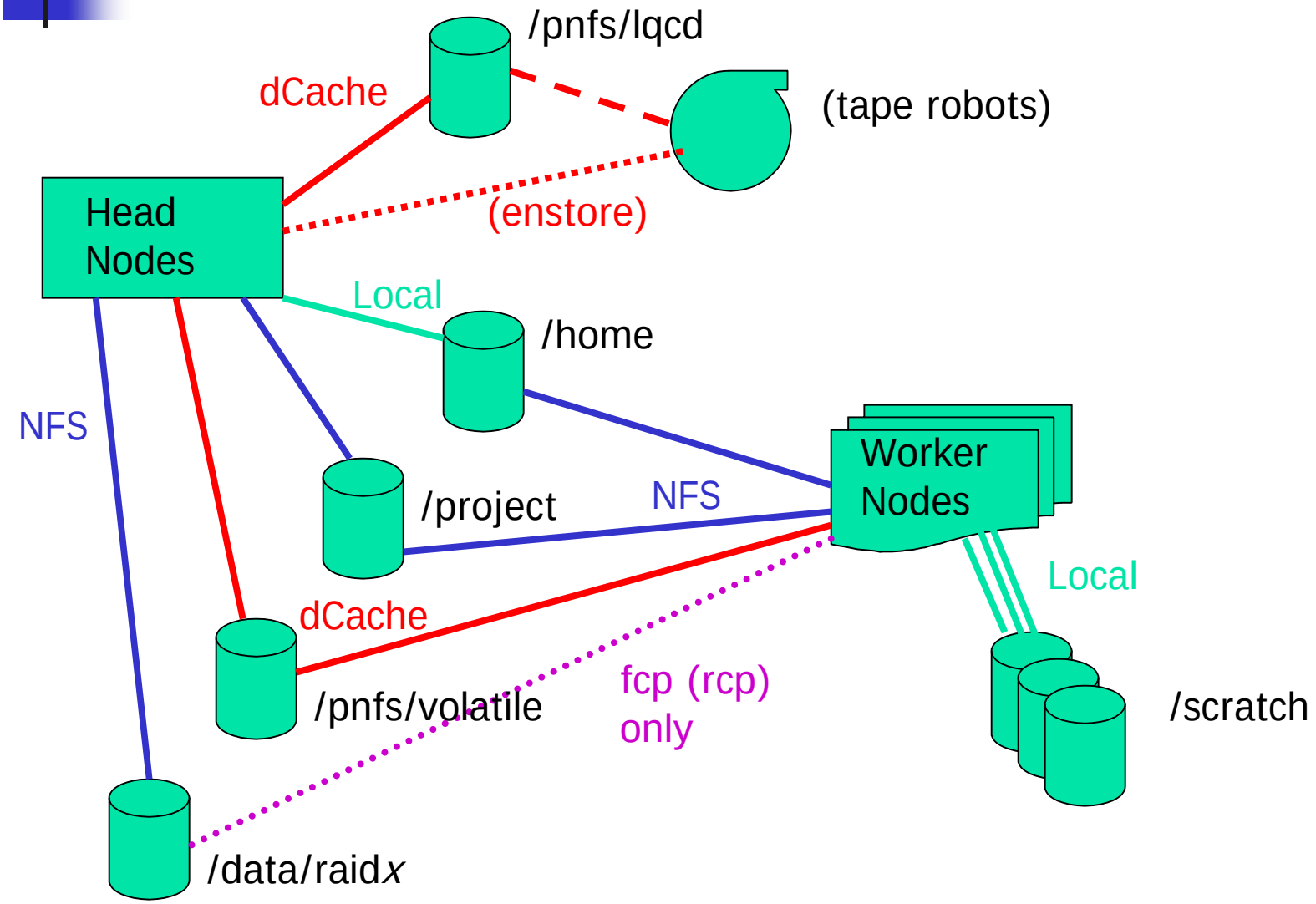
Kaon Statistics

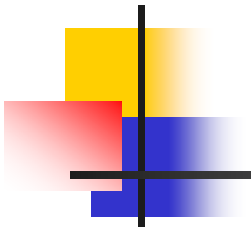


Kaon Statistics



Storage

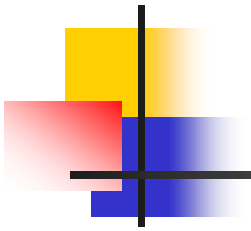

/pnfs/lqcd
dCache
(tape robots)
Head Nodes
(enstore)
Local
/home
NFS
/project
NFS
Worker Nodes
Local
dCache
/pnfs/volatile
fcp (rcp) only
/scratch
/data/raidx



Mass Storage

"Enstore"

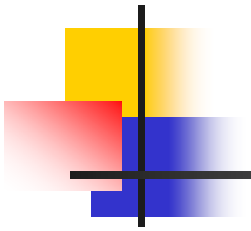
- Robotic, network-attached tape drives
- Files are copied using "encp src dest"
- 15 MB/sec transfer rate per stream
  - Increasing to > 40 MB/sec this summer
- Currently using ~160 Tbytes of storage



Mass Storage

"Public" dCache (/pnfs/lqcd/)

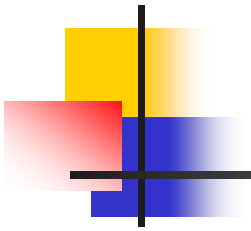
- Disk layer in front of Enstore tape drives
- All files written end up on tape ASAP
- Files are copied using "dccp src dest"
  - Pipes allowed
  - Also, direct I/O allowed (posix/ansi)
- On writing, hides latency for tape mounting and movement
- Can "prefetch" files from tape to disk in advance



Local Storage

“Volatile” dCache (/pnfs/volatile/)

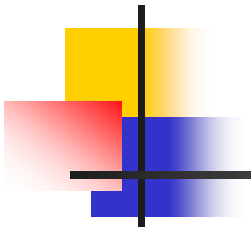
- Consists of multiple disk arrays attached to “pool nodes” connected to Infiniband network
- No connection to tape storage
- Provides large “flat” filesystem
- Provides high aggregate read/write rates when multiple jobs are accessing multiple files on different pools
- Supports file copies (via dccp) and direct I/O (via libdcap: posix/ansi style calls)
- About 27 Tbyte available
- No appends. Any synchronization between nodes in a job (MPI collectives) may lead to deadlocks.



Local Storage

Disk RAID arrays attached to head node

- /data/raid*x*, *x* = 1-8, total ~ 10 Tbytes
- Also, /project (visible from worker nodes)
- Data files must be copied by user jobs via fcp (like rcp) to/from server node
- Performance is limited:
  - By network throughput to/from server node
  - By load on server node



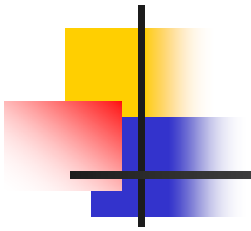
Local Storage

/scratch

- Each worker node has a local disk (30 GB on QCD and Pion, 80 GB on Kaon
- 30-40 Mbyte/sec sustained rate per node
- Cleaned at the beginning of each job
- Suitable for QIO “multifile” operations

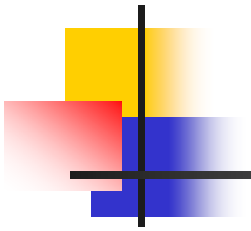
Properties of Filesystems

Name	Type	Visibiilty	Integrity	I/O Restrictions
/home	NFS	Global	Backed up nightly	Limited data rate
/project	NFS	Global	Backed up nightly	Limited data rate
/scratch	Local disk	Each worker has own	Erased at beginning of each job	High scalable data rate
/data/raidx	NFS	Head nodes only	RAID hardware but not backed up	Limited rate, use fcp to access
/pnfs/volatile	dCache	Global	Not backed up, oldest files deleted on demand	Scalable rate, no appends
/pnfs/lqcd	Enstore	Head nodes only	Data are on tape	No appends



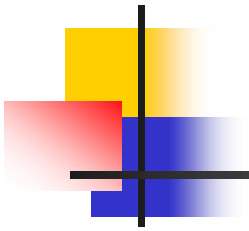
Security

- Kerberos
  - Strong authentication (instead of ssh)
  - Use Kerberos clients or cryptocards
  - Linux, Windows, Mac support
  - Clients are much easier than cryptocards – we're happy to help you learn
- Transferring files
  - Tunnel scripts – provide "one hop" transfers to/from BNL and JLab
  - See web pages for examples



User Support

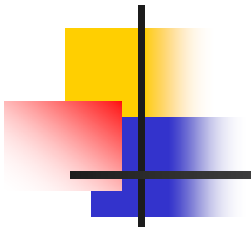
- Mailing lists
  - Lqcd-admin@fnal.gov
  - Lqcd-users@fnal.gov
- Level of support
  - 10 x 5, plus best effort off-hours
- Backups
  - /home, /project are backed up nightly from lqcd and kaon1; restores are available for up to 12 months
  - /data/raidx, /pnfs/volatile are not backed up – users are responsible for data integrity



User Support

Fermilab points of contact:

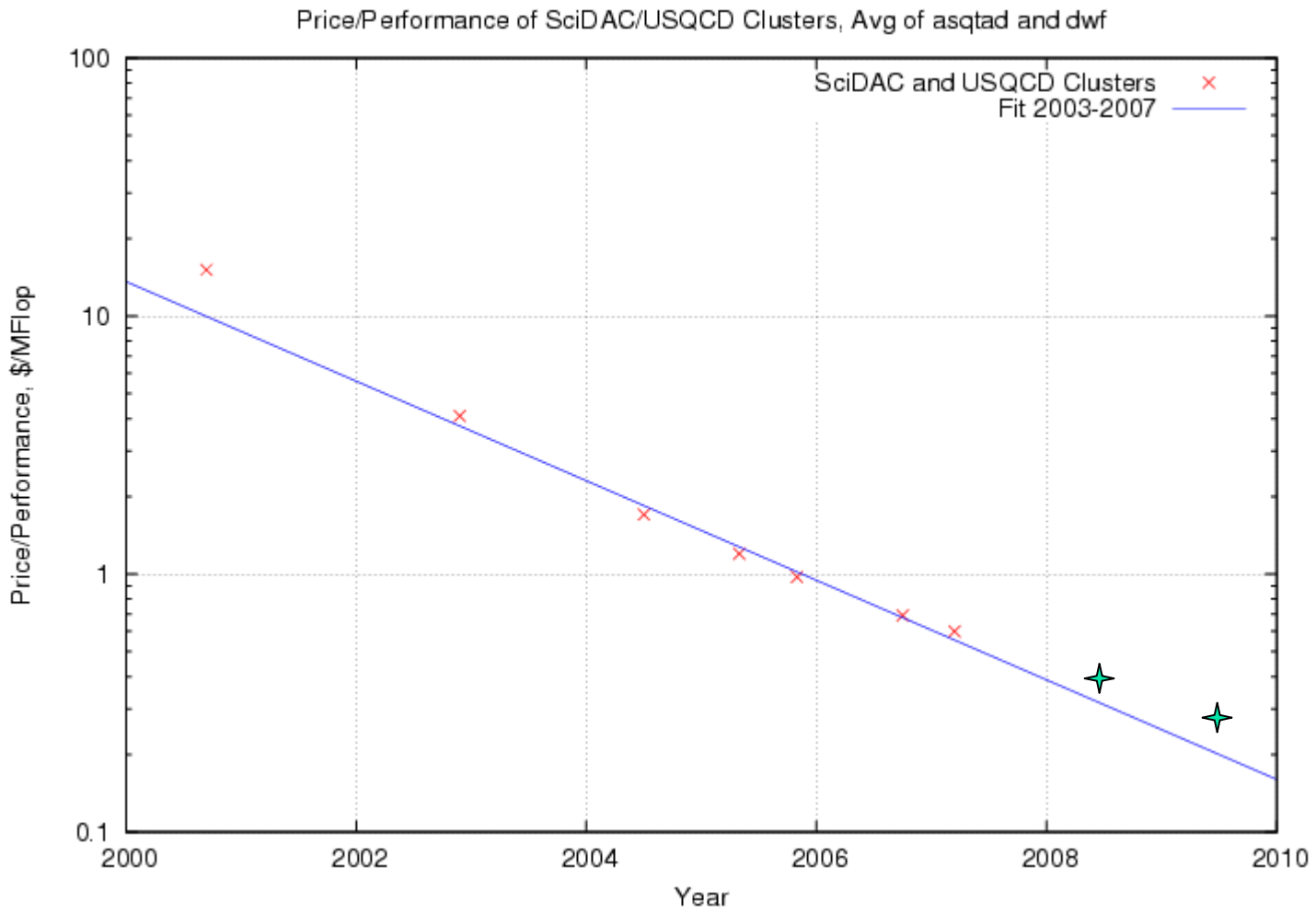
- Don Holmgren, djholm@fnal.gov
- Amitoj Singh, amitoj@fnal.gov
- Kurt Ruthmansdorfer, kurt@fnal.gov
- Nirmal Seenu, nirmal@fnal.gov
- Jim Simone, simone@fnal.gov
- Jim Kowalkowski, jbk@fnal.gov
- Paul Mackenzie, pbm@fnal.gov

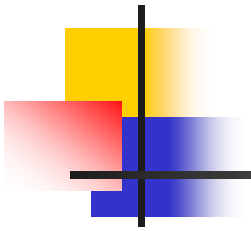


FY08/FY09 Procurement

- Plan of record (OMB Exhibit 300):
  - FY08: 4.2 TFlops system released to production by June 30, 2008, $1,630K ($0.39/MFlop)
  - FY09: 3.0 TFlops system released to production by June 30, 2009, $798K ($0.27/MFlop)
- Many potential advantages to combining FY08 and FY09 purchases into a larger buy in FY08
- Subject to negotiations

Price/Performance Trend



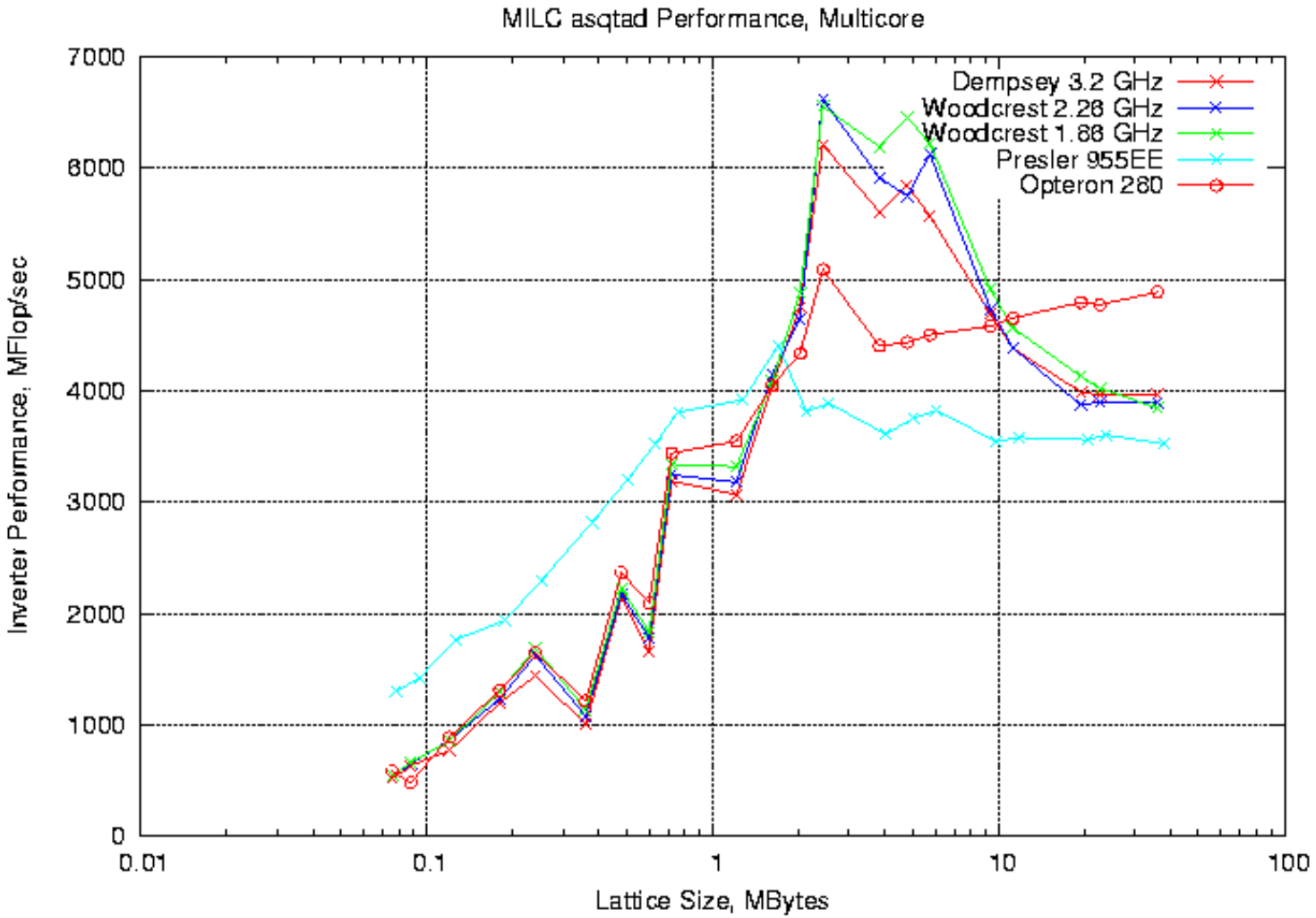


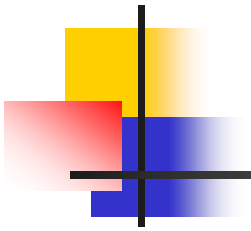
FY08/FY09 Procurement

Candidate processors:

- Opteron – quad core, better floating point and memory bandwidth than Kaon, possibly with L3 cache
- Xeon – quad core, new chipset, faster memory bus, possibly with large L3 cache
- Pentium – quad core, single socket, low cost if Infiniband is integrated

CPU Performance

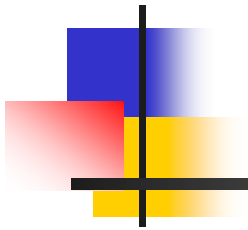




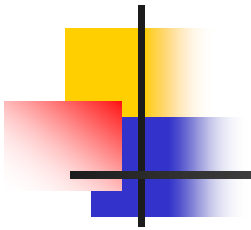
FY08/FY09 Procurement

- Meeting TFlops goals will be a challenge
  - New generation of Intel processors (“CoreDuo”) have been hampered by memory bandwidth
    - We are not the only govt customers to complain
    - FBDIMMs should be doing better – first chipsets may have been the culprit
  - Help from SciDAC multicore optimizations?
  - Help from L3 caches?
  - Infiniband improvements + next generation PCI Express may also help
    - Quad data rate + improved bus → latency to 1 µsec

Questions?



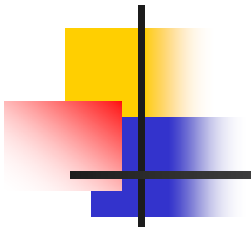
Backup Slides



Hardware

Current clusters:

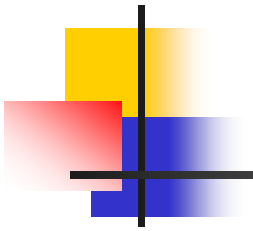
- "QCD"
  - 127 nodes, 2.8 GHz Pentium 4, 1 GB memory
  - Myrinet (128th connection is to I/O gateway)
  - Online since June 2004 → last full year of operation
  - Performance (64 node runs):
    - DWF: 1400 Mflops/node
      Ls=16, average of 32x8x8x8 and 32x8x8x12
    - Asqtad: 1017 Mflops/node
      14^4 local lattice/node
    - Total capacity: ~ 150 Gflops



Hardware

Current clusters (cont'd):

- "Pion"
  - 518 nodes, 3.2 GHz Pentium 640, 1 GB memory
  - Infiniband (single data rate)
  - Full cluster online since December 2005
    - First half online since June 2005
  - Performance (64 node runs):
    - DWF: 1729 Mflops/node
      Ls=16, average of 32x8x8x8 and 32x8x8x12
    - Asqtad: 1594 Mflops/node
      14^4 local lattice/node
    - Total capacity: ~ 860 Gflops



Hardware

Current clusters (cont'd):

- "Kaon"
  - 600 nodes, 2.0 GHz Opteron 240, 4 GB memory
    - Dual core, dual processor → 2400 cores available
  - Infiniband (double data rate)
  - Online since October 3, 2006
  - Performance (128 core runs = 32 nodes):
    - DWF: 4703 Mflops/node
      Ls=16, average of 32x8x8x8 and 32x8x8x12
    - Asqtad: 3832 Mflops/node
      14^4 local lattice/node
    - Total capacity: ~ 2.56 Tflops