LQCD-ext Technical Performance and Remaining Plans for FY10/FY11 Deployments

Don Holmgren

Fermilab

djholm@fnal.gov

SC LQCD-ext Annual Progress Review

Fermi National Accelerator Laboratory

May 10-11, 2011

Outline

- Overview of SC LQCD-ext acquisitions
- Computational requirements and cluster design
- FY10 cluster deployment and performance
- FY11 cluster and GPU deployments

Overview of SC LQCD-ext Acquisitions

- Plan on approximately five acquisitions
  - Usually one per year in FY10-FY14, but some years will have both conventional and GPU-accelerated cluster purchases
  - FY10 and FY11 conventional cluster buy is “across” the fiscal year boundary so that we employ a single contract
- Guiding principle: procure the systems that will be the most effective for the planned science, given the portfolio of operating SC LQCD-ext and other machines at that time
  - FY10/FY11 – we have deployed a commodity cluster, and in late FY11 we will purchase a GPU-accelerated cluster
  - FY12 – BlueGene/Q (BG/Q), commodity cluster, GPU-accelerated cluster, or some combination (see next talk)
  - FY13/FY14 – perhaps BG/Q in FY13, otherwise a combination of commodity and GPU-accelerated clusters

Overview of SC LQCD-ext Acquisitions

Computational capacity goals by year for SC LQCD-ext:

	FY2010	FY2011	FY2012	FY2013	FY2014
Computing hardware budget (not including storage)	$1.60M	$1.69M	$1.875M	$2.46M	$2.26M
Planned/Achieved Capacity of new cluster deployments, Tflop/s	11 / 12.5	9	24*	44*	57*
Planned GPU Deployment Count	–	128	**	**	**

- FY2011 original acquisition plan for 12 Tflop/s was changed to 9 Tflop/s plus a GPU-accelerated cluster with 128 NVIDIA “Fermi” GPUs

* FY2012-FY2014 Tflop/s cluster capacities will likely be reduced with some of the budget shifted to GPU-accelerated clusters

** FY2012-FY2014 GPU deployment counts TBD (see next talk)

Cluster Design

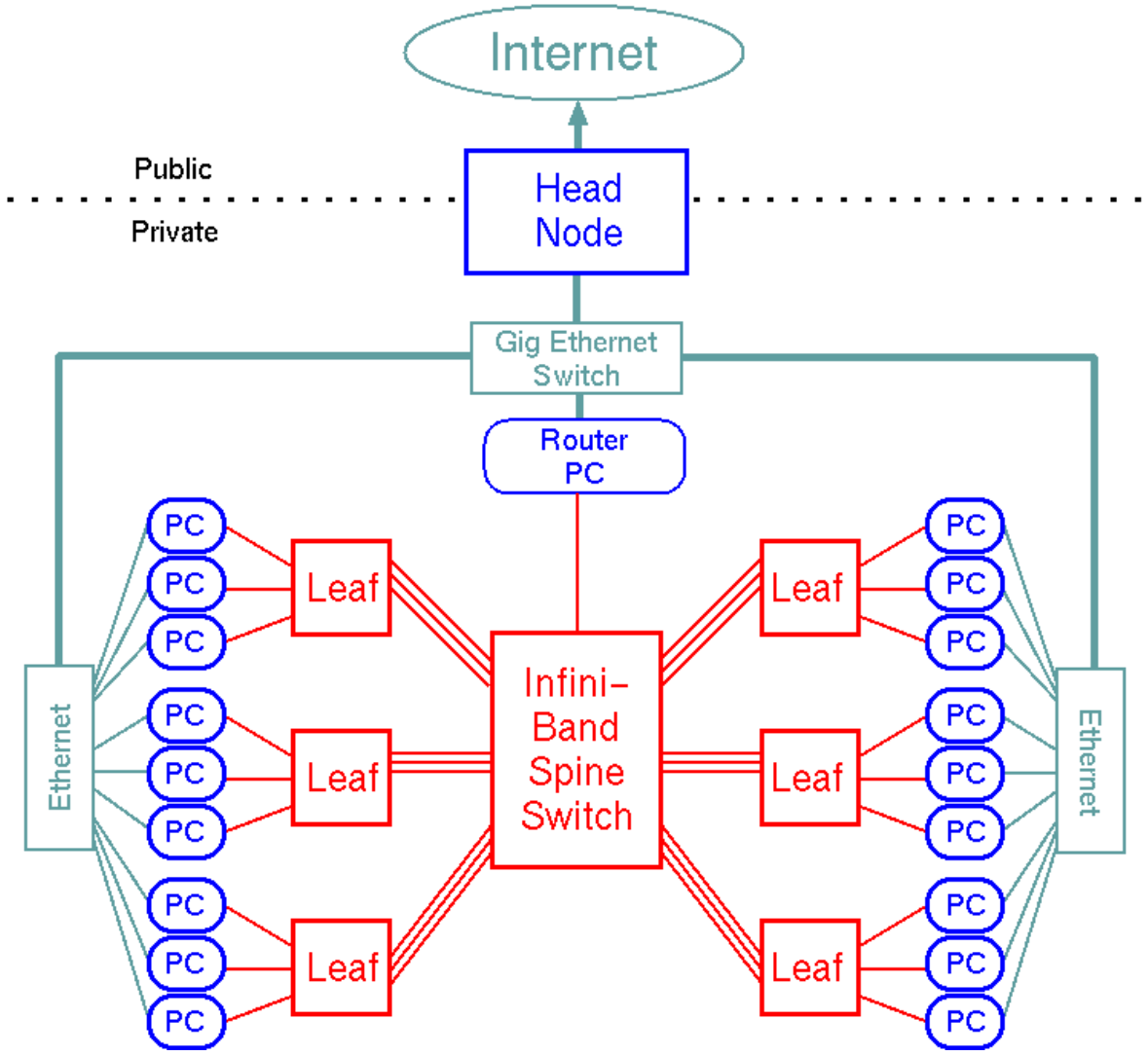
Computational Requirements

- Either memory bandwidth, floating point performance, or network performance (bandwidth at message sizes used) will be the limit on performance on a given parallel machine
- On single commodity nodes memory bandwidth in the constraint that limits performance
  - GPUs deliver more memory bandwidth per dollar than conventional CPU's, but can only be used for some of our calculations
- On current parallel computer clusters, the constraint is either memory bandwidth or network performance, depending upon how many nodes are used on a given job
  - Network performance limits strong scaling:
    Surface area to volume ratio increases as more nodes are used, causing relatively more communications and smaller messages
  - GPUs require higher network bandwidth than CPUs

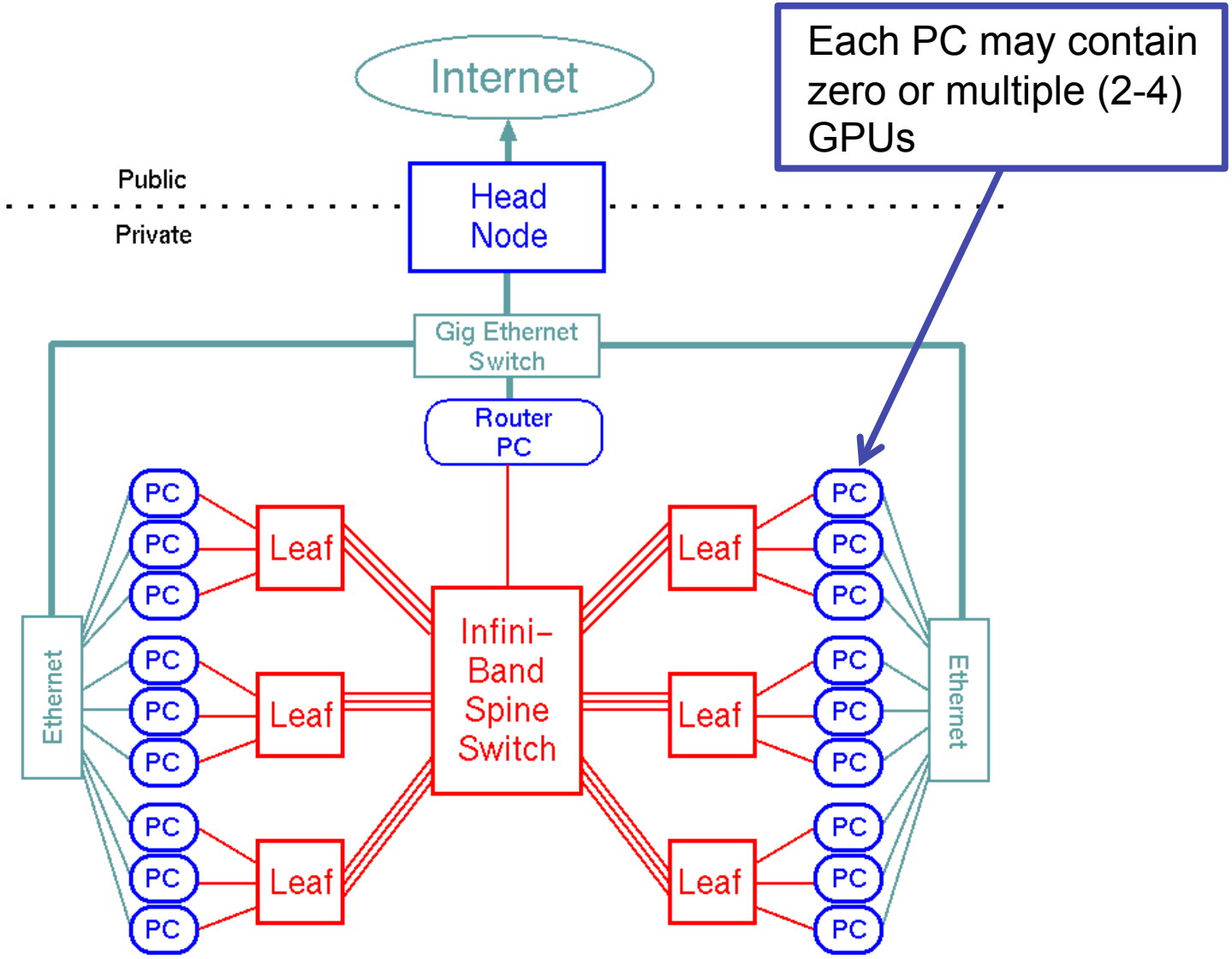
Computational Requirements

- We design and/or buy systems that as part of our hardware portfolio will most effectively carry out the current and anticipated scientific programs
- This means:
  - Systems matched to the type and size of LQCD calculations that will be performed
  - Systems with the best price/performance for LQCD applications
  - Machines with the best memory bandwidth
  - High performance interconnects
  - Networks balanced to single node capacities and anticipated job sizes

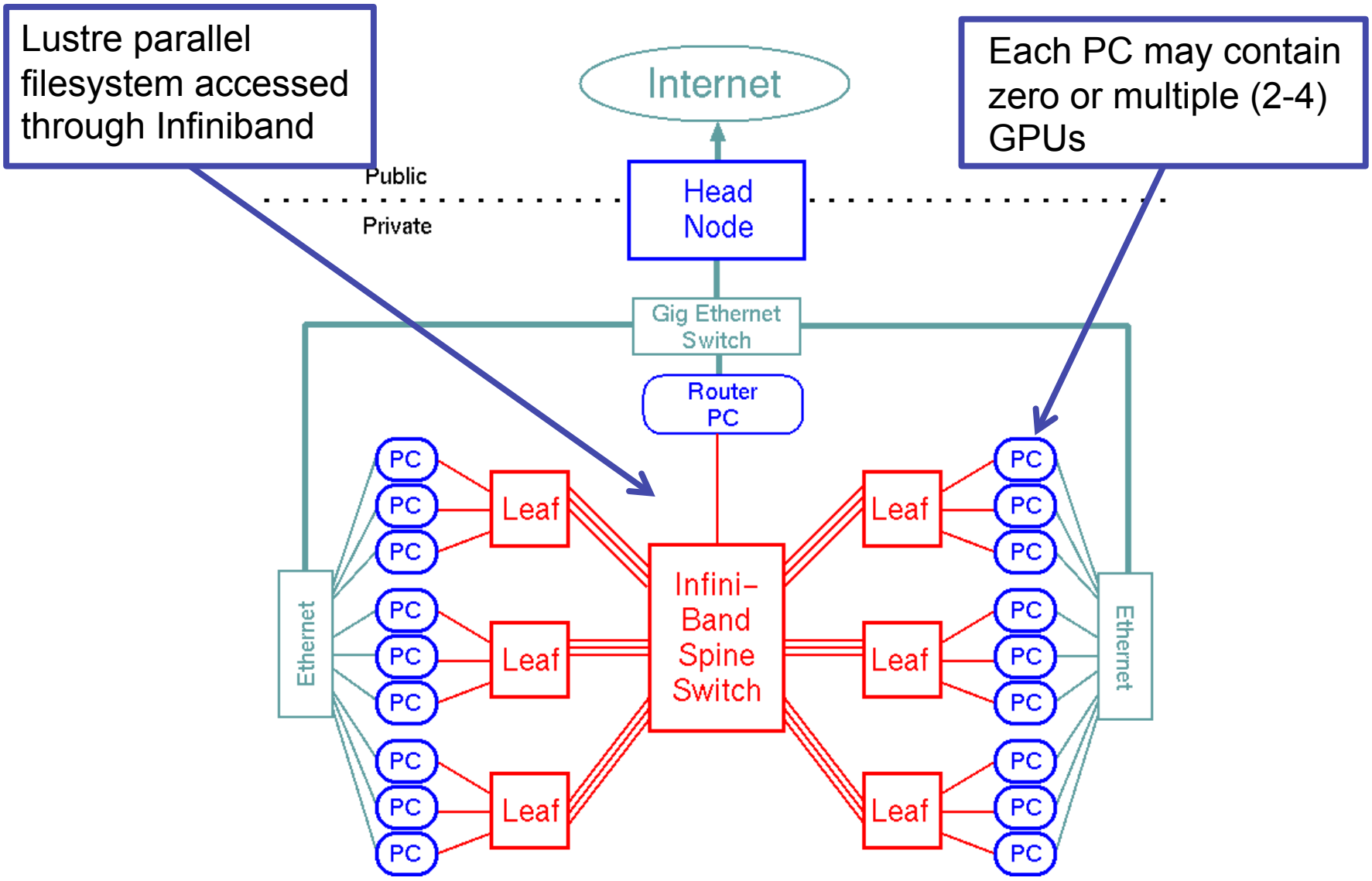
Typical LQCD Cluster Layout



Typical LQCD Cluster Layout



Typical LQCD Cluster Layout



Rating LQCD-ext Computing Facilities

- Definition of the sustained capacity of LQCD-ext computing hardware:
  - The performance of the improved staggered ("asqtad") and domain wall fermion ("DWF") conjugate gradient inverters are measured using parallel jobs spanning a significant number of processors (128 cores on clusters)
  - The average of the asqtad and DWF values (per core) multiplied by the number of available cores gives the defined sustained Tflop/s capacity
- Although the inverter is only part of the computing load, and other actions besides DWF and asqtad are used, on clusters and leadership machines the asqtad-DWF average has been predictive of overall computing throughput
- The asqtad-DWF average is not known to be predictive for GPUs
  - Neither DWF, HISQ, nor all of asqtad are in production so performance measurements have not been available
  - For some job types, execution times are not dominated by the inverter
  - Therefore requests and allocations for GPU resources are in "GPU-hours"
  - We are tracking "cost-equivalent" GPU capacities by comparing performance and costs of actual jobs run on both conventional and GPU hardware, and plan to use these data to guide future purchase decisions

FY10 Deployment and Performance

The FY10 Ds Procurement

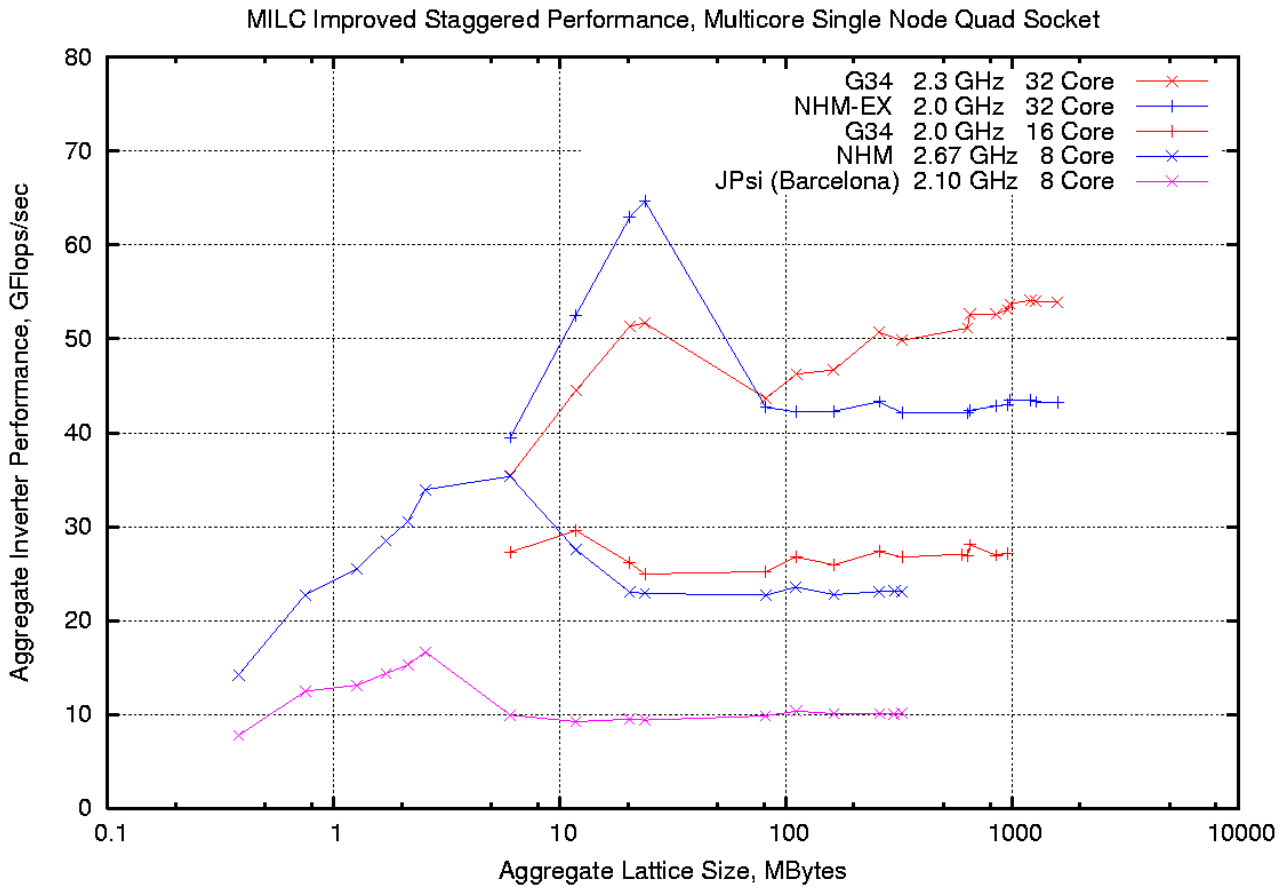
2010

- Feb 12 – RFI released to vendors
- Mar 19 – RFI responses received from vendors
- Apr 19 – RFP released to vendors
- Apr 29,30 – DOE FY10 Progress Review
- May 27 – RFP bids received from vendors
- Jun 17 – RFP award recommendation to purchasing department
- Jun 29 – Purchase order to vendor (commit FY10 funds)
- Oct 15 – Delivery of FY10 equipment complete (started Aug 20)
- Nov 1 – Friendly user period begins
- Dec 1, 2010 – Release to production FY10 portion (12.5 TFlops)

Ds Details

- Award was to best value bid, based on price, LQCD application performance, power efficiency, space efficiency, vendor qualifications and past performance
- Hardware details:
  - Quad-socket eight-core AMD 2.0 GHz “Magny-Cours” processors
  - 64 Gbytes memory per node
  - QDR Infiniband with 2:1 oversubscription
  - 245 worker nodes, plus head nodes
  - $1.51M including G&A ($1.43M for worker nodes + Infiniband)
- Performance
  - Asqtad:DWF 51 Gflop/node (128-process MPI runs)
  - 12.50 Tflop/s → $0.114/Mflop

Performance of Candidate Processors



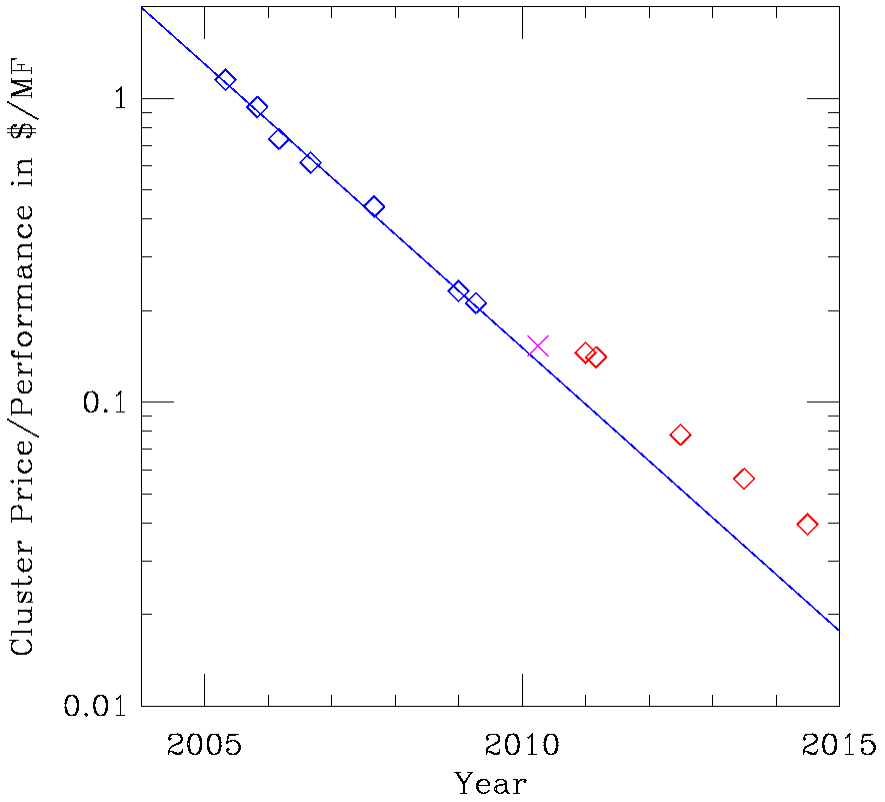
AMD Magny-Cours, 8 cores per socket. Top curve is 4 socket system, bottom curve is 2 socket system

Intel Nehalem EP and Nehalem EX. Top curve is 4 socket system, bottom curve is 2 socket system

Reference: SC LQCD J/Psi cluster, AMD "Barcelona" 4 cores per socket

Four socket versions of Intel and AMD processors show essentially perfect scaling over two socket versions

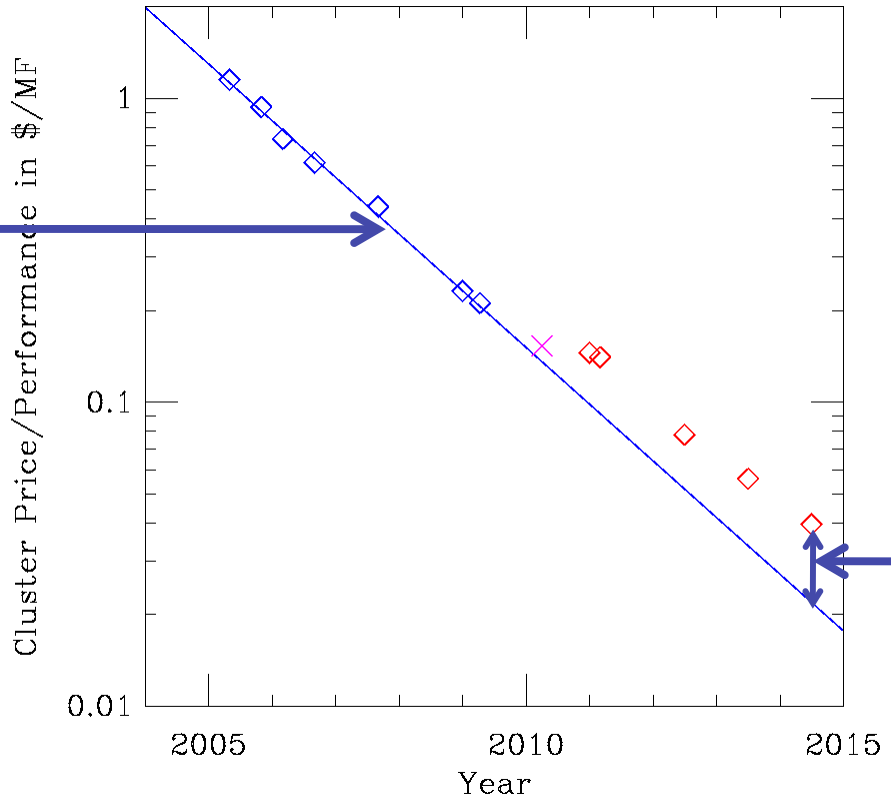
Cost and Performance Basis



Cluster	Price per Node	Performance/Node, MF	Price/Performance
Pion #1	$1910	1660	$1.15/MF
Pion #2	$1554	1660	$0.94/MF
6n	$1785	2430	$0.74/MF
Kaon	$2617	4260	$0.61/MF
7n	$3320	7550	$0.44/MF
J/Psi #1	$2274	9810	$0.23/MF
J/Psi #2	$2082	9810	$0.21/MF
10q	$3461	22667	$0.15/MF

Cost and Performance Basis

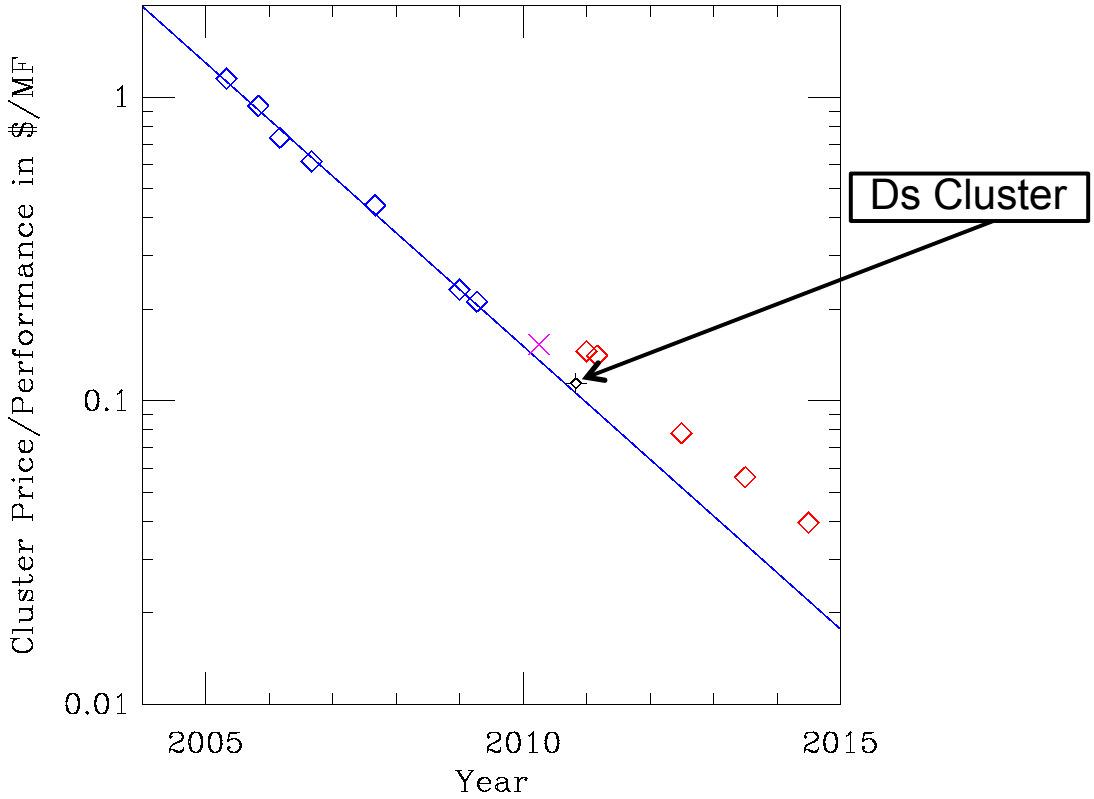
Fit is to the blue diamonds, slope gives halving time of 1.613 years



Contingency

Year	Deploy Date	Price/Perf. Goal	Price/Perf. Trend	Goal (TF)	Contingency (TF)	Contingency (TF %)
2010	2011.0	$0.15/MF	$0.098/MF	11	4.4	40%
2011	2011.2	$0.14/MF	$0.098/MF	12	4.4	36%
2012	2012.5	$0.078/MF	$0.052/MF	24	11.9	50%
2013	2013.5	$0.056/MF	$0.034/MF	44	26.8	61%
2014	2014.5	$0.040/MF	$0.022/MF	57	42.6	75%

Cost and Performance Basis



Cluster	Price per Node	Performance/Node, MF	Price/Performance
Pion #1	$1910	1660	$1.15/MF
Pion #2	$1554	1660	$0.94/MF
6n	$1785	2430	$0.74/MF
Kaon	$2617	4260	$0.61/MF
7n	$3320	7550	$0.44/MF
J/Psi #1	$2274	9810	$0.23/MF
J/Psi #2	$2082	9810	$0.21/MF
10q	$3461	22667	$0.15/MF
Ds	$5810	50810	$0.114/MF

The FY11 Ds Procurement

2011

- Jan 19 – Advice from the Executive Committee on the split between Ds expansion and a separate Ds cluster
- Mar 7 – Purchase order to vendor for first half of the expansion
- Apr 15 – FY11 Full-year C.R. signed into law by the President
- May 4 – Delivery of nodes
- May 16 – Delivery of remaining equipment for first half
- May 31 – Purchase order to vendor for second half of the expansion
- Jun 1 – Release to production of first half (4.5 TFlops)
- Aug 15 – Delivery of second half
- Sep 1 – Release to production of second half (4.5 TFlops)

FY11 Budget Delays

- Continuing budget resolutions have affected FY11 purchases
  - 8 total C.R.'s, with the final thru end FY11 signed into law 4/15
  - As of May 2 final FY11 budget guidance had not been received by Fermilab
- Fermilab throttled the rate of all spending in response
  - By February, enough funds (4/12th) for the project to requisition half of the planned "Ds" expansion, and 25% of the proposed GPU-accelerated cluster
  - Lab approval for "Ds" expansion delayed to early March, and approval for GPU cluster indefinitely delayed
  - As soon as approved we will proceed with remaining "Ds" purchase, and with GPU buy
  - Half of "Ds" expansion, and GPU cluster, will not meet schedule milestone
  - Project has also delayed purchasing additional storage until summer

FY11 GPU Cluster Design

- Existing USQCD GPU-accelerated cluster designs:
  - JLab: 4 GPUs per host, mixture of Tesla and graphics GPUs of two generations (GTX280, GTX480, S1070, C2050), restricted bandwidth Infiniband (QDR in x4 PCIe slots)
  - FNAL: small (16 GPU) deployment of S1070 Teslas with 2 GPUs per “JPsi” host machine, DDR Infiniband
  - Running to date has been primarily single GPU jobs, and also parallel jobs that cut problems along the time dimension (anisotropic clover lattices) using up to 32 GPUs
- Recent enabling software developments:
  - Support for DWF, asqtad (latter incl. force terms necessary for evolution)
  - For asqtad, support to cut problems along multiple dimensions

FY11 GPU Cluster Design

- In consultation with SciDAC software committee, project identified these requirements:
  - GPU count high enough to do large-scale configuration generation: jobs of at least 64 GPUs, and as high as 128; this would require software for cutting along multiple dimensions
  - Sufficient PCIe and Infiniband bandwidth to support these jobs
  - NVIDIA Tesla based, because:
    - ECC memory capability, necessary for non-inverter code
    - Warranty issues with non-Tesla hardware that have errors in numerical calculations but pass graphics tests
    - Direct GPU to GPU, and GPU to IB, communications only to be supported on Tesla
    - Larger memory per GPU (3 or 6 GB, vs 1.5 GB in graphics space)
    - Complementary to JLab ARRA resources, and so will help balance USQCD GPU resources against requirements

FY2011 GPU Schedule

2011

- May 31 – RFP released to vendors
- June 30 – Bids received
- July 15 – Purchased order released (commit FY11 funds)
- Aug 1 – Sample unit received
- Sept 15 – Delivery of all items
- Oct 19 – Acceptance test complete
- Oct 31 – Release to production

Storage

- Operational disk storage is provided in several flavors:
  - "Scratch" disk, reused for each job, local to each worker node, 100-200 GB available per node (varies by cluster)
  - "Home" areas, served by NFS, local to each cluster, about 4 GB available to each user (stores binaries, batch scripts and logs)
  - Large disk storage, served by Lustre FNAL and JLab
    - High performance, with parallel capabilities
- Based on 2009 allocation requests, we estimated an annual need for large disk storage of about a 6 TB addition per 1 TFlop of computing capacity

Storage Requirements and Budget

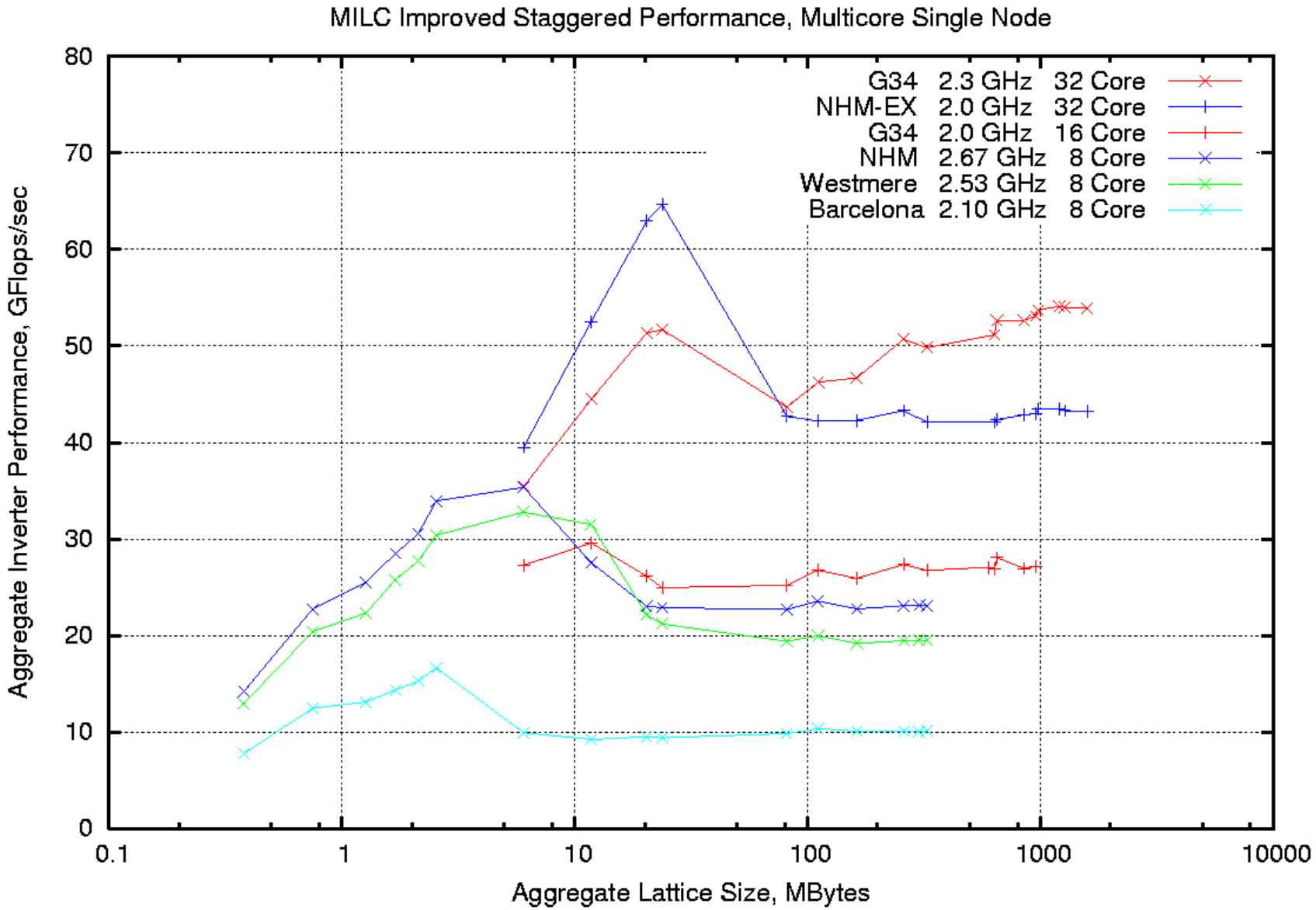
- All submitted proposals for time on USQCD resources must include estimates for storage
  - Allocation year starts July 1
  - Users are charged in equivalent node-hour “currency” for each Tbyte of disk and tape used on their project; tapes can be returned for a pro-rated credit
  - In the 2011 USQCD allocation proposals, users requested a total of 12.3M JPsi-core hours in storage (330 TB disk, 915 TB tape); depending upon how GPU flops are counted, this is still consistent with 6 TB disk per TF capacity (3.3 TB/TF with GPUs counted at 63 TF, 8.25 TB/TF not counting GPUs)
  - FNAL + JLab together now operate about 675 Tbytes of disk (375+300, BNL 24 Tbytes
  - FNAL will add up to an additional 205 Tbytes before end of FY11 (leveraging CMS purchases); JLab is bringing online now an additional 192 TB on ARRA funds
- Plan sets storage budget at 5% of annual computing hardware budget
  - Sufficient for FY11
  - Sufficient for FY12-FY14 if needs continue to be proportional to capacity

Summary

- The FY11 procurement schedule has been impacted by the series of continuing resolutions, but will be on budget
  - The price of the Ds expansion is fixed by contract and is within the project budget
  - Half of the Ds expansion will be released 1 month ahead of the project milestone (June 30); the second half will be delayed at least 2 months past the milestone
  - The GPU cluster will be delayed at least 4 months past the milestone
- FY11 purchases consist of a Ds expansion by 9 TFlops, and an accelerated cluster with 128 GPUs
- Existing storage together with purchases using FY11 funds is sufficient for the 2011 allocation requests

Questions?

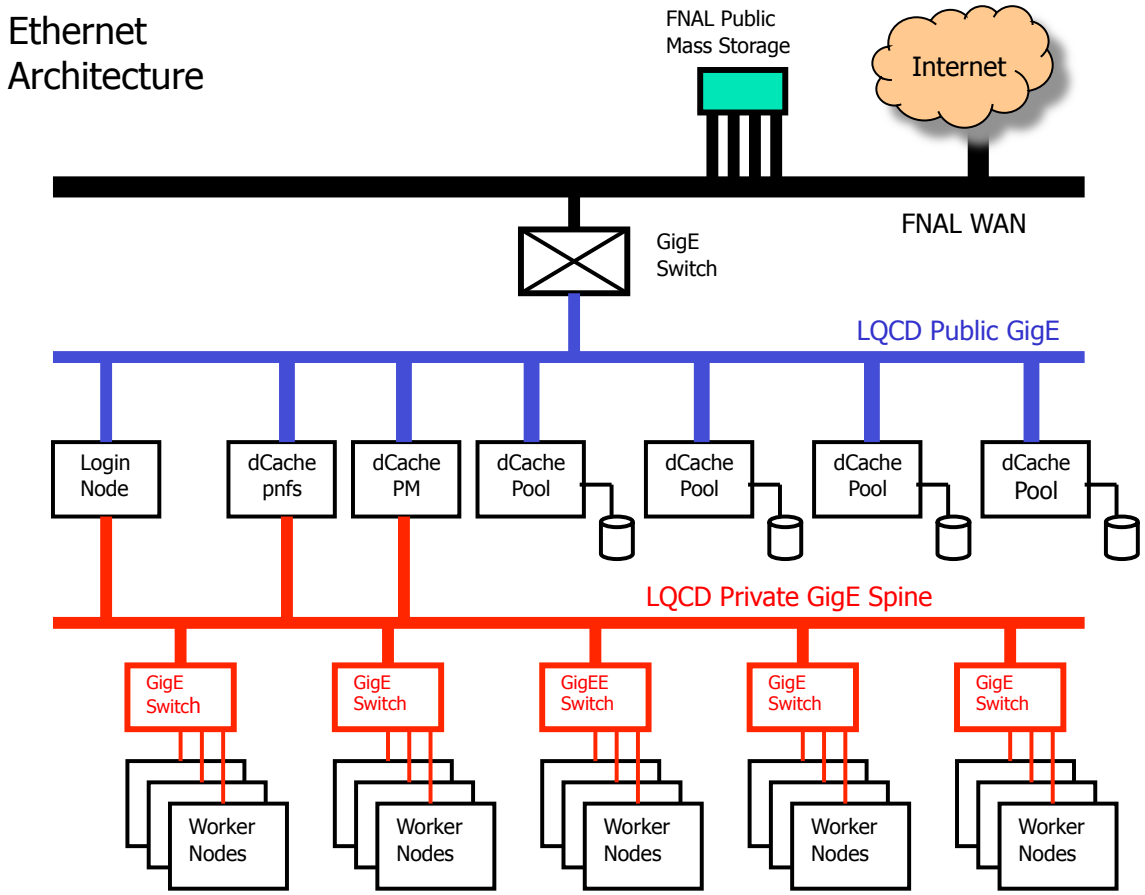
Backup Slides


MILC Improved Staggered Performance, Multicore Single Node
Aggregate Inverter Performance, GFlops/sec
Aggregate Lattice Size, MBytes
0
10
20
30
40
50
60
70
80
0.1
1
10
100
1000
10000
G34 2.3 GHz 32 Core
NHM-EX 2.0 GHz 32 Core
G34 2.0 GHz 16 Core
NHM 2.67 GHz 8 Core
Westmere 2.53 GHz 8 Core
Barcelona 2.10 GHz 8 Core

SC LQCD Clusters

Name (FNAL/ TJNAF)	6N	Kaon	7N	J/Psi
Speed Processor (Socket Count) (Cores/CPU)	3.0 GHz Pentium D (256) (2)	2.0 GHz Opteron (1200) (2)	1.9 GHz Opteron (792) (4)	2.1 GHz Opteron (1712) (4)
Memory Bus Speed	800 MHz	1066 MHz	1066 MHz	1066 MHz (1333 MHz memory)
Single or Dual Socket	Single	Dual	Dual	Dual
Interconnect Fabric	Infiniband (SDR)	Infiniband (DDR)	Infiniband (DDR)	Infiniband (DDR)
Performance	0.6 Tflop/s asqtad:DWF	2.6 Tflop/s asqtad:DWF	2.9 Tflop/s asqtad:DWF	8.4 Tflop/s asqtad:DWF
Date in Production	3/2006	10/2006	6/2007 (Upgrade 11/2007)	1/2009 (FY08) 4/2009 (FY09)

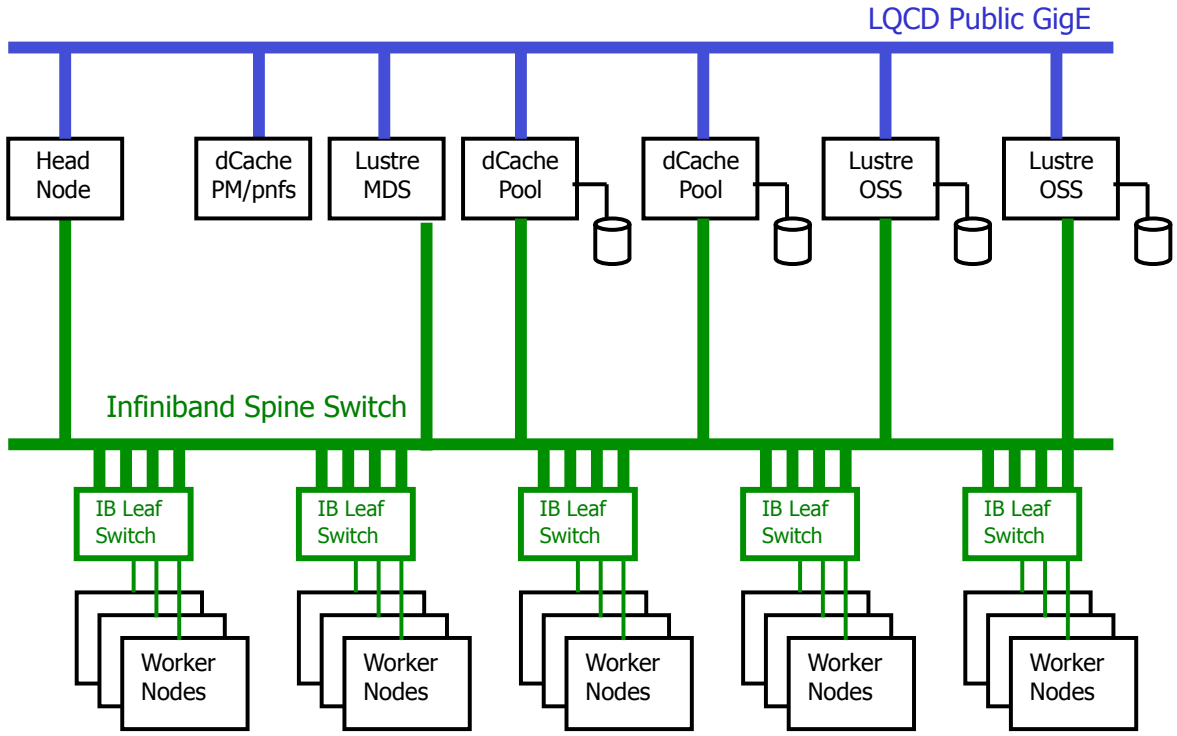
Cluster Layout – Ethernet and Mass Storage



Cluster Layout – Infiniband and Mass Storage

Infiniband

Architecture



Fermilab J/Psi

- Total cost: $1.89M
  - Includes nodes, Infiniband, ethernet, racks, all incidental cabling
  - 856 nodes, 8.4 TF/s sustained on LQCD code
  - 37.4 TF/s Top500, 57.5 TF/s peak
    - → $32.9K/peak TF
    - → $50.5K/Top500 TF
    - → $225K/LQCD Sustained TF

Fermilab Ds

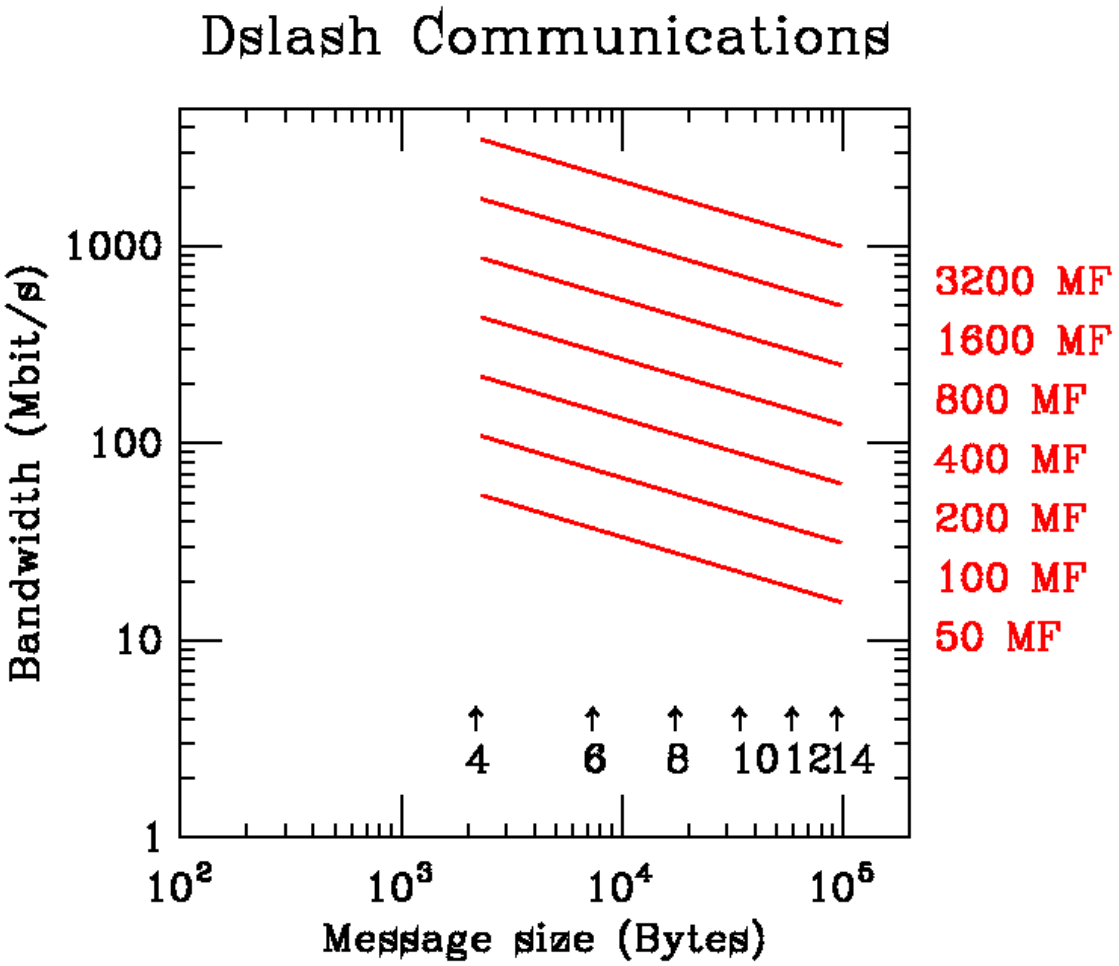
- Total cost: $1.43M
  - Includes nodes, Infiniband, ethernet, racks, all incidental cabling
  - 245 nodes, 12.5 TF/s sustained on LQCD code
  - 43.05 TF/s Top500, 61.44 TF/s peak
    - → $23.3K/peak TF
    - → $33.2K/Top500 TF
    - → $114K/LQCD Sustained TF

TJNAF 7n

- Total cost: $1.33M
- Includes nodes, Infiniband, ethernet, racks, all incidental cabling
  - 396 nodes, 2.98 TF/s sustained on LQCD code
  - 13.46 TF/s Top500, 24.1 TF/s peak
    - → $55K/peak TF
    - → $99K/Top500 TF
    - → $446K/LQCD Sustained TF

Balanced Design Requirements

Communications for Dslash



Modified for improved staggered from Steve Gottlieb's staggered model: physics.indiana.edu/~sg/pcnets/

Assume:

- L^4 lattice
- communications in 4 directions

Then:

- L implies message size to communicate a hyperplane
- Sustained MFlop/sec together with message size implies achieved communications bandwidth

Required network bandwidth increases as L decreases, and as sustained MFlop/sec increases

Communications Requirements