FY2014 Hardware Plan
and
Cluster Deployment Status

Don Holmgren

Fermilab

djholm@fnal.gov

SC LQCD-ext Annual Progress Review
Fermi National Accelerator Laboratory

May 15-16, 2014

Outline

- Overview of SC LQCD-ext acquisitions
- Computational requirements
- How we measure hardware performance
- Results of FY14 analysis of alternatives
- Status of FY14 cluster deployment

Overview of SC LQCD-ext Acquisitions

- Plan on approximately five acquisitions
  - Usually one per year in FY10-FY14, but some years have both conventional and GPU-accelerated cluster purchases
  - Every year also consider non-cluster hardware such as IBM BlueGene
- Guiding principle: procure the systems that will be the most effective for the planned science, given the current portfolio of machines operated by LQCD-ext
  - FY10/FY11 – we deployed a conventional cluster, purchased across the fiscal year boundary, and in FY11 we also purchased a GPU-accelerated cluster
  - FY12 – Combination of conventional and GPU-accelerated clusters
  - FY13 – Half-rack of BG/Q (BNL) and a conventional cluster (FNAL)
  - FY14 – Combination of conventional and GPU-accelerated clusters

Overview of SC LQCD-ext Acquisitions

Baseline LQCD-ext computational capacity deployment goals :

	FY2010	FY2011	FY2012	FY2013	FY2014
Computing hardware budget (excluding storage)	$1.60M	$1.69M	$1.875M	$2.46M	$2.26M
Planned capacity of new cluster deployments, TFlop/s	11	12	24	44	57

- Baseline computing hardware budgets are shown
- Budget adjustments:
  - Starting in FY2013, LQCD-ext began to operate LQCD-ARRA machines at JLab
  - FY2013 and FY2014 hardware budgets were reduced to increase the operating budget to accommodate the additional hardware
  - The budget allocated to storage was increased
  - In FY2014, some funds may be held over to FY2015
    - To cover operations if there the follow-on project is delayed or does not occur
    - To cover storage costs (to be determined by end of June)

Overview of SC LQCD-ext Acquisitions

Modified LQCD-ext computational capacity deployment goals:

	FY2010	FY2011	FY2012	FY2013	FY2014
Capacity of new cluster deployments, TFlop/s Planned/Revised/Achieved	11 / 12.5	12 / 9 / 9	24 / 10-15 12.8	44 / 15-22 34.6	57 / 22-33
Million “Fermi” GPU-Hrs/Yr Planned/Revised/Achieved	0	0 / 1.02 / 1.22	0 / 2.9-4.3 2.1	0 / 4.6-6.9 0	0 / 7.5-11.2

- FY2011 baseline plan for 12 Tflop/s was modified to 9 Tflop/s plus a GPU-accelerated cluster with “Fermi” GPUs
- FY2012-FY2014 revised goals reflect 40%-60% ranges in budget allocated to conventional and accelerated clusters. GPU capacity range was extrapolated from the FY2011 purchase using the observed Moore’s Law halving time for conventional hardware (this was much too optimistic)
- FY2013: project did not deploy GPUs, rather a BG/Q half-rack (21.9 TF) and a commodity cluster (12.7 TF)
- FY2014: ranges shown were estimated from baseline budget ($2.26M). Current EQ budget is $1.80M, and up to $150K additional funds will be used depending on FY2015 holdover. The ranges for $1.80M are conventional: 17.5 - 26 TF/s, GPU: 6 - 8.9M GPU-Hrs/yr

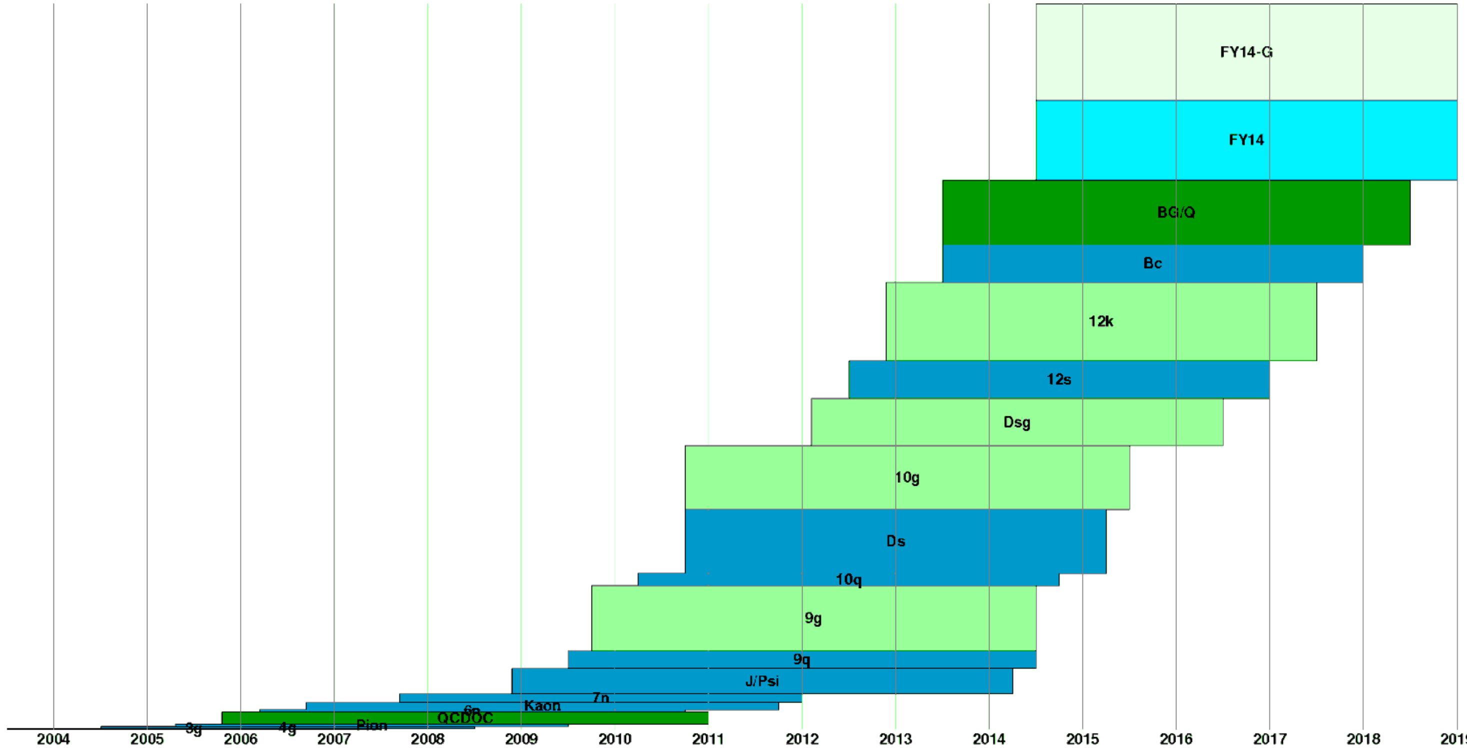
Computational Requirements

- 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

Computational Requirements

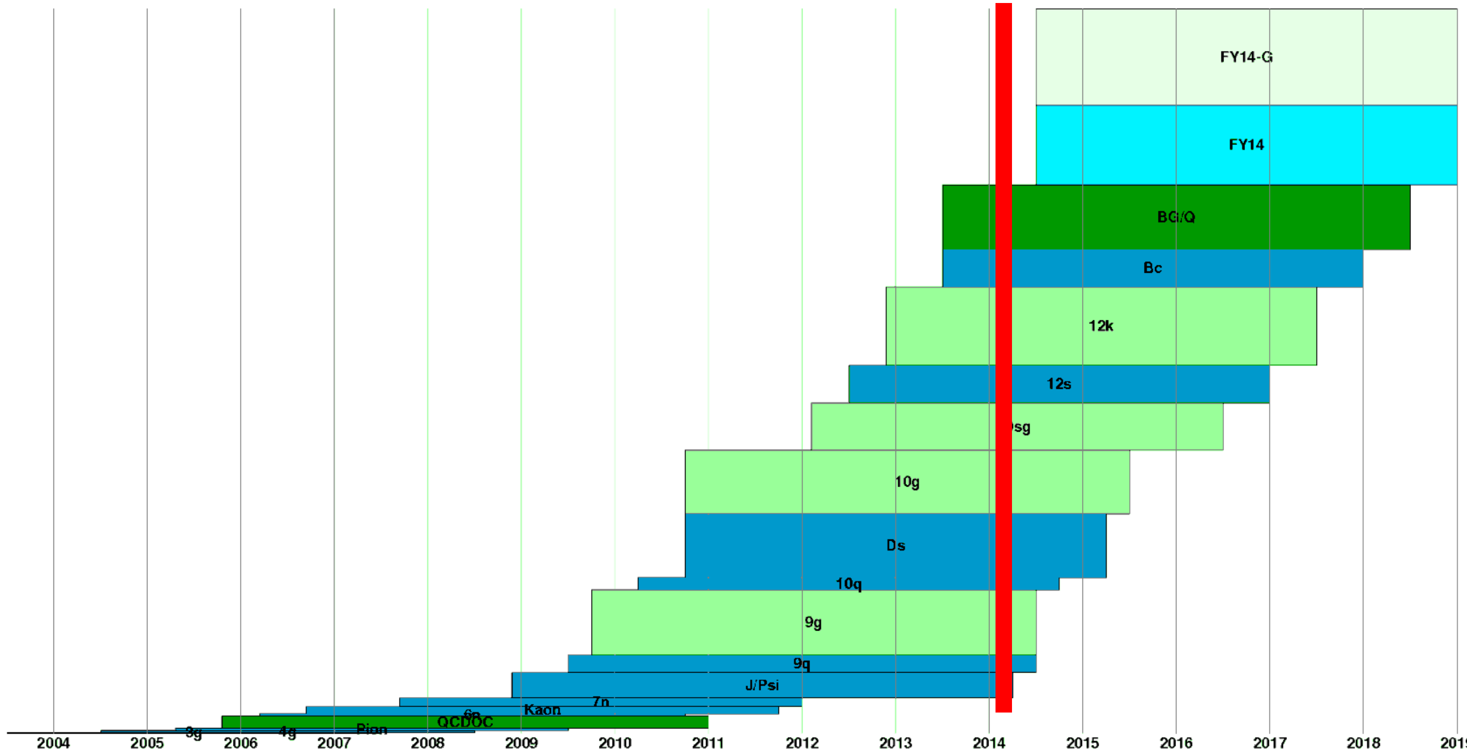
- We design and/or buy systems that as part of USQCD's entire portfolio (dedicated + LCF + other outside resources) will most effectively carry out the current and anticipated scientific programs
- This means:
  - Systems matched to the type and size of LQCD calculations that will be performed (*e.g.* mixture of conventional vs. accelerated)
  - Systems with the best price/performance for LQCD applications
  - Machines with the best memory bandwidth
  - High performance networks balanced to single node capacities and to anticipated job sizes

Dedicated Hardware Portfolio



- Blue = Conventional Cluster. Light Green = GPU Cluster. Dark Green = BlueGene Family. The height of each rectangle is proportional to the computing capacity in TFlop/sec
- Since 2006, the LQCD, LQCD-ext, and the proposed LQCD-ext II projects have operated / will operate an evolving portfolio of machines. Some of the LQCD-ext systems (Dsg, 10g, 12s, 12k, Bc, the BG/Q half-rack, and the FY14 conventional and GPU-accelerated clusters) will be operated by LQCD-ext II.
- The BG/Q rectangle is a half-rack at BNL, used for very large analysis tasks and for the generation of gauge configurations of small volume (too small to conform with LCF job-size policies)

Dedicated Hardware Portfolio



- In any year, a vertical line cuts through the resources that are operating at that time. Systems are operated from 3 to 5 years (while they remain cost effective). Today (red line), six clusters, four GPU-accelerated clusters, and one BG/Q resources are in production.
- Every year a new resources is brought online that takes advantage of decreasing price/performance and new technologies (could be a mixture of conventional and accelerated hardware)
- The project tailors annual purchases to meet the evolving needs and so can maximize cost effectiveness

Rating LQCD-ext Computing Facilities

- Definition of the sustained capacity of a conventional LQCD-ext computing cluster:
  - The performance of improved staggered (“asqtad”) and domain wall fermion (“DWF”) conjugate gradient inverters are measured using parallel jobs spanning a significant number of processors (currently 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 improved staggered are used, on clusters and leadership machines the asqtad-DWF average has been predictive of overall computing throughput
- GPU-accelerated clusters require a different rating methodology
  - For many jobs, execution times are not dominated by the inverter, and are influenced by sections of the code that must run on the host CPUs
  - Currently requests and allocations for GPU resources are in “Fermi GPU-hours”
    - The current NVIDIA GPU architecture, “Kepler”, has 1.3 to 2.0 times the throughput of “Fermi”, depending upon the mix of single and double precision
  - We often use an “effective TFlop/s” rating based on the walltime acceleration of a representative mix of LQCD jobs. Based on 2012 actual usage, a “Fermi” GPU is rated 140 effective TFlop/s, and a “Kepler” K20 is rated 210 effective TFlop/s

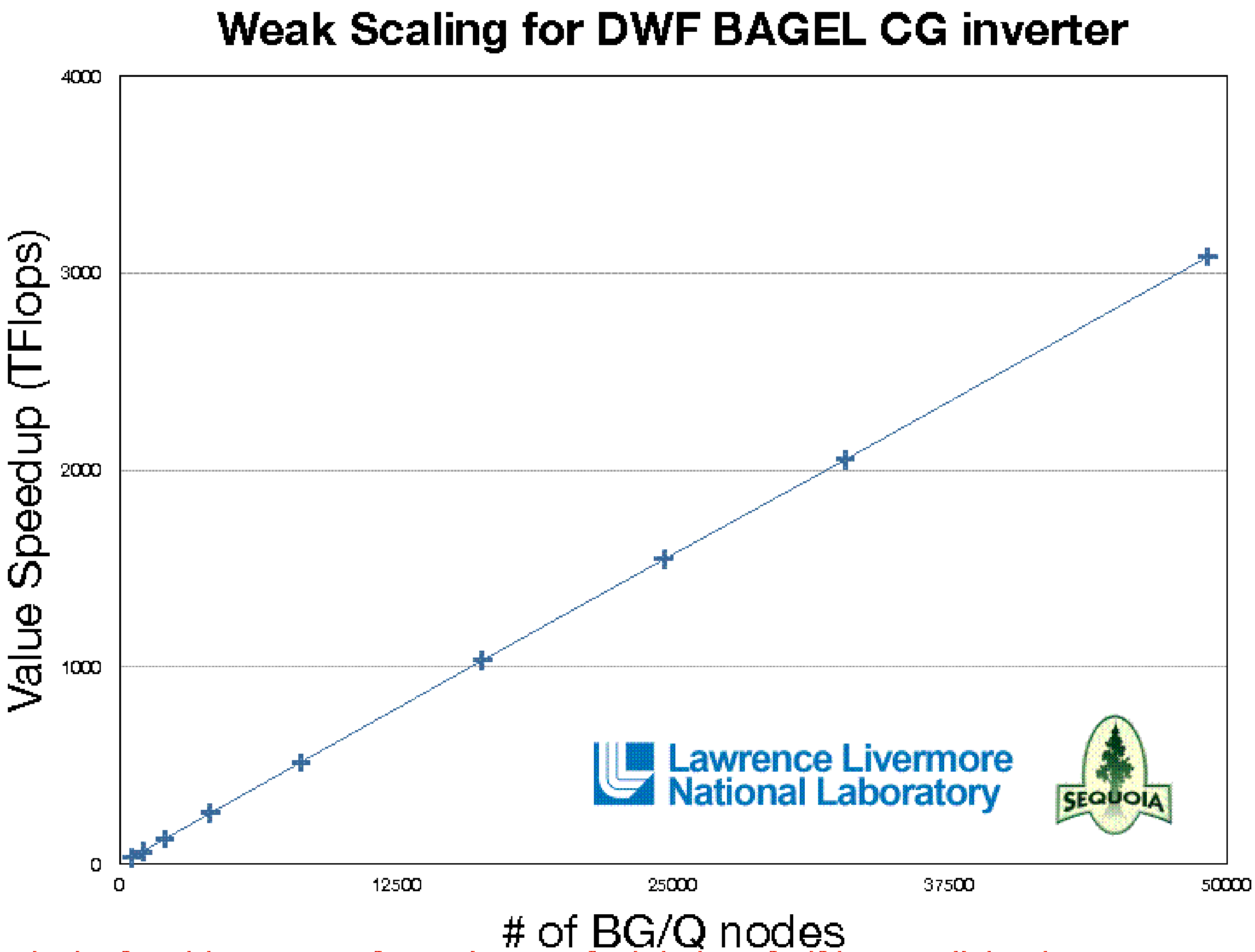
FY14 Alternatives Analysis

Available Alternatives

- For FY14 purchases, these hardware alternatives were identified:
  - A conventional Infiniband cluster based on Intel Sandy Bridge or Ivy Bridge, or AMD Abu Dhabi or Warsaw processors, at FNAL
  - A GPU-accelerated cluster based on NVIDIA Kepler GPUs, at FNAL
  - Expansion of the BlueGene/Q hardware deployed at BNL in FY13.
- Context
  - The deadline for the LQCD-ext project to determine how FY14 funds were to be distributed among the three labs was August 20, 2013
  - In FY13 BNL deployed a half-rack of BG/Q, and FNAL deployed a conventional cluster ("Bc"). An FY14 BG/Q purchase would fill out this half-rack.
  - In FY14, FNAL will retire a substantial (850 node, 8.4 TF/s) conventional cluster, "J/Psi", deployed in FY08/FY09. By node count, J/Psi comprised 38% of conventional cluster capacity (13% of conventional TF capacity).

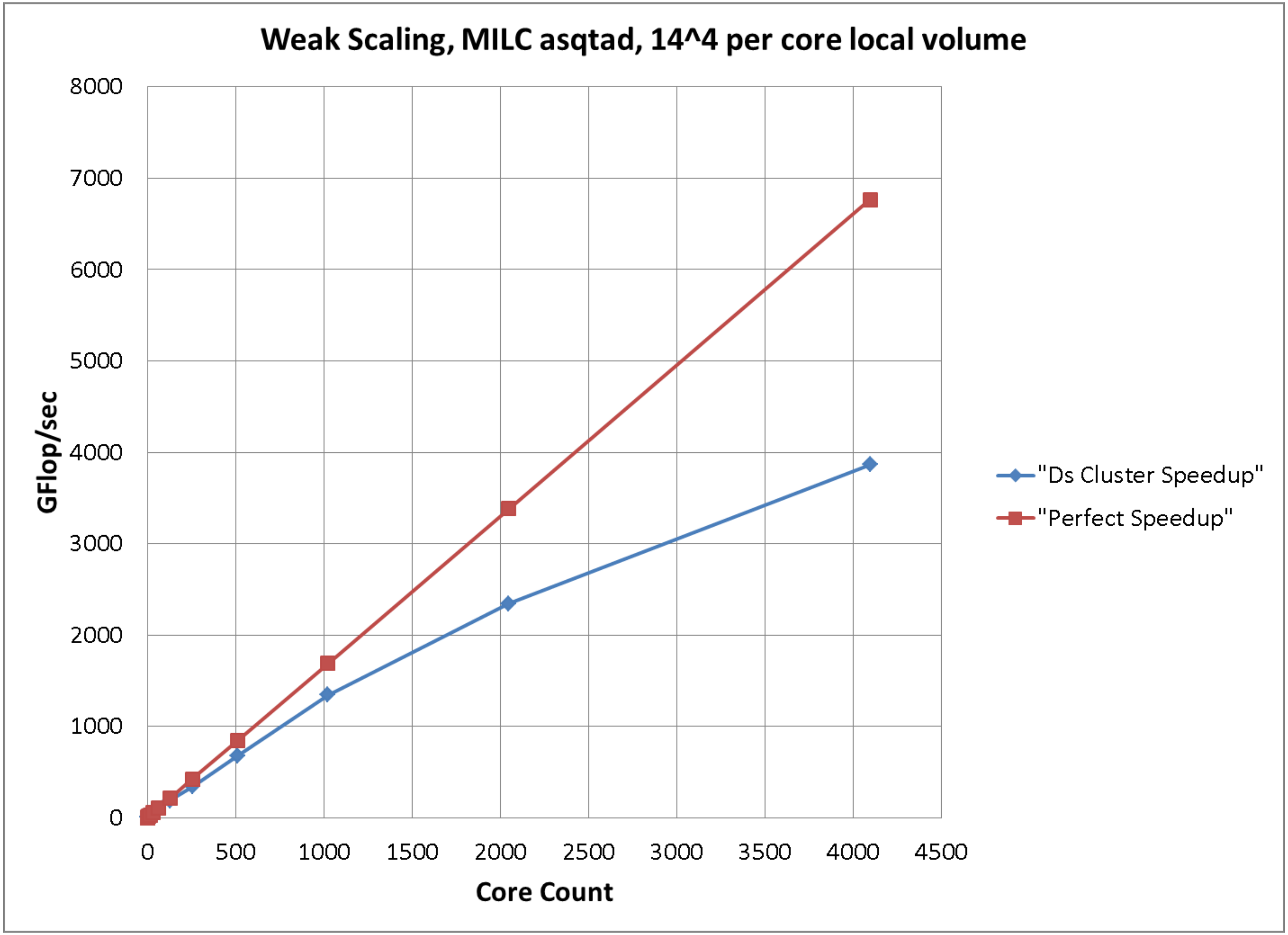
Scaling Considerations

- Weak scaling on BG/Q, as measured with DWF code at LLNL, is essentially perfect to 50K nodes



Code developed by Peter Boyle at the STFC funded DiRAC facility at Edinburgh

Conventional Cluster (Ds) Weak Scaling



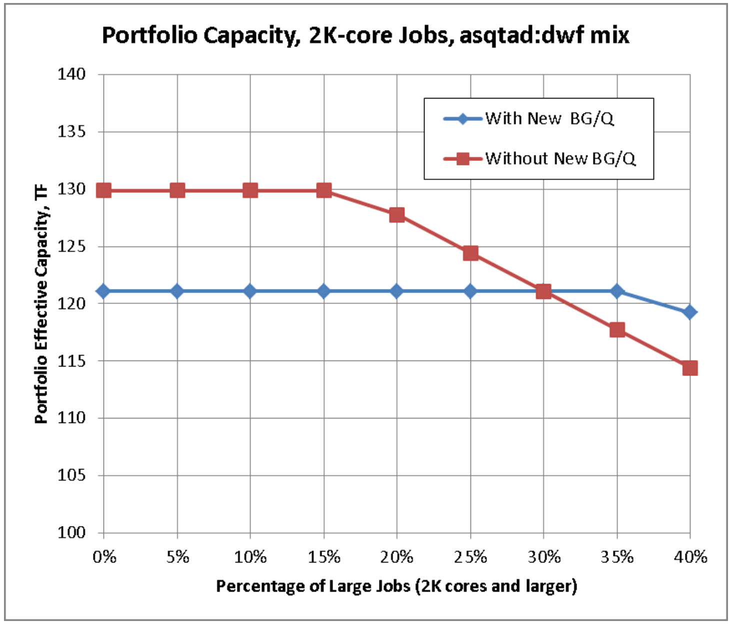
Scaling Considerations

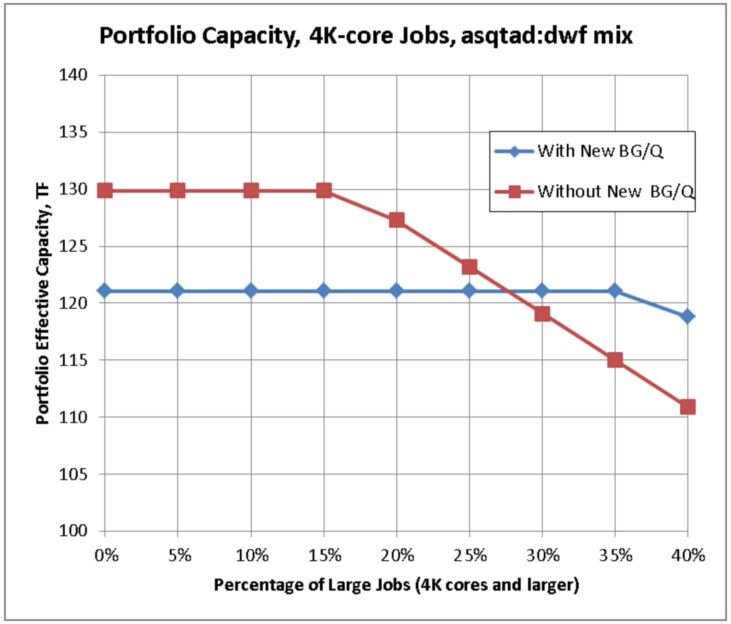
- On Infiniband clusters, weak scaling is not flat but rather performance per core drops as job core counts increase: (14^4 local volumes)
  - 128 cores: 1463 MF
  - 256 cores: 1334 MF (91.1% of 128-core performance)
  - 512 cores: 1327 MF (90.7%)
  - 1024 cores: 1314 MF (89.9%)
  - 2048 cores: 1144 MF (78.2%)
  - 4096 cores: 943 MF (64.4%)
- From Ds weak scaling data, we can estimate strong scaling performance per core, relative to a 128-core job with 14^4 per core local volume:
  - 128 cores: 0.19 TF aggregate
  - 1024 cores: 0.88 TF (59% per core relative to 128-core)
  - 2048 cores: 1.14 TF (38%)
  - 4096 cores: 1.44 TF (24%)
- GPU clusters exhibit scaling degradation similar to conventional clusters, although job node counts are typically much smaller

Alternatives Analysis

- In the FY14 LQCD-ext Alternatives Analysis document, we considered these scenarios:
  - A conventional Infiniband cluster at FNAL
  - A GPU-accelerated cluster at FNAL
  - A combination of conventional and GPU-accelerated clusters at FNAL
  - Expansion of the BG/Q half-rack at BNL to a full rack (1024 nodes, 16K cores) plus either a conventional or GPU-accelerated cluster at FNAL. Pricing and available budget dictated the size of any BG/Q expansion.
- The total computing capacity of the conventional LQCD-ext portfolio (existing machines + FY14 machines) was modeled as a function of expected fraction of jobs requiring at least 2K cores, assuming that all large jobs run on the BG/Q until that resource is exhausted

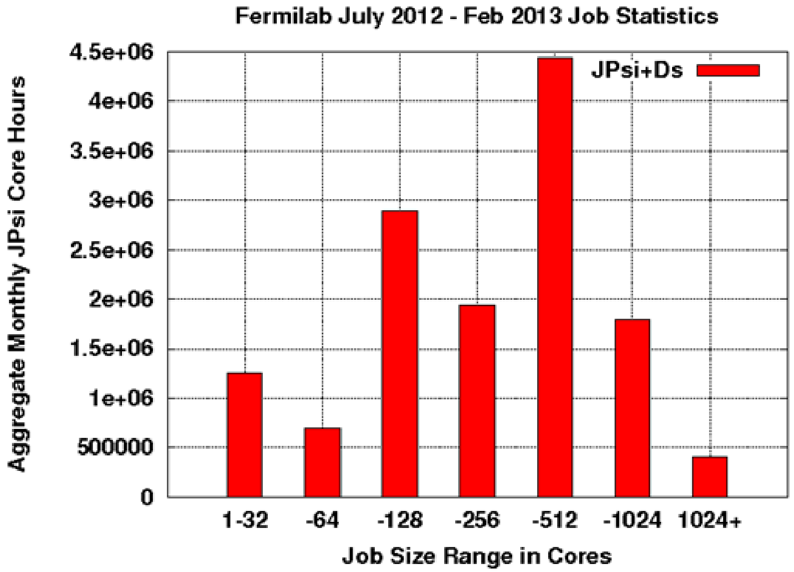
Computing Portfolio Capacity

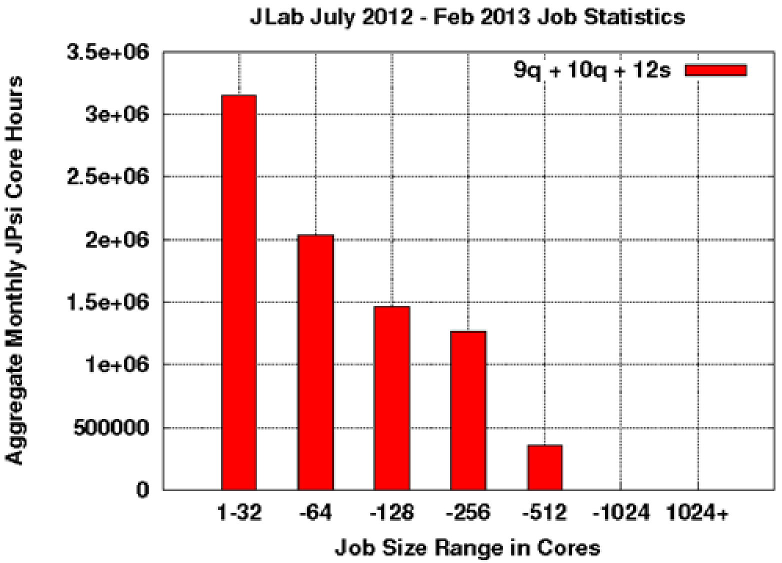




- We compared the non-GPU-accelerated portfolio capacity in two scenarios: BG/Q expansion plus a conventional cluster, or just a conventional cluster
- The overall conventional portfolio capacity is greater with BG/Q expansion if 27% of capacity is used by 4K-core or larger jobs, or 30% by 2K-core or larger jobs
- Both crossovers shift to the left for larger job sizes (8K)

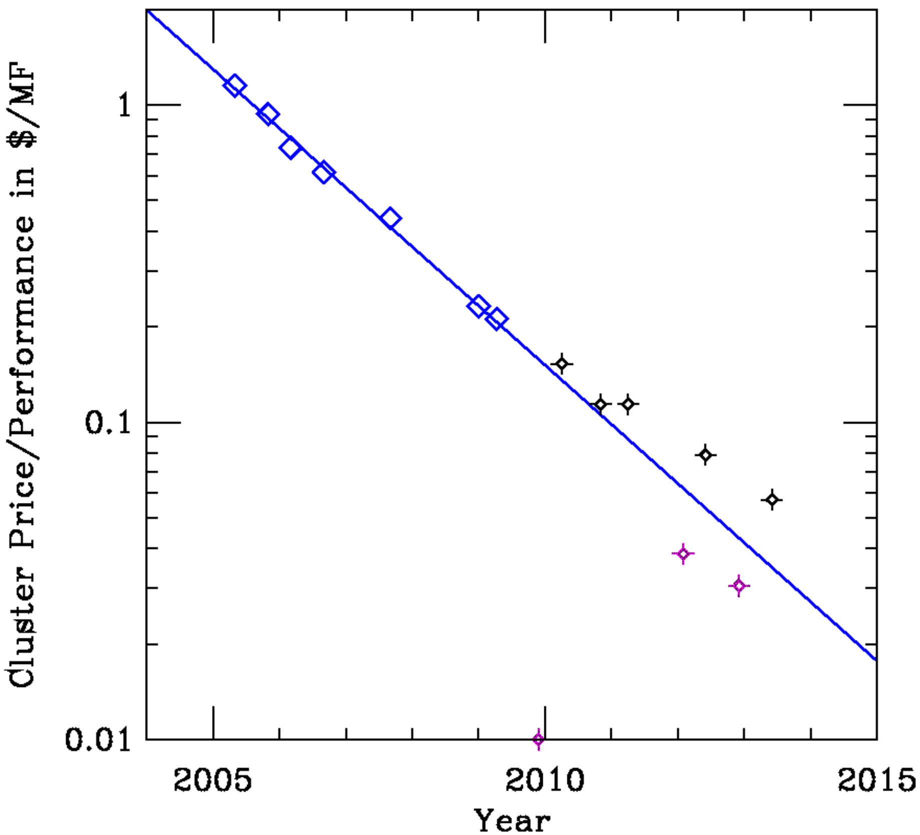
Job Statistics





- Based on actual running of jobs at FNAL and JLab during the 2013 allocation year, a negligible fraction of jobs required 2K or large core counts
- Based on allocation requests for 2013, conventional hardware was 156% oversubscribed, with BG/Q and GPU rates of 126% and 124%
- We concluded that large core-count demand was satisfied by the BG/Q half-rack plus access to LCF resources, and that the computing portfolio would be best optimized by non-BG/Q hardware

Mixed Conventional and GPU-Accelerated Deployment



Blue Diamonds: LQCD conventional clusters
Black Stars: LQCD-ext conventional clusters
Magenta Stars: GPU-accelerated clusters

- The cost effectiveness of the portion of USQCD analysis that runs on GPUs is considerably better than conventional cluster hardware
- However, much USQCD analysis uses workflows with some jobs (propagator generation) on GPUs and small “tie-up” jobs on those propagators using conventional nodes. A mixture of both types is needed.
- 2014 allocation requests have greater demand for conventional than GPU

Selection from the Alternatives

- Following the selection process approved at the May 2013 review, in consultation with the USQCD EC and the DOE HEP and NP Program Managers, LQCD-ext elected in Aug 2013 to deploy a combination of conventional and GPU-accelerated clusters at FNAL
- The decision concerning the relative budget fractions for conventional and GPU-accelerated portions was deferred until April
  - This allowed the project to wait until more information was known about the relative demand for the two resource types from the 2014 allocation requests
- Based on the relative demand, the project will deploy approximately 60% of the budget on conventional and 40% on GPU-accelerated hardware

FY14 Cluster Procurement, Deployment, and Estimated Performance

Design

- Based on on-going discussions with hardware manufacturers and system vendors, meetings at SuperComputing'13, responses to a Request for Information sent to 20 system vendors in Dec 2013, and past experience, the project wrote the specifications for the cluster design and issued a Request for Proposal in April 2014 once funds were available
- In the RFP, vendors were ask to bid a mixed conventional/accelerated system with the following constraints:
  - Conventional hosts to be AMD or Intel based, with power-of-two core counts
  - GPU hosts to be Intel based (PCIe Gen3 only available on such platforms)
  - 2, 4 or 8 GPUs per host, NVIDIA K20, K20x, or K40 GPUs
  - QDR or faster Infiniband, no higher than 2:1 oversubscription
  - Systems to be integrated into racks with power consumption per rack of 13-14 KW
  - After an initial minimum purchase (10 TF conventional, 29 TF accelerated) vendors must give pricing for options in FY14 and FY15 that could each add as much as 10 TF conventional and/or 29 TF accelerated

FY14 Procurement Schedule
(FY14 Cluster Acquisition Plan/Achieved)

2013

- Aug-Dec (completed Dec) – benchmark processor alternatives
- Oct 1 – Continuing Budget Resolutions begin (after Oct 1-16 shutdown)
- Mid-Nov (Dec 2) – Release Request for Information
- Late-Dec (Apr 11) – Release Request for Proposal

2014

- Jan 15 – Omnibus Appropriations Act of 2014 passed
- Mar 5 – Award Purchase Order – goal: May 28
- March 10 – FY14 financial plan in place at FNAL and budget codes available for requisition
- Apr 30 – Hardware received and integrated – estimate July 21
- Early May – Friendly user period begins – estimate Aug 18
- May 9 – Vendor bids received
- Mid-to-late Jun – Release to production – estimate Sep 15

Cluster Details

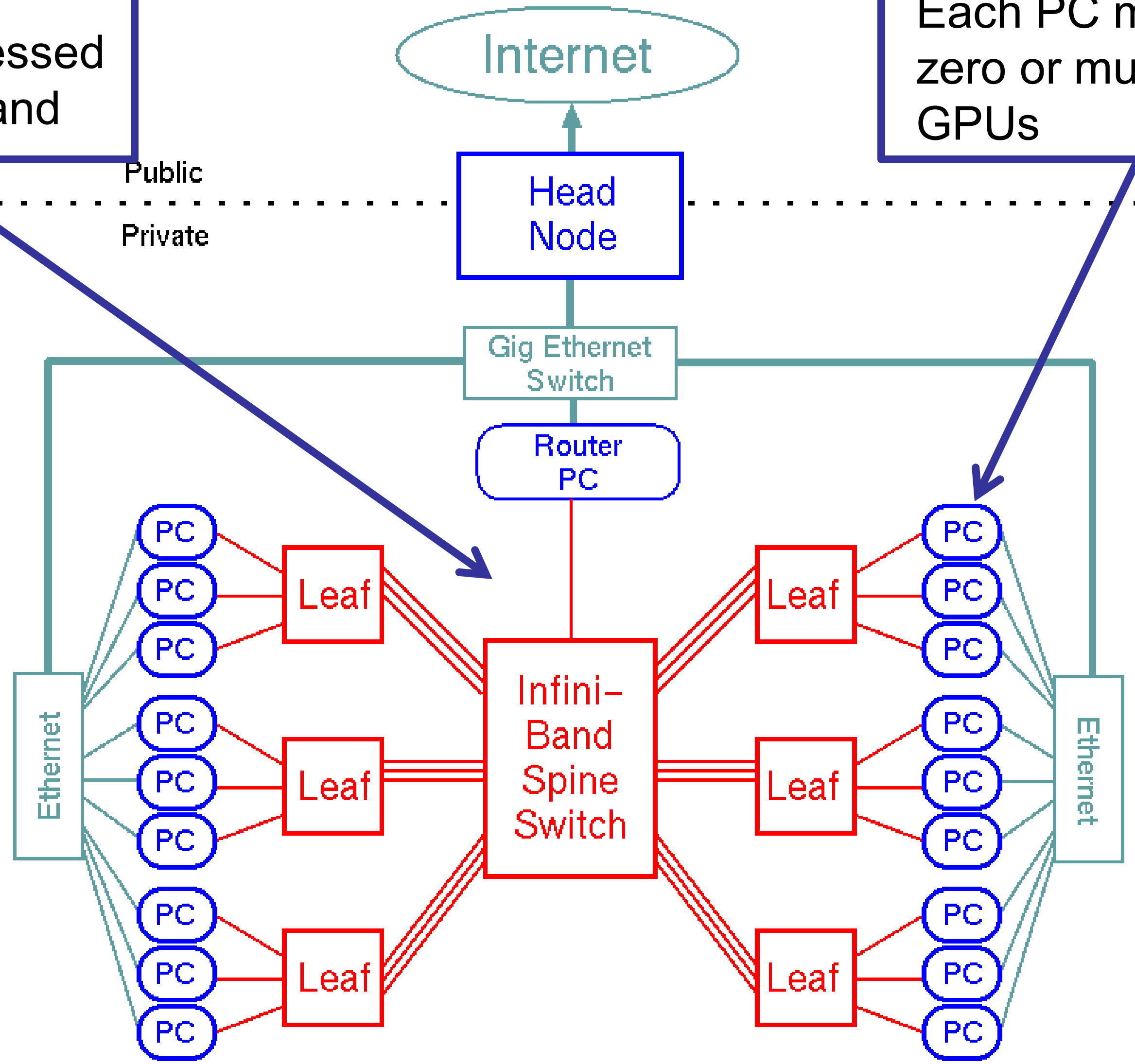
- Award will be to the best value bid, based on price, LQCD application performance, power cost, space efficiency, and aspects of the design of the GPU portion (GPUs per host, Infiniband configuration)
- Six vendors responded with 8 configurations based on AMD Abu Dhabi and Intel Ivy Bridge processors, and on NVIDIA K20x and K40 GPUs
- Bids are under evaluation now
- Based on an early reading of all proposals, estimated performance and costs as follows (Intel Ivy Bridge + NVIDIA K40):
  - Single conventional node: 67 Gflop/node (based on 128-core runs)
  - Single quad GPU node: 1040 Gflop/node (based on dual GPU runs)
  - Estimate 20.4 TF ($0.053/Mflop) and 2.2M GPU-hrs/yr (31.3 effective TF)
  - Target range for $1.80M was 17.5 - 26 TF/s, 6 - 8.9M GPU-Hrs/yr (GPU extrapolations from 2011 price/performance were too optimistic, and Moore's Law for conventional continues to slow)

Questions?

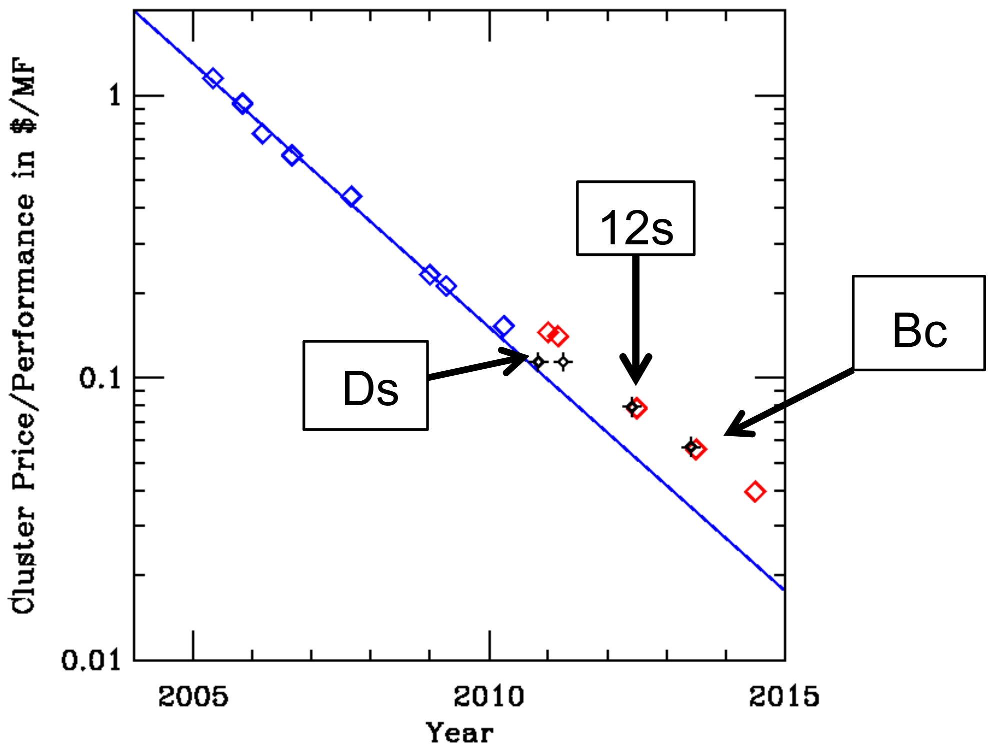
Typical LQCD Cluster Layout

Lustre parallel filesystem accessed through Infiniband

Each PC may contain zero or multiple (2-4) GPUs



Cost and Performance Basis



Cluster	Price per Node	Performance/Node, MF	Price/Performance
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
12s	$3972	50118	$0.079/MF
Bc	$3219	*56281 est.*	*$0.057/MF est.*