LQCD ARRA Computing Project

Overview, Management & Performance

Chip Watson

May 10, 2011

ARRA LQCD Project Definition

The LQCD ARRA Computing Project directly supports the mission of the DOE's Nuclear Physics Program "to foster fundamental research in nuclear physics that will provide new insights and advance our knowledge on the nature of matter and energy...".

The Project also supports the Scientific Strategic Goal within the DOE Strategic Plan to "Provide world-class scientific research capacity needed to: advance the frontiers of knowledge in physical sciences... [and] provide world-class research facilities for the Nation's science enterprise."

The project scope, management structure, and milestones are defined in the Project Execution Plan, a 17 page document submitted after award, and amended to incorporate the evolution in the plans to exploit GPUs.

The technical scope and computing architecture will be addressed in the next talk.

Management, schedule and cost, change control, and operational context (safety and cyber security) will be covered in this presentation.

ARRA Project Context

The LQCD ARRA project is complementary to the LQCD-ext project

- In 2009 USQCD collaboration requested $24M for "LQCD 2" (5 years)
  - LQCD-ext was funded at $18M
  - LQCD ARRA was funded at $5M, enabling funding of nearly the target amount
- Project co-ordination:
  - Jefferson Lab was to have received the next LQCD cluster in FY 2010
  - A collective decision was made to put LQCD ARRA resource at Jefferson Lab, and to re-locate the FY2010 LQCD-ext machine to Fermilab, shifting it later in the year to create the possibility of a combined FY2010-11 larger machine

The LQCD SciDAC project provides the necessary software for both of these computing projects

INCITE and NSF allocations address capability computing

The ARRA & LQCD-ext projects target ***high end capacity*** (many jobs < 1 Tflops)

Project Goals & Phasing

Performance goal: to double USQCD's resources, at that time 17 Tflops.

As an ARRA project, another goal was to move as quickly as possible to get funding into the economy.

- The project was structured as 2 phases
  - Phase 1, $1.78M in hardware to be awarded by the end of FY2009
  - Phase 2, $1.70M in hardware, to follow by ~3 months
- Evolution to include GPUs
  - By the time the project started, it was clear GPUs would be ready for exploitation by LQCD, enabling a significant performance increase
  - Each phase was adjusted to include a GPU component, so that deployment could match the community's uptake of the GPU
  - Phase 2 was eventually adjusted to accommodate the timeline for the new NVIDIA Fermi GPU

As a result of the GPUs, the final delivered performance was effectively 72 Tflops!

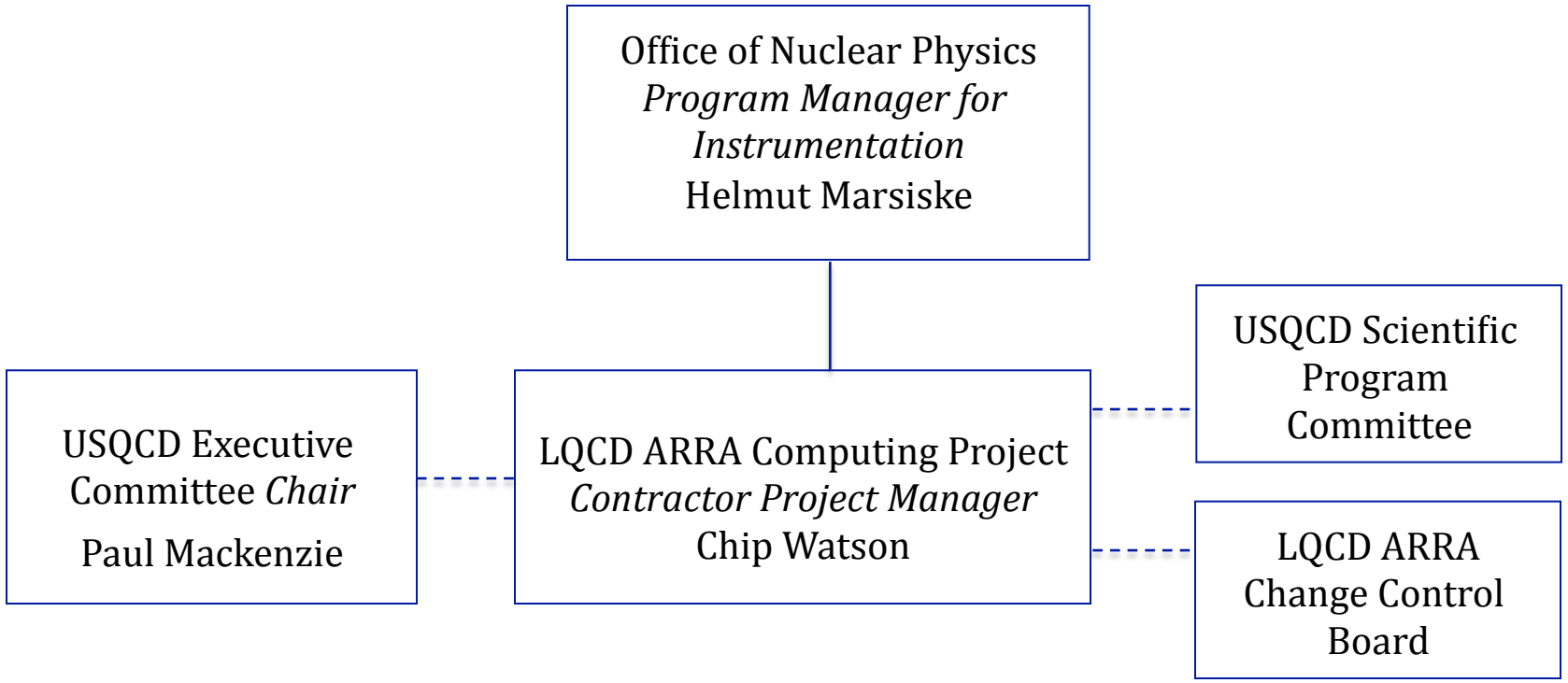
Management Strategy

The ARRA LQCD project is in many ways modeled after the LQCD Computing project, and re-uses the following components or management approaches:

- Relationship to the USQCD Executive Committee
- Relationship to the Scientific Advisory Committee (computing allocations)
- Approach to hardware selection and alternatives analysis, to achieve the greatest performance for dollars invested
- Approach to benchmarking
- Cost model for operations (FTE planning)
- Change control process (but simplified since there is only one site)

Because of the lower total project cost, a single site, and fewer deployment cycles, management is intentionally lighter weight.

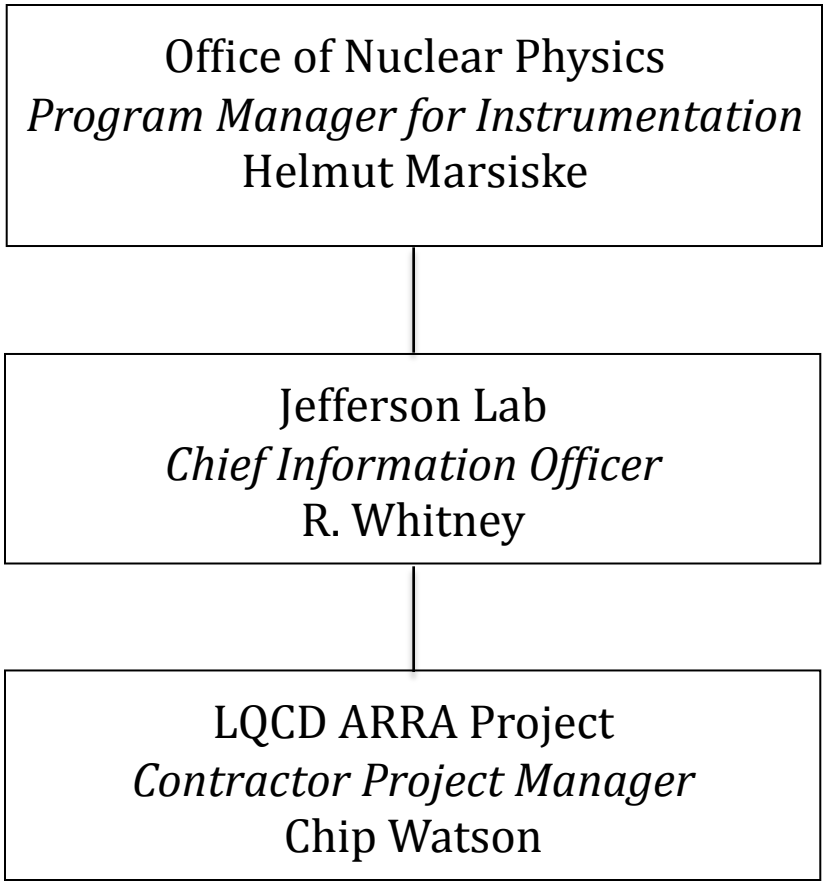
Management Organization



Management Organization Chart for the LQCD Computing Project.

Vertical lines indicate reporting relationships. Horizontal lines indicate advisory relationships

Management Authorities

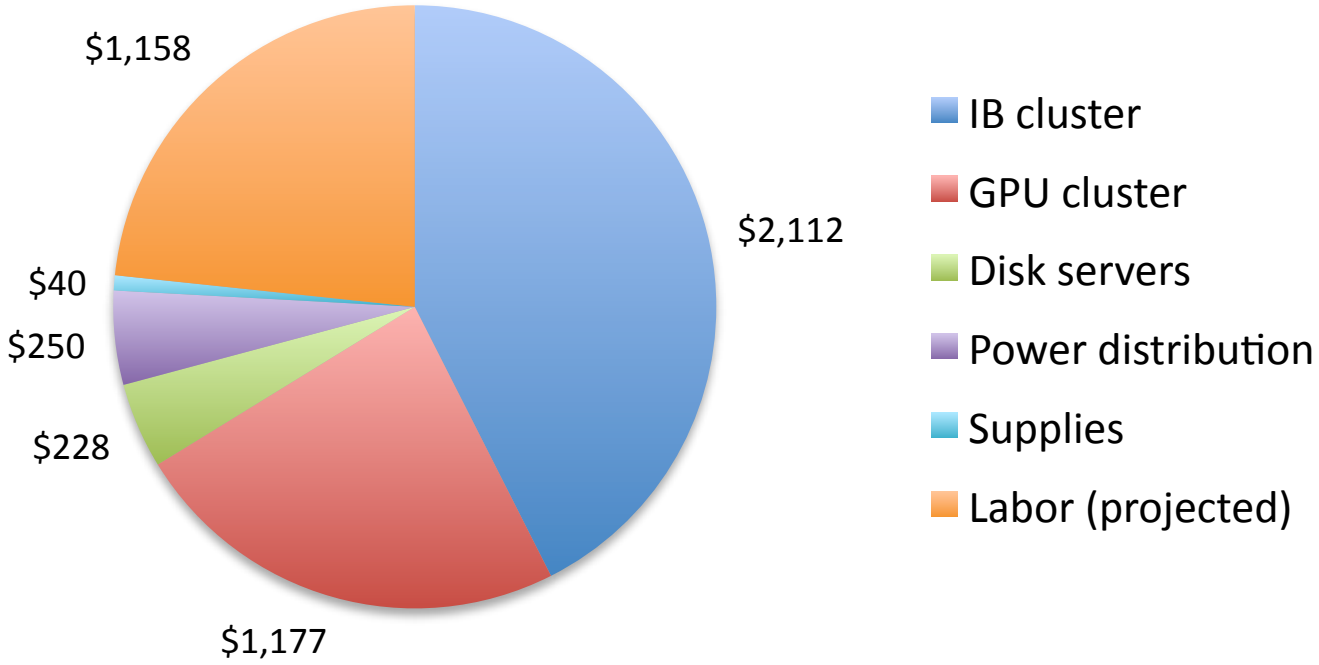


Management Authorities Chart

Budget Overview

- 70% of the funds for hardware (Infiniband & GPU clusters, disk servers)
- 5% for power conditioning and distribution
- 25% for labor

(Numbers show are actual costs, with labor projected to end of project)



Work Organization and Budget

For budgeting purposes, a simple WBS was adopted (data from FY2011 Q2):

WBS	Item	Baseline Total Cost (AY$)	Costed & Committed (AY$)	Estimate To Complete (AY$)	Estimated Total Cost (AY$)	Baseline Contingency (AY$)	Remaining Contingency (AY$)
1.01	FY09 Mgmt	26	26	-	26	0	0
1.02	FY10 Mgmt	25	25	-	25	1	0
1.03	FY11 Mgmt	15	12	8	20	1	1
1.04	FY12 Mgmt	15	-	14	14	1	1
1.05	FY13 Mgmt	16	-	15	15	1	1
2.01	Site Prep	250	250	-	250	16	0
2.02	Phase 1	1,970	1,967	-	1,967	0	0
2.03	Phase 2	1,816	1,811	5	1,816	0	0
3.01	FY10 Operations	121	145	-	145	6	0
3.02	FY11 Operations	228	162	112	278	11	10
3.03	FY12 Operations	218	-	200	200	11	10
3.04	FY13 Operations	207	-	180	180	10	10
	Totals	4,907	4,398	522	4,932	58	33

Contingencies on the Phase 1 and 2 deployments are zero as they are build to cost systems.

Milestones

Formal project milestones are focused upon deployment of resources:

Milestones	Date
Project Start	6/10
Issue Request for Proposal (RFP) for disk and compute clusters	8/10
Place order for disk cluster and first compute cluster	9/10
Begin early use on first cluster	12/09
Production running on first cluster	1/10
Place order for Phase 2a Infiniband cluster	1/10
Begin early use of Phase 2a Infiniband cluster	4/10
Place order for Phase 2b next generation GPU cluster	4/10
Production running on Phase 2a Infiniband cluster	5/10
Complete Annual Peer Review of LQCD program (June of each year)	6/10
Begin early use of Phase 2b next generation GPU cluster	8/10
Production running on all resources	9/10

All of these milestones have been met.

Staffing

Since the project is embedded in the Scientific Computing group, it is able to fund fractional people out of a larger pool of people.

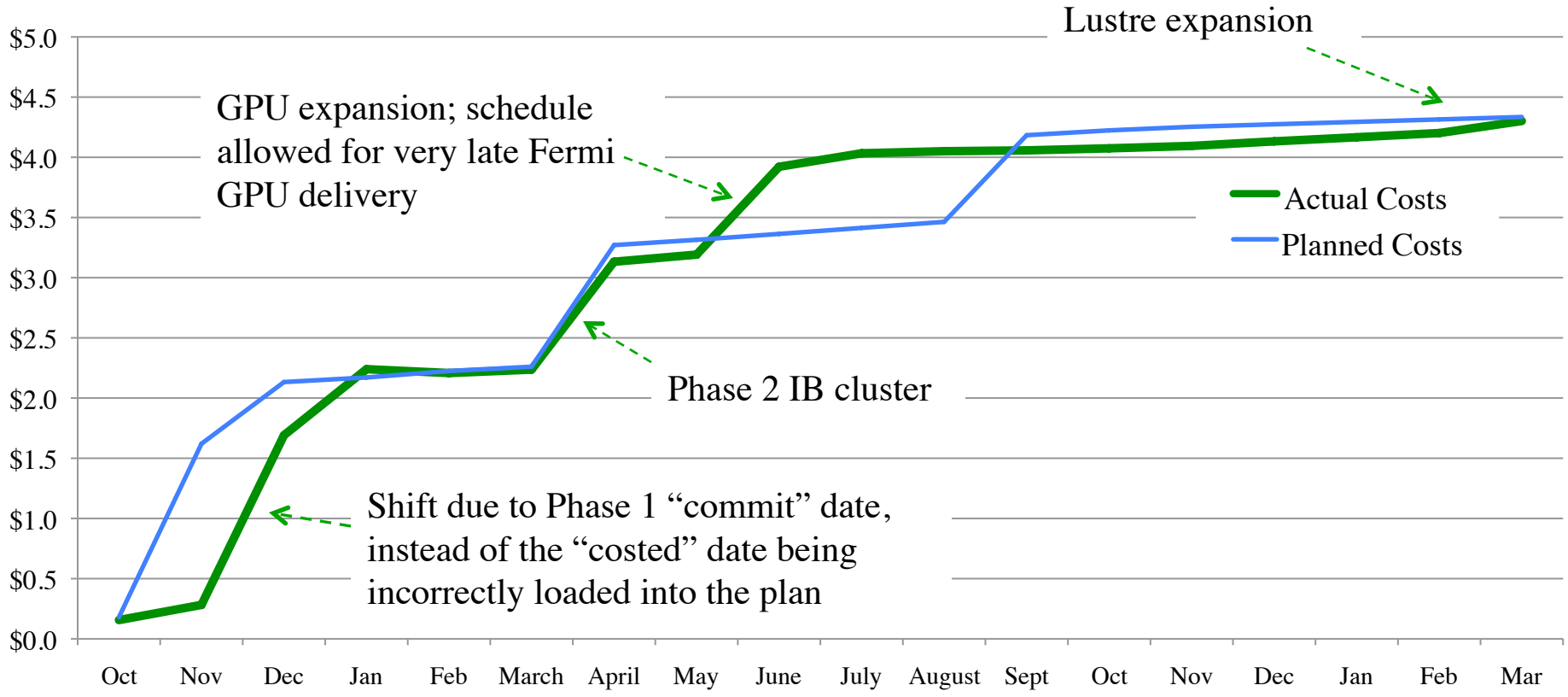
LEVEL OF EFFORT
(FTE-yrs)

	FY09	FY10	FY11	FY12	FY13
Project Management	0.10	0.09	0.05	0.05	0.05
Site Management		0.05	0.05	0.05	0.05
Steady-state Operations Support	-	0.60	1.20	1.10	1.00
Deployment Planning	0.08	0.10	-	-	-
Deployment Support	0.16	1.20	-	-	-
Project Management	-	-	-	-	-
Total	0.34	2.04	1.30	1.20	1.10

Costs through March 2011 ($K)

FY09 Budget	FY09 Actual[1]	FY10 Budget[2]	FY10 Actual	FY11 Budget	FY11 Available	FY11 Committed	FY11 % Committed
$2,246	$55	$1,962	$3,980	$254	$427	$153	36%

[1] Does not include $1,890 committed for computers and power distribution
[2] After Dec 2009 milestone & level 2 budget changes



Managing Disruptive Technology

The most difficult management challenge was the process for deciding how quickly to deploy and expand the GPU resources.

Key elements to guide the Phase 1 decision included early prototyping work and interactions with users, including the GPUs for Lattice QCD Workshop, held at Jefferson Lab August 19-21, 2009.

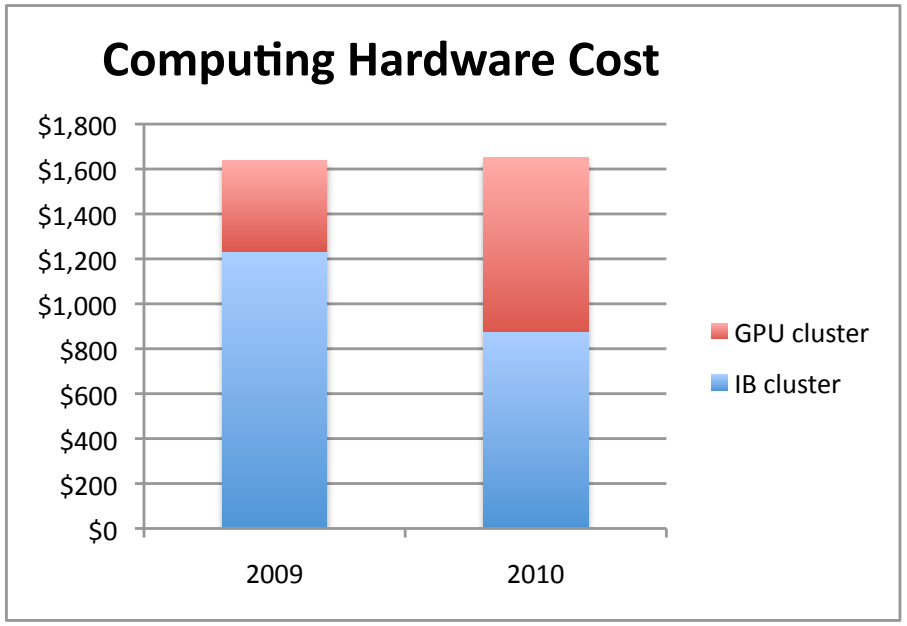
This adaptive process was again followed in March 2010 to decide the split between GPUs and conventional clusters for the Phase 2 procurement, and resulted in a large fraction going to GPUs.

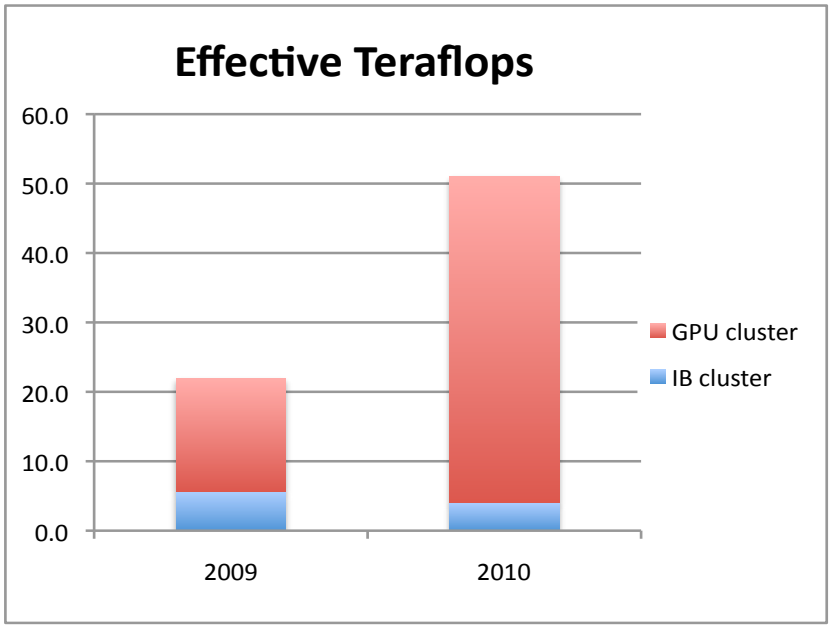
Change Control

Level	Cost	Schedule	Technical Scope
DOE Program Manager (Level 0)		> 1-month delay of a Level 1 milestone date	Change of any WBS element that could adversely affect project performance specifications
LQCD ARRA CCB (Level 1)	A cumulative increase of more than $200K in WBS Level 2	> 1-month delay of a Level 1 milestone date	Any deviation from technical deliverables that does not affect expected project performance specifications.
LQCD ARRA Contractor Project Manager (Level 2)	Any increase of > $50K in the WBS Level 2	> 1-month delay of a Level 2 milestone date	Technical design changes that do not impact technical deliverables.

To purchase the NVIDIA Fermi GPUs in FY 2010 (Phase 2) required a schedule change of > 3 months, exercising this change control process once. That also allowed GPU software to further mature, enabling a larger GPU procurement for this phase.

CPU / GPU Split





Doubling the fraction going into GPUs in 2010 allowed us to increase the total deployed capacity by more than 50%.

The GPUs yield ~10x as much computing power per dollar compared to standard IB clusters, but cannot run all of the users' software. Consequently the rate of deployment must match evolving software readiness (additional details in the next talk).

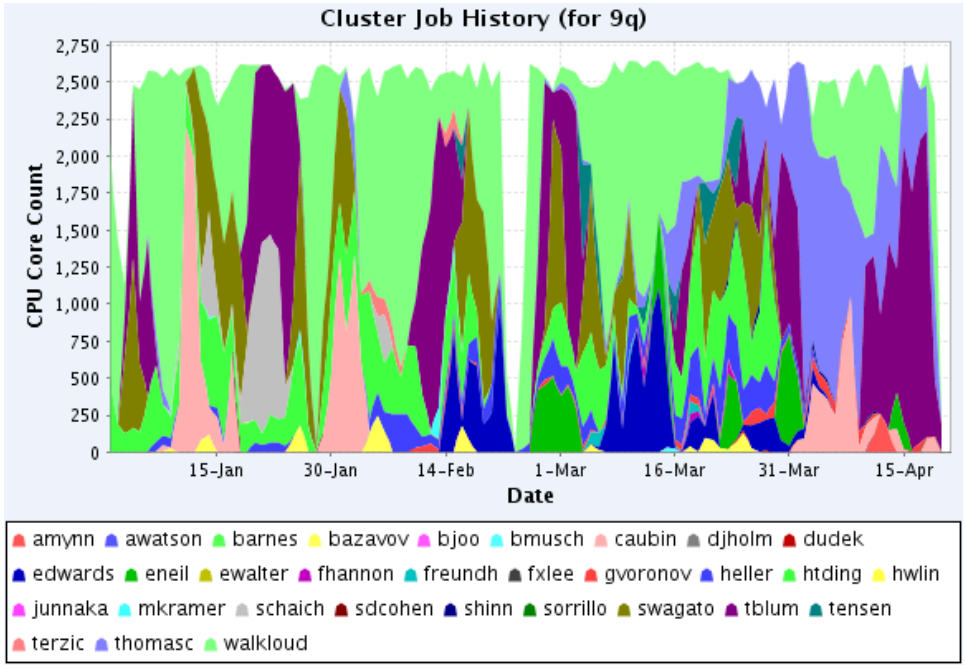
Infiniband Running Status

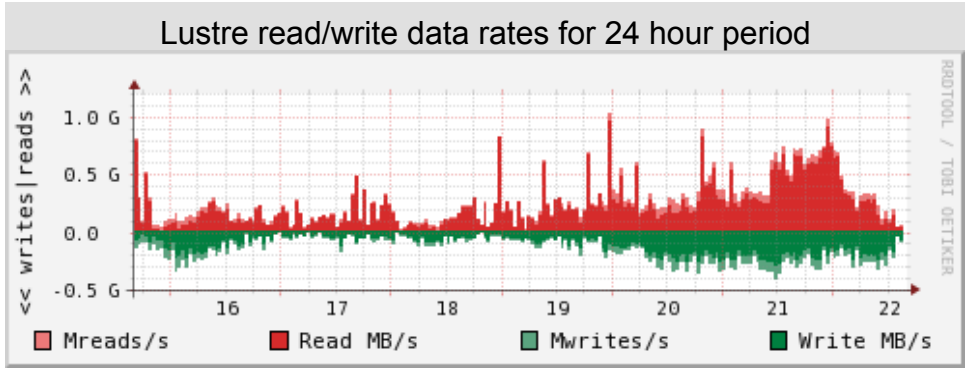
Infiniband Cluster:

Cluster runs at > 99% up except for a construction related power or cooling disruptions, or major upgrades (dip in February was to upgrade the Lustre Meta Data Server)

File Servers:

- 410 Terabytes on 18 servers
- Currently at about 60% of capacity due to a recent 80% capacity increase (growing slowly)
- Read/write spikes of up to 2 GB/s are observed, with plenty of headroom



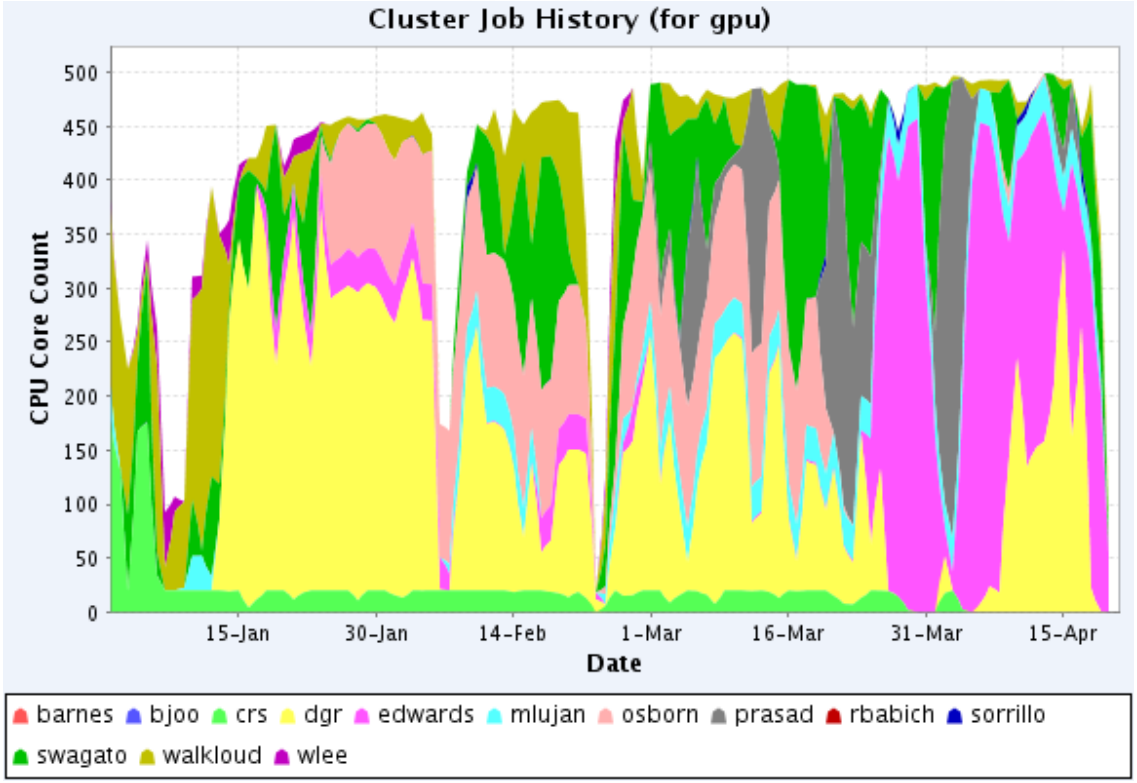


Running Status: GPU Nodes

GPU Cluster

- Multiple groups are in production, and the system is usually running at >99% utilization and uptime
- The growing capacity is the result of bringing online additional GTX-480 cards, as poor cards were replaced and re-tested (the early dip 2011 was a full re-testing of all cards); additional details in next talk

Additional operational charts and graphs can be viewed at http://lqcd.jlab.org/



Current Status (Highlights)

Hardware

98% of hardware is in production use

Last 2%, originally held as contingency, was spent on dual use R&D / production hardware, and is now being commissioned

Software (SciDAC)

Software for the GPUs has progressed remarkably, gaining performance and new capability every 6 months

Budget

On track. Project was build to cost, so budget risk was small.

Milestones

All hardware milestones have been met.

Remaining operations milestones are all on track.

GPUs: A Large Capacity Resource

504 GPUs in production at Jefferson Lab (May 2011)

- ★ 200,000 cores (1,600 million core hours / year)
- ★ 600 Tflops peak single precision
- ★ **100 Tflops aggregate sustained performance** in the Clover inverter (mixed half / single precision)
- ★ 63 Tflops "effective cluster performance", as measured by production application clock time acceleration compared to non-GPU clusters, weighted by actual usage (correctly accounts for Amdahl's Law)
- ★ Significant increase in dedicated USQCD resources

Additional details in this afternoon's technical talk.

Overview Summary

The LQCD ARRA project completed its deployment phase on budget and on schedule, and is now in stable operations.

The project achieved a 4.5-fold increase in total delivered Tflops capacity computing (63+9 Tflops vs 16 Tflops) by moving aggressively to exploit GPUs.