

Response to Technical Recommendations from the 2008 Review

Bill Boroski
Contractor Project Manager

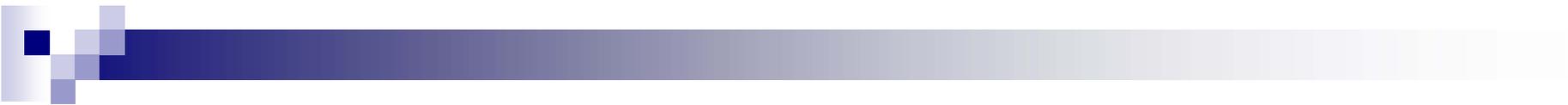
2009 Annual DOE Progress Review
SC LQCD Computing Project
Fermi National Accelerator Laboratory
June 4-5, 2009



Overview

- The 2008 progress review resulted in 14 recommendations
 - 10 recommendations were associated with the scientific program and were addressed by Paul Mackenzie, Chair of the USQCD Executive Committee
 - 4 recommendations were associated with technical and/or user aspects of the computing project, and will be addressed in this presentation.

- Recommendations were not uniquely numbered in the review report, so we've adopted the following numbering scheme when tracking resolution.
 - SectionNumber.RecommendationNumber



Recommendation 4.1:

The schedule contingency and risk associated with the uncertainty in the availability of the Nehalem technology should be clarified.

Response:

The 2008 progress review was held May 13-14, 2008, at which time there was significant uncertainty in the availability of the Intel Nehalem technology for the FY08/09 procurement. This was noted during the review and documented in our acquisition strategy.

On July 11, an RFP was issued for an integrated Infiniband-based cluster with 4.2 Tflops computing capacity (the FY08 “base” purchase). The RFP also requested pricing for options to buy up to 15 additional server racks plus required network equipment (the FY09 “options” purchase), valid through March 31, 2009.

The RFP did not specify processor type. Rather, it allowed proposals specifying either Intel or AMD processors.

(continued on next slide...)



Response to Recommendation 4.1 (cont'd):

Vendor responses were received by COB on August 11, 2008.

A total of six bids were received, none of which chose to propose an Intel solution. All vendors chose to propose Opteron-based systems.

All proposals were evaluated by a committee using technical criteria and cost, and scored via figures of merit assigned for various parameters such as price/performance on LQCD codes, normalized power consumption and footprint, etc.

In summary...

By specifying performance and schedule requirements in the RFP (as opposed to processor technology), selecting the winning proposal through a “Best Value” selection process, and specifying an option clause in the purchase order, we maintained the flexibility to take advantage of the Nehalem technology if it became available in a timely manner, without incurring risks associated with Nehalem production release schedules.



Recommendation 5.1:

LQCD should determine the dollar amounts of this budget change in Finding 5.3 and it should be presented to the USQCD executive board for approval. The funds would have to be taken from the project's hardware acquisition budget, and LQCD should verify their claim that the findings change will not seriously compromise the hardware performance of the planned cluster.

Response:

Finding 5.3 noted that we stated that an additional 0.65 FTE of systems admin support was needed at both Fermilab and JLab to better support operations.

Before reducing the hardware budget, we reviewed staffing needs at all three sites across all personnel categories (e.g., site mgmt, sys admin) as part of our FY09 budget planning process.

We identified several areas in which we were able to adjust the level of support for various personnel categories, to better meet staffing needs, while staying within the baseline budget envelope.

(continued on next slide...)



Response to Recommendation 5.1 (cont'd):

Adjustments in Staffing Profile:

Reduced site management from 0.25 to 0.15 FTE at all sites ($\Delta = -0.30$ FTE)

Reduced sys admin support at BNL from 0.75 to 0.25 FTE ($\Delta = -0.5$ FTE)

Reduced deployment support at FNAL from 0.75 to 0.50 FTE ($\Delta = -0.25$ FTE)

Increased level of operations sys admin support at FNAL and JLab from 1.1 to 1.9 FTE per site ($\Delta = 0.8$ FTE/site, or 1.6 FTE total)

Net result of all adjustments = increase of 0.55 FTE in salary support.

However, based on differences in fully-loaded salary costs for the various positions at the various sites (e.g., site mgmt vs. sys admin), the corresponding cost increase of the effort adjustment was only \$8K (from \$896K to \$904K).

Therefore, we were able to increase the level of sys admin support in FY09 to better meet project needs without reducing the hardware budget, and without negatively impacting deployment schedules or operating performance.



Recommendation 5.2:

The informal contributions of power and space that the labs make to LQCD should be tracked quantitatively and, if necessary to ensure stability of the project, should be formalized through the amendment of the present MOUs.

Response:

Formal signed MOUs are in place between the LQCD Project and each host laboratory. The MOUs define anticipated in-kind contributions.

Excerpt from Fermilab MOU (note that all MOUs contain similar language):

7.3. Facilities and Equipment

Adequate facility infrastructure will be made available to the LQCD project to carry out the implementation and operation of the LQCD computing system at the Fermilab site. Fermilab agrees to pay for all facility and utility costs, such as the power needed to support the computing and HVAC systems.

The LQCD Site Managers maintain close working relationships with the compute facility managers at their respective laboratories to ensure that LQCD power and space needs are factored in to future facility planning.

Response to Recommendation 5.2 (cont'd):

The Site Managers maintain records of the power requirements for LQCD compute facilities at their respective sites.

For example, the following table shows the in-kind power contribution for clusters deployed at FNAL and JLab.

Cluster Name	Date	Node Cnt	Power/Node (W)	Total Compute Nodes Power (KW)	Cooling Power Factor	Total Compute Facility Power (KW)
<i><u>FNAL Deployments</u></i>						
QCD	Jun-04	128	147	18.8	1.5	28.2
Pion	Dec-05	520	176	91.5	1.5	137.3
Kaon	Oct-06	600	275	165.0	1.5	247.5
Jpsi FY08+FY09	Apr-09	864	300	259.2	1.7	440.6
<i>FNAL Sub-total</i>		<u>2,112</u>		<u>534.5</u>		<u>853.6</u>
<i><u>JLab Deployments</u></i>						
6n	Jan-06	260	180	46.8	1.5	70.2
7n	Jun-07	396	300	118.8	1.5	178.2
<i>JLab Sub-total</i>		<u>656</u>		<u>165.6</u>		<u>248.4</u>

Note: The power totals shown in the column "Total Compute Nodes Power" are for the compute nodes directly, and do not include ancillary items such as UPS power loss, nor power used by the A/C system. Multiplying these values by the Cooling Power Factor (CPF) provides a conservative estimate of total power required. The CPF for J-Psi is slightly higher because that computer room is only partially occupied and both the cooling and UPS systems are not as efficient as they will be once additional systems are installed.



Response to Recommendation 5.2 (cont'd):

In addition to the cluster deployments, power requirements for the 12,288-node QCDOC machine deployed at BNL are as follows.

- The water-cooled QCDOC crates use 11 KW each and there are 12 crates for a total of 132 KW.
- Additional power is required for the front-end hosts, file servers, air-cooled crates, and other supporting hardware.
- Total in-kind power contribution for QCDOC is of order 200 KW.

Response to Recommendation 5.2 (cont'd):

Quantified Space Needs

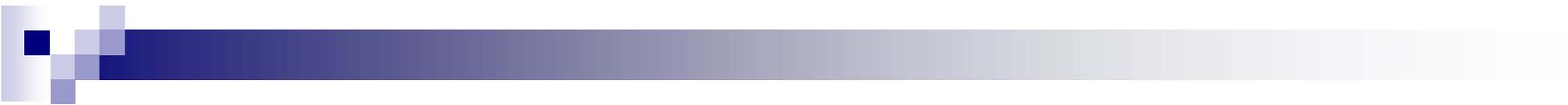
As shown in the adjacent table, cluster deployments at FNAL and JLab require approximately 1,620 ft² and 440 ft², respectively.

Note: the floor space taken by a rack position is ~5 tiles, each 4 ft².

Clusters are sited in available, suitable space. FNAL are housed in three computer rooms; clusters at JLab are housed in a single facility.

At BNL, the QCDOC machine requires approximately 100 ft² of floor space in the computer room directly. Additional floor space is required in adjacent mechanical areas for supporting equipment such as dedicated heat exchangers for the water-cooled machine.

Cluster Name	Date	Node Cnt	# of Rack Positions	Floor Area (ft ²)
<i>FNAL Deployments</i>				
QCD	Jun-04	128	6	120
Pion	Dec-05	520	22	440
Kaon	Oct-06	600	31	620
Jpsi FY08+FY09	Apr-09	864	22	440
<i>FNAL Sub-total</i>		<u>2,112</u>	<u>81</u>	<u>1,620</u>
<i>JLab Deployments</i>				
6n	Jan-06	260	7	140
7n	Jun-07	396	15	300
<i>JLab Sub-total</i>		<u>656</u>	<u>22</u>	<u>440</u>



Response to Recommendation 5.2 (cont'd):

Summary

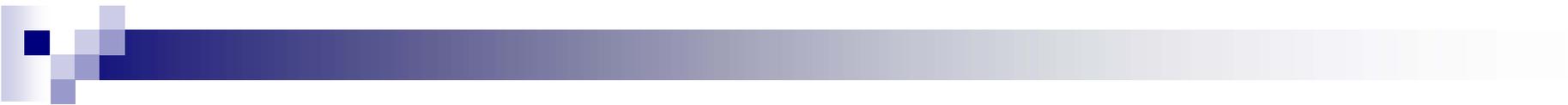
We have quantified at a reasonable level the power and space requirements for the LQCD compute facilities.

We maintain close working relationships with facility and senior management at the three host laboratories to ensure that LQCD needs are understood and factored into facility planning.

Throughout the project, all three host laboratories have shown a strong commitment to supporting project facility and infrastructure needs.

Signed MOUs are in place with the three host laboratories that define anticipated in-kind contributions for facilities and infrastructure equipment.

At this point in the current LQCD project, we believe that the agreement defined in the existing MOUs, along with the strong working relationship we have with laboratory management, adequately ensures the successful completion of the project. Therefore, we do not believe it necessary to amend the present MOUs.



Recommendation 6.1:

The user surveys indicated that the transparency of the allocation process could be improved. Additional more specific surveying should be pursued by LQCD to pinpoint the source of the problem and to remedy it.

Response:

In response to suggestions made in the 2007 Survey, proposals were more quickly and prominently linked from the USQCD web site in 2008 than in 2007.

Rotation in the membership of the Scientific Program Committee (SPC) has started to bring in groups that had not had a representative in the past.

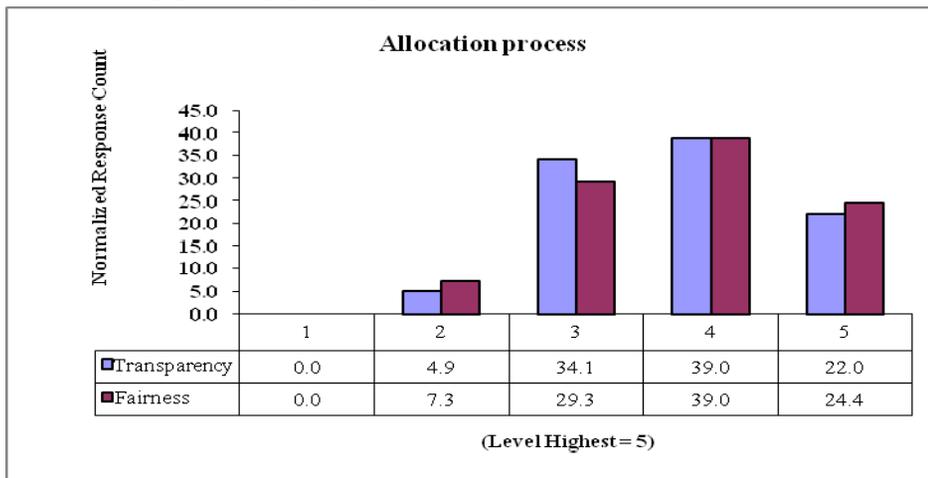
Some of the questions in the 2008 User Survey were modified to gain additional insight into the USQCD community's perception of the allocation process.

User response level was smaller in 2008 than 2007 (44 vs. 54 respondents), but results showed some movement in the satisfaction level of the allocation process.

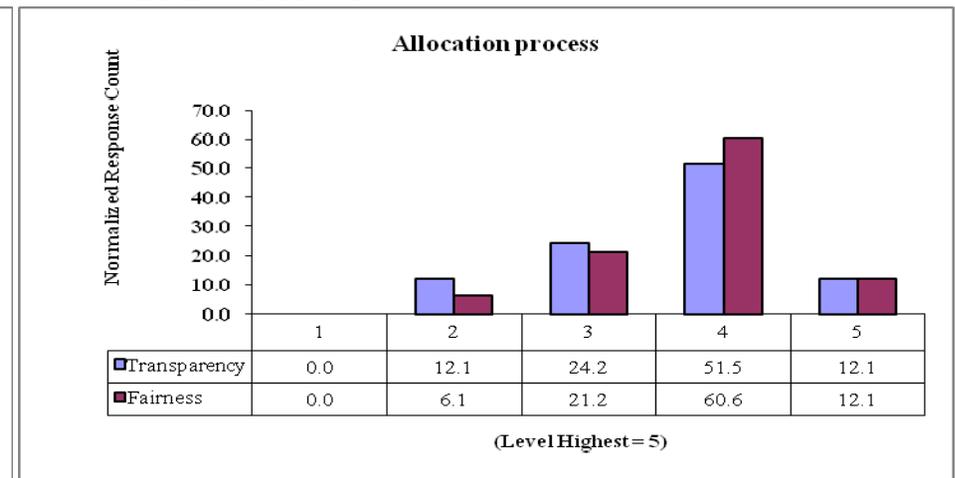
Response to Recommendation 6.1 (cont'd):

Comparison of results from 2007 and 2008 user surveys, regarding satisfaction with the allocation process.

2007 Results



2008 Results



2008 survey questions:

Transparency: Rate the transparency of the allocations process (SPC deliberations, All-Hands Meeting, e-mail communications from the SPC)

Fairness: Please rate the fairness of the allocations process (consider how the process applies to everyone, not just your own proposal.)

2008 survey results:

Transparency: 10% moved from 5 to 4, and 10% moved from 3 to 2; overall = **xx% vs. xx%** (FY07)

Fairness: 12% moved from 5 to 4 and 9% increased from 3 to 4; overall = **xx% (relatively unchanged?)**



Response to Recommendation 6.1 (cont'd):

From 44 survey respondents, 8 chose to submit free-form comments. Common threads summarized as follows:

- Notification of final allocation is via e-mail with no chance for appeal or “hearing second opinions from the PI” regarding the SPC’s decision.
- Would like to see more feedback on how allocation sizes were decided.
- Comments from SPC regarding how final allocation is adjusted would make the process less of a black box for most people.
- Allocation-based resource distribution often directly conflicts with efficient use of resources. Inflexibility often forces decisions to be made according to the dictates of the allocation, in opposition to scientific output criteria.
- Concerns over the transparency of the allocation process for BG/P resources.

In response, some private discussions have been held to better understand “lack of transparency” concerns.

Also, concerns are being addressed in this year’s allocation process, which is currently underway but not yet finished:

- For example, long discussion at the May 14-15 All-Hands Meeting regarding the distribution of resources in general. Main issue was how to adjust split between Type-A and Type-B proposals. Agreed to increase upper limit for Type-B requests in 2009/10 to take into account additional resources.