Lattice QCD Computing Project (LQCD)

Response to Recommendations from the 2008 Annual Progress Review of the LQCD Computing Project

Compiled by

Paul Mackenzie
Chair, LQCD Executive Committee

William Boroski
Contractor Project Manager, LQCD Computing Project

June 1, 2009

INTRODUCTION

On May 13-14, 2008, the U.S. Department of Energy Office of High Energy Physics and the Office of Nuclear Physics conducted an Annual Progress Review of the ongoing Lattice Quantum Chromodynamics (LQCD) Computing Project. The review was held at Brookhaven National Laboratory and resulted in a written report that contained 14 recommendations to help improve the project's effectiveness and impact. Ten recommendations were associated with the scientific program and four recommendations were associated with technical and/or user aspects of the computing project.

This document summarizes the LQCD Computing Project's response to these recommendations and where appropriate, the actions taken to implement specific recommendations. Since the recommendations were not uniquely numbered in the review report, we have adopted the following numbering scheme when tracking resolution:

ReviewReportSectionNumber.RecommendationNumber

RESPONSE TO RECOMMENDATIONS

Recommendation 2.1

USQCD should consider including experimenters and theorists from outside the lattice community on the USQCD Executive Committee and Allocations Board. This would broaden the scientific program and allow a wider community to influence the prioritization of physics topics and the allocation process.

Response: We agree that in setting scientific priorities USQCD should obtain broad input from experimenters and theorists outside the lattice community. We believe that the best way to obtain broad input is through the workshops we organize that bring lattice gauge theorists together with experimenters and/or outside theorists with similar interests. In the last two years we have held such workshops on CKM matrix elements, QCD at high temperatures and densities, physics beyond the Standard Model, and the structure and spectroscopy of hadrons. Each of these workshops has one or more organizers from outside USQCD, as well as from inside. We plan to hold workshops of this type on a continuing basis.

A key step in setting scientific directions for USQCD is the discussion of proposals and priorities held at the annual All Hands Meeting. We have decided to invite one of the outside coordinators of each workshop to make a presentation at the first All Hands Meeting following his/her workshop. This person is asked to summarize the workshop and participate in the discussion of USQCD priorities. We believe that this approach provides broader scientific input than including one or more physicists only representing a few specific subfields outside USQCD on the Scientific Program Committee or the Executive Committee. At the May 14-15, 2009 All-Hands Meeting, nuclear physicist Curtis Meyer summarized the November, 2008 JLab workshop, "Revealing the Structure of Hadrons".

Recommendation 2.2:

USQCD should continue its workshops with other segments of the high energy and nuclear physics communities. It should also continue its series of summer schools to encourage the grown of the field by attracting talented young physicists.

Response: We intend to follow both of these recommendations. We scheduled a workshop on hadron structure that was held at JLab on November 21 and 22, 2008. We have scheduled one on QCD thermodynamics at BNL on June 8-12, 2009. We also expect to hold another workshop on weak interaction matrix elements in the coming year.

We have been coordinating with Europe to organize alternating summer schools. One was held at the INT, Seattle, August 8-28, 2007. The next will be held at Les Houches, Aug. 3-28, 2009.

Recommendation 2.3:

As the accuracy of LQCD simulations have improved, small discrepancies between alternative methods and discrepancies with experimental results are becoming apparent. The source of these problems should be identified. The independence of the members of gauge ensembles should be monitored closely, and the results of such studies should be included in the stated errors of the resulting matrix elements.

Response: One of our major objectives is to look for possible discrepancies between lattice calculations and experimental results. At present, the one case in which such a discrepancy may be occurring is the determination of the leptonic decay constant of the *Ds* meson. This question is under intense investigation by members of USQCD. We are not aware of any significant discrepancies in quantities calculated with different lattice methods; however, we believe it is essential to continue to investigate this possibility by calculating a number of quantities using different methods. The independence of members of gauge ensembles is monitored closely, and correlations are taken into account in error budgets. The committee is right to emphasize the importance of doing so.

Recommendation 2.4:

USQCD should encourage planning within the community to ensure that analytic calculations in chiral perturbation theory are completed in a timely fashion.

Response: We agree that for many projects it is imperative to carry out chiral perturbation theory calculations, and to have them completed in time to be used in analyzing the data. Members of our community are well aware of this fact. Many of them already do such calculations themselves, collaborate with theorists specializing in chiral perturbation theory, or actively encourage calculations that they need. Planning for such calculations is encouraged by the fact that doing so strengthens proposals for USQCD allocations.

Recommendation 2.5:

USQCD should encourage more work on the charmonium and open-charm spectra, in light of recently discovered at the B factories. USQCD should similarly encourage spectroscopy calculations (light JPC exotics, etc.) that are relevant to the 12 GeV upgrade at JLAB, since this is currently the highest DOE NP experimental priority.

Response: We have major efforts in progress in both of these areas, which could use more resources; however, we have other high priority projects in progress that would also benefit from more resources. The critical question is how to balance resources among high priority projects. We believe that our allocation process does a good job of this, but we actively solicit input from other sources.

Recommendation 2.6:

USQCD should encourage the calculation of transport coefficients in finite temperature simulations, since these quantities are crucial to different theoretical approaches to the subject, and are central to experimental programs at the Relativistic Heavy Ion Collider (RHIC) and elsewhere.

Response: We agree that these calculations are important. One of these, which was very highly rated by the Scientific Program Committee last year, has been completed. Several new projects in this area have been approved by the Scientific Program Committee for 2009/10.

Recommendation 2.7:

In allocating time to Type C projects, USQCD should give special emphasis to exploratory work on physics beyond the Standard Model.

Response: We agree. As we indicated at the 2008 review, Class C projects are under-subscribed, so the Scientific Program Committee and the Executive Committee have worked to encourage them in all subfields. Because beyond the Standard Model (BSM) physics is a less mature sub-field of lattice gauge theory, exploratory work in it is particularly appropriate for Class C allocations. However, it should be noted that we already have several BSM projects with large (Class A and B) allocations. Allocations for BSM projects have been increasing every year for the last few years. In the 08/09 allocation year, six out of sixteen high energy physics projects were BSM projects, and they were awarded 4.1% of USQCD resources. In the 09/10 allocation year, five out of sixteen HEP projects were BSM projects, and they were awarded 7.1% of USQCD resources.

Recommendation 2.8:

USQCD uses a ``bottom up'', proposal-driven allocation process. There is, therefore, no process to guarantee that the LQCD facilities will be used to meet the priorities of the broader High Energy and Nuclear Physics communities. Several of the recommendations above address this concern in part, but USQCD might consider developing a more definite roadmap outlining actual commitments of groups to particular calculations, with projected estimates of precisions. In particular, USQCD might consider a process that has been *applied to large experimental collaborations, specifically providing allocations for some assigned activities to insure the physics community that specific high priority opportunities are not missed.*

Response: We do indeed use a "bottom up", proposal-driven allocation process, and we believe very strongly that this is the correct approach for our field. In our LQCD and LQCD-ext proposals we set out roadmaps indicating our highest priority projects and, where possible, the precision we expected to reach with a given amount of computing resources. These high priority projects involve very exciting physics, and there has never been a lack of proposals to carry them out. It is the responsibility of the Scientific Program Committee in recommending allocations and the Executive Committee in approving them, to make certain that allocations are properly balanced among high priority areas. We have obtained input from the broader high energy and nuclear physics communities regarding priorities. We have made several presentations to the High Energy Physics Advisory Panel, which has provided advice on scientific priorities and strong support for our efforts. In nuclear physics, where NSAC has agreed on a comprehensive set of national milestones, these milestones have always been considered in the Scientific Program Committee's discussions, and the national USQCD program has consistently been well aligned with them. The process outlined in response to Recommendation 2.1 will provide very useful input from the broader high energy and nuclear physics communities on a yearly basis.

Recommendation 2.9:

USQCD should become more systematic in making physical quantities (and their associated error matrices) publicly available before chiral and/or continuum extrapolations, to allow future improvements in these areas to be propagated back to earlier results.

Response: This is an interesting suggestion. We believe that it is important to discuss with members of USQCD and our experimental colleagues exactly what information would be most useful to include in publications and/or to post on the web. We believe that the workshops mentioned above would be good venues for doing so.

Recommendation 2.10:

The number of post-docs, graduates and undergraduates involved in LQCD research should be better documented, in order to understand the impact the project is having on the demographics of NP and HEP.

Response: We agree. We have compiled a new version of the USQCD membership list that includes the academic rank of each of our members. Because junior members of the collaboration change rank and institution relatively often, we will institute regular surveys to help us keep the list up to date.

***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.

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.

Adjustments in Staffing Profile:

- *Reduced site management from 0.25 to 0.15 FTE at all sites (Δ = -0.30 FTE)*
- *Reduced sys admin support at BNL from 0.75 to 0.25 FTE (Δ = -0.5 FTE)*
- *Reduced deployment support at FNAL from 0.75 to 0.50 FTE (Δ = -0.25 FTE)*
- *Increased level of operations sys admin support at FNAL and JLab from 1.1 to 1.9 FTE per site (Δ = 0.8 FTE/site, or 1.6 FTE total)*

These adjustments resulted in a net increase of 0.55 FTEs of salary support. However, due to the differences in the 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 systems 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, as noted in the following excerpt from the Fermilab MOU. Similar language is contained in the JLab and BNL MOUs.

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, which helps ensure that LQCD power and space needs are factored in to future facility planning.

Quantified Power Needs

LQCD Site Managers maintain records of the power requirements for LQCD compute facilities at their respective sites. For example, the following table summarizes the in-kind power contribution for clusters deployed at FNAL and JLab.

Table 1. Compute Facility Power Requirements for Clusters 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)
FNAL Deployments						
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
FNAL Sub-total		2,112		534.5		853.6
JLab Deployments						
6n	Jan-06	260	180	46.8	1.5	70.2
7n	Jun-07	396	300	118.8	1.5	178.2
JLab Sub-total		656		165.6		248.4

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 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.

In addition to the cluster deployments at FNAL and JLab, 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.

Quantified Space Needs

As shown in Table 2, the cluster deployments at FNAL and JLab require approximately 1,620 ft^2 and 440 ft^2, respectively. Note that the estimated floor space taken by a rack position is ~5 tiles, each measuring 4 ft^2.

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

Table 2. Floor Space Requirements for Clusters at FNAL and JLab

Cluster Name	Date	Node Cnt	# of Rack Positions	Floor Area (ft^2)
FNAL Deployments				
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
FNAL Sub-total		2,112	81	1,620
JLab Deployments				
6n	Jan-06	260	7	140
7n	Jun-07	396	15	300
JLab Sub-total		656	22	440

At BNL, the QCDOC machine requires approximately 100 ft^2 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.

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.

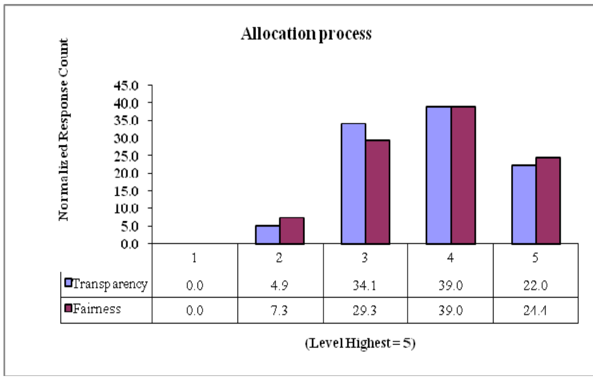
Response: In response to suggestions made in the 2007 Survey, proposals were more quickly and prominently linked from the USQCD web site in 2008 than in 2007. In addition, rotation in the membership of the Scientific Program Committee (SPC) has started to bring in groups that had not had a representative in the past.

Modifications were also made to some of the questions in the 2008 User Survey to gain additional insight into the USQCD community's perception of the allocation process. The following two questions relate to the perceived transparency and fairness of the allocation process:

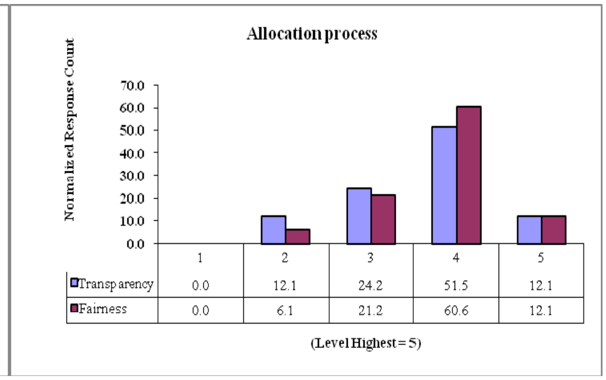
1. Transparency: Rate the transparency of the allocations process (SPC deliberations, All-Hands Meeting, e-mail communications from the SPC)
2. Fairness: Please rate the fairness of the allocations process (consider how the process applies to everyone, not just your own proposal.)

The following charts offer a comparison of the results from 2007 and 2008 user surveys with regard to satisfaction with the allocation process. Although the user response level on the FY08 survey was smaller in 2008 than 2007 (36 vs. 54 respondents), the results show movement in the satisfaction level of the allocation process.

2007 Results



2008 Results



From 36 survey respondents, 8 chose to submit free-form comments. Common threads are 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 to this feedback, some private discussions have been held to better understand some of the "lack of transparency" concerns. In addition, transparency and fairness concerns are being addressed in this year's allocation process, which is currently underway but not yet finished. For example, a lengthy discussion was held at the May 14-15, 2009 All-hands Meeting regarding the distribution of resources in general, and in particular how to adjust the split between Type-A and Type-B proposals. As a result, the SPC agreed to increase the upper limit for Type-B requests in 2009/10 to take into account additional resources.