Lattice QCD Extension II Computing Project
(LQCD-ext II)

Response to Recommendations from the 2015 DOE Annual Progress Review of the LQCD-ext II Computing Project

Compiled by

Paul Mackenzie
Chair, USQCD Executive Committee

Robert D. Kennedy
Associate Contractor Project Manager, LQCD-ext II Computing Project

November 3, 2015

This page intentionally left blank

LQCD-ext II 2015 Annual Progress Review
Response to Review Recommendations

INTRODUCTION

On May 21-22, 2015, the U.S. Department of Energy (DOE) Office of High Energy Physics and the Office of Nuclear Physics conducted an Annual Progress Review of the LQCD-ext II (LQCD Extension II) project. The review was held at the Brookhaven National Laboratory and resulted in a written report that contained no formal recommendations. However, the report did contain six suggestions to help improve project effectiveness and impact. This document summarizes the project response to these suggestions, along with subsequent actions taken.

RESPONSE TO SUGGESTIONS

Suggestion #1: Several reviewers commented that the use of TFlop/s and TFlop/s-year as measures of throughput is not that useful a metric and made it difficult to compare expected performance with realized performance and with the allocation system. It is a poor performance measure, especially for computations with significant I/O burdens, as the lattice QCD propagator calculations are. The allocations are measured in "MJ/psi", a reference to the throughput of an obsolete cluster at FNAL. Comparing TFlop/s-year with MJ/psi requires conversion factors that were not obvious. Worse, the GPU systems are measured in "effective TFlop/s", with even more cross calibration required. The collaboration should consider using the compute time, not rate, of standard algorithms to measure the performance of the machines, for example using the benchmarks used now, with standard data set sizes. Several reviewers noted that NVidia markets their GPU systems using throughput numbers for MILC and CHROMA benchmarks, not TFlop/s. The use of compute time for standard datasets makes it easier to compare the capacity of LQCD with the requested allocations and should make it easier to evaluate the effectiveness of the allocation use, once the compute time scaling properties are understood..

Report Section: LQCD-ext II Review – Progress towards Scientific and Technical Milestones

Response: The project will address this topic more directly in reviews in the future. Please permit us to offer some detailed context before addressing the suggestions directly.

There are two units of measurement for describing LQCD application throughput that are very familiar to all members of the USQCD community. The first unit is sustained floating point operations per second, with TFlop/s being the most appropriate fundamental unit in terms of size on current computing hardware. The second unit is a core hour, representing the work done per core per hour on an x86 cluster or on a BlueGene system. In the annual Call for Proposals, the computing capacity of each of the dedicated hardware resources is clearly given in both sets of units, along with the relevant conversion factors.

LQCD sustained TFlop/s are defined as the floating point throughput of the Dirac operator inverter implemented using the conjugate gradient method. For each of the principal actions of interest to USQCD (HISQ, clover-improved Wilson, and Domain Wall Fermion), there is a well-defined number of floating point operations needed at each point in the simulation lattice for each iteration ("D-slash") of a conjugate gradient inverter. Using standard lattice sizes and numbers of MPI ranks, the throughput in Flop/sec is measured for each action on each of the project's dedicated conventional clusters. The average of this throughput across the actions, normalized per MPI rank (i.e., per core) gives the LQCD sustained Flop/sec rating per core for each cluster. This rating is sensitive to the floating point performance, network bandwidth, and memory bandwidth of each machine.

The sustained TFlop/s rating is only predictive of conjugate gradient inverter performance. Many LQCD users have workflows that include portions with throughputs not proportional to inverter performance, such as sections of I/O to local or remote storage, or algorithms that differ substantially from the conjugate gradient inverter. Accordingly, USQCD allocations on the project's dedicated hardware have always been in computing time units, originally based on a 2006-era project cluster, and later node hours and then core hours from a 2009-era project cluster (J/Psi). Because performance varies across the dedicated clusters, due to different processor generations, memory speeds, and network speeds, USQCD has always maintained a table of relative performance figures for the dedicated machines based on the TFlop/s inverter ratings. Thus a "Ds" cluster core hour is equivalent in terms of inverter performance to 1.33 J/Psi core hours.

When applying for an allocation, it is expected that USQCD members will estimate and ideally measure relevant throughput numbers for their specific workflows on the clusters of interest to them. Proposals are expected to include discussions of the software to be used and the expected physics results that can be obtained by the requested computing resources. The requested computing resources include integrated computing time, expressed in the current computing time units ("J/Psi core hours"), as well as any disk and tape storage resources expressed in terms of TBytes. Users may request testing time on the clusters to measure throughputs.

GPU-accelerated clusters represent a greater challenge for the project for rating computing throughput. Typically only some portions of a physics project's workflow can be well-accelerated on a GPU, so a given application may be thought of having something analogous to the parallel and serial sections described by Amdahl's Law. For some workflows and algorithms, substantial throughput increases are observed compared to the cost-equivalent amount of conventional cluster hardware. For other workflows and algorithms, the throughput gain is more modest, and for specific applications conventional hardware can be more cost effective. There is value, however, for the project and for the funding offices to have a notion of equivalent computing capacity on the various hardware platforms. For GPU-accelerated clusters, the project uses a set of representative physics problems spanning the actions of interest. The throughput of these programs is measured on both GPU-accelerated and conventional hardware (of course, different binary codes are used, but they solve exactly the same physics problems). For each problem, the throughput on the GPU-accelerated cluster can be expressed in proportion to the throughput on the conventional cluster. Averaging across the benchmarks gives an average relative throughput. Using the LQCD sustained TFlop/s rating of the conventional cluster, an "effective TFlop/s" rating is thus defined for the GPU-accelerated cluster. This terminology was chosen in part because the actual floating point operation count for an algorithm expressed for a GPU may differ widely from the operation count of a conventional CPU implementation.

In direct answer to this Comment section, in fact USQCD does use compute time for allocations. The compute time, expressed in J/Psi-core-hour units, derives from standard datasets used in the measurement of the performance of conjugate gradient inverters for the physics actions of interest. These units, the measurements used to obtain the ratings, and the conversion factors between units and the various dedicated platforms are documented and made available to the USQCD community via the Call for Proposals and web resources referenced therein. The use of core-hour allocation units is similar to the use on the various DOE and NSF Leadership Class computing resources also employed by USQCD. Ultimately it is the responsibility of the individual PI to determine the estimated or measured throughput of his or her workflow on the dedicated resources and to describe the expected physics production given an allocation in the computing time units used by the Scientific Program Committee.

There is no easy and inexpensive solution to the end-to-end benchmarking and allocations process with a variety of codes running on disparate hardware technologies. What seems simple in one context will still require conversion factors in another context, and still may not be a good predictor of scientific productivity in practice. As codes evolve in time to take advantage of technology, or consume more resources to tackle large problems, re-measurement and rework can become prohibitively expensive. With new technologies on the horizon, this topic will become even more important for the project to address clearly to a broad audience.

<u>Suggestion #2:</u> The reviewers felt that actual scientific progress is solid, but the procedural steps to assess it, document it, and define future goals in appropriate detail, could be improved. They encouraged the project and USQCD to develop a viable way to achieve this goal.

Report Section: LQCD-ext II Review – Progress towards Scientific and Technical Milestones

Response: We agree with this suggestion and intend to actively look for ways of improving how we report goals and progress. We feel that we have achieved some success in setting goals and reporting progress for those weak matrix elements that involve a single, hadronically stable hadron in the final state. These include stable meson decay constants, semi-leptonic decays, and neutral meson–anti-meson mixing. These have been performed on the computers of the last five years at physical light quark masses and adequate volumes allowing for detailed quantitative uncertainty budgets. This in turn has allowed for estimation of future uncertainties in some detail and with decent accuracy. For many other important quantities, larger physical volumes are needed, and we are only now coming into the era when computers are powerful enough to produce ensembles at the physical quark masses and adequate volumes. These include weak matrix elements with multiple hadrons in the final state such as the K to two pion amplitudes needed for epsilon'/epsilon, and most nuclear physics quantities including multi-hadron interactions and the hadronic resonances studied at JLab. As the uncertainty budgets for these quantities become more detailed and quantitative, forecasting future uncertainties will become more straightforward. For the present, we are trying to develop ways to present our best estimates of what it will take to reach a given accuracy, while at the same time doing our best to make clear that such estimates are less robust than estimates of future progress for quantities for which current uncertainties are currently more detailed.

<u>Suggestion #3:</u> **The reviewers suggested that alternative means of measuring the user needs and satisfaction be developed by the project. These might include interviews with the heaviest users and separate interviews with new users, etc.**

Report Section: LQCD-ext II Review – Progress towards Scientific and Technical Milestones

Response: With 74% of the PI's and 50% of the active users responding to the FY14 user survey, the project believes there is a good sampling of the population of interest, though of course it could be better. The user survey accomplishes much with little additional effort each year by project management staff after the initial setup and only a modest time commitment on the part of respondents (15 minutes or less, at a time of their convenience). The project will consider how we might complement the user survey approach with interviews, "town hall" calls, or other means as suggested, within our project management budget.

<u>Suggestion #4:</u> **They (the reviewers) suggested that a few elected members of the user community participate in the allocation decisions through rotating positions on the SPC, in addition to the more senior and experience people doing the bulk of the work.**

Report Section: LQCD-ext II Review – Effectiveness of USQCD, Scientific Impact, Procedures and Related Activities

Response: See response to next suggestion.

<u>Suggestion #5:</u> **Several reviewers suggested that USQCD consider the management structures of typical HEP experimental collaborations. An Executive Committee model that works well in large experimental collaborations is to have one designated slot for an elected representative of the junior physicists in the lattice community.**

Report Section: LQCD-ext II Review – Effectiveness of USQCD, Scientific Impact, Procedures and Related Activities

Response to suggestions #4 and #5: We agree that the question of selection of the members of the leadership committees, and the question of elections specifically, are important ones and we have been discussing it in detail this year. We have also had discussions this year and in previous years on more general issues in selecting members of the leadership committees. These include rotations on the

Executive Committee, the optimal age distributions of the members of the EC and SPC and the role of young people on the leadership committees, and the selection of members of the committees including the possibility of elections. We have already made some changes as a result of these discussions. We have instituted an annual rotation of Executive Committee members. Each year, two members of the Executive Committee are considered for reappointment, with the expectation that on average about half will be reappointed and about half will rotate off, so that on average one member of the EC will rotate every year. In reappointing members representing major collaborations within USQCD, the wishes of the collaboration in question are taken into consideration. In this way, we expect that over a ten year period to have nearly complete rotation, with older members of the Executive Committee in general being replaced by younger members of the community. About half of the Executive Committee has rotated so far. We have discussed the role of elections in previous years and have investigated the governance of a number of experimental collaborations without finding one which seemed like an appropriate model for our community (which is open to all lattice theorists in the US without any particular qualification, unlike experimental groups which must define their contributions when joining a collaboration). However, we were not aware of the particular model suggested by the reviewers this year and look forward to investigating it. Likewise, we have discussed alternative ways of choosing members of the SPC with previous SPCs, but we have not previously thought through the specific suggestion of this year's review committee and we will discuss this possibility with them as well as other possibilities.

<u>**Suggestion #6:**</u> **Several reviewers thought that the SAB can be used more effectively. Perhaps the SAB members could be engaged more actively in the processes of the USQCD through participation in the All Hands Meeting.**

Report Section: LQCD-ext II Review – Effectiveness of USQCD, Scientific Impact, Procedures and Related Activities

Response: The SAB is relatively new and its role is a work in progress. We have discussed various ways of involving the SAB in the operations of USQCD more extensively and will continue to do so. We discussed the possibility of having an annual fact-to-face meeting of the SAB. We found that most SAB members were reluctant to commit to such a meeting, and additionally, the cost of flying seven people to a one- or two-day meeting annually seemed exorbitant for project with a total budget of around $3,000,000/year. (Procedures that make sense for a laboratory with a multi-hundred million dollar budget don't translate precisely to a project that is two orders of magnitude smaller.) We have also investigated involving the SAB in the allocations process. Two years ago, we invited the entire SAB to look at the proposals for that year's allocations. When we seemed to get no takers, we requested several of the most active members of the SAB to read those proposal that seemed most interesting to them and report to us. They reported that while the proposals were interesting and seemed scientifically well motivated, the SAB members didn't feel that they had enough expertise to usefully contribute to the allocations process. One SAB member commented, "I find the proposals I read mostly pretty well written, with a science justification in the intro, the abstracts are all remarkably of the same format: brief science justification, goals, requested allocation, which is pretty accessible (without being asked to judge whether the project is realistic)...I do not actually imagine that the SAB is going to have much useful feedback for you". This year, we intend to investigate the possibility of encouraging at least a couple of SAB members to participate in the All Hands Meeting. We have also begun discussions with the SAB itself to encourage the current members to think about how they believe they can be most useful.