Offices of High Energy and Nuclear Physics
Report on the

LQCD-ext II

2015 Annual Progress Review

May 21-22, 2015

Executive Summary

The Annual Progress Review of the LQCD-ext II (Lattice Quantum Chromodynamics extension II) project was held on May 21-22, 2015 at the Brookhaven National Laboratory (BNL). The purpose of the review was to assess LQCD-ext II's progress towards their overall scientific and technical goals, and to assess the role of the USQCD collaboration in governing the usage of the project's hardware. In particular, the LQCD-ext II team was instructed to address five charges:

1. The continued significance and relevance of the LQCD-ext II project, with an emphasis on its impact on the experimental programs supported by the DOE Offices of High Energy and Nuclear Physics;
2. The progress toward scientific and technical milestones as presented in the LQCD-ext II's Project Execution Plan;
3. The status of the technical design and proposed technical scope for FY 2015-2016 for the project;
4. The feasibility and completeness of the proposed budget and schedule for the project;
5. The effectiveness with which the LQCD-ext II project has addressed the recommendations from last year's review.

The USQCD collaboration addressed the charge:

> The effectiveness of USQCD in allocating the LQCD-ext II resources to its community of lattice theorists, the scientific impact of this research on the entire HEP and NP communities and the status, operational procedures and related activities of the USQCD collaboration itself.

Six expert reviewers from the nuclear physics, high energy physics and computer science communities heard presentations on project management, computing hardware acquisitions and operations, organization of the USQCD collaboration, scientific progress, allocation of resources, and dissemination of scientific results. In general, the review panel was very impressed with the technical and scientific achievements of LQCD-ext II and USQCD. The impact of LQCD-ext II simulations on experimental programs in precision measurements of the Standard Model (SM), Heavy Ion collisions and hadron spectroscopy has grown dramatically over the last few years. These developments have been driven by algorithmic improvements and the use of new hardware platforms, including LQCD-ext II's early mastery of Graphical Processing Units (GPUs). The governance of the projects by the USQCD collaboration was judged to be effective and fair. The organization of the USQCD into an Executive Committee (EC) and a Science Policy Committee (SPC) was also praised. The review panel suggested that USQCD consider electing more youthful lattice gauge theorists into the higher positions of the collaboration in light of its aging demographics. The reviewers argued that the Physics Beyond the Standard Model (BSM) efforts of USQCD are underperforming. Finally, the review panel was generally impressed by the results in the annual user survey, but they suggested other forums of user feedback to track the project's assessment by its user community.

Introduction and Background

The DOE Offices of Advanced Scientific Computing Research (ASCR), High Energy Physics (HEP) and Nuclear Physics (NP) have been involved with the National Lattice Quantum Chromodynamics Collaboration (USQCD) in hardware acquisition and software development since 2001. The Lattice Quantum Chromodynamics IT hardware acquisition and operations project ("LQCD"), which started in 2006 and ran through 2009, operated a "Quantum Chromodynamics-on-a-chip" (QCDOC) machine at Brookhaven National Laboratory (BNL), and built and operated special purpose commodity clusters at the Fermi National Accelerator Laboratory (FNAL) and the Thomas Jefferson National Accelerator Facility (TJNAF). LQCD met its 2009 project goal of providing 17.2 Teraflops of sustained computer power for lattice calculations.

The hardware acquisition strategy of LQCD was essential to its success. Each year the project's technical personnel benchmarked the kernels of the QCD code on the newest cluster and supercomputer hardware, and the winner of the price-to-performance competition became next year's provider.

The usage of the hardware procured by LQCD has been governed by the USQCD collaboration through its Executive Committee (EC) and Scientific Program Committee (SPC). In addition, the collaboration organizes the community's access to the DOE Leadership Class Supercomputers available through the INCITE (Innovative and Novel Computational Impact on Theory and Experiment) program. Members of the USQCD collaboration submit proposals through the EC for computer time, some on the Leadership Class machines for large-scale capability computing, and some on the dedicated clusters of LQCD for large scale capacity computing. Allocations on the dedicated clusters of LQCD are awarded by the SPC based on a merit system. Three classes of applications for computer time allocations on the dedicated LQCD hardware are distinguished, these being large-scale mature projects (allocation class A), mid-sized projects (allocation class B), and exploratory projects (allocation class C). Suitable computer platforms are assigned to the various projects upon approval.

Following recommendations from past reviews, a Science Advisory Board (SAB) was formed in 2013 and has participated in the USQCD allocation process. The SAB brings the perspective of the broader HEP and NP community into the high level decision making processes of USQCD and is meant to guarantee that the goals of the lattice effort reflect the diverse needs, challenges and interests of high energy and nuclear researchers. The SAB consists of seven members, four experimentalists and three theorists. They comment on the science goals of USQCD, the effectiveness and fairness of the allocation process and participate in the annual all-hands meeting.

In addition to the original hardware project LQCD, USQCD has also played a role in software development through the Scientific Discovery through Advanced Computing (SciDAC) program. USQCD was awarded a SciDAC-I grant (2001-2006) which was used to develop efficient portable codes for QCD simulations. USQCD was subsequently awarded a second "SciDAC-II" grant (2006-2011) to optimize its codes for multi-core processors and create a physics toolbox. These SciDAC grants supported efforts to provide a user interface to lattice QCD which permits the user to carry out lattice QCD simulations and measurements without the need to understand the underlying technicalities of the lattice formulation of relativistic quantum field theories and its implementation on massively parallel computers. In 2012 USQCD submitted two proposals to the SciDAC-III program, and both were funded, one through NP and ASCR, and the other through HEP and ASCR.

USQCD proposed to extend the work of LQCD beyond 2009, and submitted the proposal "LQCD-ext Computational Resources for Lattice QCD: 2010-2014" to the DOE in the spring of 2008. The scientific content of the proposal was reviewed successfully on January 30, 2008 at the Germantown facility, and the scientific vision and specific goals of the project were enthusiastically endorsed in full by a panel of scientific experts. The proposal requested funding of $22.9M over a five year period to achieve the specified scientific goals.

In the January 30, 2008, review, USQCD argued that the purchase, construction and operation of mid-scale computer hardware was a critical component of the overall strategy to extract physics predictions from lattice Quantum Chromodynamics. That strategy depends on access to the largest Leadership Class machines for the generation of large lattice gauge configurations. These configurations are then analyzed for accurate predictions of matrix elements and spectroscopy on the mid-scale computers operated by LQCD, and results of interest to the experimental and theoretical communities in high energy and nuclear physics are obtained. These mid-scale LQCD computers are also used to generate smaller gauge configurations which are critical to studies of Quantum Chromodynamics in extreme environments (e.g. high temperature and density); these are used to ultimately interpret data from the heavy ion physics program at the Relativistic Heavy Ion Collider (RHIC) at BNL, which is operated by the Office of Nuclear Physics. Many of these calculations are not suited for Leadership Class machines, but run efficiently on mid-scale platforms. Several computer scientists at the January review carefully evaluated and then endorsed the mix of computers advocated by USQCD. The review panel also assessed USQCD's claim that the accuracy of some of its predictions rival the accuracy of the present generation of experiments now running at DOE HEP and NP facilities. The review panel also analyzed USQCD's claim that the proposed project, LQCD-ext, was needed to maintain this parity in the future.

The LQCD-ext project then entered the DOE Critical Decision review process. The CD-0 Mission Need Statement for LQCD-ext was approved on April 14, 2009.

The CD-1, Approve Alternative Selection and Cost Range, readiness review occurred at Germantown on April 20, 2009. The review evaluated the LQCD-ext project's documents on conceptual design, acquisition strategy, project execution plan, integrated project team, preliminary system document, cyber security plan, and quality assurance program.

The LQCD-ext team updated its documents following recommendations from the CD-1 review panel and received formal CD-1 approval on August 27, 2009, through a paper Energy Systems Acquisition Advisory Board (ESAAB) review.

The CD-2/3, Approve Performance Baseline/Start of Construction, readiness review occurred at Germantown on August 13-14, 2009. Final approval for the project was granted on October 28, 2009.

The Offices of High Energy Physics and Nuclear Physics developed the following planning budget for the LQCD-ext CD-2/3 review:

Table 1. Planning Budget for LQCD-ext (in millions of dollars)

	FY2010	FY2011	FY2012	FY 2013	FY 2014	Total
HEP	2.50	2.50	2.60	3.10	3.20	13.90
NP	0.50	0.75	1.00	1.00	1.00	4.25
Total	**3.00**	**3.25**	**3.60**	**4.10**	**4.20**	**18.15**

The TPC of $18.15 left the LQCD-ext project $4.75M short of the figure of $22.9M which had been supported by the scientific review of January 30, 2008, and which USQCD had also included in their original proposal. This shortfall was subsequently effectively addressed by the successful application by the Office of Nuclear Physics for $4.96M of funding through the American Recovery and Reinvestment Act of 2009 (ARRA) to build a 16 teraflop commodity cluster at TJNAF and operate it for four years. Although this effort was not a formal part of the LQCD-ext project, the resulting hardware at TJNAF was governed by USQCD using exactly the same procedures that applied to LQCD-ext, and the acquisition, deployment and operation of this hardware was tracked on a monthly basis by the same team that was running LQCD-ext. In this manner the Offices of High Energy Physics and Nuclear Physics monitored the full scope of the science effort put forward in the USQCD proposal "LQCD-ext Computational Resources for Lattice QCD: 2010-2014". It was agreed that the two efforts (LQCD-ext and LQCD/ARRA) would share Annual Progress Reviews.

The LQCD-ext project team argued at the CD-2/3 review that the budget in Table 1 would support the new deployments and operations described in Table 2 below:

Table 2: Performance of New System Deployments, and Integrated Performance

	FY 2010	FY 2011	FY 2012	FY 2013	FY 2014
Planned computing capacity of new Deployments, teraflops	11	12	24	44	57
Planned delivered Performance (TJNAF + FNAL + QCDOC), teraflop-years	18	22	34	52	90

The original computing goal for the LQCD/ARRA project was 16 teraflops (sustained), from a single cluster at TJNAF. The project team initially estimated that $3.2M would be used for hardware (to be operated for four years), and that labor costs for deployment, operations and management would be $1.2M, with incidental costs for disc space, spares, travel and misc. The project would require the addition of one position at TJNAF. Subsequently, a more quantitative and detailed cost breakdown was developed, which follows in Table 3:

Table 3: LQCD/ARRA Project Funding (in dollars)

Budget	FY09	FY10	FY11	FY12	FY13	Total
Steady State Operations	-	237,406	283,279	294,370	305,905	**1,120,960**
Hardware Deployment	1,929,280	1,817,423	-	-	-	**3,746,703**
Project Management	26,000	27,040	14,061	14,623	15,208	**96,932**
Total	**1,955,280**	**2,081,870**	**297,340**	**308,993**	**321,113**	**4,964,596**

Planning for hardware acquisition for these projects was, however, strongly affected in FY 2010-11 by a "disruptive technology" development in the field of PC chips. Although the first year of acquisitions had been planned assuming commodity cluster technologies, the development of Graphical Processor Units (GPUs) for the commercial gaming industry opened new opportunities for these projects. GPUs consist of several hundred cores per chip, and are the source of the high resolution interactive graphics capabilities needed for video game entertainment. Typically GPUs are capable of an order of magnitude more processing per second than general duty desktop CPUs. GPUs however are difficult to program at this time, and are unbalanced (too little memory per core) for general purpose applications. However, certain low-memory but computationally intensive, highly parallel algorithms can take advantage of a GPU's floating point capabilities, and can thus run 10-100 times faster on GPUs than on a CPU of comparable clock period. Lattice QCD calculations are dominated by one such algorithm; typically ~ 90%+ of the CPU time in lattice QCD is expended in inverting a sparse matrix, which is the Dirac operator that describes the dynamics of virtual quarks in QCD. Anticipating these developments, LQCD/SciDAC had been developing software for several years to run lattice algorithms on GPUs, and the fruits of that effort were then apparent in the GPU hardware acquired for LQCD/ARRA. Two complete physics projects ran on a GPU cluster at TJNAF during the GPU cluster's first year of availability. That number grew to ~ 9 projects in the second year, and continued to increase annually. The price performance on GPUs is ~$0.0l/megaflop which compares to $0.15-0.22/megaflop for the best CPU hardware. This development constituted an important new alternative in the hardware acquisition strategy of these projects, and was assessed in detail by previous review teams. These reviews have contributed several observations about this development:

1. The success of the hardware projects is very sensitive to the continuance of the LQCD/SciDAC software grant, because this is where the software will be developed that will eventually make GPUs more generally useful to the science community;

2. A mix of CPU and GPU clusters will be needed in the short term, because many lattice scientific applications are not ready to be ported to GPUs, but will be much more productive if and when that happens;

3. The initial estimates of the teraflop rating of clusters that can be built for $22.15M proved to be considerably higher than the original planning figures shown in Table 2, but it was hard to estimate new milestones at that time;

4. The scientific output and impact of these projects might be considerably higher than was initially assumed; and

5. The risks associated with using the new GPU hardware would exceed that of the more familiar CPUs.

All of these considerations became part of the discussions regarding plans for hardware acquisitions in FY 2010-12. Several of the observations and predictions quoted above have been confirmed: The ARRA GPU cluster is sustaining ~76 teraflops on a fairly diverse set of physics projects, which exceeds the project's original milestone by a factor of 76/16 ~ 4.75. The LQCD-ext project installed a GPU cluster in 2013 at FNAL to meet the extra demand coming from proposals submitted to USQCD over the previous 12 months. At the termination of the ARRA project in 2012 the ARRA hardware was incorporated into the LQCD-ext project, bringing its budget up to $23.115M, in line with the initial proposal from USQCD.

In 2014 LQCD-ext operated a ½ rack BlueGene/Q supercomputer at Brookhaven National Laboratory (BNL) and built and operated several commodity CPU clusters and GPU clusters at Fermi National Accelerator Laboratory (FNAL) and the Thomas Jefferson National Accelerator Facility (TJNAF). This aggregate of computers sustains ~165 TeraFlops for realistic physics codes. The project was completed in 2014 and exceeded its leading scientific and technical milestones by ~10%.

Since the LQCD-ext project was scheduled to terminate at the end of fiscal 2014, an extension of the LQCD-ext project was proposed by USQCD and was described in the proposal entitled "LQCD-ext II Computational Resources for Lattice QCD: 2015-2019" dated October 23, 2013. This document presented the scientific objectives, the computational strategy, and the hardware requirements of the LQCD-ext II project. The scientific content of the proposal reviewed successfully on November 8, 2013 and the scientific vision and specific goals of the project were enthusiastically endorsed by a panel of scientific experts. The reviewers recommended full funding, $23.4M for the five year period. However, due to budget constraints, the OHEP and ONP provided budget guidance to the project team of either $14M or $18M for the five year period, well below the project's request of $23.4M. These plans became the basis for the project team's planning for LQCD-ext II. That project passed its CD-1 review on February 25, 2014 and was granted CD-1 approval on May 1. It held its CD-2/3 review on July 10 and was approved on Oct. 1, 2014.

The budget planning for the LQCD_ext II project was of some concern to the review panels of the 2014 and the 2015 Annual Review. The original five year budget of $23.4M ($4.68M per year) proposed by the collaboration and endorsed by the November 8, 2013 Science Review resulted in the following anticipated Teraflops profile for new deployments from FY2015 to FY2019:

Full Funding Scenario ($23.4M)	FY2015	FY2016	FY2017	FY2018	FY2019
Planned computing capacity of new deployments, TeraFlops	165	233	330	467	660

However, if the project is funded at the $14M level, with the following profile:

	FY2015	**FY2016**	**FY2017**	**FY 2018**	**FY 2019**	**Total**
HEP	1.0	2.0	2.0	2.0	2.0	9.0
NP	1.0	1.0	1.0	1.0	1.0	5.0
Total	**2.0**	**3.0**	**3.0**	**3.0**	**3.0**	**14.0**

as proposed by the Offices of High Energy Physics and Nuclear Physics in light of constrained, anticipated federal funding, then the estimated Teraflops profile for new deployments is reduced to:

Reduced Funding Scenario ($14.0M)	FY2015	FY2016	FY2017	FY2018	FY2019
Planned computing capacity of new deployments, TeraFlops	0	107	160	244	358

which is a 53% reduction in compute power compared to the full funding scenario. This reduction in computing capacity challenges USQCD to maintain its productivity, its balance with its Leadership Class computing allocations and its international standing. The 2014 review panel commented on these developments since they influence the use and productivity of the FY 2014 hardware acquisitions they endorsed. The 2015 review panel also commented on the extra challenges that constrained funding places on the project and they noted that any additional funding would directly increase the project's hardware acquisition plans.

The Annual Progress Review of LQCD-ext II took place at BNL on May 21-22, 2015. The review consisted of one day of presentations and a second half-day of questions and answers, report writing, and a closeout session. The Appendices to this report provide additional detailed material relating to the review: App.A contains the charge letter to the LQCD-ext II management team, App.B lists the reviewers and DOE participants, and App.C contains the agenda and links to the talks.

The remaining sections of this report detail the findings, comments, and recommendations of the review committee for each of the six charge elements that the LQCD-ext II project team was asked to address.

LQCD-ext II Review

Continued Significance and Relevance

Findings

The LQCD-ext II program supports activities in four research areas:

1) Intensity Frontier. Precision calculations which are relevant to the determination of standard model parameters extracted from heavy quark processes have been a major element in lattice calculations for several years. Calculations of decay constants and form factors which are essential for the extraction of CKM elements from experimental data and for looking for hints of new physics are continuing with ever increasing precision. Strong interaction matrix elements and scattering processes that are relevant to experiments at the Intensity Frontier, including the muon g-2 and the muon to electron conversion experiments at Fermilab, numerous kaon physics processes which are used to extract fundamental Standard Model parameters from various decay rates and scattering amplitudes, and low energy neutrino-nucleon cross-sections which are crucial to extracting results from neutrino oscillation experiments in progress at Fermilab, are new focus areas of lattice calculations. Andreas Kronfeld summarized this subfield of lattice gauge theory at the review. He emphasized the alignment of the lattice calculations with the growing set of experiments and projects in the near term Intensity Frontier program. He explained that algorithmic improvements in the muon g-2 calculational program may produce a sufficiently accurate lattice calculation to improve the theory prediction before the experiment's data analysis scheduled for 2019-20. The recent productivity of the Intensity Frontier lattice

effort has been good with 11 Physical Review Letters in the last five years that have accumulated over 300 citations.

2) Energy Frontier. Exploratory calculations based on "beyond the standard model" (BSM) theories, which in many cases are strongly coupled field theories, for which lattice gauge theory is at present the only effective technique for extracting quantitative predictions, constitute the main area of lattice calculations in this subfield. The emphasis has been on "walking" Technicolor models in which strong dynamics of new generations of quarks and gauge fields generate a composite Higgs which breaks electroweak symmetry. Unlike QCD, these theories are "almost" conformal. Calculations which accommodate the Higgs at 125 GeV/c^2 as a pseudo-Goldstone boson and predict additional states accessible to the LHC 14 TeV run were presented. Some of the composite Higgs models have natural Dark Matter candidates. Investigative studies of supersymmetry are also underway. GPU clusters are proving particularly useful in these studies. Anna Hasenfratz summarized this subfield of lattice gauge theory at the review. She emphasized that this work is exploratory and only accounts for ~9% of the total USQCD efforts. Over the last year there have been several workshops where the lattice community interacts with theorists, phenomenologists and experimentalists in the field.

3) Hadonic Spectroscopy and Form Factors. Hadronic physics quantities such as the spectrum of hadrons, form factors, moments of structure functions, hadron-hadron interactions and scattering make up this subfield. Many of these calculations are relevant to several NSAC Milestones. In addition, several of these calculational programs are well aligned with experiments planned for the 12 GeV upgrade of the Continuous Electron Beam Accelerator Facility (CEBAF) at TJNAF, including the spectroscopy of exotic mesons relevant to the GlueX project. Other calculations focus on the program planned for the Facility for Rare Isotope Beams (FRIB). The advent of peta-scale computing will lead to calculations with physical pion masses so chiral extrapolations and the attendant uncertainties will no longer be relevant. This will lead to a new era in hadronic structure and spectroscopy calculations and make lattice simulations even more relevant to NP's experimental program. Will Detmold summarized this subfield of lattice gauge theory at the review. The productivity of this lattice subfield has been strong in 2014 with two Physical Review Letters and ~50 publications in all. Recent developments include: Coupled channel spectroscopy calculations, phase shifts and inelasticities, magnetic moments of nuclei and Isospin mass splittings in QCD + QED. The job market in the field has also shown some life recently with the announcement of three lattice positions in this subfield at Michigan State University.

4) Extreme Environments. Calculations of the properties of QCD at finite temperature and baryon density, which is explored experimentally in relativistic heavy ion collisions, are critical to this subfield. These simulations are having an impact on the run plans of RHIC at BNL. The equation of state of the quark-gluon plasma is an essential input into the analysis of experimental data and the development of phenomenological models of final states. Recent calculations have focused on the critical temperature for the formation of the quark-gluon plasma, the critical point, the freeze-out lines, the velocity of sound and its temperature dependence, susceptibilities, and thermal dileptons. Calculations of the Equation of State and the Transition Temperature are

now considered "mature" and definitive. Several lattice calculations address questions posed in the NSAC Long Range Plan 2007. As lattice calculations become more accurate and ambitious, they are having an ever larger impact on the experimental NP program at RHIC and other worldwide facilities. Frithjof Karsch summarized this subfield of lattice gauge theory at the review. He emphasized the large number of workshop and conferences involving lattice contributions over the past year. Considerable progress in exploring the QCD phase diagram using charge fluctuations has been achieved in 2014, with an emphasis on the computation of freeze-out lines that should impact the next set of runs scheduled for RHIC.

Concerning the RHIC program, the reviewers commented that the crossover transition in the QCD phase diagram at vanishing baryon chemical potential, μ_B, is a mature result and is now well understood. More recently, the Beam Energy Scan (BES) program at RHIC puts the spotlight on the region of finite temperature, T, and non-zero baryon chemical potential, μ_B. The push by the LQCD community into this region is currently well controlled up to about $\mu_B / T \sim 2$. Fluctuation observables that are accessible to experiment have been calculated. LQCD has the potential to predict a critical point and its location on the QCD phase diagram.

An important future goal is to extend the highest investigated μ_B / T values from ~2 on up to ~3. Strangeness fluctuations are being calculated, and these are only now beginning to get experimental attention, in addition to the usual net-baryon fluctuations (or proxies). Freeze-out curves on the phase diagram are normally extracted in a hybrid experiment + model approach from fitting experimental particle spectra using statistical models, but lattice now offers an independent and fundamental theoretical prediction of such freeze-out lines. Transport coefficients and spectroscopy of heavy flavors at top RHIC energies are also areas where lattice is expected to make progress in the future. The reviewers encouraged such efforts.

Additional RHIC physics predictions in the areas of in-medium properties of hadrons and calculations of thermal di-leptons should be possible in the near future.

The reviewers were impressed by the fact that there are many LQCD projects in areas relevant to HEP and NP experimental programs, even relevant to low energy nuclear physics and fundamental symmetries. USQCD holds workshops at least annually to connect lattice QCD research to experimentally relevant questions. The goals of USQCD, expressed in regularly revised whitepapers, are developed in consultation with experimenters. The Scientific Program Committee (SPC) makes relevance a part of the allocation process for computer time.

The deployment of LQCD hardware often leads to design ideas and prototyping that is useful to other programs at labs and universities. One example that was given during the review presentations is that at Fermilab for several years Ds-type machines (quad-socket Opterons) have been the standard used for Run 2, FermiGrid, and CMS Tier 1. Other example are the GPUs and Intel MIC architecture, for which there seems to be interest from other programs at the labs.

Comments

The reviewers were generally very impressed by the past progress and future promise of the entire LQCD-ext II project, and they supported it enthusiastically.

In general, the reviewers were very impressed by the scientific productivity and focus of the USQCD users that run simulations and analysis efforts on the LQD-ext II hardware.

While extending in new directions, the project's significance in more traditional lattice QCD applications, e.g.in support of searches for new physics in the flavor sector, remains as strong as ever. Accuracy of the calculations of some quantities has reached the precision required by the present experimental data. However, accuracy of calculations of some other key quantities e.g. related to mixing in the B meson sector, or to |Vub| determination from exclusive B decays, is still significantly behind the accuracy of the experimental results. New experimental methods e.g. determination of |Vub| from B decays, call for a matching effort on the LQCD side. In the near future, the DOE supported Belle II project will bring a new level of experimental accuracy to some of the observables in the flavor sector, with complementary results from the NSF supported LHCb. Thus, the relevance of lattice QCD calculations in the flavor sector will remain strong for many years to come and will directly impact the discovery potential of these important new physics searches.

The relevance of LQCD calculations in support of cold and hot nuclear physics is strong and growing.

Several reviewers were worried that new hadron spectroscopy results from high energy physics experiments related to four-quark phenomena observed in the heavy-heavy-light-light sector is not receiving enough attention from the LQCD community. While the high-energy experiments don't articulate these results as their top priority, they are equally important as searches for exotic states among lighter hadrons being carried out by the nuclear community. The present status of phenomenology of four-quark states clearly calls for input from the fundamental theory of strong interactions. While putting such states on the lattice is difficult, it is worth engaging in such a program as more experimental data is guaranteed to come from Belle II and on-going LHC experiments (LHCb and CMS have published important results in this sector). The existence and forms of four-quark, or six-quark states impacts not only high energy physics but potentially astrophysics.

In RHIC physics the reviewers commented that lattice theory complements other theoretical approaches. However, it can predict only a subset of the experimental observables. Compared with other approaches like Boltzmann transport and hydro models, it is more fundamental and less of a "black box". Those alternative theory approaches have various ad hoc parameters and assumptions. So these considerations confer a major advantage to LQCD.

The physics goals in the LQCD area are very important to the RHIC community and are highly deserving of DOE support. The detailed hardware plans and milestones of the LQCD project are as well defined as could be expected given the natural uncertainties associated with planning in any area where there is so much rapid progress in technology. However, in contrast, the physics prediction goals of USQCD, and especially their relative priorities, are vague in the documentation presented to the review committee and several reviewers suggested that specific physics goals, in priority order, should be defined. In the Q&A sessions at the review, there was some information provided along these lines, but it would be helpful to future reviewers to have this kind of information, across all areas of LQCD physics, prominently featured and kept up-to-date in the written year-by-year record. Some rather dated material buried in whitepapers does not do the job.

The BSM computations are speculative and have not developed at a rate commensurate with the other scientific areas covered by USQCD. Composite models of the Higgs mechanism are not popular outside LQCD because there is no hint for such physics from the high precision experiments in the subject. The challenge to the field is not just to predict a 125 GeV Higgs. In addition, the Higgs must have the decay modes of the elementary scalar field of the textbook Standard Model, now measured by the LHC to fair and improving precision, and there must be no statistically significant "anomalies" in heavy quark experiments performed at the LHCb or B factories. Data from astrophysics now favor weakly interacting Dark Matter (DM) sectors, in marked contrast to the expectations from composite Higgs models, and there are improving constraints on DM-DM cross sections from dwarf galaxy observations. These experiments will be improved over the coming years and constrain models of DM. Much of this physics, as well as composite Higgs models' long standing inability to explain the fermion masses of the Standard Model, especially the top quark, are being ignored by the lattice BSM efforts. This USQCD effort, and the presentation at the review in particular, were graded poorly by several of the reviewers.

With the exception of the lattice BSM effort, the reviewers felt that the USQCD collaboration and the LQCD project remain very relevant to the experimental programs in HEP and NP. The progress of USQCD toward quantitative predictions in these fields is impressive and clearly dependent on the LQCD hardware projects. The possibilities of calculating nuclear excitation spectra from lattice QCD, as mentioned by Paul Mackenzie, is very exciting. It will require long-term support for this project to achieve this challenging goal.

Several reviewers were very impressed with the concept of the LQCD hardware facility, which leverages the infrastructure and manpower resources at 3 different DOE laboratories. They were equally impressed with the technical and managerial skills of the existing personnel supported on the project. Their multiyear experience in building and operating the hardware facilities was apparent. Their focus on maximizing scientific output of the users of the facilities, spanning different particle physics and nuclear communities is commendable. The facilities are a key asset in the broad experimental program supported by the DOE; interpretation of the data from several experimental undertakings relies on continued parallel efforts by the lattice community to provide numbers bridging the measurements with their interpretations. The facilities have stimulated growth of the lattice QCD field which is now even more relevant to a broader class of measurements. Since the hardware budget has been substantially reduced for this funding cycle, the productivity of the field will likely be limited by available resources, especially when the older equipment is retired. While this may be unavoidable in this era of declining science budgets, supplementing the hardware allocation should be seriously considered if additional funds become available.

All the reviewers commented that it is "disheartening" to see that the funding has decreased from previous 5 years, despite the impact that Lattice QCD calculations have on the field. The hardware resources seem to be well below the needs of the community, which clearly has an impact on science productivity. Since any additional funding would go directly to hardware, increased support, if available at any time in this 5 year cycle, would have a significant impact on the science output.

The reviewers were also impressed by the dedication and sophistication of the lattice physicists playing key roles in the USQCD collaboration, which manages the resource allocation to individual PIs and theory groups using the facilities.

Recommendations

None.

Progress towards Scientific and Technical Milestones

Findings

Bill Boroski, the LQCD-ext contractor project manager, presented the management and performance information for the project. He presented 1. The project scope, organization, and budget, 2. Performance measures and metrics, 3. FY15 year-to-date performance results, 4. FY15 year-to-date financial results, and 5. The project summary.

The User Survey results were presented by the associate contract manager at Fermilab, Rob Kennedy.

In addition, Bill Boroski also presented the closeout report of the LQCD-ext project which ran from FY 2010 through FY 2014. He presented the 1. The project scope, organization, and budget,
2. Performance measures and metrics, 3. FY14 performance and financial results, 4. Total project performance and financial results, and 5. A Project Summary.

The hardware selection strategy and acquisition plan for the FY 2015-16 deployment was presented by Chip Watson, the technical contractor project manager at JLAB.

The reviewers heard that the technical goals of the LQCD-ext II are to acquire and operate the best mix of affordable high performance computing for the portfolio of jobs to be done. Success depends on a balance of DOE/NSF leadership computing and LQCD resources. This project is building on many years of previous projects; LQCD-ext II is just beginning, planning for continued operations of existing and new facilities with an anticipated lifetime of 4 to 5 years.

The reviewers commented that the technical milestones are documented in exemplary detail in the review materials.

Comments

Many technical (hardware) metrics have been presented, and the great majority have been met or exceeded. The occasional exception has been studied carefully and remediation steps taken. Since this is an IT hardware project, these technical milestones are arguably the crucial ones.

The standard design is to build systems of many inexpensive, high performance nodes, connected by high performance interconnections like Infiniband. This is an excellent solution with great scalability in the face of fluctuating budgets. In addition, such architectures permit very effective use of any additional funds that become available in immediate expansion of existing systems.

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.

One reviewer pointed out that physics deliverables are the ultimate objective, and that science milestones were conspicuous by their absence at the review. Notwithstanding the challenges of defining physics milestones in the special context of this hardware "facility-like" project and assessing progress towards them, more can be done. 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.

Several reviewers commented on the good scores that the project received from the users in their annual user survey. However, since the number of users is relatively small, the user survey results are difficult to interpret in some instances and meaningful annual trends are hard to uncover. The reviewers suggested that alternative methods 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. These in depth approaches might yield more useful measures of success and failure compared to the short answers obtained on electronic surveys.

Recommendations

None.

Technical Design and Scope for FY2015

Findings

Amitoj Singh, who assists Don Holmgren at the Fermilab site, presented the technical performance of the FY 2014 cluster deployment.

The technical design and scope for FY 2015 under this project is limited to operations and planning for the future. The Pi0 cluster expansion at FNAL in 2015 is based on FY14 carryover from the LQCD-ext project. As planned, the project will lose 45TFlop/s-yr between FY15 and FY16.

Although there were no projected procurements in FY 2015, the team is preparing for procurements in the upcoming fiscal years. The procurement/architecture decision process contains several stages. First, they perform research, where they are tracking what NVIDIA, Intel, AMD, etc are doing. To do so effectively some of the members of the team have non-disclosure agreements with the manufacturers. Second, they use forward looking designs. This is the stage where a cost benefit analysis is being done. One of the possibilities that they consider is also the option of not purchasing anything in some of the fiscal years. They do alternative option analysis, and compare it to the default plan. Their main aim is "if we buy this, which science code could be run on this", i.e. they are interested in optimizing science. They also consider whether they should invest in some other technologies. For instance the extra half-rack of BlueGene/Q in FY13 was put at BNL since the relevant expertise was already in place there.

In the final stage they perform adaptive optimization. This is the final tweaking of procurement at the last minute so that the science would be optimized (for instance they have doubled memory per chip reducing number of teraflops, but increasing the science of the Pi0 cluster). This also involves the configuration of network, etc. This is at the level of 10% of the cost, fine tuning. A smaller part of the purchase price are also the warranties. For the FY14/15 hardware acquisition the warranty cost was 3% of the cost ($75,600).

In short the project decides year by year what is the best for science. BNL has recently expressed interest in hosting a project-funded Intel Knights Landing computer in FY 2017. This will be factored in for the replacement of their BlueGene/Q. It may be also possible that the BlueGene/Q could be run beyond FY18, if this is more cost effective.

Comments

The reviewers were impressed with the attention to detail in the review materials and the considerable expertise of the project technical personnel.

The current work at FNAL to maintain existing 4-year-old capacity computing platforms is a useful effort to deal with the lower funding in FY 2015 and the delays in new acquisitions. Reasonable efforts should be made to preserve the overall capacity computing of the collaboration. The Pi0 upgrade of 6 TF/s-yr with FY14 carryover funds helps to alleviate some of the stress.

Recommendations

None.

Feasibility and Completeness of Budget and Schedule

Findings

The budgets for each year of the project are very constrained and challenging. The allocation of $2M for FY15 supports operations only. The operations staff for LQCD at each site is shared with the larger compute infrastructure within the host laboratory. The procurement strategy is to acquire large systems in one year with options to extend the procurement in succeeding years. Two such acquisitions are intended. In years FY 2015, FY2016 and FY 2017, hardware capacity increases will not keep pace with the loss due to capacity computing retirement from age and reduced reliability until FY18.

The project team is doing a balancing act to optimize science with the existing budget. They have very impressive access to vendors and are trying to leverage their expertise. For instance, NVIDIA is paying a person to do part of the software development for their platforms. The team is also considering other saving measures (e.g., reusing the racks).

Comments

Budgets are clearly allocated with great care, and no effort is spared to maximize physics deliverables per dollar spent.

The project budget is very tight, being mostly needed for operations of existing capacity. The operations budgets are very lean and efficient, with little room for further economy. The operations staff strategy is very effective a providing technical depth at a minimal cost.

The project management's careful planning will keep the program going with substantial added capacity by FY19. The procurement strategy is a very effective use of the funds provided for hardware acquisition. Any additional funds will be effectively used to increase the facilities' overall capacity.

Recommendations

None.

Effectiveness of Management Structure and Responsiveness to past Recommendations

Findings

The management structure of the project consists of the contractor project management at FNAL, and the site managers at FNAL, BNL and TJNAF, closely interacting with the USQCD Executive Committee and Scientific Program Committee. Each of the site managers has a fully capable deputy, and there is an Associate Contractor Project Manager at FNAL.

Comments

The management structure is effective at keeping this project performing well. The depth of the management team appears to be good and well planned.

The project responses to the recommendations from 2014 were positive and reasonable.

Recommendations

None.

Effectiveness of USQCD, Scientific Impact, Procedures and Related Activities

Findings

Scientific impact of this project is strong and has been living up to its promised expectations.

The function of USQCD was compared with the operation of the FNAL fixed target beam facility.

A number of workshops are organized or attended by the USQCD executive members, which connect the organization with various experimental and theoretical communities.

USQCD collaboration has developed an elaborate resource allocation procedure, which weighs scientific merits of each request and maintains a good balance between different sub-communities using the facility. Matching the requests with the most suitable hardware is also performed. The process includes additional interactions with the proponents before the allocation decisions are made.

The requested computational resources always exceed what's available, but generally by less than a factor of two. So far, the Scientific Program Committee (SPC) has always approved at least a fraction of the request in each proposal, with the fraction being close to one in the most meritorious cases. Coming constraints, in view of the slower growth and possible near term reduction of available compute nodes may require more difficult decisions.

Past reviews (over several years) have urged a less top-down management structure for the Executive Committee (EC).

The Scientific Advisory Board members offer opinions on the white papers produced by USQCD. They choose not to be involved in the allocation process. USQCD is not required to take their advice.

Comments

The reviewers pointed out that the USQCD organization is well connected with the relevant groups of the experimental and theoretical programs.

The reviewers agreed that the project's resource allocation process is mature and very reasonable.

The reviewers had several comments on the makeup and rollover of the management of USQCD.

They 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 experienced people doing the bulk of the work. This would help train the younger generation in this difficult task and make the process more transparent to the broader community. Such positions could be open to self-nominations and to nominations from other lattice QCD physicists in the U.S. All PIs who were previously allocated resources at the LQCD facility could be allowed to cast election votes.

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. (A typical definition of "junior" is anyone less than 5 years past their PhD.) This slot could rotate on a different time schedule from "normal" members of the EC. This will place some burden on just one young researcher, but generally speaking, there are always many junior physicists who are happy to volunteer to stand for election to this kind of position.

Several reviewers thought that the SAB can be used more effectively. For example, at accelerator facilities, there is an annual face-to-face meeting between SAB members, the management and PAC, at which advice is sought about the program and, perhaps, consensus is sought. Perhaps the SAB members could be engaged more actively in the processes of the USQCD through participation in the All Hands Meeting.

The SPC follows a careful and deliberative process for assessing submitted proposals and allocating the available resources. Surveys verify that the users have few complaints and are generally satisfied with the fairness and transparency of the decisions. However, several reviewers felt that more specific and quantitative measures of allocation effectiveness should be developed. The review committee felt that they could not answer several simple but important questions concerning allocation effectiveness using just the material presented at the review. These questions included: How efficiently are the processors used by the grants? How is research output related to the size of the allocations? Perhaps USQCD could develop metrics to quantify the success of the SPC's allocation process. These metrics could be incorporated into the presentations at the Annual Progress Reviews.

Overall, the reviewers concurred that this project is managed very effectively and the community gets an excellent physics return on the DOE resources invested. They were impressed by the careful and comprehensive approach taken to maximizing the computing power year-by-year, as well as the month-by-month recording of the availability of the various computing platforms at all three labs.

Recommendations

None.

APPENDIX A

Charge Letter to the LQCD-ext II Project Team

Dr. W. Boroski
LQCD Contractor Project Manager
Fermi National Laboratory
Mail Station: 127 (WH 7W)
P.O. Box 500
Batavia, IL 60510-0500

Dear Dr. Boroski:

The Department of Energy (DOE) Office of High Energy Physics and the Office of Nuclear Physics plan to conduct an Annual Progress Review of the Lattice Quantum Chromodynamics (LQCD-ext II) Computing Project on May 21-22, 2015, at the Brookhaven National Laboratory (BNL). A review panel of experts in high energy physics, nuclear physics, project management and computer science is being convened for this task.

John Kogut of the Office of High Energy Physics is responsible for this review; he will be assisted by Elizabeth Bartosz and Ted Barnes of the Office of Nuclear Physics.

Each panel member will evaluate background material on the LQCD-ext II project and attend all the presentations at the May 21-22 review. The focus of the 2015 LQCD-ext II Annual Progress Review will be on understanding:

- The continued significance and relevance of the LQCD-ext project, with an emphasis on its impact on the experimental programs' support by the DOE Offices of High Energy Physics and Nuclear Physics;
- The progress toward scientific and technical milestones;
- The status of the technical design and proposed technical scope for FY 2015;
- The feasibility and completeness of the proposed budget and schedule;
- The effectiveness of the proposed management structure, and responsiveness to any recommendations from last year's review.

Since LQCD-ext II provides computer cycles that are distributed by the USQCD collaboration, the panel members will also consider:

- The effectiveness of USQCD in allocating the LQCD-ext II resources to its community of lattice theorists, the scientific impact of this research on the entire HEP and NP communities and the status, operational procedures and related activities of the USQCD collaboration itself.

The two days of the review will consist of presentations and executive sessions. The later half of the second day will include an executive session and preliminary report writing; a brief close-out will conclude the review. Preliminary findings, comments, and recommendations will be presented at the close-out. You should work with John Kogut to generate an agenda which addresses the goals of the review.

Each panel member will be asked to review those aspects of the LQCD- project listed above which are within their scope of expertise and write an individual report on his/her findings. These reports will be due at the DOE two weeks after completion of the review. John Kogut, the Federal Project Manager, will accumulate the reports and compose a final summary report based on the information in the letters. That report will have recommendations for your consideration that you and USQCD should respond to in a timely fashion.

Please designate a contact person at BNL for the review panel members to contact regarding any logistics questions. Word processing, internet connection and secretarial assistance should be made available during the review. You should set up a web site for the review with relevant background information on LQCD-ext II, links to the various LQCD-ext II sites the collaboration has developed, and distribute relevant background and project materials to the panel at least two weeks prior to the review. Please coordinate these efforts with John Kogut so that the needs of the review panel are met.

We greatly appreciate your willingness to assist us in this review. We look forward to a very informative and stimulating review at BNL.

Sincerely,

James Siegrist
Associate Director
Office of High Energy Physics

Timothy Hallman
Associate Director
Office of Nuclear Physics

APPENDIX B

Reviewers for 2015 LQCD-ext II Annual Review (BNL May 21-22)

HEP Phenomenology

Jure Zupan jure.zupan@uc.edu.

HEP Experiment

Tomasz Skwarnicki tskwarni@syr.edu

HEP Computer Specialist

R. Jeff Porter (LBNL) RJPorter@lbl.gov

NP theory

Eric Swanson swansone@pitt.edu

NP Computer Specialist

Robert Varner varnerrl@ornl.gov

NP Experiment (Heavy Ions)

Declan Keane Keane@kent.edu

List of attending DOE program managers

J. Kogut (HEP, LQCD-ext II Federal Project Director)

T. Barnes (NP)

E. Bartosz (NP, LQCD-ext II NP Project Manager)

APPENDIX C

Review Agenda

DOE Annual Progress Review of the
Lattice Quantum Chromodynamics (LQCD) Computing Project
LQCD-Ext II

May 21-22th, 2015

Brookhaven National Laboratory
Berkner Hall, Room B

Agenda

—

May 21
08:30 Welcome (15 min) – ***Robert Tribble, Deputy Director for Science and Technology, BNL***
08:45 Executive session (45 min)
09:30 Logistics and Introductions (5 min) – ***Bill Boroski***
09:35 LQCD-ext II Overview, USQCD Governance (incl. Allocations) (55 min) – ***Paul Mackenzie***
10:30 Break (15 min)
10:45 Science Talk 1: Cold Nuclear Physics (30 min) – ***Will Detmold***
11:15 Science Talk 2: QCD Thermodynamics (30 min) – ***Frithjof Karsch***
11:45 Science Talk 3: Beyond the Standard Model (20 min) – ***Anna Hasenfratz***
12:05 Lunch / Executive Session
1:05 Science Talk 4: QCD for Particle Physics (40 min) – ***Andreas Kronfeld***
1:45 LQCD-Ext: Technical Performance of FY14 Cluster Deployment (15 min) –

Amitoj Singh
2:00 LOCD-Ext: Project Close-out (30 min) – *Bill Boroski*
2:30 Coffee Break (15 min)
2:45 LOCD-Ext II: Project Management and Performance (60 min) – *Bill Boroski / Rob Kennedy*
3:45 LOCD-Ext II: FY16 Hardware Acquisition Planning (30 min) – *Chip Watson*
4:15 Executive Session (60 min)
5:15 Committee request for additional information – *Review Committee & Project Leadership*
6:00 Adjourn
7:00 Dinner

May 22
8:30 Executive Session (30 min)
9:00 Committee questions and discussion (60 min)
10:00 Break (15 min)
10:15 Executive Session / Preliminary Report Writing
12:00 Lunch
1:00 Closeout (60 min)
2:00 Adjourn