Offices of High Energy Physics and Nuclear Physics
Report on the

LQCD-ext II
2016 Annual Progress Review

June 28-29, 2016

Executive Summary

The Annual Progress Review of the LQCD-ext II (Lattice Quantum Chromodynamics extension II) project was held on June 28-29, 2016 at the Thomas Jefferson National Accelerator Facility (TJNAF or "JLAB"). 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:

6. 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.

There were also two special charges for this year's review:

7. Proposal to re-combine separate CPU and GPU performance metrics into a single metric to accommodate new "hybrid" hardware architectures, and
8. Proposal to add Brookhaven National Laboratory (BNL) as a third cluster-hosting site.

Both of these special charges are formal Change Requests to the LQCD-ext II program managers and require special actions in accord with project management best practices.

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, dissemination of scientific results, and the two Change Requests listed above. 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. Previous review panels have suggested that USQCD consider electing more youthful lattice gauge theorists into the higher positions of the collaboration in light of its aging demographics, and these recommendations are being enacted. The reviewers did, however, have several suggestions that the project team should address:

1. The reviewers argued that the Physics Beyond the Standard Model (BSM) efforts of USQCD should focus on general properties of strongly coupled extensions of the Standard Model instead of concentrating on various particular models which are unlikely to be realized in nature.

2. The reviewers suggested that more emphasis be placed on investigations of new simulation algorithms for non-zero chemical potential in light of the limitations of expansion methods which the BNL group has specialized in.

3. The review panel was generally impressed by the results in the annual user survey, but they suggested adding questions to involve the users in the annual hardware choices.

4. The review panel encouraged the project to communicate more with several of the ASCR, NERSC and SciDAC efforts on allocation and acquisition procedures, as well as software and porting experiences.

5. Several reviewers suggested that USQCD management should continue their efforts to persuade their users across all LQCD areas to regularly address the project's physics milestones and to document the progress toward these long range milestones more systematically.

Further guidance and essential specifics of these suggestions can be found in the body of the review report.

Finally, the reviewers unanimously endorsed both Change Requests. The first Change Request was seen as a convenient and sensible change in light of new hardware platforms, such as Intel's Knight's Landing (KNL), which straddle CPU and GPU architectures. Presentations on the second Change Request showed that the project had negotiated favorable terms with BNL management so that BNL's in-kind contributions to the project more than offset the cost of adding another national laboratory to the cluster-hosting element of the project.

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 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 were 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 would be 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 FY2014 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 TJNAF on June 28-29, 2016. 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 present 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 in their charge letter plus the two special timely Change Requests. The first Change Request consists of the proposal to add BNL as a cluster-hosting site. In the past BNL has hosted special purpose hardware constructed by the Columbia group and BlueGene/Q's purchased from IBM. BNL is in the process of expanding and broadening its involvement in scientific computing and has made an attractive offer to the LQCD_ext II project to join Fermilab and JLAB and begin designing, constructing and operating clusters. The second Change Request is the proposal to replace separate performance metrics for CPU and GPU clusters with a single simpler metric. The motivation for this proposal is the fact that commercial vendors are producing chips which incorporate the advantages of CPUs and GPUs on a single board and these new architectures are prime candidates for hardware acquisitions.

LQCD-ext II Review

Continued Significance and Relevance

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

1) QCD for Particle Physics. 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, for the second year running. 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 Standard Model theory prediction before the experiment's data analysis scheduled for 2018-19. Recent improvements in the most difficult part of the calculation, light-by-light scattering, have been very promising. The recent productivity of the Intensity Frontier lattice effort has been good with 3 Physical Review Letters in the last 12 months and a total of 14 publications in referred journals.

2) Beyond the Standard Model. 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 composite Higgs models, composite models of Dark Matter and lattice versions of Supersymmetry. Several of the most interesting models are "almost" conformal although they employ familiar gauge groups (SU2, SU3, SU4,…) but have many species of massless "quarks" in various representations of the gauge group. Calculations which accommodate the Higgs at 125 GeV/c^2 as a pseudo-Goldstone boson also predict additional states accessible to the LHC 14 TeV run. Investigative studies of supersymmetry are also underway. GPU clusters are proving particularly useful in these studies. Ethan Neil summarized this subfield of lattice gauge theory at the review. He emphasized that this work is exploratory and only accounts for 5-10 % of the total USQCD efforts. Over the last year there have been five publications in referred journals in this subfield. In addition, there was a workshop at Argonne National Laboratory on April 21-22 where the lattice community interacted with theorists, phenomenologists and experimentalists in the field.

3) Cold Nuclear Physics. 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. William Detmold summarized this subfield of lattice gauge theory at the review for the second year running. Recent developments include: New developments in the calculations of parton distribution functions in both longitudinal fraction and transverse momentum, studies of the origin of the proton spin, higher order color exchange effects in heavy quark systems and nucleons, pentaquark and tetraquark calculations, neutrinoless double beta

decay calculations and neutrino-nucleon cross-sections. The productivity of this lattice subfield has been strong in the last 12 months with two Physical Review Letters and 10 publications in referred journals in all. The job market in the field has also shown considerable recent improvements with five junior faculty appointments in the 2015//2016 timeframe.

4) QCD Thermodynamics. 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. Considerable progress in exploring the QCD phase diagram using charge fluctuations has been achieved since 2014, with an emphasis on the computation of freeze-out lines that should impact the next set of runs (Beam Energy Scan II) scheduled for RHIC. It is hoped that the next RHIC run will explore that portion of QCD parameter space in chemical potential and temperature where the critical endpoint of the first order transition line between the quark-gluon plasma and the cold hadronic phase exists. S. Mukherjee summarized this subfield of lattice gauge theory at the review. There have been two workshops and conferences over the last twelve months focusing on the Beam Energy Scan II. The productivity of the subfield has been good with 5 publications in referred journals over the past year.

The reviewers reported findings on these four scientific areas:

Findings

QCD for Particle Physics

The LQCD work on quark-flavor physics allows for the extraction of fundamental constants of nature, the CKM elements, that cannot be measured in any other way. It also provides a way for experimental results, primarily of the Belle II and LHCb experiments, and also in some cases for CMS and ATLAS, to be used to test for physics beyond the standard model indirectly, by seeing if measurements of the CKM elements, mixing and CP violation are consistent or if new physics is needed. These indirect measurements are sensitive to masses of new particles in the 10-1000 TeV range, depending on the process. In addition, a precision calculation of the form-factors in $B \to D^* \tau \nu$ and $B \to D^* \mu \nu$ could be decisive in deciding whether current experimental indications of a violation of lepton universality in the decay ratio are due to new physics or merely to QCD corrections.

Beyond the Standard Model

Strongly coupled field theories may play a role in new physics beyond the SM e.g., dark matter or weak scale physics related to the hierarchy puzzle.

Cold Nuclear Physics

Understanding the spectrum of QCD (e.g. are there narrow penta-quark states, are there excited states of glue?) is very important scientifically. Only lattice gauge theory can make defensible contributions to this subfield.

QCD Thermodynamics

The crossover transition in the QCD phase diagram at zero baryon chemical potential (μ_B) is a mature result that emerged from the LQCD work of past years. The current best value for the temperature T at the crossover in the limit of $\mu_B = 0$ is determined from lattice calculations to be about 155 MeV, with an uncertainty under 10 MeV.

The Beam Energy Scan (BES) program at RHIC has put the spotlight on the region of non-zero μ_B. As the beam energy is scanned below the top RHIC energies, produced baryon-antibaryon pairs no longer overwhelm the initial-state baryonic matter, and the region of non-zero baryon chemical potential becomes increasingly important. The push by the worldwide LQCD community into this region has resulted in a promising description up to $\mu_B /T \sim 2$. For example, in a 2016 paper by members of USQCD [A. Bazavov et al., PRD 93, 014512 (2016)], an NLO Taylor expansion is compared with net-charge fluctuation data from RHIC experiments, with a focus on the particle freeze-out line in the phase diagram. A noteworthy conclusion is that for center of mass energies per nucleon pair around 27 GeV and above ($\mu_B /T \sim 1$ and below), good agreement in the freeze-out properties is found. This paper further supports the inference that the search for the possible critical point should be pursued below 27 GeV.

An important thrust of current LQCD research is to obtain useful predictions at values of μ_B /T as high as ~3, which will extend coverage down to the lower energies where experimental results from Phase-I of the RHIC Beam Energy Scan hint that new physics may be found. Terms as high as 8th order will be calculated. Important progress towards this goal is expected during the coming year.

Comments

QCD for Particle Physics

The burgeoning effort in computing the hadronic vacuum polarization and hadronic light-by-light contributions to the anomalous magnetic moment of the muon is very welcome. Lattice QCD is clearly crucial for the success of this DOE experimental program.

Beyond the Standard Model

Since BSM physics is still speculative and any particular model has only a small chance of being relevant, the emphasis in this area should be on the structure of strongly coupled field theories (QCD with different representations for the matter fields, conformal field theories, etc.) rather than particular models.

The funding that goes into this area is a small fraction of the total LQCD budget but nonetheless this research is important and should continue.

Cold Nuclear Physics

Using lattice QCD to reduce uncertainties in nuclear matrix elements relevant for neutrino-less double beta decay experiments will be impactful to the experimental program of the DOE.

QCD Thermodynamic

Lattice theory complements other theoretical approaches in this subfield. It can directly predict a subset of experimental observables, such as various fluctuations. High moments of fluctuation observables are involved in one of the most promising signatures of a potential QCD critical point. Compared with other approaches like Boltzmann transport and hydrodynamic models, which involve various ad hoc parameters and assumptions, lattice QCD is more fundamental and less of a black box. These considerations all confer a major advantage on LQCD. In addition, LQCD provides essential inputs to Boltzmann transport and hydro models, and to hybrid models combining these two approaches.

At present, the DOE is getting ready to invest ~ $180 million in Phase-II of the RHIC Beam Energy Scan (BES-II), beginning in 2019. The goal is to search for a line of 1st order phase transitions and especially a QCD critical point. Already decisions have been made that have narrowed the focus of BES-II in terms of the beam energies to be covered. LQCD at non-zero μ_B has been very important in helping to shape those decisions about energies, and also other decisions about BES-II. LQCD will continue to be crucial between now and 2019, and then on until the end of BES-II running and subsequent analysis. This is the key point that strongly justifies the current LQCD ext-II project from the perspective of RHIC experimentalists.

All of the above means that future progress on LQCD at finite μ_B and the "sign problem" are extremely important. The USQCD community is perceived by some observers as perhaps being too cautious and relying too heavily on the "brute force" approach of Taylor expansions, whereas other practitioners devote more effort to exploration of higher-risk approaches like the complex Langevin method and the Lefschetz thimble. Those speculative investigations have not met with any real success so far, and so the conventional approach of the USQCD community has in this sense been vindicated. Nevertheless, it sometimes happens that a clever new idea comes to light, and suddenly an unproductive direction turns into a breakthrough. Some alternative explorations are indeed occurring within USQCD, but their visibility is very low. It would be good if the USQCD community raises the visibility of their more speculative/high-risk efforts in this area, so that the broader community can be assured that they will be ready to pivot if they, or a non-USQCD group, suddenly makes a breakthrough.

The computation of transport coefficients relevant to the RHIC experiment appears to be a new component of the hot QCD program. This is an interesting and potentially impactful development.

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. FY16 year-to-date performance results, 4. FY16 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.

The hardware selection strategy and acquisition plan for the FY2016-17 deployment was presented by Chip Watson, the technical contractor project manager at JLAB.

The project has received a total of $5M in funding in FY15 and FY16. The project’s five year base budget is $14M.

A significant portion of the $5M (over 68%) has gone into personnel. Only about 17% is/will be spent on hardware procurement in FY16.

The FY15 actual cost ($2.442M) was higher than the FY15 funding received ($2M). The project used a portion of its $0.8M carryover from FY14 to accomplish this. The spending trend in FY16 is on-track.

The hardware performance achieved during the first two years of the project is exceeding the baseline goals, both on the conventional clusters and GP accelerated clusters.

The hardware utilization of almost all clusters has been high. The exception is the DSG, which is a GPU-accelerated cluster at FNAL, but its shortfall is due to the expected failure of aging hardware.

The FY15 survey results indicated that the users were satisfied with the project's hardware choices and their operations.

Technical milestones are documented in exemplary detail in the review materials. Many precisely-defined technical (hardware) metrics have been defined for the current report period, and they have been met or exceeded.

Comments

Overall, the progress of the LQCD-Ext II project is good. The project has met their technical hardware milestones flawlessly.

The team should be commended for being able to deliver more performance than the original goal.

Only 50% of the active users responded to the survey. The reviewers would like to see an effort by the project team to increase this percentage.

Physics deliverables are the ultimate objective of the project. The definition and documentation of science milestones seem to be quite good in some LQCD areas, and lacking in others. The project should develop procedures to document scientific milestones uniformly over all the LQCD areas so that the project can track their annual progress more quantitatively.

Recommendations

None.

Technical Design and Scope for FY2016

Findings

Chip Watson presented the technical performance of the FY2016 cluster deployment.

The acquisition process includes an analysis of available architectures and bench-marking activity.

A decision has been made to acquire a cluster based on the Intel Knights Landing (KNL) chips.

The KNL cluster is expected to have 192 compute nodes with 192 GB of main memory per node. There is an option to expand the cluster to have a total of 256 compute nodes in FY17.

New Lustre file servers are in the procurement plan.

Comments

The FY16 acquisition plan presented at the review was very thorough, detailed and impressive.

The project team described the choice of hardware to be "organic". It appeared to rely on one or more members of the USQCD collaboration to explore new architectures. This seems to be somewhat ad hoc and it lacks input from the user community at large.

The reviewers would like the project team to seek broader engagement with the users on future architecture needs. One possibility would be to add specific questions in the user survey to seek input on, for example, currently available clusters and hardware preferences for future procurements.

The architectures currently considered by the project team are very similar to those available at the ASCR NERSC Center and the two Leadership Computing Facilities (LCFs). Since those three centers have strong track records in hardware procurement, allocation processes, and user support, the project team should develop closer communications and interactions with them.

There appears to be a considerable burden on the users to port their codes to new architectures delivered by the project. While the LQCD-ext II project is entirely a hardware project and doesn't provide support for software development, it should consider, as part of user support, a role in connecting the users with ASCR researchers in order to advance algorithmic development and code tuning/optimization. The SciDAC program provides a mechanism to achieve this. In fact, the project's documents and presentations referenced the SciDAC Program several times. The SciDAC Program funds computational mathematicians and computer scientists through the SciDAC Institutes. Perhaps the LQCD-ext II team should invite members from the SciDAC Institutes to participate in the All-Hands Meetings in order to facilitate interactions and collaborations between LQCD and math/computer science communities.

The technical team is experienced and well-organized. Thorough vetting of the proposed KNL cluster has been made with an examination of a variety of alternative acquisition scenarios.

Recommendations

None.

Feasibility and Completeness of Budget and Schedule

Findings

The total budget of the LQCD-ext II Project is $14M over 5 years, with $9M and $5M coming from the Offices of HEP and NP, respectively.

The project started in 2015, so the project is in the second year of its operations.

About 57% of the budget goes into personnel, which includes system administration, engineering and technical labor, site management, and project management.

About 33% of the budget goes into compute/storage hardware.

There is a detailed breakdown of the hardware budget over the lifetime of the project and a well-planned path of hardware procurement and the development of performance goals.

There was no hardware procurement in FY15.

Hardware procurement spending in FY16 is less than each remaining year of the project which ends in FY19.

Comments

The project team is managing the budget with care.

The performance goal dropped between the first and second years of the project, because of the retirement of aging clusters. It is encouraging to see that the planned performance goal in the final year of the project is expected to be twice that of the first year.

The budget allocation for hardware in the last year of the project (FY19) is significantly larger than other years. There doesn't seem to be a good rationale for this decision. The reviewers questioned if it is a wise decision to have a large procurement during the final year of the project.

Recommendations

None.

Effectiveness of Management Structure and Responsiveness to past Recommendations

Findings

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

The roles and responsibilities of various parts of the management structure are clearly documented.

The role of 'site architect' has been added to the organizational structure of the collaboration.

Comments

The management structure is clearly defined. It is functioning well.

One particularly impressive aspect of the project management is their responsiveness to changes in hardware availability. They are on top of the latest developments and able to make the right choices to maximize the physics output for the project.

Recommendations

None.

Effectiveness of USQCD, Scientific Impact, Procedures and Related Activities

Findings

The Executive Committee has ten members, most of them senior. The committee is balanced by lab, group, and field. The Scientific Program Committee (SPC) reviews proposals and allocates hardware. The Scientific Advisory Board (SAB) provides informal reviews and general feedback on physics.

This year, USQCD management has put in place a new process whereby a junior researcher is elected by his or her peers to serve a two-year term as their representative on the Executive Committee. Will Detmold has been duly elected.

The computational resources that USQCD members request in their proposals always exceed what is available, but generally by less than a factor of two. So far, the Scientific Program Committee (SPC) generally approves a fraction of the request of each proposal, with the fraction being close to one in the most meritorious cases. Averaging over all requests during the past year, about two-thirds of the proposed time was approved. The allocation decisions are based on scientific impact and the readiness of a project to start its calculation. While balance over the computing resources is driven by the proposals, balance among scientific areas is explicitly considered in the allocation discussion.

Proposals for large allocations are scrutinized to make sure that they address USQCD goals and DOE programs in HEP and NP.

The scientific impact of this project is strong and is meeting, and in some cases exceeding, its expectations.

There were three workshops in 2015, and one each in 2014, 2013, and 2012.

Three to four new US lattice faculty positions have been created annually over the past several years.

Comments

This year all proposals were accepted, but with a reduction of about 70% in the number of cycles for most requests.

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.

The ASCR NERSC Center and the Leadership Computing Facilities have a great deal of experience in the time allocation process. It may make sense to leverage their experience by, for example, adding a member from one of these facilities to the Scientific Program Committee.

The project should consider adding an open-ended question concerning desired architecture or resources (memory, queue management, ticketing, etc) to the annual user survey. The idea is to allow for more bottom-up information flow within the collaboration on the hardware acquisition process.

The current management structure is very effective. One change for this year is the presence of an elected “younger" member of the Executive Committee. There were 4 nominations, with one coming from each scientific area. This action was in response to recommendations from previous reviews and is a very good development. If it works well, the reviewers would encourage increasing this to one or two more elected representatives.

One reviewer did not favor the decision to exclude graduate students from the electorate. He would support extending the franchise to all students who have been contributing long enough to have their names on a published USQCD paper. In addition, the nominees in the recent election were all assistant or associate professors. Very few people would classify an associate professor as a “junior scientist”. Prof. Detmold is of course an excellent choice, but as a 2002 PhD, he easily warrants being a regular member of the Executive Committee.

The idea of more specific physics milestones has been raised in past reviews. There has definitely been progress in this aspect of documentation, but the improvement has not been uniform across the full spectrum of LQCD physics topics. USQCD management should continue their efforts to persuade their users across all LQCD areas to pay more attention to physics milestones. (See also comments under charge bullet #2.)

The creation of the Acquisition Review Committee was a good idea.

The procurement procedure appears somewhat ad-hoc; however, it is very thorough, is conducted by personnel with extensive experience, and is responsive to changing conditions.

Recommendations

None.

Change Request: Replace separate performance metrics for CPUs and GPUs with one unified metric

Findings

This change request was presented by Rob Kennedy.

It would consolidate separate performance metrics for CPUs and GPUs.

Comments

The change request presents a better way to track performance than the previous method, considering the mix of conventional CPU's and GPU's and the development of "hybrid" architectures.

The change request adds valuable flexibility and clarity, with only upsides and no apparent downsides.

Recommendations

Each reviewer recommended the acceptance of this Change Request.

Change Request: Add cluster-hosting at BNL to the baseline project plan

Findings

This proposal was presented by Bill Boroski and the BNL proposal and the BNL computing environment was summarized by Bob Mawhinney.

The advantages, funding adjustments, risk assessment, and benefits of this Change Request are well documented in the review materials. The project developed a "straddle plan", meticulously devised and documented by Rob Kennedy, to bring BNL into the allocation process with FNAL and JLAB with no compromise to their hardware funding plans.

BNL is executing a new Computational Science Initiative, which includes the acquisition of an institutional cluster.

BNL has proposed contributing a portion (equivalent to ~40 nodes on average) of its new institutional cluster for LQCD calculations as an in-kind contribution to the project. The BNL institutional cluster has similar architecture to the project's "pi0g" cluster, which should ensure its immediate usefulness.

The BNL contribution of a portion of its institutional cluster will balance the additional overhead in maintenance and support costs due to adding a new site to the project. The new contributions from BNL are worth roughly $1 million over the life of the project.

The risks associated with creating a third cluster host have been carefully evaluated and mitigation strategies have been articulated, and in many cases, written into the MOU with BNL.

Comments

The downside of this proposal is increased administrative costs. However BNL has offered the project access to a portion of its institutional cluster, so the net cost to the project is expected to be zero.

The upside of this proposal is that it brings additional computing expertise to the effort from BNL and additional support from a local community of very involved and capable users.

The project team stated that the BNL institutional cluster would be acquired using funds from non-DOE sources. This fact should be fully documented to avoid any potential misunderstandings or interpretations that the LQCD-ext II project acquired ~40 nodes on the BNL institutional cluster using additional DOE funds from outside the LQCD-ext II project.

There is some concern about the capability of BNL staff to handle a large number of new users. The project should carefully monitor this issue.

Recommendations

Each reviewer recommended the acceptance of this Change Request.

APPENDIX A

<u>Charge Letter to the LQCD-ext II Project Team</u>

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 June 28-29, 2016, at the Thomas Jefferson National Accelerator Facility (TJNAF). 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 June 28-29 review. The focus of the 2016 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 2016;
- 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 TJNAF 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 TJNAF.

Sincerely,

James Siegrist
Associate Director
Office of High Energy Physics

Timothy Hallman
Associate Director
Office of Nuclear Physics

APPENDIX B

Reviewers for 2016 LQCD-ext II Annual Review (JLAB June 28-29)

HEP Phenomenology

Mark Wise (California Institute of Technology)

HEP Experiment

Sheldon Stone (Syracuse University)

HEP Computer Specialist

Craig Tull (LBNL)

NP theory

Eric Swanson (University of Pittsburgh)

NP Computer Specialist

Esmond Ng (LBNL and NERSC)

NP Experiment

Declan Keane (Kent State University)

List of attending DOE program managers

J. Kogut (HEP, LQCD-ext II HEPProject Manager)

T. Barnes (NP, Theory and Computation)

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

APPENDIX C

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

LQCD-EXT II

June 28-29, 2016

Thomas Jefferson National Accelerator Facility
CEBAF Center, Room F224-225

Agenda

June 28

8:30 Executive Session (45 min)
9:15 Logistics and Introductions (5 min) – *Bill Boroski*
9:20 Welcome (15 min) – *Amber Boehnlein, CIO, Thomas Jefferson National Accelerator Facility*
9:35 LQCD-ext II Overview (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) – *Swagato Mukherjee*
11:45 Science Talk 3: Beyond the Standard Model (20 min) – *Ethan Neil*
12:05 Lunch / Executive Session

1:05 Science Talk 4: QCD for Particle Physics (40 min) – *Andreas Kronfeld*
1:45 LQCD-Ext II: Project Management and Performance (45 min) – *Bill Boroski / Rob Kennedy*
2:30 LQCD-Ext II: Change Requests: Overview, Unified Performance KPIs (15 min) – *Rob Kennedy*
2:45 Coffee Break (15 min)
3:00 LQCD-Ext II: Change Requests: 3-Site Cluster Hosting (30 min) – *Bill Boroski*
3:30 LQCD-Ext II: BNL Computing Overview (15 min) – *Bob Mawhinney*
3:45 LQCD-Ext II: FY16 Hardware Acquisition (30 min) – *Chip Watson*
4:15 Executive Session (60 min)
5:15 Committee request for additional information – *Review Committee & Project Leadership*

6:00 Adjourn
6:30 Dinner: *10 min drive from TJNAF*

June 29

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 (105 min)
12:00 Lunch
1:00 Closeout (60 min)
2:00 Adjourn