DOE FY2017 Annual Progress Review

Response to Technical Questions

Bill Boroski, Gerard Bernabeu, Rob Kennedy,
Chip Watson and Alex Zaytsev

for the

LQCD-ext II Project Team
and USQCD Collaboration
May 17, 2017

DOE FY17 Annual Progress Review
Fermilab
May 16-17, 2017

2) Continued need for GPUs: are there codes that really need GPUs or can they be run on KNL? Suppose only KNL in the future?

- GPUs remain the most cost effective solution for perhaps 20% of our portfolio, and equal to KNL for another 20%. While at the time of the JLab procurement this parity or better was perhaps only half this (a total of 20%), today the Pascal GPUs are much lower in price. This pricing change is due both to competition (good all around) and due to the imminent release of the Volta chip.
- Specifically, for pure configuration generation (smaller or specialty ensembles) and for the generation of propagators and perambulators (flops intensive intermediate products), GPUs are still a very good architecture.
- While it is no doubt true that it may no longer be worthwhile to make additional large investments in GPU software (since KNL is a more easily used advanced architecture), nevertheless, we have a large code base that is capable of exploiting GPUs and for that portion of our portfolio (50% of USQCD in-house flops?) we can continue to run well on GPUs. Furthermore, the next large increase in US capacity will be the GPU-accelerated Summit machine at ORNL, and LQCD will want to be able to fully use that machine's capability over the next few years, thus we have a compelling reason to keep that portion of our software base in good shape.

2) Continued need for GPUs: are there codes that really need GPUs or can they be run on KNL? Suppose only KNL in the future? (2)

- Depending upon the future pricing of the Volta GPU, and the price and timing of the release of the Knights Hill variant of the Xeon Phi, it is possible that within a year GPUs might again be a compelling procurement option. Twelve to fifteen months from now, JLab will likely retire its large 2012 complement of GPUs, opening the door for additional GPU purchases. This might be especially so following KNL clusters at JLab and BNL.
- Software investments to run on GPUs have been made and keeping that code up to date is a small price to pay for the flexibility it brings, even beyond the ability to exploit Summit next year.

3) Computing resources are currently distributed across 3 sites. Would a single site operating a large institutional cluster effectively meet the needs of the collaboration?

The distributed 3-site hosting model has served the USQCD collaboration very well since before 2006. We have a proven operating model that we have refined and optimized based on real operating experience and it has enabled significant scientific output and advances, on both our dedicated hardware as well as LCFs.

The three-site operating model has been so productive that we submitted a baseline change request (CR) in 2016 to maintain the three-site model through the end of FY19.

- Proposed CR to continue deploying hardware at BNL was vetted and supported by the 2016 review committee
- CR was approved by our Change Control Board and Federal Project Director

There are some obvious costs associated with operating three sites:

- Slightly higher staff costs (~1 FTE/site, based on operating experience)
- Higher overhead costs due to redundancy (e.g., user documentation, tracking allocations and utilization, system monitoring systems, etc.)
- Diversity of interfaces presented to users
- Switching costs for users whose allocations move from one site to another

3) Computing resources are currently distributed across 3 sites. Would a single site operating a large institutional cluster effectively meet the needs of the collaboration? (2)

There are also significant benefits to a 3-site model:

- 3-site model is representative of the distribution of experimental programs / needs
- Symbiotic relationship between experimental programs, lattice theorists, software development and the computing project (concentrated hubs of related activities)
  - Three healthy programs with effective, productive ecosystems in place
- Ability to attract critical and very competent software developers tied closely to lab scientific programs
- Broader hardware portfolio to support software development on future LCF platforms

Modest level of cost savings (~1-2 FTE/yr) would result in a significant loss of value to the lattice program, which would result in reduced scientific output:

- Loss of shared knowledge/experience -> increased downtime; missed opportunities
  BNL-JLab: KNL shared experience; JLab-FNAL: Lustre storage shared experience
- Loss of technology and infrastructure diversity (ability to explore new platforms)
- Loss of infrastructure for developing code bases for future running on LCFs
- Loss of talented HPC developers working closely with lattice scientists
- Loss of geographical separation

3) Computing resources are currently distributed across 3 sites. Would a single site operating a large institutional cluster effectively meet the needs of the collaboration? (3)

Down-select of sites is not an obvious choice

- Lack of experience with institutional clusters for long periods at large scale
- Long-term alignment of a single lab's strategic direction with changing LQCD needs
- No written agreements regarding the long-term commitment to design and dedicate a significant fraction of an institutional cluster to meet specific LQCD computing needs

A single site operating a large institutional cluster would not meet the needs of the collaboration. The costs and risks to the scientific program are too great.

- At a minimum, we could consider a two-site model
- Cost-benefit analysis to quantify cost savings against value loss (2 vs. 3 sites)

We should also investigate other potential cost-saving measures:

- Reduced level of user support
- Increased tolerance for reduced uptime
- Increased level of support from the host labs (staffing, hardware, storage, etc.)

We should begin investigating options now as we start planning for post-FY19 activities.