



# LQCD-ext II Project Change Request CR16-01

Bill Boroski  
Contractor Project Manager

DOE FY2016 Annual Progress Review  
Thomas Jefferson National Accelerator Facility  
June 28-29, 2016

# Outline

- ▶ Background and Motivation
- ▶ Proposed Solution
- ▶ Approach and Outcome
- ▶ Benefits to the Scientific Program
- ▶ Strategic and Technical Benefits
- ▶ Summary

# Background and Motivation

- ▶ Project baseline was set on August 28, 2014, when LQCD-ext II received CD-2/3 approval.
  - Cluster Acquisitions at JLab (FY16-17), FNAL (FY18-FY19)
  - Operate BNL BG/Q until retirement in FY17-FY18
- ▶ After the project was baselined, the LQCD Project Office learned of BNL's interest in hosting clusters.
  - BNL expressed interest in cluster hosting in spring 2015
  - BNL underwent a reformulation as part of their M&O contract bidding process. During this time, the BNL Computational Science Initiative began taking form and was formally announced in December 2015. Significant investment at BNL.
- ▶ A 3-site hosting model has served USQCD well since the LQCD Computing Project's inception in 2006.
  - The opportunity to work with and leverage the BNL Computational Science Initiative and RHIC/Atlas-experienced staff at BNL in a manner comparable to that at FNAL and JLab has many advantages for the project and science program.
    - Maintaining three active labs in the project keeps the project stronger and more technically astute.
    - Greater breadth and depth in technical team - keep up with rapidly evolving technologies
    - Improved buffer against staff transitions, and smoother transitions when they do occur

# Proposed Solution

- ▶ **CR16-01:** Add cluster-hosting at BNL to the project in a manner that fits within the approved funding profile and maintains the total computing delivered by the project.
- ▶ **3-Site FY-Straddle Plan**
  - Recommended approach for new hardware hosting and corresponding distribution of acquisition, deployment, and operations funds

Plan Name	FY16	FY17 Deployments	FY18 Deployments	FY19 Deployment
Baseline	JLab	JLab (FY16 options)	FNAL	FNAL (FY18 options)
3-Sites Fiscal Year Straddle	JLab	1/3 JLab (FY16 options) 2/3 BNL	2/3 BNL (FY17 options) 1/3 FNAL (initiate procurement)	FNAL (execute procurement)

- ▶ Through a very detailed analysis of options, this plan stands out as the most cost-effective means of increasing value to USQCD within the existing funding envelope
  - Keeps project costs within approved funding profile
  - Distributes acquisition funds about evenly across three host sites
  - Takes into account aging systems, available space, power, cooling, etc. at host sites
  - Achieves the total Delivered Computing goal for the project (annual levels do shift)
  - Increases value through additional strategic and technical benefits that enhance the scientific program.

# Approach

- ▶ Documented the resources required to effectively host LQCD clusters, based on 10 years of operating experience
  - **Storage:** disk, tape; **Network:** LAN, WAN; **User Support:** from few to many LQCD users
- ▶ Assessed the operations cost impact of adding third cluster host site
  - Developed revised cost forecast based on extending the 3-site operating model
  - Forecast the gap in Delivered Computing as a result of the required shifting of budget from hardware acquisition & deployment to steady-state operations.
- ▶ Negotiated with BNL Mgmt commitments and in-kind contributions to fill gap, host clusters. Personal involvement of:
  - **Kerstin Kleese van Dam**, Associate Lab Director and Director, Computing Sciences Initiative
  - **Nick D’Imperio**, Chair, Computational Science Laboratory
  - **Tony Wong**, Head, RACF\* Processing and Facility Operations and new LQCD Site Manager for BNL
  - **Shigeki Misawa**, Head, RACF\* Mass Storage and General Services Department and new LQCD Co-Site Architect for BNL

\*RHIC & Atlas Computing Facility

# Impact Study – Performance Forecasts

## Performance Forecast

- ▶ 3-Site FY-Straddle, just shows *level of detail treated*
- ▶ Extends forecast model used in Project CD process + cluster ratings updates
- ▶ Documents expected life cycle for clusters, and KPIs calculations (separate sheet)
- ▶ Future cluster performance: Baseline scaled by Acquisition Budget change
- ▶ Calculates min integral BNL-IC nodes needed to maintain total Delivered Computing

3-Site FY Straddle Scenario with BNL IC																					
6/10/16		Compute Equip Buy Only				FY15			FY16			FY17			FY18			FY19			FY15-FY19
Machine	Base TFlops	Base \$	Alt \$	Alt TFlops	Fraction	Uptime	Delivered	Fraction	Uptime	Delivered	Fraction	Uptime	Delivered	Fraction	Uptime	Delivered	Fraction	Uptime	Delivered	Aggregate	
DD2	4.38			4.38	1.00	91.3%	4.38	1.00	91.3%	4.38	1.00	91.3%	4.38	0.00	91.3%	0.00	0.00	91.3%	0.00	13.14	
BG/Q	21.90			21.90	1.00	91.3%	21.89	1.00	91.3%	21.89	1.00	91.3%	21.89	0.00	91.3%	0.00	0.00	91.3%	0.00	65.68	
Ds	21.03			21.03	1.00	91.3%	21.02	0.95	91.3%	20.00	0.00	91.3%	0.00	0.00	91.3%	0.00	0.00	91.3%	0.00	41.03	
10g / 11g - average	17.09			17.09	0.75	91.3%	12.81	0.75	91.3%	12.81	0.00	91.3%	0.00	0.00	91.3%	0.00	0.00	91.3%	0.00	25.63	
Bc	12.73			12.73	1.00	91.3%	12.73	1.00	91.3%	12.73	1.00	91.3%	12.73	1.00	91.3%	12.73	0.00	91.3%	0.00	50.91	
12s	12.80			12.80	1.00	91.3%	12.80	1.00	91.3%	12.80	0.00	91.3%	0.00	0.00	91.3%	0.00	0.00	91.3%	0.00	25.59	
Pi0	13.10			13.10	1.00	91.3%	13.10	1.00	91.3%	13.10	1.00	91.3%	13.10	1.00	91.3%	13.10	1.00	91.3%	13.10	65.48	
Dsg	15.90			15.90	1.00	91.3%	15.90	0.49	91.3%	7.80	0.00	91.3%	0.00	0.00	91.3%	0.00	0.00	91.3%	0.00	23.70	
12k	26.40			26.40	1.00	91.3%	26.39	1.00	91.3%	28.11	0.00	91.3%	0.00	0.00	91.3%	0.00	0.00	91.3%	0.00	54.50	
Pi0g	25.10			25.10	1.00	91.3%	25.09	1.00	91.3%	25.09	1.00	91.3%	25.09	1.00	91.3%	25.09	1.00	91.3%	25.09	125.47	
15C - Pi0 Expansion	6.12			6.12	0.46	85.0%	2.61	1.00	91.3%	6.12	1.00	91.3%	6.12	1.00	91.3%	6.12	1.00	91.3%	6.12	27.09	
15G - NONE	0.00			0.00	0.00	0.0%	0.00	0.00	0.0%	0.00	0.00	0.0%	0.00	0.00	0.0%	0.00	0.00	0.0%	0.00	0.00	
16C	10.30	\$389,523	\$496,088	13.12	0.00	0.0%	0.00	0.00	85.0%	0.00	1.00	91.3%	13.11	1.00	91.3%	13.11	1.00	91.3%	13.11	39.34	
16G	38.70	\$389,523	\$496,088	49.29	0.00	0.0%	0.00	0.00	85.0%	0.00	1.00	91.3%	49.27	1.00	91.3%	49.27	1.00	91.3%	49.27	147.82	
17C	13.80	\$512,235	\$359,624	9.69	0.00	0.0%	0.00	0.00	0.0%	0.00	0.22	85.0%	1.98	1.00	91.3%	9.69	1.00	91.3%	9.69	21.36	
17G	52.00	\$512,235	\$359,624	36.51	0.00	0.0%	0.00	0.00	0.0%	0.00	0.22	85.0%	7.48	1.00	91.3%	36.50	1.00	91.3%	36.50	80.47	
18C	28.20	\$534,177	\$504,329	26.62	0.00	0.0%	0.00	0.00	0.0%	0.00	0.00	0.0%	0.00	0.34	85.0%	8.30	1.00	90.8%	26.46	34.76	
18G	106.10	\$534,177	\$504,329	100.17	0.00	0.0%	0.00	0.00	0.0%	0.00	0.00	0.0%	0.00	0.34	85.0%	31.23	1.00	90.8%	99.56	130.79	
19C	36.10	\$685,152	\$448,102	23.61	0.00	0.0%	0.00	0.00	0.0%	0.00	0.00	0.0%	0.00	0.00	0.0%	0.00	0.25	85.0%	5.49	5.49	
19G	136.20	\$685,152	\$448,102	89.09	0.00	0.0%	0.00	0.00	0.0%	0.00	0.00	0.0%	0.00	0.00	0.0%	0.00	0.25	85.0%	20.73	20.73	
BNL Inst Cluster	0.00	\$0	\$0	36.00	0.00	0.0%	0.00	0.33	100.0%	12.00	1.00	100.0%	36.00	1.00	100.0%	36.00	1.00	100.0%	36.00	120.00	
	597.9515		\$3,616,286	560.636	169.0			175.2			191.9			244.1			343.9			1118.99	
Delivered Computing Total			-\$625,889		176.6	95.7%	168.7	212.6	82.4%	176.8	227.9	84.2%	191.2	328.4	74.3%	241.1	428.4	80.3%	341.1	1117.11	
Conventional						88	88.53		102	91.02		83	73.31		72	63.04		82	73.97	1.88	
GPU				rating/node	0.9	Tflops/node	92	80.20		102	85.82		126	117.84		186	178.10		275	267.15	0.2%
				n_nodes	40		180	168.72		204	176.84		209	191.16		258	241.14		357	341.12	Same CPU c
				total nodes	200	Delivered			PEP Set Values		PEP Rounded Values		PEP Rounded Values		PEP Rounded Values		PEP Rounded Values		PEP Rounded Values		BNL IC time

# Outcome

## ▶ Negotiated Commitment from BNL

- Time-averaged allocation of 40 BNL-IC nodes from late FY16 through FY19
- Disk and Tape-based data storage in several tiers
  - 100–200 TB of disk storage on the 1 PB Institutional Cluster storage system.
  - 0.5 PB of older disk storage for scratch and/or intermediate-term storage on the existing storage systems for BG/Q systems at BNL.
  - ~\$80k of disk storage in FY17 and FY18 that project would have purchased as part of the FY17–18 “acquisition year” disk storage
  - Use and access of BNL mass storage systems, including tape robot with performant intermediate disk caching system.
- Brings BNL LQCD cluster hosting to similar level as FNAL and JLab.

## ▶ Result:

- Shifts ~\$1M from the overall hardware budget to cover increased staff costs in FY17–FY19.
- Added overhead of 3<sup>rd</sup> site cluster operations is largely offset by BNL in-kind contributions
- No gap in achieving the total Delivered Computing goals in the approved baseline plan.
- Significant increase in value and benefits to USQCD.

# Benefits to the Scientific Program

- ▶ **Close coupling between theorists and computing professionals**
  - Having more lattice theorists in close contact with the project and the project's computing professionals allows theorists to better keep the project abreast of their needs and plans.
  - Also allows the project to better keep theorists abreast of the capabilities of the project's computational resources.
  
- ▶ **Access to the new BNL Institutional Cluster (IC)**
  - Although the negotiated allocation is for a defined number of nodes, allocation on this large machine provides flexibility to run bigger jobs (for fraction of year).
  - Capability to periodically run very large jobs for fractions of the year is very valuable to USQCD.
  - We have experience resource-sharing with physics programs at JLab. This arrangement provides the opportunity for a similar arrangement with BNL.

# Strategic Benefits

- ▶ Maintains active involvement of executive management from three national laboratories heavily involved in high energy / nuclear physics
  - LQCD has always benefitted greatly by having three lab CIOs/Associate Directors directly engaged and intimately involved with the project [Tom Schlagel (BNL), Vicky White (FNAL), and Roy Whitney (JLab)]. As part of the integrated management team and members of the Change Control Board, they regularly provided critical insight and support.
  - The situation at BNL is markedly different than it was 2 years ago when the project was baselined.
    - In Nov 2014, DOE awarded the BNL 5-year M&O contract to Brookhaven Science Associates. BNL operations have been reformulated, new members of the leadership team are in place, and the lab's agenda includes expanding BNL's role in scientific computing.
  - Maintaining a 3-site operating model allows the project to continue to benefit from the direct engagement of executive management from three premier labs involved in scientific domains that benefit directly from Lattice QCD, and heavily interested in advancing scientific computing.
    - FNAL: Rob Roser, CIO. *Drove the addition of Advanced Computer Science, Visualization, and Data to Fermilab's list of Core Competencies acknowledged by DOE.*
    - JLab: Amber Boehnlein, CIO. *Leading and driving computational initiatives focused on LQCD, since LQCD results are integral to all aspects of JLab's physics program.*
    - Kerstin Kleese van Dam, Assoc. Lab Director and Head, Computational Sciences Initiative. *Leading and driving the Computational Sciences Initiative.*

# Strategic Benefits...cont'd

- ▶ Leverages BNL RHIC–Atlas Computing Facility along with comparable organizations and efforts at FNAL and JLab
  - As part of BNL’s reformulation, lab wide scientific computing expertise is placed into a single unit. The expertise of the RHIC–ATLAS Computing Facility (RACF) is being leveraged to unify all scientific computing operations into one organization.
  - RACF has provided centralized computing services for BNL experiments for over 20 years. Broad base of relevant technical expertise.
- ▶ Transitions LQCD to the new organizational structure at BNL
  - USQCD benefitted greatly from the relationship with the Information Technology Division, which operated and maintained the QCDOC during LQCD and into LQCD–ext, and the BG/Q during LQCD–ext and into LQCD–ext II.
  - Going forward, USQCD stands to benefit greatly from a relationship with the Computational Science Initiative, which has committed to operating and maintaining the LQCD BG/Q through end of life, providing USQCD with access to a new Institutional Cluster, and operating and maintaining LQCD clusters in the future.

# Technical Benefits

## ▶ Broader Project Team and Experience

- Expanded depth in LQCD-related architecture, design approaches, industry contacts
  - Leverages the resources of the BNL Scientific Data and Computing Center; had been supported by the BNL Information Technology Division
  - Brings additional expertise to track technical advances and augment future hardware evaluations
  - Technical experts from all three labs have worked together on hardware evaluations since 2009, when we first began considering the BG/Q as an alternative
  - Architects from all 3 sites actively worked together on the Alternative Analysis and subsequent efforts associated with the FY16 acquisition
- Greater sharing of administrative experience, potential possibilities for economies of scale
  - FNAL and JLab shared Lustre upgrade experiences, will be able to contrast to BNL's GPFS experience
- Enhanced sharing of best practices for User Support
  - FNAL and JLab have over 10 years of experience operating LQCD clusters and supporting the USCQD community
  - BNL has experience operating clusters for RHIC & Atlas; provides tremendous opportunity to share best practices and improve operational effectiveness and efficiencies.

# Summary

- ▶ [CR16-01 Three-Site Cluster Hosting](#): Add cluster-hosting at BNL to the project in a manner that fits within the approved funding profile and maintains the total computing delivered by the project.
- ▶ The BNL commitment offsets the impact of shifting acquisition funds to cover increased operations costs; baseline goal for total Delivered Computing will still be met or exceeded
- ▶ Added benefits of the proposed change and implementation approach
  - Leverages capabilities at BNL that were not as readily available or accessible when the project was baselined.
  - Maintains the ability of the project to meet or exceed established performance goals for total Delivered Computing.
  - Provides greater breadth and depth in technical team – keep up with rapidly evolving technologies and smoother staff transitions over time.
  - Augments the level of knowledge and best practices sharing amongst the sites.
  - Establishes closer ties between broader set of theorists and the Project team so that each may better understand and anticipate each other's needs
  - Helps the project more effectively meet the scientific needs of USQCD.
- ▶ Approved by the Project Manager and Change Control Board. Now under consideration by the LQCD Federal Project Director.

# Supporting Slides

# Impact of the Proposed Change

## ▶ Cost Impact

- Shifts ~\$1M from the overall compute and hardware procurement budget to cover increased Personnel and M&S costs in FY17–FY19.
  - Offset mostly by the in-kind allocation of BNL Institutional Cluster nodes.

## ▶ Schedule Impact

- Adds BNL acquisitions to the project WBS and alters acquisition timing within year per proposed acquisition plan.

## ▶ Performance Impact

- Total Delivered Computing is increased slightly (~2%) as a result of the BNL IC in-kind contribution.
- Total Deployed Computing is reduced by 7%
  - We will be deploying new computing resources earlier, but less overall. Time integral (Delivered Computing) is a bit higher.

## ▶ Change Level per Project Change Process

- Level 4: Federal Acquisition Executive approval required
  - Threshold: “> \$250k in budget distribution between DME and SS O&M costs”

# CR16-01: Impact (1)

- ▶ What is the Gap in Delivered Computing due to Third Site Ops?
- ▶ Models Compared: to track updated info *and* impact of 3 vs 2 Sites
  - Baseline: Exactly the output of the project CD process.
  - Reference: “Baseline” with FY15 actuals included.
  - 2-Sites: “Reference” with all changes to future plans included, such as overhead rates.
    - This is the updated and improved forecast of the project if it proceeds *without* CR16-01.
  - 3-Sites FY-Aligned: Adds cluster-hosting at third site using a one-site-per-year rotation
  - 3-Sites FY-Straddle: Adds cluster-hosting at third site using “3-sites in 4 years” rotation
- ▶ Acquisition Plans in These Models:

Plan Name	FY16	FY17 procurement	FY18 procurement	FY19 procurement
Baseline	JLab	JLab (FY16 options)	FNAL	FNAL (FY18 options)
Reference	JLab	JLab (FY16 options)	FNAL	FNAL (FY18 options)
2-Sites	JLab	JLab (FY16 options)	FNAL	FNAL (FY18 options)
3-Sites FY-aligned	JLab	BNL	FNAL	JLab
3-Sites FY-straddle	JLab	1/3 JLab (FY16 options); 2/3 BNL	2/3 BNL (FY17 options); 1/3 FNAL (slide to FY19)	FNAL

- *3-Sites FY-Straddle is the proposed Acquisition Plan going forward with this CR.*

# CR16-01: Impact Study - Cost Forecasts

## Cost Forecasts

- ▶ 3-Site FY-Straddle, just shows *detail level treated*
- ▶ (Bot): Project Budget Detail
  - Sheets for Sites, Mgmt Reserve
- ▶ (Top): Compute Funds Split
  - Excerpt from Staffing Model
- ▶ Staffing Model captures costs of cluster hosting, based on 10+ years experience.
- ▶ Broader "budget" to better forecast performance
  - Treats past project carry-over with this project's funding.
  - = why Total Budget Profile looks different here from ext II Project Funding profile.
- ▶ Cost Forecast then drives the Performance Forecast via the Acquisition Budget.

### BUDGET (\$K)

	(closed) FY15	(allocated) FY16	FY17	FY18	FY19	Total
<b>Steady-state Operations</b>						
Personnel	1,543,737	1,451,490	1,543,813	1,447,666	1,408,561	7,395,267
Travel	11,000	11,000	11,000	11,000	11,000	55,000
M&S (hardware, repairs, etc.)	281,000	297,786	146,000	120,000	120,000	964,786
Sub-total (SS Ops)	1,835,737	1,760,276	1,700,813	1,578,666	1,539,561	8,415,053
<b>New Hardware Deployment</b>						
Personnel	-	198,800	296,144	141,354	244,374	880,673
Travel	-	-	-	-	-	-
Equipment (compute)	495,000	992,175	719,248	1,008,658	896,203	4,111,284
Equipment (storage)	60,000	50,359	39,452	55,327	77,931	283,069
Sub-total (New Deployment)	555,000	1,241,334	1,054,844	1,205,339	1,218,508	5,275,026
<b>Project Management</b>						
Personnel	110,298	127,351	131,172	135,107	139,160	643,089
Travel	6,000	6,000	6,000	7,000	7,000	32,000
M&S	2,000	2,000	2,000	2,000	2,000	10,000
Sub-total (Project Mgmt)	118,298	135,351	139,172	144,107	148,160	685,089
<b>Total Project Cost</b>						
Personnel	1,654,035	1,777,641	1,971,129	1,724,127	1,792,095	8,919,028
Travel	17,000	17,000	17,000	18,000	18,000	87,000
M&S	283,000	299,786	148,000	122,000	122,000	974,786
Equipment (compute)	495,000	992,175	719,248	1,008,658	896,203	4,111,284
Equipment (storage)	60,000	50,359	39,452	55,327	77,931	283,069
Management Reserve	45,964	83,039	105,173	71,890	93,767	399,833
<b>Total</b>	<b>2,555,000</b>	<b>3,220,000</b>	<b>3,000,002</b>	<b>3,000,002</b>	<b>2,999,996</b>	<b>14,775,000</b>
CD-2/3 Budget Guidance Profile	2,000,000	3,000,000	3,000,000	3,000,000	3,000,000	14,000,000
Carry-over Funds from Past Project	555,000	220,000	-	-	-	775,000
<b>Total Budget Profile</b>	<b>2,555,000</b>	<b>3,220,000</b>	<b>3,000,000</b>	<b>3,000,000</b>	<b>3,000,000</b>	<b>14,775,000</b>

CPU Funds Split	FNAL	JLab	BNL
FY15	0	0	0
FY16	0	1	0
FY17	0	0.30	0.70
FY18	0.30	0	0.70
FY19	1	0	0
<b>Total Compute H/w</b>	<b>\$1,198,801</b>	<b>\$1,207,949</b>	<b>\$1,209,534</b>

# CR16-01: Timeline

- ▶ October 2015: Formal CR work begun, Fed Project Dir informed
- ▶ May 20, 2016: Preliminary Approval by Contract Project Manager
- ▶ June 13, 2016: CR Documentation Finalized, Approved by CPM
- ▶ June 21, 2016: CCB Meeting
- ▶ June 28-29, 2016: DOE Review – Reviewers advise Federal Director
- ▶ July 2016: Fed Prj Director & Fed Acquisition Exec consider CR
- ▶ August 2016: Formally Implement CR in Project BEFORE sending Allocation Request to DOE. *Schedule constraint.*

# CR16-01: Affected Project Documents

## ▶ Affected Controlled Project Documents

### ◦ Site MOUs

- LQCD-BNL MOU: boilerplate updated, adds in-kind contributions for this CR
- LQCD-FNAL MOU: boilerplate updated
- LQCD-TJNAF MOU: boilerplate updated

### ◦ Project Execution Plan

- Updated boilerplate. Includes FY15 PEP Addendum for Site Manager/Site Architect roles, etc.
- Updated project organization charts for staff transitions at BNL and FNAL
- Adjusted funding levels for acquisitions, new Deployed/Delivered KPIs

### ◦ WBS: (see modest schedule impacts listed on Impact summary slide, Schedule Impact bullet)

## ▶ Affected Other Project Documents

### ◦ Cost Forecasts

- Annual process: reconciled at end FY15 and begin calendar 2016 (when carry-over settled)

### ◦ Performance Forecasts

- A reconciliation of performance tracking in project was done as part of this CR preparation.