

**Acquisition Strategy  
for the  
SC Lattice QCD Computing Project Extension  
(LQCD-ext)**

*Operated at*  
Brookhaven National Laboratory  
Fermi National Accelerator Laboratory  
Thomas Jefferson National Accelerator Facility

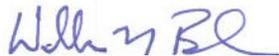
*for the*  
U.S. Department of Energy  
Office of Science  
Offices of High Energy and Nuclear Physics

Version 2.0

April 21, 2011

PREPARED BY:  
Don Holmgren, FNAL

CONCURRENCE:



May 2, 2011

---

William N. Boroski  
LQCD Contractor Project Manager

---

Date

**Acquisition Strategy  
Change Log**

<b>Revision No.</b>	<b>Description/ Pages Affected</b>	<b>Effective Date</b>
Revision 0.9	LQCD II plan draft drawn from LQCD FY08/FY09 Plan	Dec 30, 2008
Revision 1.0	Pre-ARRA version.	March 19, 2009
Revision 1.1	Combined FY2010/FY2011 buy. LQCD II → LQCD-Ext	April 13, 2009
Revision 1.2	Added performance basis, Section 508 and M-07-11 compliance, and cyber security sections.	April 15, 2009
Revision 1.3	Final Version for CD-1	April 15, 2009
Revision 1.4	Reorganize to present strategy and plan separately	June 18, 2009
Revision 1.5	New Version for CD-2 / CD-3	August 3, 2009
Revision 1.6	Expanded storage discussion	August 17, 2009
Revision 1.7	Updated discussion of Intel and AMD processors	April 1, 2010
Revision 2.0	Updates for FY11 – acquisition plan details removed to a separate document. Focus on FY12 decision procedure in this version.	April 17, 2011

## **Table of Contents**

Introduction.....	1
Previous LQCD Computing Project and LQCD-ext Project FY10 and FY11 .....	2
Overview of LQCD-Ext Project Deployments .....	2
Design Considerations and Strategies for LQCD-Ext Deployments.....	3
Compute Nodes.....	3
High Performance Network .....	5
Service Networks .....	5
Network Plan .....	5
File I/O .....	5
Procurement Strategy.....	6
Strategy for the FY2012 LQCD-Ext Deployment.....	7

## Introduction

The Lattice QCD Computing Extension Project (LQCD-Ext) develops and operates new and existing systems in each year from FY2010 through FY2014. These computing systems are deployed at Fermilab (FNAL), Jefferson Lab (JLab), and Brookhaven (BNL). During FY2010, the project operated the 4.2 Tflop/s US QCDOC supercomputer at BNL, as well as the clusters developed during the last two years of the prior Lattice QCD Computing Project (FY2006-FY2009) at FNAL and at JLab. During FY2011, the project will continue to operate the QCDOC supercomputer until approximately July 2011. Table 1 shows the actual and planned total computing capacity of the new deployments, and the actual and planned delivered (integrated) performance. In FY2011, the LQCD-ext project will for the first time deploy a mixture of conventional and GPU-accelerated clusters; in FY2012-FY2014, the split between these types of clusters will be determined based upon a number of factors, including cost effectiveness, availability of software, demand, and scientific impact. Currently the project uses GPU-yr as the metric for delivered computing capacity on GPU-accelerated clusters; this unit may change in the future to better reflective scientific production. In all discussions of conventional cluster and supercomputer performance, unless otherwise noted, the specified figure reflects an average of the sustained performance of domain wall fermion (DWF) and improved staggered (asqtad) algorithms.

	<b>FY 2010</b>	<b>FY 2011</b>	<b>FY 2012</b>	<b>FY 2013</b>	<b>FY 2014</b>
<b>Actual</b> (FY10) and <i>Planned</i> (FY11-FY14) Computing capacity of new Deployments, Tflop/s	<b>12.5</b> <i>11</i>	9	24*	44*	57*
<b>Actual</b> (FY10) and <i>Planned</i> (FY11-FY14) delivered Performance (JLab + FNAL + QCDOC), Tflop/s-yr	<b>19.2</b> <i>18</i>	22	34*	52*	90*
<i>Planned</i> (FY11) GPU Deployment, number of Nvidia GPUs	-	128	**	**	**
<i>Planned</i> delivered GPU computing capacity, GPU-khrs	-	922	**	**	**

**Table 1 – Performance of New System Deployments, and Integrated Performance** (DWF+asqtad averages used). Integrated performance figures use an 8000-hour year. \* FY2012 through FY2014 deployments and delivered performance will be modified via change control to take into account GPU-accelerated cluster acquisitions. \*\* FY2012 through FY2014 GPU-accelerated cluster acquisitions and operations are to be determined.

All LQCD-Ext Project cluster hardware procurements will utilize firm, fixed-price contracts. Cluster purchases will use contracts with vendors specializing in COTS hardware. The steady-state operations of the project computing facilities are performed by the three host laboratories, each of which is a government-owned contractor-operated facility.

In each year of the project, the hardware that best accomplishes the scientific goals for LQCD calculations will be purchased. In FY2010, an Infiniband cluster was deployed at Fermilab. In FY2011, the project determined that two deployments, consisting of the expansion the FY10 cluster at Fermilab, plus a GPU-accelerated Infiniband cluster, would best optimize the scientific capabilities of the portfolio of hardware operated by the LQCD-ext project. In FY2012, the IBM BlueGene/Q supercomputer will be an important hardware candidate in addition to conventional and GPU-accelerated clusters. In all years, project

personnel will also consider alternative hardware designs suitable for LQCD computing that may become available.

In the rest of this document we first discuss the design considerations and strategies that we will use for all of the procurements of the LQCD-Ext Project. We then discuss the process by which the project will determine the best hardware combination for the FY12 acquisition. This document will be updated each year to concentrate on the upcoming year's hardware acquisition.

### **Previous LQCD Computing Project and LQCD-ext Project FY10 and FY11**

From FY2006 through FY2009, the DOE High Energy Physics (HEP) and Nuclear Physics (NP) program offices funded the DOE Office of Science LQCD Computing Project (SC LQCD). The total project cost of \$9.2M for SC LQCD funded the deployment of four clusters at Jefferson Lab and Fermilab, plus the operations of these clusters, the QCDOC supercomputer at Brookhaven, and several SciDAC LQCD clusters at JLab and FNAL acquired in 2003 through 2005.

The clusters developed during SC LQCD were as follows:

- “6n”, at JLab in 2006, based on single-socket dual-core Pentium processors and single-data-rate Infiniband
- “Kaon”, at FNAL in 2006, based on dual-socket dual-core Opteron processors and double-data-rate Infiniband
- “7n”, at JLab in 2007, based on dual-socket dual-core Opteron processors, upgraded to quad-core processors, and double-data-rate Infiniband
- “J/Psi”, at FNAL in 2008 and 2009, based on dual-socket quad-core Opteron processors and double-data-rate Infiniband

The “J/Psi” cluster was procured using the funds from both FY2008 and FY2009. The FY2008 piece of “J/Psi” was awarded late in the fiscal year under a purchasing contract that allowed, via an option, additional compute nodes and network hardware of the same configuration to be purchased in the first half of FY2009. The FY2008 portion of the cluster was released to physics production at the beginning of January, 2009, and the FY2009 portion from the exercise of the purchase option was released to physics production in mid-April of 2009.

By executing a combined purchase, a single request for information (RFI) and a single request for proposal (RFP) were used, reducing project labor costs, laboratory labor costs, and the overhead (G&A) charged to the project.

### **Overview of LQCD-Ext Project Deployments**

When the LQCD-Ext project began in FY2010, the most effective hardware for the calculations performed on the existing SC LQCD project compute resources in FY2009 were commodity clusters built using Intel or AMD x86\_64 processors and an Infiniband interconnect. We predicted this to be the case in both FY2010 and FY2011, and so proposed a combined cluster purchase in FY2010 and FY2011 at Fermilab similar to the SC LQCD FY2008/FY2009 purchase of “J/Psi”. In 2010, FNAL purchased the “Ds” cluster, based on quad-socket eight-core Opteron processors and quad-data-rate Infiniband. In FY2011, the “Ds” cluster will be expanded from 245 nodes to a total of 421 nodes, utilizing an option in

the FY2010 purchase contract. This expansion will consume approximately 60% of the FY2011 project hardware budget. Based on the successes of the JLab GPU-accelerated clusters procured in 2009 and 2010 via ARRA (American Recovery and Reinvestment Act of 2009) funding, the LQCD-ext project will design and acquire a GPU-accelerated cluster, consuming the remaining approximately 40% of the FY2011 project hardware budget.

In late FY2012 the next generation of the IBM BlueGene series, BlueGene/Q, may be available for purchase. This machine, like its BG/L and BG/P predecessors, should perform very well on LQCD applications and may be competitive with commodity cluster and GPU-accelerated cluster hardware. The LQCD-Ext project will therefore evaluate BG/Q for deployment at BNL in FY2012. Significant BlueGene expertise resides at BNL. If other hardware, such as a conventional cluster or a GPU-accelerated cluster, or a mixture of the two, is determined to be more cost effective and would better meet scientific demands in FY2012, the project will instead deploy that alternative.

In FY2013 and FY2014, the project will procure one or two additional systems, using the most effective hardware as determined by the anticipated usage.

All procurements will be performed by the host laboratory chosen for the particular hardware deployment. Such purchases will utilize firm, fixed-price contracts. The typical sequence for new deployments will be:

1. In consultation with USQCD community (Executive Committee, Scientific Program Committee) determine usage profile for new deployment (*e.g.*, distribution of job types and sizes, file I/O requirements)
2. Complete preliminary design
3. Issue a Request for Information (RFI) to likely vendors
4. Evaluate the RFI responses and complete a final design
5. Obtain host laboratory purchase approvals via the local requisition process
6. Issue a Request for Proposal (RFP) to likely vendors
7. Evaluate RFP responses and award purchase contract
8. Approve sample node and sample scalable unit (rack)
9. Test and approve vendor-integrated final system
10. Operate final system in “friendly user” mode and tune the configuration
11. Release the final system to users

## **Design Considerations and Strategies for LQCD-Ext Deployments**

### ***Compute Nodes***

Lattice QCD codes are floating point intensive, with a high bytes-to-flops ratio (1.45 single precision, 2.90 double precision for SU(3) matrix-vector multiplies). When local lattice sizes exceed the size of cache, which is nearly always the case, high memory bandwidths are required.

Commodity processors that were available at the time of the FY2010 deployment with the greatest memory bandwidths were Intel x86\_64 processors with 1066 and 1333 MHz front side buses (Xeon “Nehalem-EP”, “Nehalem-EX”, and “Westmere”) and AMD Opteron processors (“Istanbul” and “Magny-Cours”). Xeon and Opteron processors can be used in dual and quad processor systems. In past years, the total cost of quad processor systems of both types, including the cost of the high performance

network, exceeded the cost of two dual processor systems with network because of the high cost of the quad-capable processor variants. However, in 2010, the AMD “Magny-Cours” processors did not have a special quad-socket variant, but rather could be used in either dual or quad-processor systems. Quad processor systems based on these processors were as cost effective, or even more cost effective, as dual processor systems.

Since late 2006, Intel and AMD have switched all new processors of relevance to lattice QCD to multi-core (initially dual core, now quad core, eight core, and twelve core). The JLab “7n” and Fermilab “J/Psi” clusters purchased and deployed in 2007 and 2008, respectively, use quad core processors; both use motherboards that accommodate two Opteron “Barcelona” processors. The Fermilab “Ds” cluster purchased and deployed in 2010 uses eight-core AMD processors. Lattice QCD production on these clusters has shown that multi-core processors scale very well on MPI jobs when the cores are treated as independent processors. Multi-core processors typically have lower clock speeds than the older analogous single core processors; however, the degree of scaling on MPI jobs is sufficient to make these processors a more cost effective choice. Roadmaps from both Intel and AMD indicate that all forthcoming designs will be multi-core.

In 2007, all commodity dual processor Xeon motherboard designs used a single memory controller to interface the processors to system memory. As a result, the effective memory bandwidth available to either processor was half that available to a single processor system. At that time, AMD Opteron processors had integrated memory controllers and local (to the processor) memory buses, with a high-speed link (HyperTransport) allowing one processor to access the local memory of another processor. This NUMA (Non Uniform Memory Access) architecture made multiprocessor Opteron systems a better choice for lattice QCD codes because of increased memory bandwidth. In the 2007 and 2008 LQCD project acquisitions of the “7n” and “J/Psi” clusters, and the 2010 LQCD-ext project acquisition of the “Ds” cluster, multiprocessor Opteron systems were chosen, as these were the most cost effective designs. Since 2009, Intel systems, beginning with the “Nehalem” processor family, have also used NUMA architectures. The 2009 and 2010 JLab ARRA “9q” and “10q” clusters use processors from this Intel family.

Starting in 2008, various LQCD scientists implemented codes to run on Nvidia graphics processing units (GPUs) using the “CUDA” extensions to the C and C++ languages. Typically these codes accelerated part of the overall computational work performed during LQCD configuration generation and analysis. Although very labor intensive to implement, these codes greatly accelerated those portions of LQCD computations, and GPU-accelerated clusters clearly can have greater cost efficiency than conventional Infiniband clusters for some of the calculations of interest. In early 2009, the LQCD project acquired a small GPU-accelerated cluster based on “J/Psi” host nodes and Nvidia Tesla S1070 GPUs, with a total of 16 GPUs deployed across 8 “J/Psi” hosts. In 2009 and 2010, via an ARRA-funded project, the “9g” and “10g” GPU-accelerated clusters were purchased and deployed at JLab. The “9g” cluster utilizes NVidia GPUs based on the GeForce 200 series (such as the GTX-280, GTX-285, Tesla C1060, and Tesla S1070), and the “10g” cluster utilizes Nvidia GPUs based on the “Fermi” GeForce 400 and 500 series (such as the GTX-480, GTX-580, and Tesla M2050).

To repeat, for those calculations to which GPUs can be applied, significant accelerations have been observed. Comparing the “J/Psi” cluster with the “9g” and “10g” clusters, users in the fall of 2010 reported relative throughputs of between 5:1 and 15:1 when comparing a single GPU to a single “J/Psi”

node. This represents a significant increase in cost effectiveness. From 2011 forward, the LQCD-ext expects to use portions of the hardware budget in each year for GPU-accelerated systems. The portion to be used in any year will depend upon the scientific demand for this hardware, which in turn is related to the fraction of the LQCD calculations that can take advantage of GPU acceleration because of the availability of code.

### ***High Performance Network***

Based on LQCD SciDAC prototypes in FY 2004 and FY 2005, the “6n” and “Kaon” clusters purchased in FY2006, the “7n” cluster purchased in FY2007, and the “J/Psi” cluster purchased in FY2008 and FY2009, Infiniband was the preferred choice for the first LQCD-Ext cluster “Ds”, and is the likely choice for any later clusters. These clusters will use quad data rate (QDR) or faster Infiniband parts.

Current QDR switch configurations from multiple Infiniband vendors include 24, 36, 108, 144, 216, 288, 324, and 648-port switches. For the large clusters to be built in this project, leaf and spine designs are preferred. Because QDR 4X HCA bandwidths exceed the requirements for lattice QCD codes, oversubscribed designs can be used. A 2:1 design, for example, would have 24 computers attached to a 36-port switch, with the remaining 12 ports used to connect to the network spine.

### ***Service Networks***

Although Infiniband supports TCP/IP communications, we believe that standard Ethernet will still be preferred for service needs. These needs include booting the nodes over the network (for system installation, or in the case of diskless designs, for booting and access to a root file system), IPMI access (IPMI-over-LAN) for remote hardware control and management, serial-over-LAN, and NFS access to “home” file systems for access to user binaries. All current motherboard candidates support two or more embedded gigabit Ethernet ports.

In our experience, serial connections to each computer node are desirable. These connections can be used to monitor console logs, to allow login access when the Ethernet connection fails, and to allow access to BIOS screens during boot. Serial-over-LAN (standard with IPMI 2.0) will be used to provide these serial connections.

### ***Network Plan***

For LQCD-Ext Project clusters, we will replicate the network layout currently used on all of the FNAL and JLab lattice QCD clusters. In these designs all remote access to cluster nodes occurs via a “head node”, which connects to both the public network and to the private network that forms the sole connection to the computer nodes. Secure ID logon (Kerberos at FNAL, ssh at JLab) is required on the head node. “R-utility” (rsh, rlogin, rcp) or host authenticated ssh are used to access the compute nodes.

### ***File I/O***

Particularly for analysis computing, large aggregate file I/O data rates (multiple streams to/from diverse nodes) are required. Data transfers over the high performance Infiniband network, if reliable, will be preferred to transfers over Ethernet. Conventional TCP/IP over Infiniband relies on IPoIB (“IP over IB”,

one of the protocols supported by the Open Fabrics Enterprise Distribution, or OFED, Infiniband software stack).

NFS has not proven to be reliable on our prior lattice QCD clusters for extensive file reading and writing, though it has been reliable for access to binaries and for smaller writing activities, such as job log files. Instead, command-based transfers using TCP, such as *rcp*, *scp*, *rsync*, *bbftp*, *etc.*, have been adopted for the transfer of large data files. On the JLab and FNAL clusters, multiple raid file systems available at multiple mount points have been used. Utility copy routines have been implemented to throttle access, and to abstract the mount points (*e.g.*, copy commands refer to */data/project/file*, rather than */data/diskn/file*).

FNAL and JLab use *Lustre* as an alternative to NFS. *Lustre* provides a POSIX-compliant file system visible from all worker nodes and from the cluster head node. *Lustre* has the property that the storage volume and aggregate performance (instantaneous rate of data movement summed across all active transfers) can be scaled upwards by adding additional storage server nodes (known as *OSS* nodes, which serve *OST* disk volumes). Each new storage server node adds additional independent spindles of disks to the file system.

The LQCD-ext project will carefully watch developments in the parallel file system area for changes that can impact the deployed systems. LQCD-ext will leverage work in this area performed by the large high energy physics experiments such as Atlas and CMS at the Large Hadron Collider. Relevant issues in this area include concerns over the long term viability of *Lustre* given the pending acquisition of Sun by Oracle, and the emergence and/or maturation of parallel file systems such as GPFS, pNFS (the parallel version of NFSv4), and Hadoop.

### ***Procurement Strategy***

LQCD-Ext will procure at most five separate lattice QCD computing systems, one in each of the five years of the project; here we are considering a mixed conventional and GPU-accelerated cluster purchase to be a single procurement, as these would take place at a single host laboratory. The guiding principal of all of these procurements is that the most cost effective hardware will be deployed, where effectiveness is judged by the quantity of science (and of course, quality of science in terms of the reliability of the numerical results) that will be produced during the lifetime of the individual lattice QCD system. In addition to commodity hardware and GPU-accelerated clusters, similar to those deployed during the preceding LQCD Computing Project and in the JLab LQCD ARRA project, we will evaluate alternatives such as the IBM BlueGene family of computers, traditional supercomputers such as the Cray XT series, purpose built machines such as the QCDOC, and other hardware suitable for lattice QCD calculations that may emerge.

At each of the annual project progress reviews, scheduled in or about the month of May of each fiscal year, LQCD-Ext will present the plans for the deployment that will occur in the next fiscal year. For example, in spring of calendar year 2010, the project presented the plans for the procurement that will occur in FY2011. The only exception to this schedule occurred during the first year of LQCD-Ext; the plans for the FY2010 acquisition were presented at the CD2/CD3 (Critical Decision 2 / Critical Decision 3) review in August of 2010. The annual presentation of procurement plans will include the selection of

hardware designs that will be considered, or the procedure that will be used to determine this selection, cost and performance estimates and their justifications, and a detailed schedule.

All procurements will utilize a multistep process:

1. Identify and characterize candidate computer and network hardware
2. Create a preliminary system design
3. Solicit vendor feedback on the preliminary system design through an RFI (Request for Information) solicitation
4. Create a system design based on vendor feedback and any new information that has emerged
5. Solicit vendor cost proposals for the system design through an RFP (Request for Proposal) solicitation
6. Evaluate RFP responses and award purchase order(s) to the winning vendor(s), issuing a final system design as necessary
7. Accept or reject the delivered system(s) based on acceptance testing

Both the preliminary system design and system design may include two or more selections of hardware; for example, in a given year, both commodity clusters and GPU-accelerated clusters may be included. Throughout the five years of the project, LQCD-Ext personnel will actively monitor the market, identifying and characterizing through benchmarking candidate hardware for upcoming procurements. Project personnel will also interact closely with computer and network manufacturers to understand product features and schedule roadmaps.

The evaluation and selection of hardware for the preliminary system design, and the evaluation of vendor responses to the RFP, will rely on the projected performance of the anticipated lattice QCD applications that will be run on the hardware during its lifetime. The particular mixture of lattice QCD applications to be used will be determined by LQCD-Ext Project staff in consultation with the USQCD Executive Committee and the USQCD Scientific Program Committee.

All awards will utilize firm, fixed-price contracts. Vendors will be encouraged to include modifications to the system designs in their RFI and RFP responses that would maximize the value of the delivered systems. Purchase awards will be based on best value evaluations that will include factors such as price/performance, quality of the vendor, quality of the proposed hardware, power consumption of the proposed hardware, impact on the facility infrastructure of the host laboratory, and usability of the delivered system.

LQCD-ext will procure storage for *Lustre* and NFS file systems separately from the computing systems. The amount of storage purchased will be determined in part from the requests that are required for all proposals to the Scientific Program Committee for allocations of time. The incremental storage added at each site annually will be at least as great as the sum of requested storage in the annual allocations proposals. Further, the storage will be deployed using a sufficient number of servers to meet the anticipated I/O bandwidth needs of the coming allocation year.

### **Strategy for the FY2012 LQCD-Ext Deployment**

As discussed in the sections above, in FY2012 the hardware candidates are an IBM BlueGene/Q system at BNL, a conventional Infiniband cluster at JLab, a GPU-accelerated Infiniband cluster at JLab, or some mixture of the three.

As of the time of the preparation of this document, the schedule for the availability of BlueGene/Q (BG/Q) hardware is not known. However, it is believed that BG/Q hardware may be available by the end of FY2012. Also, because BG/Q hardware is still under development, performance data are not available. Performance data for the principal LQCD algorithms (asqtad, clover, DWF), through arrangement with IBM, may be measured by the USQCD collaboration using prototype hardware by August 2012. Given the expertise of BNL in the operation of BlueGene hardware and the laboratory’s interest in hosting BG/Q systems, should this hardware be selected for the FY2012 purchase the LQCD-ext budget will be altered to provide for the hardware purchase and for operations support at BNL during the operational lifetime of the BG/Q (FY2012 through FY2014). Should BG/Q hardware not be chosen in FY2012, the current LQCD-ext budget, which calls for hardware funds to be allocated to JLab for a 2012 cluster acquisition and its subsequent operations, will be executed.

The LQCD-ext strategy for determining the hardware for FY2012 will take into account the availability of hardware, pricing, hardware performance, and full life-cycle costs. Because the project must request the distribution of FY2012 funds among the three laboratories by mid-August 2011, a sequence of information gathering steps will occur as listed in Table II below, culminating in the selection of the laboratory to host the FY2012 hardware. The possible hardware acquisition scenarios include:

- A full rack (1024) nodes of BG/Q hardware at BNL
- A half rack of BG/Q hardware at BNL, and a combination of conventional and GPU-accelerated clusters at JLab
- A combination of conventional and GPU-accelerated clusters at JLab

If it is determined in August 2011 that hardware will be procured at JLab, the budget breakdown between a conventional and a GPU-accelerated cluster will be determined by January 2012. Hardware at JLab will be released to production by the end of September 2012.

**Table II – FY2012 Acquisition Planning Process**

Step	Description	Target Due Date
1	The LQCD-ext Computing Project team (i.e., “the Project”) will provide the LQCD Executive Committee (EC) with data summarizing the distributions of job types and sizes during the prior year on the hardware operated by the Project (Infiniband clusters, GPU-accelerated clusters, and the QCDOC). The Project will request that the EC provide the anticipated scientific program requirements for various architectures (i.e., leadership-class machines, BG/Q rack or Infiniband cluster, and GPU-accelerated cluster). Information on USQCD hardware usage will be presented to the collaboration at the 2011 All-Hands Meeting May 5-6.	Apr 15
2	The Project will prepare the F12 Acquisition Strategy document for presentation and review at the FY2011 DOE Annual Progress Review. The Acquisition Strategy will outline the various options under consideration and the proposed process for selecting the mix of computing hardware that will be procured and deployed in FY12 using project funds.	May 10-11
3	The Project will request that the BNL site manager prepare a plan for procuring and	Jun 1

	operating a BG/Q rack, detailing estimating hardware, storage, deployment, and operations costs.	
4	The EC, with input from the Scientific Program Committee (SPC), will provide the Project with the anticipated scientific program requirements for various architectures (i.e., leadership-class machines, BG/Q rack or Infiniband cluster, and GPU-accelerated cluster). A helpful way of conveying this information would be for the EC to provide an estimate of the relative fractions of “analysis core-hours” and “cost-equivalent GPU-hours” needed to support the scientific program over the next 1 to 2 years. Ideally, the EC will provide the Project with anticipated needs on a per year basis for FY12 and FY13.	Jun 15
5	The BNL site manager will provide the Project with a preliminary plan for procuring and operating a BG/Q, including estimated costs and schedule.	Jul 1
6	The BNL site manager will provide the Project with a final plan for procuring and operating a BG/Q, including costs (hardware, storage, costed manpower for deployment and operations) and schedule.	Jul 22
7	The Project will review the technical landscape, conduct an alternatives analysis of the various options, and propose a cost-effective solution for the FY12 hardware deployment. When considering viable options, the Project will need to factor in the total cost of ownership (TCO) for each solution. In addition to hardware and deployment costs, TCO also includes on-going operations and support costs. Hardware costs will include any necessary storage acquisitions. For solutions involving Infiniband clusters and GPU-accelerated clusters, an operations cost model already exists. For a BG/Q option, the Project will need to understand the cost model for operating a BG/Q at BNL. Information on cost and availability of production BG/Q hardware will also be needed. Results of the analysis and an overview of the proposed solution will be summarized in the Alternatives Analysis document. The Project will verify the host laboratory’s ability and willingness to provide the necessary space, power, and cooling for each alternative.	Jul 29
8	The EC will review the Alternatives Analysis document and proposed FY12 hardware solution, and will provide advice on how to proceed to the Project Manager.	Aug 10
9	The Project will analyze the advice of the Executive Committee as well as any new data that might have been obtained, and will produce the final plan for the FY12 hardware deployment. The Project Manager will advise the EC, the host laboratories, the Federal Project Director, and Project Monitor of the planned FY12 hardware acquisition.	Aug 15
10	The Project Manager will revise the project budget as necessary to accommodate the FY12 hardware solution. Depending on the alternative selected, changes may be required in the planned allocation of funds across the three host laboratories.	Aug 20
11	The Project Manager will provide the Federal Project Director with the FY12 Financial Plan, containing the requested distribution of project funds to the three host laboratories.	Aug 20 (est.)
12	The Project will develop a detailed acquisition plan, with timeline, based on the approved FY12 architecture solution.	Sep 30,2011
13	The Project will execute the FY12 acquisition plan in a manner that meets approved performance goals and milestones.	Sep 30, 2012