Acquisition Plan – FY2008 and FY2009
for the
Lattice QCD Computing Project (LQCD)

at

Brookhaven National Laboratory, Brookhaven, New York
Fermi National Accelerator Laboratory, Batavia, Illinois
Thomas Jefferson National Accelerator Facility, Newport News, Virginia

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

Date: April 30, 2008

PREPARED BY:
Don Holmgren, FNAL

CONCURRENCE:

Bill Boroski — 5-8-2008
Bill Boroski — Date
Contractor Project Manager

Acquisition Plan - FY2008 and FY2009
Change Log

Revision No.	Description/ Pages Affected	Effective Date
Revision 1.0	Final version for May 14, 2007 DOE External Review	May 14, 2007
Revision 1.1	Applied standard project template to revision 1.0	August 14, 2007
Revision 2.0	Final version for May 13, 2008 DOE External Review	May 13, 2008

Table of Contents

Introduction

The Lattice QCD Computing Project develops and operates new systems in each year from FY2006 until 2009. These computing systems are deployed at Fermilab (FNAL) and at Jefferson Lab (JLab). In addition, the project operates the 4.2 Tflop/s US QCDOC supercomputer at Brookhaven National Lab (BNL), as well as the prototype clusters developed under the SciDAC program at FNAL and at JLab in 2004-2006. Table 1 shows the planned total computing capacity of the new deployments, and the planned delivered (integrated) performance as stated in the FY2009 OMB Exhibit 300 for the project. For FY2006 and FY2007, the table shows the achieved figures as well as the planned numbers. In FY2006, FNAL deployed the "Kaon" cluster with 2.3 TFlop/s capacity, and JLab deployed the "6N" cluster with 0.32 TFlop/s capacity. In FY2007, JLab deployed the "7n" cluster with 3.0 TFlop/s capacity. The planned integrated performance figures assume, at the beginning of FY2008, 1.9 Tflop/s of total capacity from the SciDAC prototype clusters at JLab and FNAL, and 4.2 Tflop/s capacity from the US QCDOC at BNL. Note that in all discussions of performance, unless otherwise noted, the specified figure reflects an average of the sustained performance of domain wall fermion (DWF) and improved staggered (asqtad) algorithms. On clusters, DWF code sustains 20% to 40% greater flop/sec than asqtad code.

	FY 2006	**FY 2007**	**FY 2008**	**FY 2009**
Planned computing capacity of new Deployments, Tflop/s (FY06 and FY07 show achieved value in parentheses)	2.0 (2.6)	2.9 (3.0)	4.2	2.0
Planned delivered Performance (JLab + FNAL + QCDOC), Tflop/s-yr (FY06 shows achieved value in parentheses)	6.2 (6.27)	9 (9.67)	12	15

Table I – Performance of New System Deployments, and Integrated Performance (DWF:asqtad averages used)

In FY2007, the project procured and deployed a cluster at JLab. In FY2008 and FY2009, the new deployments will occur at FNAL. The Project plans to execute a FY2008 acquisition with an option to expand the resulting system early in FY2009 with additional identical hardware. Such an acquisition would use a single subcontract and two delivery dates, one in October 2008 with the hardware and services purchased using FY08 funds, and the second smaller delivery in December 2008 with the hardware and service purchased using FY09 funds, executed by means of a contract option. By combining the two planned purchases into a single subcontract, a single larger homogenous cluster could be deployed if scientifically advantageous. Combining the purchases would reduce total labor costs associated with the procurement, deliver integrated scientific computing capacity at a greater rate during the first three years of operation of the combined system, and result in a larger homogenous lattice QCD (LQCD) cluster. On the other hand, if a separate

FY09 purchase is scientifically advantageous, the project will not exercise the contract option and will instead perform a separate purchase.

All LQCD Computing Project hardware procurements utilize firm, fixed-price contracts with vendors specializing in COTS hardware. These procurements are awarded on the basis of best value. The steady state operations of the project clusters are performed by the three host laboratories, each of which is a government-owned contractor-operated facility.

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, high memory bandwidths are required.

The currently available commodity processors with the greatest memory bandwidths are Intel IA32/Intel64 processors with 1333 to 1600 MHz (effective) front side buses (Xeon "Woodcrest", "Clovertown", and "Penryn", Pentium "Conroe" and "Kentsfield"), and the AMD x86_64 processors (Athlon 64 FX, Phenom, and Opteron). The Pentium, Athlon, and Phenom processors can only be used in single processor systems. The Xeon and Opteron processors can be used in dual and quad processor systems. The total cost of quad processor systems of both types, including the cost of the high performance network, exceeds the cost of two dual processor systems with network. At the current cost of Infiniband and competing high performance networks, quad processor systems are not as cost effective as single or dual processor systems.

Since late 2006, Intel and AMD have switched all new processors of relevance to lattice QCD to dual or quad core. The JLab "6n" and Fermilab "Kaon" clusters purchased and deployed in 2006 use dual core processors; "6n" uses single-socket dual-core Pentium D 830 motherboards, "Kaon" uses dual-socket dual-core Opteron 270 motherboards. JLab's "7n" uses dual-socket quad-core Opteron 2347 motherboards. 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. The multi-core versions 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 multicore, moving predominantly to four or more cores in 2008 and beyond.

All current commodity dual processor Xeon motherboard designs use a single memory controller to interface the processors to system memory. As a result, the effective memory bandwidth available to either processor is half that available to a single processor system. Opteron processors have 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 makes multiprocessor Opteron systems viable

for lattice QCD codes. In the 2006 and 2007 project acquisition of the "Kaon" and "7n" clusters, multiprocessor Opteron systems were chosen, as these were the most cost effective designs.

In 2006, Intel began selling dual processor systems with dual independent memory buses connecting the processors to a memory controller in turn connected to multiple DIMM channels. This memory subsystem is based on FBDIMM ("fully buffered DIMM") technology. To date, this technology has not proven to deliver as much memory bandwidth as the integrated memory controllers on AMD Opteron systems. However, in 2007 and 2008 Intel planned to separately introduce two new generations of chipsets and processors fabricated using a new 45-nanometer process.

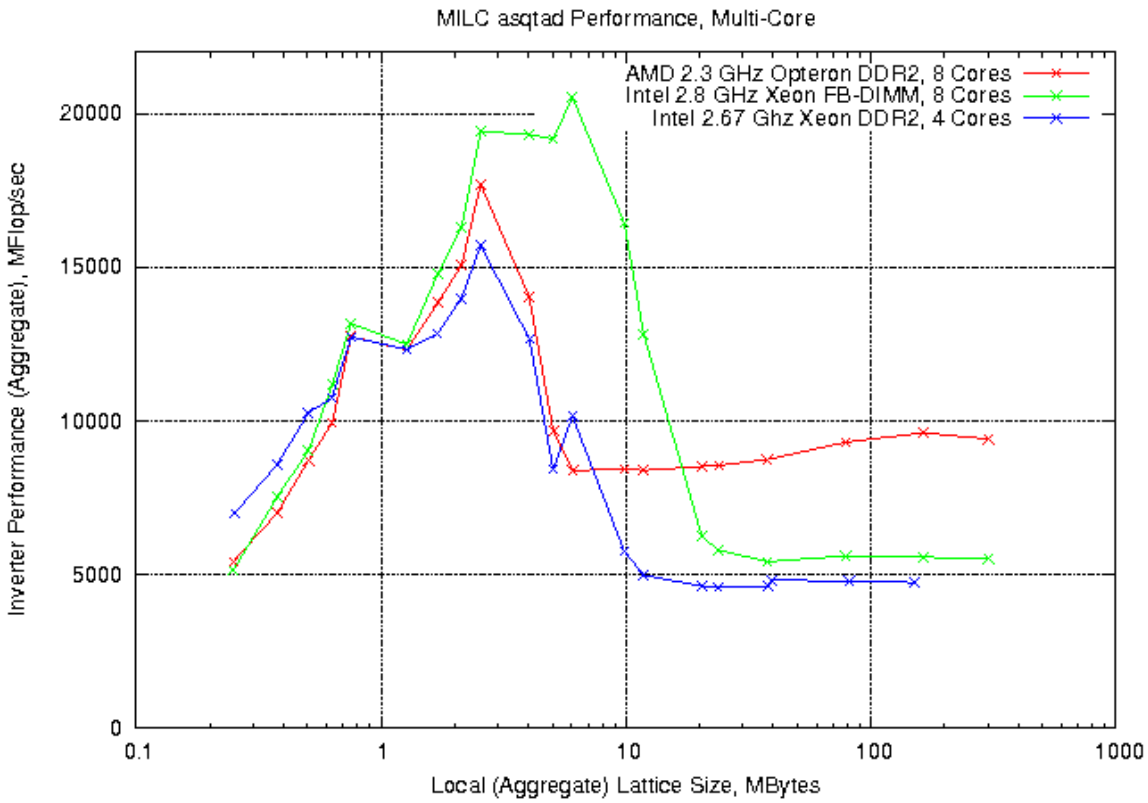
The first of the two generations has code-names "Penryn" for the processor family, and "Seaburg" for the chipset family. The combination of the new processors and chipsets allows for faster memory buses, up to 1600 MHz from the current 1333 MHz, as well as improved achievable memory bandwidth due to modifications of the memory controller design and the "snoop filter." Improvements in sustainable memory bandwidth of 25% over the current designs are predicted. Further, because of improvements in controlling leakage currents, the "Penryn" family will be capable of higher clock speeds, up to 3.2 GHz compared to the current generation's 2.67 GHz maximum speed. Systems incorporating Penryn/Seaburg entered the market in fall of 2007.

The second of the two generations has code-names "Nehalem" for the processor family, and "Tylersburg" for the chipset family. According to industry rumors, substantiated by conversations with Intel, "Nehalem" will incorporate a new memory controller design and will use DDR3 memory rather than FB-DIMMs. Estimates of improvement of in sustainable memory bandwidth over their 65nm designs are 6 to 7-fold aggregate for a two-socket system. If realized, these improvements would greatly impact the performance of LQCD code. Intel schedules indicate that the first "Nehalem" processors should be available for the LQCD project to benchmark by approximately May 2008. More recent schedules indicate that "Nehalem" systems will not ship in volume until January or February of 2009.

The latest AMD quad core processors, code-named "Barcelona", were available in time for the FY2007 7N acquisition. Preliminary LQCD code benchmarking in late April on engineering samples of the Barcelona family showed very strong performance, leading to the decision to deploy "Barcelona" processors on the "7n" cluster.

The graph below shows the relative performance of nodes built using Intel and AMD processors on the MILC asqtad application code. The data shown were generate by running this MPI application with one process per core. The horizontal axis corresponds to the aggregate lattice size (*i.e.*, the sum of the per-core local lattice sizes), and the vertical axis corresponds to the aggregate performance (*i.e.*, the sum of the per-core performance values). The performance shown is the MFlop/s sustained in the sparse matrix inverter. The three curves correspond to a single processor Intel system based on DDR2 memory, a dual processor Intel system based on FB-DIMM memory, and a dual processor AMD system based on DDR2 memory. In all cases four cores per processor were used. This plot

clearly shows the advantage in performance on LQCD applications currently enjoyed by the AMD architecture because of its superior memory bandwidth.



High Performance Network

Based on SciDAC prototypes in FY 2004 and FY 2005, the "6n" and "Kaon" clusters purchased in FY2006, and the "7n" cluster purchased in FY2007, Infiniband is the preferred choice for the FY08/FY09 cluster. The cluster will use double data rate (DDR) or quad data rate (QDR) 4X Infiniband parts.

Current switch configurations from multiple Infiniband vendors include 24, 96, 144, and 288-port switches. For the large clusters to be built in this project, leaf and spine designs are required. Because DDR 4X HCA bandwidths exceed the requirements for lattice QCD codes, oversubscribed designs will be used. A 2:1 design, for example, would have 16 computers attached to a 24-port switch, with the remaining 8 ports used to connect to the network spine.

A 2:1 oversubscribed design supporting 1024 compute nodes would employ 64 24-port leaf switches, and either six 96-port, four 144-port, or two 288-port spine switches. A 5:1

oversubscribed design supporting 1024 compute nodes would employ 52 24-port leaf switches, and either two 144-port or one 288-port spine switch.

Starting in mid-2008, new Infiniband switch configurations are expected based on 36-port rather than 24-port building blocks. These switches will be considered for the FY2008/FY2009 purchase.

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), serial-over-LAN, and NFS access to "home" file systems for access to user binaries. All current motherboard candidates support two 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. Either serial-over-LAN (standard with IPMI 2.0) or serial multiplexers will be used to provide these serial connections.

Network Plan

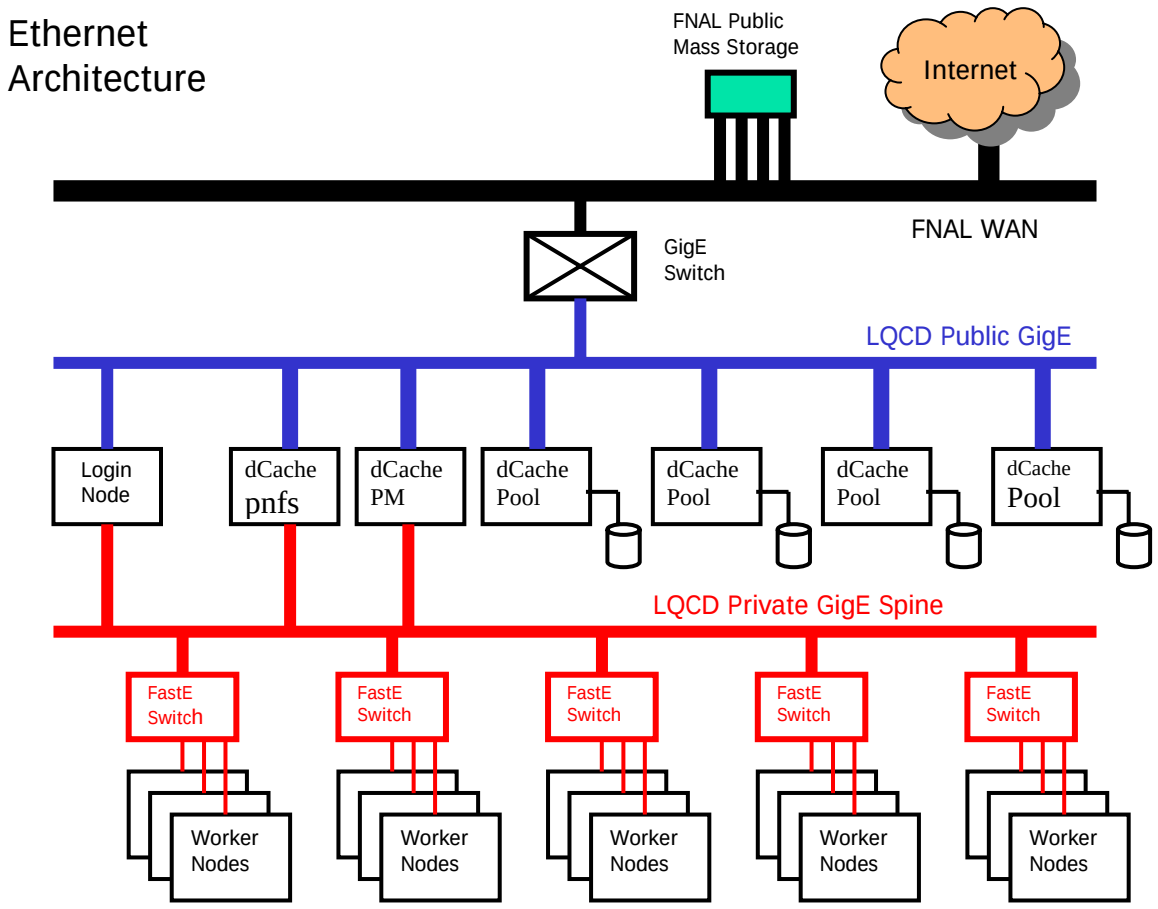
We will replicate the network layout currently used on all of the FNAL and JLab lattice QCD clusters. In these designs all remote access to cluster nodes occurs via a "head node", which connects to both the public network and to the private network that forms the sole connection to the computer nodes. 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, with SDP (Socket Data Protocol) available as an attractive alternative that incurs less processor overhead.

NFS has not proven to be reliable on our prototypes 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/disk*n*/file*. FNAL uses *dCache* as an alternative. *dCache* provides a flat file

system with scalable, throttled (reading), and load balanced (writing) I/O; additionally, it supports transparent access to the FNAL tape-based mass storage system.



Ethernet Network Architecture Diagram and Description

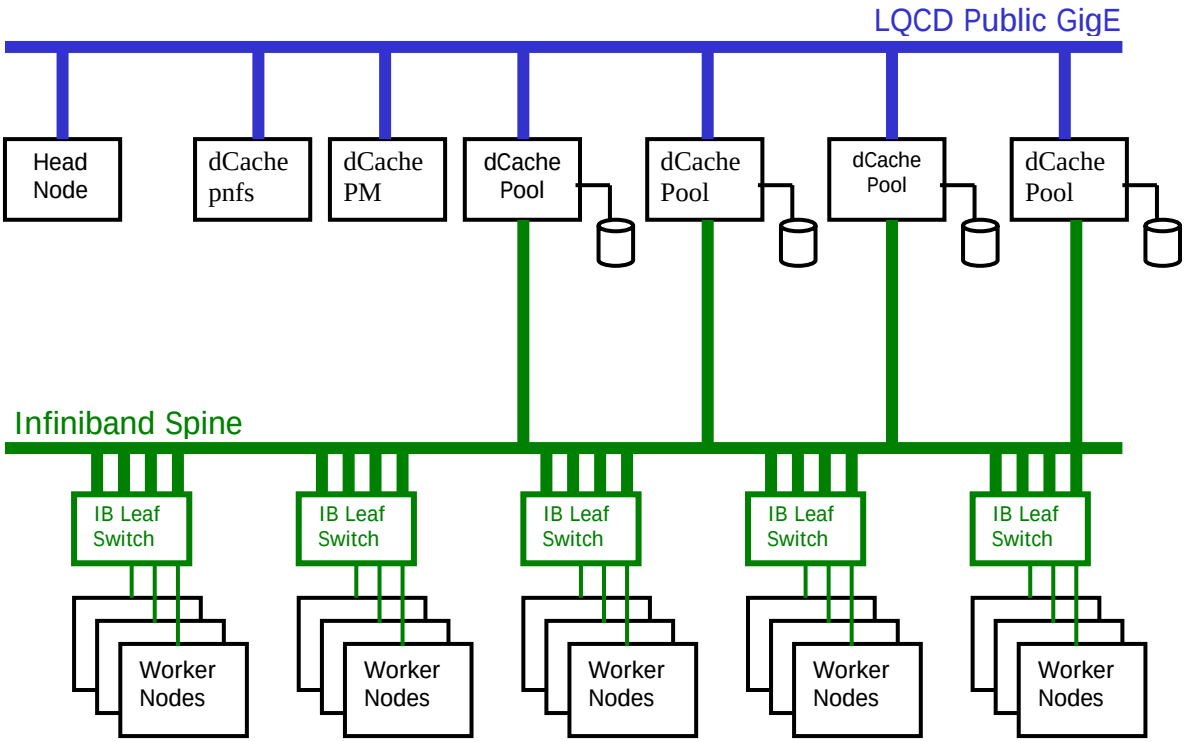
The diagram above shows the Ethernet network architecture of the Kaon cluster installed at Fermilab in FY2006. A similar architecture will be used for the FY08/FY09 cluster. Public and private gigE networks will be used, as shown in the diagram. The public gigE network will connect via a Cisco

switch to FNAL's wide area network via a set of four channeled gigabit Ethernet connections. The FY08/FY09 facility will access the Fermilab mass storage facility via the Fermilab WAN. Within the mass storage facility are multiple *tape mover nodes*, each attached to an STK 9940B tape drive (in the future, LTO-3 drives will be available as well). Also within the mass storage facility are multiple *dCache pool nodes*, which provide a disk cache on top of the tape storage.

Users of the FNAL facilities login to the head (login) node; the scheduler (*Torque* plus *Maui*) runs on either this node or another dedicated node. Approximately 10 Tbytes of local disk are attached to the login node.

The worker nodes will be connected via fast Ethernet switches with gigabit Ethernet uplinks to a private spine gigabit Ethernet switch. The head node will communicate via this private network with the worker nodes. This network is used for login access to the worker nodes by the scheduler (using *rsh*). Each worker NFS-mounts the /home and /usr/local directories from the head node. Binaries are generally launched from the /home directory. Each worker node has considerable (120 Gbytes or greater) local scratch space available. High performance I/O transfers to and from the worker nodes utilize the Infiniband network (see drawing below); the head node is bridged to the Infiniband network via a *router node* (not shown in drawing) that is connected to both the private gigabit Ethernet network and the Infiniband fabric.

Infiniband
Architecture



Infiniband Architecture Diagram and Description

The diagram above shows the Infiniband architecture used on the Fermilab Kaon cluster. A similar architecture will be used on the FY08/FY09 cluster.

On the FY08/FY09 cluster, similar to the JLab 6N and FNAL Kaon clusters, a leaf and spine approach will be used. Each set of worker nodes will be connected to a 24-port or 36-port leaf switch. Multiple links connect each leaf switch to a central spine switch (or switches, depending upon implementation). The Infiniband fabric will be used for internode communications for LQCD applications via MPI (*mvapich* and *openmpi* versions will be available). The Infiniband fabric will also be used for high performance file I/O via TCP, using IPoIB. A high performance I/O path from the head node to each worker node will be available through the use of a *router node* which will route packets from gigabit Ethernet to the Infiniband fabric and vice versa (not shown in diagram). Infiniband fabric connections to each of the NFS server nodes at JLab will allow for high performance I/O by the worker nodes.

Because the FY08/FY09 cluster will be housed in a different building than Kaon, but will share the same dCache pool nodes, a dedicated, private 10-Gbit/sec Ethernet link between the two buildings will be used to provide a high performance path for access to the pool nodes from all Fermilab LQCD clusters.

Software Deployment and Other Integration Tasks

To bring the FY08/FY09 cluster into production, the following integration tasks will be necessary (order may vary from that shown):

1. Prepare system installation images for worker nodes (Scientific Linux). These images will include the Infiniband software stack (OpenIB, or commercial) as well as the SciDAC LQCD shared libraries.
2. Install system images on all worker nodes.
3. Unit test worker nodes. These tests will include memory tests, multiple reboot and power cycle tests, disk tests, and LQCD single node application testing and performance verification.
4. Unit test worker racks. The vendor will be required to configure each rack after assembly as an independent Infiniband fabric, and to provide the results of a set of LQCD application parallel benchmarks run on various subsets of the nodes in the rack. Our unit test will repeat these measurement and will verify that the results match those provided by the vendor.
5. Integrate worker racks. This requires the interconnection of the individual racks to the Infiniband and gigabit Ethernet spine fabrics, and the configuration of the Infiniband subnet manager and monitoring facilities. The vendor will be responsible for this step.
6. Configure IPMI facilities on all worker nodes; this includes initializing BMC network parameters (IP addresses, subnet masks, ARP and gratuitous ARP configuration).
7. Test IPMI facilities on all worker nodes.
8. On head node, deploy commercial compilers (Intel, Portland Group, Pathscale as requested by user community).
9. On head node, build and deploy SciDAC libraries.
10. On head and worker nodes, deploy SciDAC common runtime environment.
11. On head and worker nodes, deploy and configure batch system (*Torque* plus *Maui*).
12. On head and worker nodes, create authorized user accounts.
13. Test batch system.
14. Test LQCD applications.

Computing Room Facility for the FY2008/FY2009 Cluster

Fermilab will house the FY08/FY09 cluster in the "GCC-C" computer room. GCC is the acronym for "Grid Computing Center". Currently GCC consists of 2 computer rooms, GCC-A, and GCC-B, each of which provides 840 KW of usable power with cooling for computer systems. GCC-C will similarly provide 840 KW of power. The building for GCC-C exists, and the construction project that will build out GCC-C began in March 2008. Construction of the facility, based on the earlier GCC constructions, will take three to six months, with beneficial occupancy planned for mid-June, and access planned for mid-July. To take into account the risks of a late construction start due to a continuing budget resolution and to delays associated with any other budget constraints, the LQCD Project assumed that GCC-C construction would start in March 2008 and last for six months. With the inclusion of an additional month of float, access was assumed to be available on October 1, 2008. The LQCD Project schedule will track these external dependencies.

Schedule

The FY08/FY09 procurement will consist of a number of phases. In the first phase, running from June 2007 through April 2008, LQCD Project staff will evaluate the performance of LQCD codes on the various hardware options detailed in this document. These evaluations will determine the configurations of processor, chipset, and networking hardware that will be in the competitive range for vendor bidding. In the second, slightly overlapping phase, a Request for Information (RFI) will be used to obtain information from prospective vendors about their ability to design and build the cluster. In the third phase, a Request for Proposal (RFP) will be used to solicit vendor bids and to award the subcontract for the FY08/FY09 cluster. This subcontract will be a firm, fixed-price contract competitively awarded based on best value. The subcontract will stipulate two delivery dates, the first corresponding to hardware funded by FY08 project monies, and the second optionally delivery date corresponding to hardware funded by FY09 monies through the exercise of a contract option. In the fourth phase, the vendor will deliver a single machine. Next, the vendor will integrate a full rack (approximately 40 machines with an Infiniband and an Ethernet leaf switch, capable of running production LQCD jobs). Fermilab personnel will visit the vendor site to inspect the full rack and to verify operations and performance. Next, vendor will proceed to build the remaining racks of the cluster. Fermilab will approve the hardware at each step using functional tests before releasing the vendor for the next increment. In the final phase, the FY08 equipment will be integrated and tested by Fermilab, and the FY09 portion of the subcontract will be executed. The FY09 hardware will be integrated with the FY08 hardware, and after a period of "friendly user" production the full system will be released to production.

The abbreviated schedule for the FY08/FY09 cluster deployment is shown below:

- Summer, 2007: Evaluate Intel “Penryn” processor family and “Seaburg” chipset family for price/performance for LQCD codes. Evaluate AMD “Barcelona” processor family for price/performance.
- Fall 2007 - Spring 2008: Test Intel “Nehalem” processors and “Tylersburg” chipsets for price/performance for LQCD codes. Depending upon availability from Intel, this testing may occur as late as April 2008.
- February 15, 2008: Preliminary Cluster Design Document completed
- March 15, 2008: Request for Information (RFI) released to vendors.
- April 15, 2008: Evaluation of RFI responses complete and documented.
- May 30, 2008: Request for Proposal (RFP) released to vendors.
- July 9, 2008: RFP responses evaluated and award recommendation complete.
- July 23, 2008: Purchase subcontract awarded, committing FY08 project funds.
- August 8, 2008: Approval of sample unit.
- Sept 8, 2008: Delivery of remaining equipment.
- October 15, 2008: Exercise of option in subcontract for additional identical hardware, committing FY09 project funds.
- October 15, 2008: “Friendly User” production on hardware integrated from the September delivery.
- December 1, 2009: Release to production of FY08 cluster.
- December 15, 2008: Delivery of FY09-funded hardware.
- January 15, 2009: Release to production of full combined FY08/FY09 cluster.