# The Hopper System: How the Largest XE6 in the World Went From Requirements to Reality

**Katie Antypas, Tina Butler,** and **Jonathan Carter**
*NERSC Division, Lawrence Berkeley National Laboratory*

**ABSTRACT:** *This paper will discuss the entire process of acquiring and deploying Hopper from the first vendor market surveys to providing 3.8 million hours of production cycles per day for NERSC users. Installing the latest system at NERSC has been both a logistical and technical adventure. Balancing compute requirements with power, cooling, and space limitations drove the initial choice and configuration of the XE6, and a number of first-of-a-kind features implemented in collaboration with Cray have resulted in a high performance, usable, and reliable system.*

**KEYWORDS:** Deployment, Testing, XE6

## 1. Introduction

The National Energy Scientific Computing (NERSC) Center, located at the Lawrence Berkeley National Laboratory, is the flagship computing center for the US Department of Energy, serving the whole range of science sponsored by DOE Office of Science. The current systems deployed include a more than 9000 node Cray XT4, Franklin, a 6400 node Cray XE6, Hopper, and a 400 node IBM iDataplex, Carver. The systems share a common high-performance filesystem, based on GPFS, for permanent data storage and locally attached storage for temporary data.

Deployment of our next generation computing system, named after Admiral Grace Murray Hopper, has taken place in two phases. Phase 1, a Cray XT5, entered production on March 1, 2010. The Phase 2 system was deployed in the fall of 2010 and was be provisioned with the Cray Gemini interconnect and nodes having two 12-core AMD Opteron 6100 series (Magny Cours) chips. The Phase 2 system entered production of May 1, 2011.

## 2. Requirements and System Selection

The Request For Proposals (RFP) was formulated based on user requirements during the spring and summer of 2008. At that time, NERSC conducted a formal requirements gathering exercise producing a document called the Greenbook describing future computing needs from all of the major science domains served. In addition to the hardware and software characteristics needed in the next generation system, NERSC identified an unmet need in a set of users whose productivity was severely limited by lack of allocated time. In other words, the only things preventing a large class of users from moving to higher concurrencies and running larger, more accurate, or more encompassing simulations was a lack of installed resources at NERSC.

The RFP used the Best Value Source Selection (BVSS) method, where a set of Minimum Requirements and Performance Features lay out a list of requirements that must be met for a proposal to be considered further, and a list of desirable features respectively. Vendors are free to propose additional features that they perceive will add value to their bid, and these features may be evaluated in selecting the best proposal. The RFP consisted of 13 Minimum Requirements, for example “An application development environment consisting of at least:

standards compliant Fortran, C, and C++ compilers, and MPI and MPI-IO libraries" and 38 Performance Features, for example, "Support for advanced programming languages such as UPC, CAF, the emerging HPCS languages, shared memory abstractions such as Global Arrays through one-sided messaging (e.g., put/get remote memory access semantics), efficient RDMA support, and/or global addressing". Energy efficiency was targeted via the minimum requirement to use 480-volt power at the cabinet, and performance feature to operate at higher levels of thermal range.

In addition to Minimum Requirements and Performance Features, the RFP featured a comprehensive set of benchmarks ranging from kernels, mini-applications through applications for conducting full scientific simulation. The full applications were drawn from the entire NERSC workload including: atmospheric modeling, computational chemistry, fusion, accelerator physics, astrophysics, quantum chromodynamics, and materials science. A summary description is shown in Table 2.1.

| Application | Domain | Algorithm Space | Notes |
|---|---|---|---|
| CAM | Climate | Navier Stokes, CFD | F90+MPI |
| GAMESS | Chemistry | BLAS2, BLAS3 | F90+DDI (application specific communication API) |
| GTC | Fusion | PIC, finite difference | F90+MPI |
| IMPACT-T | Accelerator | PIC, FFT | F90+MPI |
| MAESTRO | Astrophysics | AMR, block-structured grid | C++,F90+MPI |
| MILC | QCD | Conjugate gradient, sparse matrix | C+MPI |
| PARATEC | Materials | BLAS3, FFT | F90+MPI, ScaLAPACK |

The evaluation process also considers supplier attributes, such as ability to produce and deliver a system, commitment to the HPC space, and corporate risk. Finally the cost of ownership, including site modifications, ongoing power and cooling costs, are factored in.

### *2.1 Sustained System Performance*

The performance of the full applications at the highest concurrency benchmarking runs is averaged to determine the sustained system performance (SSP). The benchmark times are used with a flop count measured on a reference system to determine a flop rate per core. The geometric mean of the per-core flop rates multiplied by the number of cores on the system produces an extensive performance measure to compare systems between each other.

The Cray proposal was judged to be the best value having the best application performance per dollar (as judged by the SSP metric), the highest sustained application performance of all proposals, and the highest sustained performance for the power consumed.

## 3. Hopper System

A two-phase deployment was negotiated to provide a balance of early increase in computational power at NERSC, combined with a later system comprised of mostly new technology.

### *3.1 Phase 1 System*

The Phase 1 system is based on XT5 technology, with 668 dual socket nodes, each with 2.4 GHz quadcore AMD Opterons and 16 GB of DDR2 RAM. There are a further 46 internal service nodes to provide MOM, Lustre router, DVS server, network, boot, and system logging functionality. Finally, 12 nodes (on compute blade hardware) provide shared-library access via DVS.

The Phase 1 Factory Test was conducted in Fall 2009 at the Cray manufacturing plant in Chippewa Falls, WI. The system was complete aside from half the IO equipment, which was late in being delivered. The system completed a 72-hour stability test without interruption. All functionality tests passed or only had minor issues. The Phase 1 Acceptance test was conducted during the last part of 2009 and into 2010 after the system was installed at NERSC. Early users were added at the end of November, with all users enabled by the beginning of a 30-day availability test. The system demonstrated 99% availability during the test.

### *3.2 Phase 2 System*

The Phase 2 system is based on the new Cray XE6 platform, with 6384 dual-socket compute nodes, each with a 2.1 GHz 12-core AMD Opteron and 32 GB of DDR3 RAM. To provide some flexibility for applications that need more memory per core, 384 nodes have 64 GB of DDR3 RAM. In addition, there are 83 internal service nodes to provide Lustre router, DVS server, network, RSIP, boot, and system logging functionality. The DVS server nodes are used to connect to the NERSC Global Filesystem (NGF), a GPFS installation that serves all NERSC systems. Finally, 56 nodes (on compute blade hardware) provide MOM (head node) functionality and shared-library access via DVS. The ECOphlex liquid cooling delivered with the Phase 2 system enabled it to be located in a tighter space than would have been possible with an air-cooled system.

The Phase 2 Factory Test was conducted during the summer of 2010 in the Cray manufacturing plant in Chippewa Falls, WI. Only 8 cabinets were available for testing, with no external filesystem available as all equipment had been delivered to NERSC with Phase 1. NERSC also shared access to a 20-cabinet system being built for delivery to the ACES Sandia/LANL partnership.

The Phase 2 system was slowly integrated into the NERSC computing center. Four esLogin nodes and one of the scratch filesystems were moved from the Phase 1 system, upgraded to be compatible with the newly delivered hardware and linked to the first Phase 2 shipments. When the full Phase 2 compute portion was delivered the remaining scratch filesystem was also moved over, while the Phase 1 system continued to operate using scratch space provided by the NERSC Global filesystem. The goal was to minimize disruption in the operation of Phase 1 and allow as smooth a transition as possible.

The system completed a 72-hour stability test without interruption. Cray demonstrated “warm swap” – removing a section of compute nodes from a running system with bringing the system down and replacing them and re-introducing them into service. The Phase 2 Acceptance test was conducted from December 2010 through April 2011 after the system was installed at NERSC. Early users were added starting in November, with all users enabled by the end of December. The 30-day availability test ended in April, after the test completed successfully with 97.7% availability. Hopper was placed No. 5 on the Top-500 List at SC’11 with a HPL performance of 1.05 PFlop/s.

### *3.3 External Services*

During the contract negotiation for the Hopper system, one of the goals was to architect a system based on a Cray technology offering that could incorporate the many advantages of a tightly coupled XT system, but that could address some of issues seen in day-to-day production. As a solution, Cray Custom Engineering designed a configuration consisting of a set of external servers, based on Dell hardware, to augment the tightly coupled compute pool hosted on XT hardware.

A set of external servers augmented the Phase 1 and subsequently the Phase 2 compute system. The Lustre filesystem is not hosted on the XT, but on a subset of the external servers (esFS nodes). An Infiniband network, DDR with Phase 1 upgraded to QDR for Phase 2, provides the connectivity between the OSS/MDS and Lustre router nodes. Eight of the external servers act as login hosts (esLogin) where users can build applications, submit batch jobs, and run serial pre- and post-processing steps.

The esFS nodes are divided between 4 MDS and 52 OSS nodes. Filesystem storage is provided by 26 LSI 7900 storage subsystems, each with dual controllers and 120 TB of disk configured as 12 RAID6 (8+2) LUNs for a total of more than 2 PB of user accessible storage. In the current Phase 2 system, the OSS are connected via QDR Infiniband fabric to 56 internal service nodes serving as Lustre routers. The esFS serves two independent, identically-sized scratch filesystems. Both the esFS nodes and internal router nodes run the same generation Lustre software, although support for the external server Lustre 1.8.3 installation is provided by Oracle and the internal Lustre 1.8.4 is provided by Cray.

Finally, four servers act as data movers (esDM) providing offload from the login nodes for high bandwidth traffic to the NERSC mass storage system. External node characteristics are shown in Table 3.1 and a schematic of the system layout is shown in Figure 3.1.

| Server Type | Count | Model | CPU | Memory |
|---|---|---|---|---|
| esLogin | 8 | Dell R905 | 4 x AMD Opteron Quadcore 2.4 GHz | 128 GB |
| esFS (OSS + 3 MDS) | 52+4 | Dell R805 | 2 x AMD Opteron Quadcore 2.6 GHz | 16 GB |
| esDM | 4 | Dell R805 | 2 x AMD Opteron Quadcore 2.6 GHz | 16 GB |
| esMS | 1 | Dell R710 | 4 x Intel Xeon Quadcore 2.67 GHz | 48 GB |

Table 3.1: External Server Configuration

## 4. Working with External Servers

In this section we describe the challenges of working with external servers, and the benefits that we have ultimately derived from their deployment. Here, we confine the discussion to esLogin which have provided the greatest benefits to the reliability and usability of the system. Our experiences in using external Lustre and esFS, and also using DVS to connect to the GPFS NERSC Global Filesystem are described in [1], and previous work in deploying external services are described in [2].

Our aim was to increase the usability of the login environment by ensuring faster compiles, and allowing for more complex workflows where a user might run some pre or post processing steps on the login nodes. The esLogin nodes must provide all the functionality of an internal login node as transparently as possible, including: developing and building applications; obtaining system status and statistics; running jobs and interacting with the batch system.

Developing and building applications for the XE is taken care of via the Cray CADE/CADES package that provides cross-compilers and libraries and can be used on many Linux servers [3]. Similarly, obtaining system status and node use statistics is enabled by the set of commands such as xtstat, etc., that are bundled into the Cray eswrap package. These commands contact the XT compute system and display results as if they were run on an internal service node.

For the batch system, each esLogin node runs a Torque job submission server with a set of routing queues that normally forward the job on to the central Torque server running on the system database node of the XT system. In the event that the central server is unavailable, the jobs are routed to local queues on each node and held. When the central server is available, these local queues are started and the jobs are routed to the central server. Since the numerous Torque servers can be confusing to users, a set of batch command wrappers were written by NERSC staff and provide batch environment continuity for the users by hiding the implementation details.

The added computational power and large memory available on the esLogin nodes has elicited the most positive response from users. One esLogin server has 4 quad-core processors and 128GB of memory compared internal service nodes which have two dual-core processors and 8GB of memory. This represents and increase of 4 times the computational power and 16 times more available memory than an internal service blade could offer. We typically see 100-200 users logged on to a single esLogin node during normal working hours, so this is a significant advantage. Users on the Hopper system are able to compile applications, run python scripts, and post-processing applications such as IDL with much less interference from other users. Additionally, because the esLogin nodes are external to the main XE6 system, they are often kept available when Hopper is taken down for maintenance. This provides continuity for users who can login, compile applications and submit jobs which are held locally on the esLogin nodes until the XE6 becomes available again.

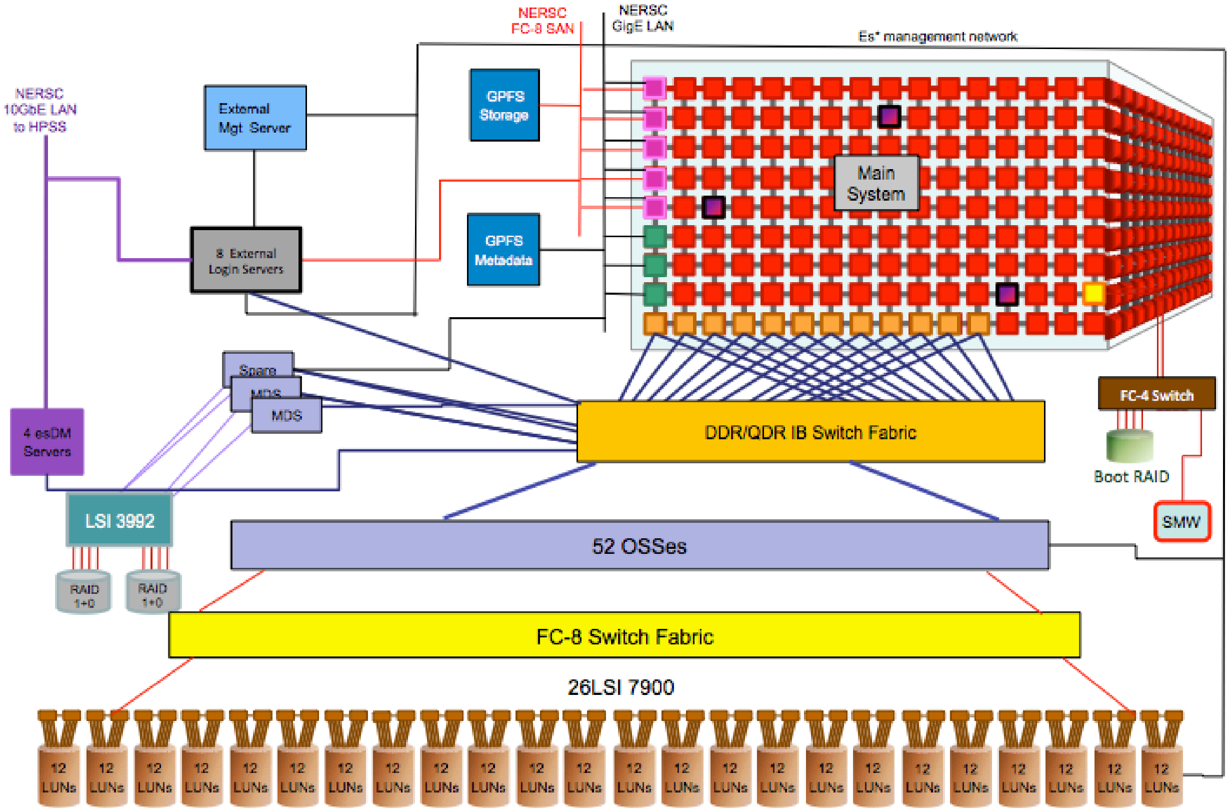

Figure 3.1 External Server Configuration

The greatest challenge we encountered installing and configuring the esLogin nodes was maintaining software consistency both across the pool of esLogin nodes, as well as between the esLogin nodes and the internal MOM nodes. esLogin node configuration is not managed by the "shared root" feature of the internal compute portion of the XT and must be managed by some external mechanism. After evaluating several alternatives, Cray custom Engineering chose Bright Cluster Manager (BCM) [4] to manage consistency across the esLogin pool. An esLogin node is designated as the master image and then that image is synced to the other esLogin nodes. BCM is a new tool for Cray and so the primary challenge was getting familiar with a new tool and deciding which components needed to be the same across all esLogin nodes and which components needed to be unique to each esLogin node.

In addition to software consistency amongst the esLogin nodes, it is also important for the software versions to be the same on the esLogin nodes as on the internal service nodes. This is because while users compile applications on the esLogin nodes, jobs are launched from the internal service nodes and users rightly assume that software versions will be the same. Currently there is no supported way to verify that the esLogin node versions are compatible with the internal service node software. Software for the esLogin nodes and internal service nodes are packaged and distributed separately and are installed individually, and in addition must be verified by hand. While different software versions is less of a problem with statically built applications, now that the XT platform is supporting dynamic and shared library applications, software that is available on the esLogin nodes must also be available on the internal service nodes.

The esFS servers are not directly accessed by the users, but only used when accessing the filesystem – consequently the issues here are much simpler. However, maintaining software and configuration consistency across the pool of esFS nodes was initially problematic, until this was also brought under BCM. One of the key benefits of the esFS is to provide highly reliable and performing data storage, and to complement the benefits of the esLogin nodes. The combination of the two enables a consistent, highly reliable interface to be presented to users. In the event that the compute portion of the system is down for maintenance or failure, users can still access data, compile applications, and submit jobs.

## 5. Early Use

During the early use period, nearly two-thirds of all NERSC projects carried out substantial computation on the Phase 2 Hopper system. The science domain breakdown of this early usage is shown in Figure 5.1.

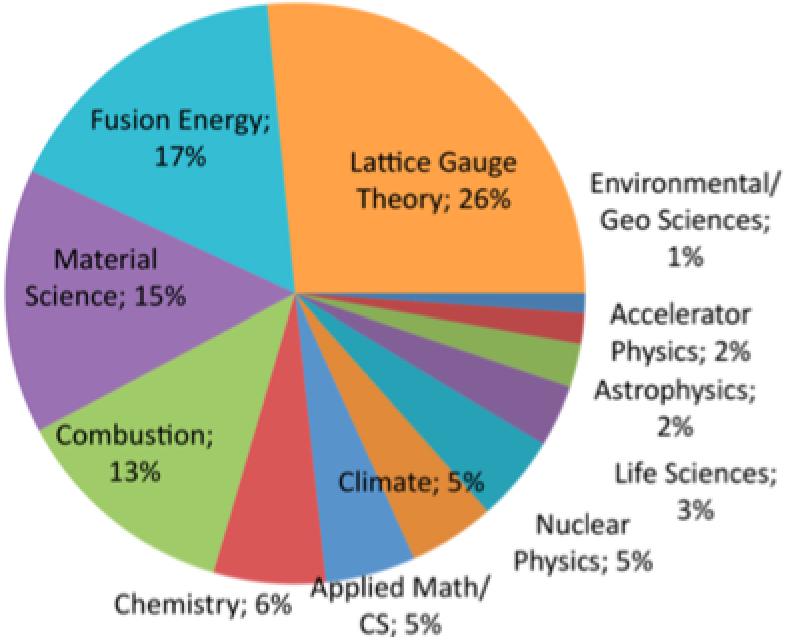

Figure 5.1: Breakdown of Early Use Time by Science Domain

Having such a broad set of science applications on the system early on allowed us to evaluate system stability, but also the usability of the system from the point of view of both the programming environment, and also how easily users could adapt to the hardware on the system. Compared with previous NERSC systems, Hopper has many more cores per node with less memory per core (1.3 GB per core) than on other systems. For a segment of the workload with both substantial memory requirements and a flat MPI programming model this could represent a problem. As part of the deployment, a joint NERSC/Cray Center of Excellence (COE) for Programming Models was formed to examine these issues for a set of representative applications. While the full scope of the COE activities is beyond the scope of this paper, and is discussed full in [5], the goals so far have been to convert MPI-only applications to a mixed MPI+OpenMP programming paradigm. This has in most cases considerably reduced the memory footprint per node, while leading to performance gains or the same performance as the original MPI-only application. These studies have been presented at user training events to show what can be achieved with some code re-engineering.

Despite any concerns we might have had, the concurrency of jobs run during the early access period has exceeded our expectations considerably. Almost half of all computing hours has gone to jobs running at 64K cores or higher, and with a second very substantial fraction going to jobs running with 16K to 64K cores. Users show little or no indication that there are any issues in scaling up their applications.

Turning to raw performance, after analysing the performance across the benchmark suite and comparing to the NERSC XT4, Franklin, we see better or similar performance across the entire suite – see Figure 5.2.

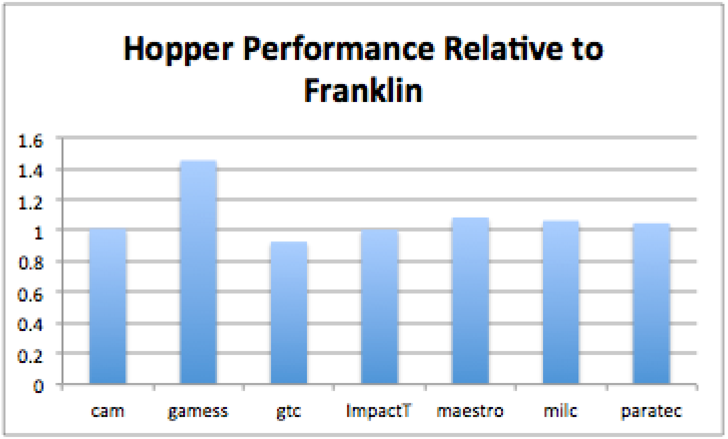

Figure 5.2: Hopper Performance on NERSC-6 Benchmarks compared to Franklin (higher indicates Hopper is faster)

Even though the clock speed is slightly slower on Hopper than on Franklin (2.1 GHz vs. 2.3 GHz), the cache size is larger on the AMD Magny-Cours Opteron as opposed to the older Budapest Opteron, and the Gemini interconnect provides lower latency and better bandwidth. Even with the increased number of cores, the memory bandwidth is roughly similar to the quadcore nodes in Franklin.

## 6. Summary

Judging from the early successes across all science domains, and the speed at which users have started to utilize Hopper, we are optimistic that the system will become a major resource to deliver scientific simulation results at all scales for the DOE Office of Science. The XE6 is showing itself to be a very stable and performing computer platform, with many of the issues concerning external services that we encountered in our initial deployment having been resolved.

## References

[1] "DVS, GPFS and External Lustre at NERSC - How It's Working on Hopper", T. Butler, Cray User Group Meeting, Fairbanks, AK, May 2011.
[2] "External Services on the NERSC Cray XT5 System", K. Antypas, T. Butler, J. Carter, Cray User Group Meeting, Edinburgh, UK, May 2010.
[3] "Cray Application Developer's Environment Supplement Installation Guide, S-2485-17, Cray Inc.
[4] Bright Cluster Manager - Advanced Linux Cluster Management Software http://www.brightcomputing.com/Bright-Cluster-Manager.php
[5] "The NERSC-Cray Center of Excellence: Performance Optimization for the Multicore Era", N. Wright, H. Shan, A. Canning, L.A. Drummond, F. Blagojevic, J. Shalf, K. Yelick, S. Ethier, K. Fuerlinger, M. Wagner, N. Wichmann, S. Anderson, and M. Aamodt, Cray User Group Meeting, Fairbanks, AK, May 2011.

## Acknowledgments

This work was supported by the Director, Office of Science, Office of Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The authors would like to thank: Manuel Vigil and the ACES team for valuable discussions and test time at the factory; Nick Wright and

Helen He for providing early COE results; and the Cray on-site and Cray Custom Engineering staff for valuable discussions.

## About the Authors

Katie Antypas is the Group Leader of the User Services Group at the NERSC Center, and was co-lead of the Hopper Implementation team. Tina Butler is a systems analyst in the Computational Systems Group at NERSC, and was co-lead of the Hopper Implementation team. Jonathan Carter is Deputy Director of Computing Sciences at Lawrence Berkeley National Laboratory and was the Hopper Project Manager.