EDUCAUSE Center for Applied Research

Research Bulletin

Volume 2008, Issue 5

March 4, 2008

Developing and Extending a Cyberinfrastructure Model

Rosio Alvarez, Lawrence Berkeley National Laboratory

4772 Walnut Street, Suite 206 ◆ Boulder, Colorado 80301 ◆ www.educause.edu/ecar/

Overview

Increasingly, research and education institutions are realizing the strategic value and challenge of deploying and supporting institutional cyberinfrastructure (CI). CI is composed of high-performance computing systems, massive storage systems, visualization systems, and advanced networks to interconnect the components within and across institutions and research communities. CI also includes the professionals with expertise in scientific application and algorithm development and parallel systems operation. Unlike “regular” IT infrastructure, the manner in which the components are configured and the skills required to perform these configurations are highly specific and specialized. Planning and coordinating these assets is a fundamental step toward enhancing an institution’s research competitiveness as well as the return on personnel, technology, and facilities investments.

Coordinated deployment of CI assets has implications across the institution. Consider the vice chancellor for research whose new faculty in the life sciences are now asking for simulation systems rather than wet labs, or the provost who lost another faculty candidate to a peer institution that offered computational support for research, or the vice chancellor for administration who has seen a spike in power and cooling demands from many of the labs and office spaces that are being converted to house computing systems and instrumentation. These are just some of the issues that research institutions are wrestling with as research becomes increasingly computational, data-intensive, and interdisciplinary. This bulletin discusses these issues and presents an approach for developing a CI model that was successfully developed at the Lawrence Berkeley National Laboratory (LBNL) and extended to the University of California, Berkeley (UCB).

Highlights of a Cyberinfrastructure Model

Data-intensive research, interdisciplinary research, and interinstitutional research are quickly becoming the standards in most scientific disciplines. Computation is central, if not fundamental, to this research. As many have suggested, computation has recently become the third pillar of science, joining both theory and experiment. At a recent talk, the undersecretary of science for the Department of Energy went further to argue that in the very near future, computation, or more specifically simulation, will to some degree replace theorization, where first principles of theory do not yet exist, and experimentation, where direct experiment is not possible, affordable, or desirable.[1] If we think about how discovery happened in the past, discoveries accrued to those who had access to unique instruments and their data. But the growing costs and complexity of tools prohibit the proliferation of instruments. Instead, we have a very few unique instruments producing data (for example, the Large Hadron Collider on the outskirts of Geneva), and that data is housed in a few locations across the globe. The result is that discovery is now based on asking the right questions rather than on having access to unique tools. We are moving from hypothesis-driven science—“I have a question, so I will collect or find data”—to exploration-driven science—“I have lots of data, so what can I glean from it?” Research in many disciplines has become much more about access, analysis, movement, and management of very large amounts of data. Indeed, the availability of data and computational resources has created the rise of new areas of study, such as synthetic biology, genomics, and bioinformatics. They have prompted the National Science Foundation (NSF) to issue a report stating that "[d]igital computation, data, information, and networks are now being used to replace and extend traditional efforts in science and engineering research, indeed to create new disciplines."[2] But computation has also become important in fields not traditionally considered "hard" science, such as the social sciences and the arts and humanities.

Researchers Respond

Researchers have responded to the need for computation in much the same way they did in the late 1970s and early 1980s when they needed network access but the wiring was slow in coming to buildings and research labs. The common solution then involved science departments running cables, setting up electronics, and providing networking services to individuals for whom network access was a fundamental requirement for doing research. Similarly, some 25 years later, many researchers have deployed their own clusters of high-performance computers, housing them outside of the central institutional data center. A cluster is a single system composed of interconnected computers that communicate with one another. There is usually a master node and many (even hundreds) of compute nodes. While CI encompasses much more than clusters, over the past few years clusters have fundamentally transformed research-based computing for a few compelling reasons.

Clusters have very important socioeconomic effects. They are built from relatively inexpensive commodity PCs, and they typically use open source Linux operating systems and tools available in the public domain. Individuals and laboratories at universities—and even in high schools—believe (and have shown) that they can assemble and incrementally grow a small to midsize supercomputer. In many instances, clusters have provided the computational staging ground needed by researchers who have outgrown the power of their desktop computers but are not quite ready to compete nationally for allocations on larger supercomputers such as those available at the National Center for Supercomputing Applications, the San Diego Supercomputer Center, or the National Energy Research Scientific Computing Center. In the mid-range space, researchers are able to use clusters for staging production runs, parallelizing and optimizing code, and doing other work that prepares them for competing nationally for supercomputing allocations.

Cluster Economics

Clusters are advantageous primarily as a result of the economics. Not only are the hardware and software reasonably inexpensive, but these distributed clusters ride "free" on their lab, department, or research group's organizational overhead, which includes space, networking, and personnel. It is no surprise, then, that research and education organizations have pockets of computing strewn throughout their research areas. Much like the scenario of the late 20th century, when IT divisions began piecing together the hodgepodge of do-it-yourself networks around campus, the landscape of high-performance computing presents a similar challenge to research organizations. As perhaps the only comprehensive study on research computing shows, many campuses find that most research computing is highly decentralized.[3] While these distributed labs are able to tailor the configuration of systems to the needs of the specific science discipline and their discovery methods, there is much effort expended on the part of researchers that is hidden or implicit. Further, these resources are usually confined to a specific department and are therefore unavailable to support other areas of science within the institution.

There are hidden costs associated with assigning researchers or graduate students to spend considerable time on system administration rather than on scientific inquiry. The cost of lost research productivity of these individuals is not factored into the total costs of owning a cluster system. Additionally, researchers are not system administrators, so the time expended taking care of cluster systems can be considerably higher than it would be if the systems were administered by professionals. These costs become highly visible when there are security breaches to cluster systems because of inadequate protections. There are also hidden costs associated with losing office or teaching space to computers that are placed in areas that are not designed for them and therefore do not maximize the use of space. Most noticeably, the exponential increase in power and cooling needed for these cluster systems is a cost that, while previously hidden to some degree, is now becoming much more explicit.

Hacker and Wheeler provide a cogent example of the hidden electrical and cooling costs incurred by distributed cluster systems.[4] They suggest that a 1-teraflop (TF) system can consume about $52,416 of electrical energy in a year. In the short timeframe since the publication of their article in 2007, computing power per compute node of each cluster has increased substantially. Evidence from clusters going into production now shows that a 1-TF system consisting of fewer nodes with more processing cores will consume approximately one-tenth of that amount of energy. The demand for computation power, however, for the most part is constrained by the dollars available for purchasing systems. Therefore, 2- and 3-TF systems are not at all uncommon. One can easily find 50 and even perhaps as many as 100 of these in various research labs at a mid- to large-sized research university. Anecdotally, several large public research universities have reported these numbers. Based on these power-consumption calculations, those institutions can project that these clusters add anywhere from $750,000 to $1,500,000 to their utility bills annually. If you add to that the cost of inefficiently cooling these systems due to aging or poorly designed computer room facilities (oftentimes using fans or window air conditioners purchased at local retail stores), one can easily double the cost. In the absence of systematic data on the proliferation of clusters, these are somewhat rough estimates, but the numbers are large enough to no longer be ignored.

The projections are embellished by the anecdotes of computer disasters. The stories of overheated computers smoking and burning abound. At LBNL, a cluster was housed in a substandard space that had marginal cooling and insufficient monitoring of the environment. Because there were no proper alert systems in place, during a power outage the computers sat in 114-degree temperatures overnight. When discovered, one storage unit was permanently damaged and had to be sent to a firm specializing in last-resort data recovery. The cost of research or experimental data loss can be invaluable, not to mention the time the IT and facilities staff spend on responding to these emergencies. And for the most part, ad hoc remedies are patched together in response to an emergency, and attention is not paid to building a more coherent cost- and energy-efficient plan.

The implicit and explicit costs associated with the deployment of distributed high-performance clusters is perhaps one of the more compelling arguments to be made in support of a more planned and rational model. The following section describes how LBNL developed such a model and then extended it in support of scientific research at the University of California, Berkeley.

Scientific Cluster Support: Phase I

Computing has been part of scientific research for the past 50 years. At LBNL, scientific research computing had evolved from centralized supercomputers and timesharing systems to powerful desktop computing in the mid-1990s. Many scientists' computational needs exceeded the power offered by desktop devices, however, and they were finding themselves at a competitive disadvantage compared to their peers at other institutions. Moreover, many were interested in allocations at the national supercomputing user facility housed at LBNL, but only a select few were chosen to obtain them, leaving a gap between high-end supercomputing and low-end desktop computing. This mid-range computing gap was identified by LBNL, and as a consequence, a mid-range working group of scientists from a variety of fields was formed in early 2000 to identify how to address the gap. After several months of work, the working group put forward a proposal that opposed the idea of purchasing an institutional computing resource and instead supported the idea of a central cluster support program. That is, scientists would obtain funding for their clusters, but the services to support the clusters would be made available by the information technology (IT) division, which had developed a high degree of expertise in supporting researchers over the years. The relationships that the IT division staff had developed with the scientists; the high degree of interest, need, and support within the research community; and the commitment of key management were factors in getting a centrally managed cluster support program started.

Specifically, the cluster support program staff would provide the following:

- **Prepurchase consulting:** Based on scientific problems to be solved, determine the hardware, interconnect, operating system, compilers, and application software.
- **Procurement assistance:** Develop a budget and specification for RFP and acceptance criteria, and evaluate bids.
- **Cluster integration:** Install and configure hardware, operating systems, cluster software, applications, and computer security.
- **Ongoing systems administration and cybersecurity:** Maintain system and upgrades, monitor clusters, troubleshoot hardware, maintain cluster software stack, manage and support resource and scheduler, and maintain user accounts.

- **Data center space, networking, power and cooling:** Host clusters in the data center to ensure adequate cooling, power, and networking.

The Scientific Cluster Support (SCS) program, as it was called, was allocated $1.3 million from central funds for the first four years. This would fund the services listed above as well as two full-time employees. At the end of the four-year period, SCS was to have developed a full recharge business model. Ten research groups with funds to purchase a cluster were selected for support in the first four years.

The SCS program developed a cost-effective methodology with hardware and software standards that facilitated the scaling of systems administration support. They also used open source software, such as Linux. Because the 10 pilot projects came from a variety of scientific fields, the computational needs of some varied from others, so requests for exceptions to the standards were not uncommon. In these instances, a small steering committee composed of stakeholders and outside technical expertise proved invaluable. They helped enforce the standards with their peers by vetoing requests for these exceptions, understanding that if too many were allowed, the costs of the program would increase for the scientists who would pay after the pilot period.

For the key area of cluster management, the IT division was not able, at that time, to find a toolkit that would allow for a scalable method of supporting the clusters. Therefore, IT developed a cluster management toolkit called Warewulf that greatly simplified installation and management of clusters. Simply put, Warewulf allows compute nodes to boot from a shared image on the master node of each cluster so that a system administrator need only support the master node. In effect, the compute nodes are “stateless.” In cases where clusters can have 100 or more compute nodes but only one master, this was an immensely valuable labor-saving solution.

Toward the end of the four-year pilot, several business models were developed and vetted with various advisory boards and members of the scientific community at LBNL. The model that was found to be acceptable to both the administration and to researchers was a partially funded SCS program. An allocation of $350,000 was given to the program (with promised cost-of-living increases), and all power and cooling costs would also be subsidized centrally. This dramatically dropped the cost for researchers. The program continued to grow rapidly. In an 18-month period, the number of clusters supported increased by 58%. By fall 2007 the SCS program was supporting approximately 30 clusters representing approximately 3,000 processors.

Scientific Cluster Support: Phase II

As successful as the SCS program was, it had its challenges. Intra-cluster management was efficient, but inter-cluster management was not. That is, each cluster was based on standards and was a system unto itself. Even the slightest variation required the unique configuration of each master node. Also, many researchers who outgrew an initial cluster would purchase another, but inter-cluster sharing of computation or other resources was not possible for them. As a consequence, some of the clusters were running at 50–60% utilization. A member of the IT division developed a new cluster management toolkit called Perceus that enabled much larger scalability by creating a "meta-cluster." In a sense, this flattened the cluster management.

The Perceus toolkit builds on the existing management systems' approach, but it goes further in that master nodes are also stateless. Most of the management is moved to a central Perceus appliance that contains standard software images from which all the clusters boot. In effect, the clusters in this model were aggregated in such a way that they appear as one meta-cluster requiring dramatically less time to manage. Rather than managing some 30-plus unique systems, the system administrators now manage one meta-system consisting of groups of nodes. Because the clusters were integrated in this way, they also allow researchers with two or more clusters to now move seamlessly across them when executing code or moving data. This can result in higher utilization of cluster resources overall. Additionally, spare cycles on any cluster can be harvested and allocated to other researchers. In a sense, this new toolkit functions much like a local grid system that allows researchers to use computational resources across the grid.

An additional component of the second phase of this CI model includes the implementation of a shared institutional cluster—even though four years earlier the idea had been considered too ambitious and was rejected by the working group. The IT division surveyed scientists and discovered that 38% of them depend on clusters for their research. Of this group, 70% said that they would be interested in purchasing cycles from an institutional cluster. A market analysis of commercial offerings, when normalized to a standard level of performance, showed that the cost to run an average 400 CPU-hour job ranged from $148 to $400. The IT division once again assembled a science team to help assess the viability of various business models for an institutional cluster. They analyzed the break point of buying cycles versus buying a cluster. After several financial iterations, they settled on $.10 per CPU-hour. At this price, the comparable 400 CPU-hour job would run $40. This price was a compromise between the more expensive commercial offerings and the free if-you-can-get-it allocations from the national supercomputing facilities. At this price point, it was shown to be more cost-effective to buy time on the institutional cluster unless one was prepared to purchase a 50-node cluster or larger. So for researchers with needs under 50 nodes (the average size of a cluster at that time was about 40), it was better to buy time. After putting forth a strong business case, the IT division received $1 million to purchase the institutional cluster. LBNL also provides competitive research funding that is designed to purchase time on the cluster for researchers who are awarded funds and require computational support. There is also discussion of purchasing time on the cluster to award to newly recruited scientists who have needs for computation.

Not unlike many other major research universities, UCB was facing an increase in the proliferation of clusters and demand for cluster support services. With a long tradition of collaboration (indeed LBNL was born at UCB as an organized research unit before becoming a Department of Energy–funded laboratory), the two institutions partnered to provide cluster services to UCB faculty. The partnership was supported both in spirit and financially by the vice chancellor for research and the provost at UCB. Both of these had witnessed the proliferation of requests to NSF and other funding agencies with budgets for clusters that later required infrastructure support that was a drain on their budgets and usually not the most cost-effective method for supporting the equipment. Like several of the other UC campuses and many other research universities across the country, UCB did not provide a formal offering of high-performance computing cluster services for researchers who may have preferred institutionally maintained clusters. Drawing on the strengths of each institution, UCB's IT division will house the clusters in its new data center, and the LBNL IT division will provide support services similar to those offered at the lab. Perhaps the most challenging aspect of extending the CI model to UCB was developing a financial model that would adhere to UC financial standards as well as DOE standards and oversight. As the two institutions wrestled with their disparate financial systems and overhead rates, researchers at UCB lined up for cluster services. When high-performance cluster services at UCB was kicked off in the fall of 2007, three research projects were already in the queue with another awaiting notice from a grant proposal.

What It Means to Higher Education

There are many aspects of CI, as mentioned earlier, but certainly high-performance computational clusters are one key component that is transforming research and impinging upon the physical infrastructure of universities in doing so. To some degree, clusters are the "last mile" in CI. That is, much like broadband technology (cable or DSL modems) have been the last mile into the home, cluster technology is the last mile into the research lab. As research and education institutions become increasingly aware of this, there are some important trends that should be of concern to higher education.

There is a dearth of comprehensive aggregated data on CI in higher education that can adequately paint a picture of the state of CI. The exception is ECAR's seminal study, *IT Engagement in Research: A Baseline Study*. Some interesting findings from this study show us that for the institutions surveyed, roughly two out of three project an increase in high-performance computing and high-performance networking, whereas only one out of two said they had seen an increase over the previous three years.[5] Clearly, CIOs and the like across campuses are expecting growth like they have not had in the past and should be preparing their organizations to meet this projected growth. Yet approximately half of those surveyed currently have less than one FTE assigned to provide research computing support, and 60% project that staffing will remain the same in that area.[6]

Another area of concern to higher education is funding of CI. Slightly more than half of the institutions surveyed in the ECAR study spend less than $100,000 per year on support for research computing. Not surprisingly, only 35% believe that they have a sustainable budget to support research IT, and 57% believe that the biggest barrier to funding is lack of institutional commitment.[7] Other institutional priorities take precedence over CI in the eyes of university budget decision makers. Yet, underfunding CI has very costly consequences, as mentioned above, in terms of utilities and loss of research productivity, for existing faculty as well as the lost recruits who may have gone to competing university or research organizations where CI support is an institutional priority.

There is no one-size-fits-all model for supporting CI on any campus. A strategy and program must be contextual and participatory, with the goal of making a direct impact on research. Institutions must ask some key questions and collaboratively find answers in order to begin to assess CI interests and develop a strategy to address those interests. With this type of information in hand, university administrators can begin to engage researchers who are interested in computation in developing a plan that leverages the local (discipline-specific) needs with central core IT services in ways that make financial and technical sense for the institution.

Key Questions to Ask

- What is the magnitude of the current local situation—how many clusters are run in local centers, and what other type of CI services are offered?
- How many research grants are going out with requests for computational funds or computing allocations, and how will these impact the potential for immediate growth of CI on campus?
- How much and what type of CI support does central IT offer? What are the potential gaps?
- Does IT have formal or informal relationships with researchers that can be leveraged to develop and implement a CI plan?
- What is the commitment of senior management and key decision makers to support CI?

Where to Learn More

- Blustain, Harvey, Sandra Braman, Richard N. Katz, and Gail Salaway. *IT Engagement in Research: A Baseline Study* (Research Study, Vol. 5). Boulder, CO: EDUCAUSE Center for Applied Research, 2006, available from http://ww.educause.edu/ecar.
- Klingenstein, Ken, Kevin Morooney, and Steve Olshansky. *Final Report: A Workshop on Effective Approaches to Campus Research Computing Cyberinfrastructure*. Arlington, VA, April 25–27, 2006. http://middleware.internet2.edu/crcc/docs/internet2-crcc-report-200607.html.
- Edwards, Paul N., Steven J. Jackson, Geoffrey C. Bowker, and Cory P. Knobel. “Understanding Infrastructure: Dynamics, Tensions, and Design.” http://hdl.handle.net/2027.42/49353.
- Bottom, James R., Patrick Dreher, and Bonita M. Neas. “The Challenge of Campus Cyberinfrastructure.” (presentation at the EDUCAUSE Annual Conference, Dallas, Texas, October 2006). http://www.educause.edu/Library/DetailPage/666?ID=EDU06063.

- Bell, Gordon, and Jim Gray. "High Performance Computing: Crays, Clusters and Centers. What's Next?" *Communications of the ACM*, January 2002.
- Scientific Cluster Support, at LBNL website, http://scs.lbl.gov.
- Perceus website, http://www.perceus.org.

Endnotes

1. "The American Competitiveness Initiative: Role of High End Computation," a talk given by Raymond L. Orbach, DOE Undersecretary of Science at SC07 International Conference for High Performance Computing, Networking, Storage and Analysis, Reno, NV, November 14, 2007.
2. Daniel E. Atkins et al., "Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure," January 2003, 4. http://nsf.gov/od/oci/reports/CH1.pdf.
3. Harvey Blustain, Sandra Braman, Richard N. Katz, and Gail Salaway, *IT Engagement in Research: A Baseline Study* (Research Study, Vol. 5) (Boulder, CO: EDUCAUSE Center for Applied Research, 2006), 15, available from http://www.educause.edu/ecar.
4. Thomas Hacker and Bradley Wheeler, "Making Research Cyberinfrastructure A Strategic Choice," *EDUCAUSE Quarterly* 30, no. 1 (2007): 21–29, http://connect.educause.edu/library/abstract/MakingResearchCyberi/40013.
5. Blustain, *IT Engagement in Research*, 52.
6. Ibid., 68.
7. Ibid., 76.

About the Author

Rosio Alvarez (ralvarez@lbl.gov) is Chief Information Officer at the Lawrence Berkeley National Laboratory.

Copyright

Copyright 2008 EDUCAUSE and Rosio Alvarez. All rights reserved. This ECAR research bulletin is proprietary and intended for use only by subscribers. Reproduction, or distribution of ECAR research bulletins to those not formally affiliated with the subscribing organization, is strictly prohibited unless prior permission is granted by EDUCAUSE and the author.

Citation for This Work

Alvarez, Rosio. "Developing and Extending a Cyberinfrastructure Model" (Research Bulletin, Issue 5). Boulder, CO: EDUCAUSE Center for Applied Research, 2008, available from http://www.educause.edu/ecar.