By Michael P. Cummings and Jeffrey C. Huskamp

Grid Computing

In the face of stagnant or decreasing budgets, unfunded mandates, and growing demands for services, colleges and universities are coming under increased pressure to deliver additional computational power for research and teaching in order to better position their institutions to attract the most productive faculty and the "best and brightest" students. As evidence of this trend, the 2005 EDUCAUSE survey of the top-ten issues confronting information technology managers today lists funding as the most pressing IT issue.[1] From a demand perspective, a June 2005 study by the President's Information Technology Advisory Committee (PITAC) points out that the lack of coordinated support for computational science, as well as insufficient resources for computational science research, is endangering U.S. competitiveness.[2]

The common institutional approach—purchasing more computers to satisfy increased demand—has yielded mixed results. Although computers represent the largest class of equipment purchased in higher education, most computers remain underutilized, processing only approximately 5 percent of the time. Furthermore, these existing computers already have associated investments in place to handle the space, electricity, environmental control, and systems support costs, which represent the major costs of ownership. Being able to leverage this existing investment would demonstrate responsible stewardship of scarce funding resources.

To do so, a growing number of institutions are turning to computational grids and grid computing. The word *grid* in *grid computing* comes from an analogy to the electrical power grid. With electrical power, there is a very simple interface (i.e., a wall socket) through which one connects by simply inserting a plug. Knowledge regarding the ultimate source of the electricity or the specific transmission routes from source to socket is unnecessary for proper use of the electricity.

Following this analogy, we provide the following definition:

> *Grid computing* is a model of distributed computing that uses geographically and administratively disparate resources. In grid computing, individual users can access computers and data transparently, without having to consider location, operating system, account administration, and other details. In grid computing, the details are abstracted, and the resources are virtualized.

Like other technologies, grid computing has sometimes been oversold, leading to disappointments and frustrations. Our approach to grid computing is driven by pragmatism: we see grid computing as the most practical means to satisfy computing needs and to make the best use of existing resources. From our perspective, the promise of grid computing is comprehensive but not over-reaching: grid computing can meet the needs of many users and is achievable with current technologies.

A Perfect Match

For several reasons, higher education is a perfect match for grid computing. Colleges and universities have the means, the motive, and the opportunity to realize the promise.

- *Means:* Higher education institutions have large concentrations of computing resources, ranging from desktops to high-end clusters, high-speed networking, access to regional and national-level networks (e.g., Internet2), and quite often, expertise with parallel and distributed computing.
- *Motive:* Research and education activities result in many and diverse needs for computing resources that often exceed those available to individuals in the community. At the same time, there are limitations for centralized computing resources including funding, staffing, and space.
- *Opportunity:* Higher education institutions are in many ways large communities, each with a history of sharing and an *esprit de corps*.

The University of Maryland

The University of Maryland is a good example of how means, motive, and opportunity meet for grid computing.

- *Means:* Of the approximately 40,000 individual computers on campus, roughly 28,000 are in clusters, research laboratories, or classrooms or on the desktops of faculty and staff; the remaining 12,000 or so are in residences. These computers span a number of hardware configurations and capabilities. The campus has a mix of copper, fiber, and wireless networking. As one of the founding members of the Mid-Atlantic Crossroads, Maryland is well connected to regional, national, and international networks.
- *Motive:* The need for computing continually increases in both intensity and breadth throughout academic institutions. More research and education is adopting computer-intensive methods—such as stochastic simulation, Bayesian statistics, Markov-chain Monte Carlo sampling—to estimate an increasing number of parameters in ever larger and more complex models and data sets. The need for computing has already spread well beyond the traditionally computing-intensive areas mentioned previously into chemistry

(e.g., structure and interaction modeling), life sciences (e.g., biophysics, conservation biology, landscape ecology, neurobiology, population genetics, systematics), social sciences (e.g., combinatorial optimization problems, modeling), and many other areas.

- *Opportunity:* Numerous opportunities for better sharing of computing resources exist among the institutional groups traditionally associated with both computing resources and computing needs (e.g., computer science, physical sciences, engineering), as well as among those groups not traditionally associated with high-throughput computing needs.

A campus example of where these factors meet is the Digital Studio of the Landscape Architecture Program. The Digital Studio, composed of four individual design studios and a geographic information system/computer-aided design and drafting laboratory, provides a professional-level training environment consisting of approximately one hundred workstations. Like many computing resources, these workstations are used very intensively some of the time but very lightly or not at all most of the time. The opportunity for other campus groups to utilize these resources via grid computing came about through the cooperative interactions of a resourceful faculty member (Maile Neel), a receptive director (David Myers), and a responsive systems administrator (Lauren Whitaker). Within a few months of being integrated into a research grid system, the Digital Studio had performed several CPU years of analyses in areas such as the human population structure and computational chemistry, all without conflicting with the use by landscape architecture students.

The University of Maryland is now working toward utilizing campus grid expertise from the research community and creating a production resource that can be used routinely by students, faculty, and staff. The challenges will lie in training support staff, ensuring availability, integrating into the campus security model, and authenticating and training users. The benefit will be a potentially larger grid system that could utilize a significant fraction of the campus machines.

Making It Happen: Middleware

Middleware, software that provides for interaction between other software applications, is what makes grid computing work. The primary grid computing middleware most appropriate for academic environments (i.e., open source) includes the Globus Toolkit, Condor, and BOINC (Berkeley Open Infrastructure for Network Computing, <http://boinc.berkeley.edu/>). Some grid middleware is distributed computing middleware: the difference is the nature of the resources united and how these resources are virtualized. Grid middleware components fulfill necessary functions such as authentication, authorization, resource matchmaking, data transfer, and monitoring.

Current grid middleware had its beginning in the early efforts of the Condor Project (http://www.cs.wisc.edu/condor/) and the Globus Alliance (http://www.globus.org/), which develop and distribute open source grid middleware solutions. These groups were among the first to develop more general-purpose distributed computing middleware for high-throughput applications. Another significant middleware effort, the National Science Foundation (NSF) Middleware Initiative, or NMI (http://www.nsf-middleware.org), was started in 2001. This initiative funds the design, development, testing, and deployment of middleware for creating production middleware by utilizing open source and employing open standards. Over time, organizations like the Global Grid Forum (http://www.ggf.org/), which promotes the creation of standards and builds a sense of community, and the IEEE/ACM International Conference on Grid Computing series (http://www.gridcomputing.org) have been advocates for the wider adoption of grid technologies.

Security

Everyone involved in grid computing has concerns regarding security, although the specifics of the concerns may differ. Contributors of resources are concerned that the privacy and integrity of their systems be ensured; users of the systems are concerned that the integrity of their applications, data, and results be maintained; systems administrators are concerned that the resources be made available only to approved users. Numerous mechanisms address these concerns, including creating a "sandbox" in which the grid processes run. Jobs in the sandbox are limited in the interactions they can have with the client system and can read and write only to files within the sandbox. Digital signatures, checksums, and encryption can be used where needed. The Globus Toolkit provides specific tools for dealing with authentication, authorization, and delegation of credentials. These and other mechanisms, in addition to present system security approaches, provide robust security for grid computing.

Summary

Many colleges and universities have one or more distributed computing systems or prototypical grids that are used for research computing or may themselves be research projects. The challenge for these institutions is to migrate the existing research efforts to a production environment that is supported, maintained, and available as a turnkey resource. Although this implies a significant investment in training, education, help-desk support, and documentation for the grid, the result is that a "hidden" high-throughput resource can help solve the problem of increasing computer needs and stagnant or decreasing budgets.

Notes

1. Leslie Maltz, Peter B. DeBlois, and the EDUCAUSE Current Issues Committee, "Top-Ten IT Issues, 2005," *EDUCAUSE Review*, vol. 40, no. 3 (May/June 2005): 14–28, <http://www.educause.edu/er/ erm05/erm0530.asp>.
2. President's Information Technology Advisory Committee, *Computational Science: Ensuring America's Competitiveness* (Arlington, Va.: National Coordination Office for Information Technology Research and Development, 2005), <http://www.nitrd.gov/pitac/reports/20050609_computational/computational.pdf>.

Michael P. Cummings, Associate Professor in the Center for Bioinformatics and Computational Biology at the University of Maryland, specializes in the areas of molecular evolutionary genetics and grid computing. Jeffrey C. Huskamp, Vice President and CIO at the University of Maryland, College Park, has over thirty years of experience in high-performance computing and networking with national laboratories, universities, and the Department of Defense.