

A Collaborative IT Support Model for Research at Georgetown University

Judith A. Pirani, ECAR
Donald Z. Spicer, ECAR

ECAR Case Study 3, 2006

Case Study from the
EDUCAUSE Center for Applied Research



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

A Collaborative IT Support Model for Research at Georgetown University



EDUCAUSE is a nonprofit association whose mission is to advance higher education by promoting the intelligent use of information technology.

The mission of the EDUCAUSE Center for Applied Research is to foster better decision making by conducting and disseminating research and analysis about the role and implications of information technology in higher education. ECAR will systematically address many of the challenges brought more sharply into focus by information technologies.

Copyright 2006 EDUCAUSE. All rights reserved. This ECAR case study is proprietary and intended for use only by subscribers and those who have purchased this study. Reproduction, or distribution of ECAR case studies to those not formally affiliated with the subscribing organization, is strictly prohibited unless prior written permission is granted by EDUCAUSE. Requests for permission to reprint or distribute should be sent to ecar@educause.edu.

A Collaborative IT Support Model for Research at Georgetown University

Preface

The EDUCAUSE Center for Applied Research (ECAR) produces research to promote effective decisions regarding the selection, development, deployment, management, socialization, and use of information technologies in higher education. ECAR research includes

- ◆ research bulletins—short summary analyses of key information technology (IT) issues;
- ◆ research studies—in-depth applied research on complex and consequential technologies and practices;
- ◆ case studies—institution-specific reports designed to exemplify important themes, trends, and experiences in the management of IT investments and activities; and
- ◆ roadmaps—designed to help senior executives quickly grasp the core of important technology issues.

The 2006 ECAR baseline study on IT engagement in research used a multifaceted research methodology to gather both quantitative and qualitative data from 328 higher education institutions (315 U.S. and 13 Canadian institutions). The data provide a view of one self-selected segment of higher education's collective experience with the research enterprise as well as in-depth, institution-specific perspectives.

Literature Review

The study began with a review of the relevant literature on IT engagement with academic research to define the study's major themes and create a working set of hypotheses to be tested.

Survey

The quantitative Web-based survey was designed by ECAR fellows and was sent to 1,477 EDUCAUSE member institutions in Canada and the United States. Senior IT leaders, most of them CIOs, from 328 institutions responded to the survey. The survey questions appear on the EDUCAUSE Web site at <http://www.educause.edu/ir/library/pdf/ECAR_solvers/ESI05F.pdf>.

Interviews

We conducted qualitative telephone interviews with IT executives and managers at EDUCAUSE member institutions. To obtain depth and breadth of practice, we chose to interview respondents from institutions of varying size and mission, and we included both public and private institutions. The interviews were invaluable in helping us to understand anticipated research issues and directions over the next five years.

Case Studies

Researchers conducted this in-depth case study to complement the core study. We assume readers of this case study will also read the primary study, which provides a general context for the individual case study findings. We undertook this case study to examine Georgetown University's Advanced Research Computing (ARC) division, which developed computational resources slowly and methodically as clients' needs warranted. When ARC achieved a critical mass of clients, it leveraged its contacts and project experiences to create new research support resources that now transcend Georgetown's borders.

ECAR owes a debt of gratitude to James Bodurtha, associate professor, McDonough School of Business; David Cafaro, systems analyst; Jess Cannata, manager/systems administrator; Woonki Chung, systems analyst; Robert Clarke, professor, Lombardi Comprehensive Cancer Center; Colin Freas, programmer analyst; Ardoth Hassler, associate vice president; Paul Kennedy, programmer analyst; Chad La Joie, team leader; H. David Lambert, vice president for information services and chief information officer; Arnie Miles, senior systems administrator/architect; Stephen P. Moore, program director; Nick Marcou, systems analyst; Peter McGarvey, research associate professor of biochemistry and molecular biology and cellular biology; Françoise Seillier-Moiseiwitsch, associate professor and biostatistics department chair; Baris Suzek, senior bioinformatics scientist; Jeffrey Urbach, associate professor and physics department chair; Cathy Wu, professor of biochemistry and molecular biology, and director, Protein Information Resource (PIR); and Dave Zapple, senior systems analyst.

Introduction

Higher education institutions face a changing research environment in the 21st century. In the past, many research initiatives

operated more or less autonomously, with more emphasis on individual pursuits and less on collaboration. But a competitive funding environment, combined with the increasingly interdisciplinary and multi-institutional nature of research, is changing this view. "The idea that one works in an ivory tower is an old-fashioned way of doing research," explains Robert Clarke, professor, Lombardi Comprehensive Cancer Center. "You have to share your knowledge with colleagues, to collaborate, and to leverage your resources in order to work in a more powerful way. Often those colleagues are not down the hallway; they may not even be on your campus. We have to find ways to complete our bread-and-butter work in a more collaborative way."

One consequence is that IT support of research can no longer operate in a silo, either. "The power of a team approach cannot be overstated," believes Arnie Miles, senior systems administrator/architect. "A system administrator working in isolation within a research group is doomed to failure more often than not when trying to get a handle on rapidly changing technology, keep up to date with emerging trends, and maintain the local system. This isolation leads to increased attrition. Then when the FTE leaves, the research group suddenly finds itself bereft of IT support."

H. David Lambert recognized Georgetown's need to adapt accordingly when he became the university's first vice president for information services and CIO. Dissatisfaction described the state of IT research support when he arrived in 1998. In response, the university launched the Advanced Research Computing (ARC) division in 2000, an autonomous organization that operates under the auspices of the IT organization and focuses solely on IT support for research.

From the beginning, it was obvious "that we did not have the deep pockets to support an 'if we build it, they will come' model,"

states Miles. Instead, ARC concentrated on building a responsive, service-oriented organization. Then it added computational resources slowly and methodically as the number and needs of its clients warranted. “The cornerstone was to reach the researchers, to find out their needs and to serve them,” explains Stephen P. Moore, ARC’s program director. “Then we needed a framework and architecture, first adding Beowulf clusters, then a computational research facility, and finally grid computing support.”

Now ARC is leveraging its client base and successes to provide a multilayer and multi-institutional range of computational resources. Having received a National Cancer Institute (NCI) grant to participate in the cancer Biomedical Informatics Grid (caBIG) and launched its GridsWatch portal (<http://www.gridswatch.com>), ARC’s activities now extend well beyond Georgetown’s borders. Yet it still manages to remain true to its initial goal of supporting researchers’ needs. Its success stems from the following principles:

- ◆ Proactive networking within Georgetown’s research community achieves a critical mass of clients, creating synergy and momentum from which to develop new computational resources and grant opportunities.
- ◆ A technically diverse team of staff members combines with a creative environment to foster innovative solutions to researchers’ needs.
- ◆ A collaborative funding model enables researchers to invest in ARC’s staff time and equipment as appropriate.

“We saw an opportunity to start something that was a combination of investment and commitment by interested parties,” explains Lambert. “Once we aggregated a core of clients, it began to point to additional opportunities. The result is a path to computational research without a lot of dollar zeros. As we engage with the broader community,

one thing that surprises us is the number of institutions that look at ARC and say ‘here is a piece that we missed.’” This case study depicts ARC’s formation, characteristics, and organizational success factors.

Institutional Background

Georgetown University is the nation’s oldest Catholic university, founded in 1789 as Georgetown College with a small gathering of 12 students and a handful of professors. Today, Georgetown University encompasses several schools and colleges including Georgetown College, Graduate School of Arts and Sciences, Law Center, Robert E. McDonough School of Business, Edmund A. Walsh School of Foreign Service, School of Medicine, School of Nursing and Health Studies, and School for Summer and Continuing Education. As of 2005–2006, 1,700 faculty members teach the almost 14,000 students enrolled at Georgetown, 49 percent of whom are pursuing undergraduate degrees. Research is a defining element of university life. Georgetown’s mission statement describes the university as “a Catholic and Jesuit, student-centered research university.”¹ In 2004, the university conducted \$131,387,000 in sponsored research.

As vice president for information services and CIO, Lambert heads Georgetown’s IT organization, which is divided into two components. The first, the University Information Services (UIS) division, carries out daily technical operations, including the development and support of IT infrastructure, and assistance and support. Departments include:

- ◆ Network Computing Services, which develops and manages the infrastructure for the voice, data, and video networks;
- ◆ Academic and Information Technology Services, which operates the help desk, desktop support, student technology services, residential networking, the NetID Office, training, software, and technology classrooms; and

- ◆ Enterprise Engineering and Technology Services, which develops and manages the university's core business systems and its supporting infrastructure.

The second component, the Office of Information Services (OIS), focuses on strategic and financial planning, strategic initiatives, and IT-related policy areas. The five primary areas constituting OIS include:

- ◆ Center for New Designs in Learning and Scholarships (CNDLS), which promotes best tools and practices in traditional and technological teaching and learning;
- ◆ KeyBridge, an organization that designs and develops Web sites for Georgetown clients on a contract basis;
- ◆ Business Service Center (BSC), providing financial and regulatory management and administrative support for UIS;
- ◆ University Information Security Office (UISO), Georgetown's IT security organization; and
- ◆ Advanced Research Computing (ARC), which provides a computational and IT environment to support the goals of the main campus and the university's medical center.

Creating the ARC Division

ARC's roots can be traced to 1990 when ARC's program director, Stephen P. Moore, helped to establish a Computer and Laboratory Services division at the Georgetown University Medical Center. "When I joined the medical center, the IT organization was very fragmented," Moore recalls. "There was considerable need for desktop and networking support, and there was no formal organization to do it." Together with Dave Zapple, now an ARC senior systems analyst who also fabricates novel digital instrumentation for Georgetown scientists, he formed a private company that contracted with Georgetown to provide after-hours technical support on

nights and weekends. Eventually Georgetown bought out their company, forming the Computer and Laboratory Services division, which reported to Georgetown University Medical Center's first dean of research, Alan Faden. Employing three full-time staff members, the division used a collaborative research approach, funded by researchers' grants and contracts.

When Lambert arrived in 1998, his initial goal as CIO was to unify Georgetown's IT operations across all its campuses. One of the first issues he confronted was the lack of computational support for the university's researchers. "When I was contemplating Georgetown's CIO position, I talked with different university constituencies," Lambert recalls. "The researchers were the most disaffected community. They had written off the central IT organization, thinking their problems would never be solved." Lambert became acquainted with Moore, and subsequent discussions revolved around the issue of IT support of research. Eventually Moore proposed the creation of a single organization to serve the entire university. With his extensive research support experience and knowledge of the Georgetown community, Moore was ideally suited to lead this new group. "My years working with Georgetown researchers gave me experience in how to persuade people and to market the advantage of sharing resources," he explains.

In 2000, Lambert and Ardoth Hassler, associate vice president, launched ARC. Its Web site describes the new division's intentions: "Working with investigators in all computationally demanding fields, we help determine what specialized or customized computational resources and infrastructure are required to support the needs of the University. This is done by leading in the implementation and application of emerging technologies, and designing new technologies in collaboration with principal investigators."²

Drawing upon his medical center experience, Moore developed a “pay as you go” funding model that makes ARC a cost-effective entity for the university. Georgetown’s central funds did contribute to ARC’s start-up, but the goal was to shift to a researcher-financed model in which researchers (clients) pay ARC to manage their computational equipment or fund portions of ARC staff members’ salaries, or both. Clients get the skills of their funded person as well as access to other ARC staff members and resources. ARC adds new staff members as the workload warrants. Researchers immediately related to the funding model. “It is very similar to our cancer research model that is part fee for service, part subsidized by the cancer center’s own resources, and part funded by the NCI,” notes Clarke.

ARC is not the only research support unit on campus. Georgetown’s researchers are still free to build and maintain their own computational facilities, but they operate outside the ARC model. “We do answer phone calls and meet with non-ARC members if needed on a consultative basis,” explains Miles, “but we take no final responsibility or authority.” ARC will also help researchers secure grants if they do not have current funding to contribute to their services. “If a researcher comes to us for support and does not have any funding, we will set up an account and work with him or her with the goal of securing a grant,” explains Moore. “When he or she receives funding, he or she is free to join the ARC model or to go his or her own way.” The ARC Web site even contains a boilerplate description of the division’s resources, which applicants can paste into their grants and contracts.

Building on Success: Client by Client

From the start, ARC positioned itself as the new kid on the block to address long-standing research support issues. “I asked the research-

ers to give us a chance, to outline their requirements, and not to rehash old stories,” Moore recalls. “I presented ARC as a service-oriented organization, stressing the fact that *they* were the reason that *we* have jobs. *Our* mission is to develop solutions to *their* problems. I built my whole career at the medical center using this approach because I actually think it is the best way to do it. It resonates because you can’t build something and hope they will come. You have to work with people to help you build it and to enable them to co-own it.” Indeed, this approach enables ARC to use its resources to the researchers’ greatest advantage. “They don’t build a field of dreams or build solutions to perceived problems,” explains Clarke. “You discuss how you want to position yourself competitively in your field—currently and in the long term—and ARC will offer technology solutions for your review.”

ARC strives to become part of the research community at large. More formal communication methods are newsletters and e-mails about its activities as well as networking to keep abreast of university research activities. Additionally, the group interacts socially. Miles hosts social gatherings at his house, inviting researchers and department chairs to mingle with ARC staff members. Moore, Miles, and some faculty members play in a band and use it to publicize ARC when performing at Georgetown events. ARC staff members often eat lunch in a highly visible spot where they can be available to chat with researchers who happen to wander by.

Success breeds success. “Steve is a proactive leader who started collaborating with one group first and then gradually grew ARC’s client base,” states Cathy Wu, professor of biochemistry and molecular biology. By 2004, ARC reached a critical mass of buy-in from Georgetown’s computational scientists and currently works with 13 principal investigators. The ARC team itself has grown from three to nine staff members. Interestingly, nearly

all of the originally displeased researchers signed on as anchor clients. The sidebar, “A Client Sampler,” describes a few ARC assignments today.

Success Expands ARC’s Reach Internally and Externally

ARC’s resources and activities grew in tandem with its growing client base as its synergy generated expansion opportunities. In 2001, the division implemented Beowulf cluster resources and began to centralize pre-existing resources under a new model. “We took advantage of Linux and Beowulf technologies when they became viable options for managing clusters,” explains Lambert. In 2003, ARC established workshops to help PIs “parallelize” their code and began a grid technology awareness initiative for Georgetown’s research community. Another ARC activity leverages licensing agreements for certain software applications, such as MATLAB. “We brought MATLAB licensing under ARC so we can conduct vendor negotiations directly, creating economies of scale and greater value for the researchers,” explains Moore. “It also offloads support from the UIS help desk, enabling us to better understand the researchers’ MATLAB requirements by addressing their problems directly.”

“Gradually the research community began to realize the benefits from investing in new nodes and processors in our Beowulf cluster,” recalls Lambert. “Eventually we had a cluster with significant capacity, and we started to design an architecture that took it to the next level. Our goals were to create a secure environment for our computational equipment, to promote greater efficiencies, and to increase Georgetown’s grant and fundraising competitiveness.” The result is ARC’s Computational Core Facility, which opened its doors in 2005 to provide an intermediate resource between department clusters and national computing labs. Currently the CCF offers

five Beowulf clusters, a flexible disk storage configuration, and grid computing support, as well as application support for Blast, Charmm, Gaussian, MATLAB, NWChem, VASP, Mathematica, and internally developed applications. Georgetown provides the CCF’s space, cooling, electricity, networking, and backup generation; the researchers fund equipment and personnel costs.

Again, researchers are comfortable with the CCF’s shared resource model because it includes characteristics commonly used in research: participatory funding, full-time professional management, and an oversight management group. “We understand the core facility model because that is how we use microscopes in our research,” states Clarke. “ARC ties funding to a concept with which we are already familiar. You achieve a higher return by working together and leveraging resources rather than purchasing equipment individually. Once you made that connection, it became a natural way to think about it.”

The year 2005 brought two additional milestones. In April, ARC received its first external grant—an \$800,000 National Cancer Institute contract to design the cancer Bio-medical Informatics Grid (caBIG), a research tool to foster collaborative and interdisciplinary cancer research. ARC is designing the network architecture, developing systems that manage drug trials and integrate cancer researchers in the community through access to research material on demand and information sharing using common research tools.

The NCI grant is a direct consequence of the organizational synergy described earlier. Its origins lie with ARC’s ongoing work with Clarke and his team of 15 researchers at Georgetown’s Lombardi Comprehensive Cancer Center. When Clarke’s lab began to work in high-dimensional data spaces in genomics, the team discovered that they lacked adequate storage, security, and computational resources for their research. Clarke knew about the na-

A Client Sampler

Moore assigns a staff member on the basis of the project requirements and the staff member's expertise to be the researcher's direct contact, rendering the ARC internal organization transparent to the client. Most ARC members are physically located in the areas they serve to gain a day-to-day perspective of clients' activities and needs and to foster communication. Typical assignments include the Protein Information Resource (PIR), the biostatistics department, and the physics department.

PIR is an international consortium serving as the central sharing resource for protein analysis. It operates a Sun/UNIX server locally but houses its IBM mainframe in ARC's Core Computational Facility (CCF). The group funds a portion of ARC staff member Jess Cannata's manager/systems administrator salary for administration and maintenance. "The IBM mainframe was the starting point for our ARC relationship, but our computational needs grew to include protein comparisons, requiring a Linux cluster," explains Baris Suzek, senior bioinformatics scientist. "We have a local UNIX system administrator, and ARC became a complementary resource to administer our Linux cluster." The IBM machine is available to other researchers when unused by PIR.

Françoise Seillier-Moiseiwitsch, associate professor and chair of Georgetown's biostatistics department, currently funds 20 percent of ARC systems analyst Nick Marcou for Linux support and 25 percent of ARC programmer analyst Paul Kennedy. Seillier-Moiseiwitsch commends Marcou's responsiveness and creativity in problem solving. Marcou is currently working with Seillier-Moiseiwitsch to design and implement a teaching lab for the department's master's program. Kennedy programs algorithms and developed a Web-based randomization scheme for a clinical trial. Seillier-Moiseiwitsch hopes to increase her funding of Kennedy's time to 50 percent in the next fiscal year. "I know a little about programming and system administration, but these are in no way my areas of expertise," explains Seillier-Moiseiwitsch. "This is the best solution, because ARC hires very qualified people and the department pays for only a fraction of their time."

Jeffrey Urbach, physics chair and associate professor, looks to ARC to supplement his department's resources. "Space is a real issue for our department," explains Urbach. "We needed to move our servers to a remote location to free up department space. Why should we configure a separate room for a small number of servers when it is not much harder to set them up elsewhere? I am happy to let someone else take on the responsibility." In addition, Woonki Chung, originally the UIS systems administrator for several main campus science departments, including physics, computer science, and mathematics, moved to ARC as a systems analyst. His job description did not change; Chung still supports his main campus science departments but now has access to the ARC community and resources as well as serving as a local liaison between his respective science departments and the Georgetown IT organization.

ARC clients are expanding beyond the sciences. James Bodurtha, associate professor at the McDonough School of Business, uses ARC computational resources for his financial modeling as well as evaluating Windows and Linux computing equipment options. "ARC helped me benchmark several machines," explains Bodurtha. "Based upon their recommendations, I saved several thousands of dollars in equipment costs, choosing a dual-processor Linux machine over Windows alternatives. I then earmarked my savings for ARC support for my computational requirements."

scent caBIG project and wanted to implement a complementary solution at the cancer center. Instead of hiring an outside consultant, Clarke gave ARC an opportunity to solve some problems pro bono to establish their credibility.

As he learned more about ARC's activities, Clarke put two and two together and thought ARC should work directly with the NCI on its grid initiatives. For example, PIR, already an ARC client, had developed an authoritative resource for protein sequences and functional information that was of great interest to caBIG. Through this association, ARC could leverage its work in sharing microarray data on caBIG. "The caBIG project also brings together different groups of the university and makes us a very cohesive team," states Wu. "We were already working together in a collaborative spirit through ARC, which made it easy for us to get into the caBIG project framework." Now ARC maintains one of caBIG's nodes and is a key player in identifying and addressing data sharing and authentication issues for the project.

In September 2005, ARC reinforced its grid awareness initiative by launching the GridsWatch portal "to foster collaborative development of grid technology by assembling, exposing, and archiving grid technologies in a public forum. Our goals are to track, update, and summarize grid projects worldwide, and provide a current and ongoing information distribution channel for national and international grid projects, to include PI names and sponsors, academic and commercial affiliations, funding opportunities, descriptions of purpose and projected outcomes, reference implementations, middleware grid services, certificate authorities, authentication strategies, and news channels including alerts, conferences, and training opportunities, and a communications section for developers."³

Last, but not least, ARC's reach now extends into faculty recruitment in some departments. "ARC is not instrumental in hiring decisions, but we can meet with can-

didates to provide them insight as to how ARC's computational resources and services can assist them," explains Miles. "In turn, the conversations prepare us to meet their requirements if they join Georgetown." When Seillier-Moiseiwitsch requested to learn more about Georgetown's IT research support and programming resources during her interview process, the chair of the search committee introduced her to Moore and Miles. "I was impressed by the fact that they were willing to help me," she explained. "I knew then that my IT needs would be filled. I had no qualms about that."

ARC can assist with the new hire's equipment requirements also. "When calculating a new hire's initial equipment costs, a chair can work with ARC to build upon what is already in place instead of starting from scratch," explains Moore. "The department may reduce or avoid a \$200,000 expenditure. It develops economies of scale, and the equipment is operational the first day the new faculty member arrives."

Organizational Characteristics Foster Success

ARC's characteristics in regard to the Georgetown IT organization, within its own group and within its governance structure, all contribute to its success. All work together to create an optimal environment for ARC team members to focus intently on research support and generate new resource opportunities.

Separation of Enterprise and Research Services

From ARC's inception, Lambert made a conscious decision to separate it from Georgetown's UIS division. ARC falls under Lambert's OIS division for a specific reason. "I have an office with university-wide leadership responsibilities and an organization that delivers scalable, cost-effective, and reliable

technology services to the broadest range of people,” explains Lambert. “ARC’s constituency is the PIs; its solutions don’t need to be scalable to the entire university.”

This separation gives ARC the latitude to experiment as needed to meet researchers’ unique IT requirements. Miles explains, “We need our own playground as well as proximity to UIS. The separation enables us to take risks as needed. For example, we have the autonomy to use open source solutions even though Georgetown is a Sun/Microsoft shop. We can dabble with new technology without fear of taking down a critical service like e-mail. UIS has to understand the impact of all its changes on the entire campus environment; we don’t have that burden.”

Lambert does concede that the different missions and employee profiles of UIS and ARC can create opportunities for friction. “To say there have been no conflicts is naïve, but we have tried to address them, harness them, and create better partnerships,” he explains. Over time he has worked to get the two groups to become complementary. Professionally led encounter groups help the two areas understand each other better. Constructive debates between the two areas encourage the exchange of ideas. Additionally, ARC’s achievements are promoted as successes for the entire IT organization—for example, recognizing Network Computing Services’ role when a PI receives a grant. Technology issues, too, can promote commonality.

ARC is working closely with the University Information Security Office (UISO) IT security officer, David Smith, to address the rapidly evolving regulatory and privacy concerns in the research environment. The Computational Core Facility, with UISO guidance, will house those grants and contracts that require institutionally supported data security and electronic protected information (ePI) oversight, and the principal investigators can rely on professionals to safeguard their data.

Diversity and Collaboration Promote Creative Solutions

Moore uses his experience as a professional musician to instill diversity into the ARC team. “When you’re working in collaborative music, you look for the voices and the special things that each band member brings,” he explains. “You have to listen deeply to them. I approach building the ARC team like building a rock band. Each person brings a unique skill, and we use this uniqueness to generate synergy.” The team’s diversity enables Moore to mix and match team members as projects require, and it “creates an invaluable shared knowledge base that supports multiple research needs,” Clarke explains. “ARC serves different people with similar—but not the same—needs, so one solution to another client’s problem may fit—or at least influence—my direction.” Table 1 summarizes the ARC team’s varied capabilities.

Moore consciously keeps the organization flat to produce a collaborative environment. “I am the manager who builds relationships and resolves problems, but at the critical point, the team members do the work,” he explains. “The more I can keep myself in a lower position, the more we can get a collaborative team working.” ARC members are encouraged to debate, criticize, and share ideas and solutions either face-to-face or on their private listserv.

The ARC collaborative model stresses the importance of personal interaction, within and beyond the team. “If I have a certain problem, I know who to call on the team because everyone has a unique characteristic,” explains Marcou. “For example, if I need a shell script, I call Paul.” Cannata also describes “buddy development teams of ARC members who help me move the code. I don’t have the experience and expertise; our developers review the code for the problem and can port code across machines.” From the researcher perspective, Suzek notes, “We contact our

Table 1. ARC Team's IT Capabilities

Devices and Facilities	Systems Administration	Analysis and Programming	Other
Computational Core Facility management	Cluster design, operation, and maintenance	Database programming and maintenance	GridsWatch portal
Customized scientific device design and manufacture	Grid computing	Data manipulation and parallelizing code	Provide technology perspectives to researchers
	Assistance with local system administration	High-end computational programming	
	Linux programming	Microarrays, simulations, and modeling	

designated ARC staff member to address any problems side-by-side. It built trust.” Clarke describes, “The ARC staff may not have understood our questions in the context of how we asked them—cancer research. But they quickly understood our technology, our data characteristics, and our computational and storage requirements—how big our problems were—before we did.”

The flat organizational structure also gives staff members “the freedom to research our own solutions to problems,” states David Cafaro, systems analyst. “There is no prescribed way. We have the flexibility to try the technologies we feel best address the researchers’ needs. For example, we can use open source. If there is a product that needs to be purchased, we can purchase it with available grant money.”

The ARC staff member’s creativity is balanced by his very clearly defined job metrics. “They are not working to please me, but to facilitate the researchers’ success,” explains Moore. Staff objectives are set collaboratively and individually by Moore and each staff member. Evaluations focus on job performance, career goals, advancing ARC goals and missions, and training. Researchers contribute their unsolicited feedback throughout the year and their formal evaluations just before the staff member’s evaluation.

Moore also develops a career path for every team member. “Recruitment is based on career development,” he states. “It works because we emphasize the opportunity to learn new things as well as the team environment.” Staff members are recruited through the broad Linux community as well as within Georgetown.

The group monitors technology trends through journals, conversations, and conferences to stay two or three years ahead of the technology curve. “The researchers rely on us to stay on top of the technology,” states Moore. “It enables us in turn to offer insights into technological directions and to consider new approaches to their problems.” Each ARC staff member watches a different subject area.

Advisory Committee Provides Feedback and Direction

A mix of stakeholders—ARC staff members and researchers—make up ARC’s seven-person advisory committee. In fact, through ARC’s work with caBIG, the committee recently welcomed its first non-Georgetown members: Frank Manion, CTO of the Fox Chase Cancer Center, and Randy Ford, an expert in AI and the CTO for Sonum Technologies. The committee reviews service, resource, hardware, and financial issues, but probably

would not intervene in internal ARC matters unless it impacted one of the researchers directly. One unanticipated advantage of committee participation is peer networking. “The committee is useful for that reason alone,” states Clarke. “We discuss our own issues and solutions—as well as ARC’s—and learn from them. It is a very, very useful group of people.” Additionally, the meetings help ARC stay abreast of researchers’ concerns and requirements.

The Next Steps: Expansion into New Areas

In 2006, ARC’s synergy and momentum continue to broaden its range of resources and activities on campus and beyond.

Condor Initiative Creates Campus-Wide, Perhaps Multi-Institutional, “Grid”

A new university-wide Condor initiative follows ARC’s creed—creating computational resources with minimal investment—and builds on its grid computing work. “On the one hand, it is painful to watch computers sit in our labs underutilized,” explains Miles. “On the other hand, I know researchers who must invest \$100,000 to \$200,000 to expand their Beowulf clusters. In some cases, we can save researchers the cost of buying and maintaining another Beowulf cluster by scavenging machine cycles from these idle machines.”

Miles and Cannata conceived this plan three years ago; ARC’s rising credibility with UIS coupled with technology improvements make 2006 an opportune time to implement it. Working closely with the UIS Academic and Information Technology Services group, ARC collaborated to implement Condor on Georgetown’s lab and publicly accessed computers. The project is still in the pilot stage, but the goal is to roll it out lab by lab to create a 450-node Windows-based campus Condor

pool—or a campus-wide computing “grid.” ARC’s programming and system administration teams are evaluating appropriate applications, such as Blast or Gaussian, to run on the campus Condor implementation. Internally developed applications must pass a testing process to ensure the code will not damage the resources.

Miles sees the potential for a multi-institutional Condor initiative beyond Georgetown, appealing especially to smaller campuses. “It is an ideal tool in situations where there is no shared memory [or] parallel processing [machines], or expensive applications,” explains Miles. “User scalability is an issue with grid computing; this could be one solution.”

The Condor-Shibboleth Project

One outgrowth of ARC’s Condor initiative is Georgetown’s Condor-Shibboleth project with the University of Wisconsin. Together, the two universities are developing grid access to computational resources using role-based authorization. The Condor-Shibboleth Web site explains the project’s objectives:

Condor’s “flocking” technology allows multiple Condor installations to work together to complete large job orders. Shibboleth complements Condor by creating a customizable, secured access point that can define any desired set of user parameters to regulate user priority, access time, and resource usage, and securely make those parameters available to cooperating institutions. The merger of Condor and Shibboleth will create a scheduler software package capable of consuming role attributes in a framework that allows rapid, scalable control of the utilization of computational resources for collaborations which span administrative domains. This coupling of Condor flocks encourages inter-realm computational scenarios, increasing access to idle computational resources.⁴

ARC harnessed its networking and collaborative abilities to address a potential security risk with its Condor initiative. Condor offers a potential means for hackers and unauthorized users to access the university's computational resources. ARC team members consulted with UIS principal technologist Charles Leonhardt and with team leader Chad La Joie to determine potential solutions. They developed an idea to implement a Shibboleth-based authentication and authorization solution. ARC connected with Internet2, a consortium working to develop and deploy advanced network applications and technologies, which in turn referred the group to Miron Livny, professor of computer science and Condor project lead at the University of Wisconsin. Now the two groups are working together and seeking funding. This project not only addresses Georgetown's IT security problem but also positions the university well for the next phase of grid technology.

Other Activities

ARC is also applying its model and experiences to expand its activities in three additional ways.

- ◆ To expand ARC's research support beyond Georgetown's hard sciences programs, ARC is providing support for social science research and computational linguistics applications. Current projects include James Bodurtha's financial modeling and the work of Maxine Weinstein, distinguished professor at the Center for Population and Health.
- ◆ To increase awareness and interest in Beowulf clusters, ARC is codeveloping a four-day training event with the Washington Area Beowulf Users Group. "The impact of getting people together, creating synergistic connections, and creating awareness of research tools that are available on and off the Georgetown campus cannot be understated," says

Miles. Potential training subjects include planning and implementing Beowulf clusters, environmental issues, and potential computational applications, as well as a hands-on workshop. Interestingly, this is a paid registration event, creating a new revenue stream for ARC.

- ◆ Finally, ARC is in initial discussions with caBIG's third-party project administrator, Booz Allen Hamilton, to leverage the lessons learned from the caBIG initiative into other medical-related projects and applications outside of cancer research. "By leveraging our tools and experiences, our work with caBIG will scale quite nicely into non-cancer-related research projects to develop new business and work," states Miles.

Lessons Learned

Members of the Georgetown University community offer several lessons learned; some are general truisms, others are specific to implementing an IT research support organization.

- ◆ *Faculty must determine research IT support direction.* Researchers are under special pressures to produce results in a time frame specified in their funding grants. Left to their own devices, they may try to control as many parameters as possible. But when it comes to IT support, "doing it for themselves" is not always in the institution's—or even their own—best interest. To the degree that a support organization can listen carefully to the campus research community as well as design efficient and effective services, researchers will have more confidence in the support model. "What has made ARC work for us," explains Urbach, "is that they are flexible and willing to hear our requirements. Rather than fitting us into a different box, they make the box fit us."

The exact nature of the services will likely be context sensitive. As articulated

in this and other case studies accompanying the main ECAR research study, these services may be in the form of computation, storage, visualization, advanced networking, staff support, or most likely some combination of these. In the case of Georgetown, the emphasis has been on staff support with some elements of computation and advanced networking.

- ◆ *Look internally to the researchers to understand and support their IT needs successfully.* This might seem self-evident on its face, but it is tied and complementary to the observation above. IT organizations, even while recognizing their research supporting role, often develop an agenda of their own and measure their success accordingly. IT staff can be tempted sometimes to attempt a new project just because they can do it, not because it addresses a specific need. But a research support organization's success should be measured in terms of the academic research it facilitates. As Chung notes, "ARC works with intelligent people who suggest new technology and new directions to ARC. Many IT organizations miss opportunities for collaboration." ARC has learned this lesson well and firmly disagrees with the "build it and they will come" philosophy. Staff members are motivated to support ARC's philosophy through their job metrics and performance evaluations.
- ◆ *Look externally to monitor, evaluate, and introduce new technologies to researchers.* That being said, there is a delicate dance that goes along with setting the technological stage for current initiatives and being prepared to support researchers' evolving needs. ARC knows that it can't be successful just waiting for a researcher to request another service. In fact, researchers generally are focused on their areas of expertise and don't have a good understanding of emerging IT

trends and capabilities. Therefore, part of a research IT support organization's role is to watch technology trends and help educate researchers regarding possibilities that will aid their work. As mentioned earlier, for example, each ARC staff member is assigned to monitor a specific technology topic, such as grid computing or Linux, through journals, conferences, and conversations.

- ◆ *Communicate to build trust-based relationships.* The above activities all enable ARC members to build personal connections with the researchers. "We invest a significant amount of time into meeting the individual PIs, not as members of the individual departments, but as people and researchers," explains Moore. "If there is no effort to know and connect those people into the ARC network, there is no communication. Without communication there is no trust. A top-down approach isn't the way to proceed; you have to create a social network." Wu concurs: "You need good communication between investigators and service providers to understand the researchers' capabilities and to come up with new solutions. This over time develops trust." And this, in time, opens the door to more opportunities. "If people feel you are supporting their efforts, it is astonishing how much they are willing to trust and invest in collaborative solutions," states Lambert.
- ◆ *The quality of people makes the difference in successful research support.* Lambert summarizes one ARC success factor: "Good people are critical. They have to be knowledgeable, flexible, and executors." Indeed the researchers supported by ARC were unanimous in praising both "their" identified support person and the ARC model of bringing the skills of the entire team into play as necessary. ARC's success in recruiting talented staff has resulted in this high degree of satisfaction.

Additionally, the model allows researchers to have a sense of both personal service and getting more than what they paid for. “I am a customer, not another slip in the general inbox,” explains Seillier-Moiseiwitsch. “It creates a sense of obligation and encourages responsiveness to my needs.”

- ◆ *Give careful thought to organizational design and job functions in order to hire and retain good research IT support people.* Research support requires technical competence, creativity, and a willingness to take risks. Moore created an environment and staff that amplifies these traits. Without a doubt, the most frequently used description of an ARC staff member was the willingness and job freedom to “think outside the box.”

ARC is autonomous from, but interconnected with, Georgetown’s production and enterprise IT services organizations. The autonomy attracts staff members who are more willing to push the envelope, which correlates with the inventive nature of academic research. The interconnection gives ARC access to IT services and skills such as networking and security, so that they don’t have to re-create them within the group. In addition, Moore fosters a flat organization to promote dialogue and collaboration among team members.

Moreover, since most staff members are associated with a specific research group and are also members of the broad support team, their job skills have to be simultaneously keyed to specific research needs and complementary to those of their ARC colleagues. Moore compared this earlier to assembling musicians to create a band. This approach, too, enables Moore to “mix and match” ARC staff as situations warrant. Finally, team members must have time and resources to grow professionally, enhancing their contributions to the team.

- ◆ *Be opportunistic.* Not only does the IT research support staff require the freedom to “think outside the box,” but so does its senior leadership. The ARC leadership took a small service center and built it into an institution-wide research support unit. The leaders did this by recognizing that there was a vacuum at Georgetown in IT research support, and they formed multidisciplinary alliances with researchers to fill it. ARC continued to build upon its success, eventually expanding its activities beyond Georgetown’s borders. For example, the group developed skills in grid computing to fill the need not only locally but also for the discipline as a whole.
- ◆ *Scarcity of resources can drive innovation.* Normally, the scarcity of resources would seem a barrier to success. At Georgetown, Lambert articulated this as one of the success factors for research support. When researchers don’t have the resources to “go it alone,” they become more willing to participate in collaborative activities. The genius of the Georgetown research support model is that it allowed researchers to use relatively small amounts of grant funding to finance fractions of a support staff person, who in turn belongs to a broader team. The model served to bootstrap the growth of the group. They repurposed existing resources, used that base to make researchers more competitive in the grants process and, when a participating researcher received a grant, used a portion to add to the support group. The same principle applies to computation resources. When a research group doesn’t have sufficient funding to buy what it needs, it becomes more willing to join in building cooperatively owned but centrally managed clusters.
- ◆ *Use the familiar to introduce the new.* When Lambert and Moore developed the staffing and computation resource models above, they relied on concepts

with which researchers were familiar to promote buy-in. “We talked about our models in ways that appear similar to those already embedded in our culture,” explains Lambert. “We avoided the appearance of introducing alien structures into the research culture.”

- ◆ *There are benefits in considering open source solutions.* While many recognize the benefits in leveraging the talents of a broad community of open source developers and system support people, Clarke pointed out that proprietary applications often fill only a portion of a group’s needs, and those applications can’t be easily extended or modified. By standardizing on open source approaches, ARC has been able to build a flexible environment that leverages common staff skill sets to meet a diverse range of researcher needs. Additionally, staff members can customize tools as necessary for specific application to a research problem.

Conclusion

ARC began with a simple goal: “We make researchers successful by providing the support to help them achieve their visions,” states Moore. But what began as a means to provide local IT computational support has blossomed

into a multilayered and multi-institutional range of resources. The keys are to expand gradually by constantly building and connecting a network of contacts and past successes to create new opportunities—and to do so without forgetting the primary mission of fulfilling researchers’ IT support needs. “As a result, Georgetown is now more competitive in the race for research funds,” states Lambert. “Our challenge has been to show institutionally the real research strength that a top-25 research university requires. ARC is an asset that affirms this initiative and delivers real value to help Georgetown position itself for the future.”

Endnotes

1. Georgetown University Mission Statement. Retrieved June 19, 2006, from http://www3.georgetown.edu/president/mission_statement.html
2. Georgetown University Advanced Research Computing (ARC). Retrieved June 19, 2006, from <http://arc.georgetown.edu/>
3. Georgetown University GridsWatch. Retrieved June 19, 2006, from http://www.gridswatch.com/index.php?option=com_content&task=view&id=4&Itemid=10
4. Condor-Shibboleth Integration Project. Retrieved June 19, 2006, from <http://hendrix.arc.georgetown.edu/condor-shib/>