10
Research IT at the Inflection Point

All science is computer science.
—George Johnson

The organization of IT support for the research enterprise is at an inflection point. Since the end of World War II, research has assumed greater importance within the academic community as a source of institutional and personal prestige. As the amount and complexity of research have grown, there has been an increasing reliance on IT for the conduct of that research, a trend supported by government programs and funding. These cumulative efforts have culminated in the development of a cyberinfrastructure vision that by its integrative and boundary-bursting nature promises to transform the nature of science itself.

Central IT organizations, however, have largely forgone the opportunity to engage in the facilitation and advancement of research. While they have been instrumental in building and maintaining key elements of the infrastructure, most notably the network, they have focused their attention and resources on priorities other than research, enabling the propagation of local support mechanisms. The trends outlined in this study, however, signal the beginning of what is emerging as a fundamental shift in the role of central IT in the support of research.

Past as Prologue

For many years, the research environment has been characterized by strong centrifugal forces, the most fundamental of which is the academic culture itself. Disciplines maintain their own intellectual and methodological traditions, and the explosion of knowledge has fostered the proliferation of discrete subdisciplines. Organizationally, departments have regarded themselves as discrete bastions of expertise, and their boundaries have been well maintained and protected. The mechanisms for supporting research have mirrored this pattern, with researchers creating, capturing, and controlling the resources they need to conduct their investigations. Whether it is labs, computers, or personnel, local autonomy has been the norm.

The survey confirmed a marked pattern of locally maintained resources, especially among institutions that identified research as a primary element of their mission. Within the Research Essential institutions, responsibility for both computation and data storage lay somewhere between shared central and local, and mostly local. When we looked at 12 research IT support services, those most specific to a discipline were more likely to be

done locally. This tendency, again, was even more pronounced in the institutions that were heavily invested in research.

Much of the centrifugal force is generated by the academic culture and by the logical tendency to rely on people who know the discipline, understand the research problem, and work down the hall. Yet part of the reason, too, is that the central IT organizations, for their part, have not made research support a major organizational imperative. Three-quarters of the responding institutions, including 58 percent of both the Research Essential and Balanced institutions—and these, remember, were the institutions interested enough in research to respond to the survey—had three or fewer FTE assigned to research support and low levels of engagement with researchers.

The lack of attention and resources devoted to research by central IT organizations is not surprising, given other urgent priorities. Of the 30 choices offered by the 2006 EDUCAUSE Current IT Issues Survey, research support failed to make the top-10 lists of issues vital to strategic success, issues with the potential to become more significant, issues that IT leaders spend their time on, or issues that consume the most human and financial resources. When the issues are broken out by Carnegie class, however, research support did show up fourth on doctoral institutions' list of issues with the potential to become much more significant in the coming year; tenth on doctoral institutions' list of issues most important for strategic success; and tenth on doctoral institutions' list of issues on which the campus is spending the most financial and/or human resources (Dewey et al., 2006).

Drivers for Change

With the intense focus on security and identify management, IT funding, and administrative/ERP systems (the top-three issues on the EDUCAUSE list), IT support of research may prove to be the stealth issue of the decade. A number of factors are driving a reconsideration of the traditional every-researcher-for-himself model of support.

At the heart of the transition are changes in the nature of science itself. Scientists are posing bigger questions about more complex phenomena. Funding agencies have made clear their lack of interest in single-researcher approaches and are insisting on collaboration among researchers from different disciplines.

Remote instrumentation is generating massive amounts of data, creating storage requirements in the petabyte range. Transient graduate students no longer suffice as the custodians of file-management systems. As computational needs grow, power, space, and cooling become bigger issues. Empty storage closets, especially ones that can accommodate 16-node clusters, are harder to find these days. The scale of science exceeds the capacity of any one researcher or lab.

The trend toward multi-institutional and international collaboration creates challenges of account management, identity management, and authentication. Tensions between privacy and security on the one hand and the distributed nature of data on the other are growing. The complexity of the data itself makes custodianship more difficult. Who manages large-scale data sets when the researcher changes institutions? The answer is not always clear.

Cost is another big driver for change. Funding agencies have shown that they are not willing to have their money spent on the support of existing computational resources, preferring to commit funds to newer projects. This is driving researchers to seek alternative means of support, such as cluster condominiums, where facility and personnel costs can be shared.

Reflecting on an NSF-funded workshop on campus cyberinfrastructure,[1] Ken Klingenstein, chief technologist at the University of Colorado and director of Internet2 Middle-

ware and Security, noted that the meeting was attended by CTOs, CIOs, and heads of research IT. "The motivational factors to change are really strong out there now. Everyone was aware of the need to step back into the research computing game. That part is back on the radar screen" (personal communication, May 5, 2006).

As universities wrestle with how to do that, it is becoming clearer that the central IT organization has a crucial role to play. Over the past decades, evolving technologies have led to a seesaw of ideas about IT management, governance, and division of responsibilities. Centralization gave way to more distributed resources, and several factors are now swinging the pendulum back to a greater acceptance for central IT leadership.

Identifying the Source of Value

Leadership, especially in a decentralized academic environment, only comes when people see value in going along. Central IT organizations therefore must align their support of research with both the institutional mission and researchers' needs. Are researchers seeking access to national and regional networks? Are they looking to expand their access to high-performance computing? Are they concerned about limited IT expertise within their local support community? Are they seeking better facilities for their equipment or security for their data? The answers are probably yes, yes, yes, and yes. But while common needs exist across all types of research enterprises, researchers at each institution have their distinctive configuration of needs and priorities. James Bottum, who recently announced his appointment as chief information technology officer at Clemson University, discussed the challenges he will face in his new position:

> While research environments will have common characteristics, it is not the same everywhere. I couldn't duplicate what I did at NCSA [National Center for Supercomputing Applications at the University of Illinois] at Purdue, and I won't be able to duplicate what I did at Purdue at Clemson. You need a good understanding of your institution's research profile and the cyberinfrastructure within your university. It will be something different everywhere. (personal communication, May 5, 2006)

The central IT organization must align its activities with the priorities of the research community. As Thomas West, CEO of National LambdaRail, expressed it, "Everything ties back to the institutional mission and programs. The CIO has to understand what kind of computation, data storage, and network access is required for faculty and students, because these faculty and students maintain the stature of the institution. Everyone has to be alert and engaged. And the commitment has to be made to the infrastructure to support the institutional goals" (personal communication, January 27, 2006).

This is not just an issue for the large research-intensive doctoral universities. Bachelor's colleges will have their own distinctive set of demands. Joanne Kossuth of the Olin College of Engineering described the challenge of supporting undergraduate engineering education:

> Small-school research projects are different from large universities. There are lots of small projects, and many are curriculum related. This can lead to conflict between IT and faculty at small schools. Faculty say that research should be supported at the central level because the research feeds into the curriculum and is not supported by grants. Others say that the support has to come out of grants. This causes tension when there is a gray ground between curriculum and research. If

> the research can't be tied back to the curriculum, how do you support it?… We have not answered that at the BA level. What kind of resources will we commit? How do you encourage academic experiment and freedom without breaking resources?

Several ECAR studies have focused on how central IT organizations at Purdue, Princeton, Georgetown, and the Universities of California at San Diego and Irvine have worked with researchers to enhance their research IT capabilities. But there are other examples as well. The Internet2 workshop on campus cyberinfrastructure (see endnote 1) included presentations from a number of universities showcasing how greater coordination between central and local units leveraged research IT resources. One presentation highlighted how Information Technology Services (ITS) at Penn State University has been working with researchers to address service opportunities in the areas of computation, networking, visualization, and collaboration. Among its primary computation support activities has been the creation of condominium-style clusters of usage and acquisition, based on the value proposition that when you aren't using yours, others can; when you are using yours, you can use more than yours. Several computational clusters serve researchers from disciplines as diverse as chemistry, meteorology, entomology, engineering, physics, anthropology, biology, and mathematics. The core of ITS philosophy is to make it easy for faculty to do their research by taking on the tasks and risks that the faculty should not and do not want to assume. Extensive education and training are provided to researchers with the attitude of "what you want, when you want it, where you want it." ITS's cooperative approach has shown benefits through measures such as inroads into new departments, more frequent joint authorship and acknowledgments, a more substantive role in faculty recruitment, and greater platform diversity. As opposed to modeling support around the battle cry of "build it and they will come," ITS's strategy for supporting campus research is founded on a principle of "build it together and they are already there."

A second example comes from Rice University's Computer & Information Technology Institute (CITI), which was established in 1986 to support the development of cross-disciplinary research in the areas of information technology, high-performance computing, and computational science and engineering. Separate from the central IT organization, CITI has as its core mission to "build a *community* of scholars that engages in *collaborative* research and education covering virtually every aspect of information technology and computing." Until recently, every research group at Rice developed and managed its own computing research infrastructure. It became clear, however, that this model would neither support appropriate scaling nor be cost-effective for individual research groups seeking to manage smaller local resources built for peak demand. Accordingly, the leadership of CITI has moved the unit toward the procurement of larger shared computing infrastructure, offering opportunities to faculty to off-load the burden of maintaining local resources. Jan Erik Odegard, executive director of CITI, explains:

> In trying to solve this problem, we realized that CITI was well positioned to lead an initiative that would address a rapidly growing need for coordination around cyberinfrastructure. We are backed by a strong community consisting of more than 135 affiliated faculty members, or about 25 percent of the faculty at Rice. Rice is a small institution, with only 550 faculty members, about 200 of whom are in science and engineering. Hence, the membership in CITI constitutes a

substantial community needing computational research support. (personal communication, June 6, 2006)

To date, the funding for infrastructure procurements has primarily come from grants and industrial partnerships. Shared resources, such as system management, power, cooling, and space, are currently funded by the university. Moving forward, the issue of funding campus cyberinfrastructure on an ongoing basis remains an open question. As it has become clearer that a separate organization for supporting research computing would not be efficient for Rice, more focus has been placed on finding ways to partner with the campus-wide IT support division to develop the support model for cyberinfrastructure for research. The process was helped considerably by the arrival on campus in the spring of 2004 of a new vice president for IT who both understood the importance of access to core university infrastructure in support of the education mission and was committed to building an organization that would be able to support existing and growing needs for cyberinfrastructure focused on research. The vice president has since embarked on a process to realign the IT organization so that it is a strong advocate for faculty and has formed the Information Technology Advisory Committee as a mechanism for stakeholders to participate in setting priorities for the IT organization.

Among CITI's key accomplishments have been

- the development of a core-services service level agreement (SLA) outlining what the university will provide for every member of the institution at no direct cost;
- the development of a research-support SLA that addresses how IT will facilitate the scalability and sustainability of research computing support;
- the funding of a new campus network that will leverage city, state, and national networking initiatives such as the Research and Education Network of Houston (RENoH), the Lonestar Education and Research Network (LEARN), Internet2, and the National LambdaRail (NLR);
- plans for a new state-of-the-art data center; and
- the formation of the Research Computing Support Group, a group dedicated to supporting IT needs derived from research.

Odegard sums up CITI's opportunity in this way: "The main challenge with supporting research computing is the unpredictability of how much, what it looks like, and how it scales as the research enterprise grows. Each of our initiatives is important independent of a particular research agenda, and together these projects stand to benefit computational research significantly and will impact the way we continue to support the needs of our research community."

As these examples demonstrate, there are multiple sources of value to researchers and the people who support them—reliability, security, serviceability, cost-effectiveness, performance, staying on technology's edge, maintaining a vendor relationship—and each IT leader must discover what combination of these values holds the greatest promise for researchers.

Engaging Researchers

The only way to really understand researcher needs and the value central IT can bring is to engage with them. One of the most striking findings from our study was the limited extent to which central IT organizations currently engage the research community through formal mechanisms, long-term planning, advisory groups, participation in grants, or involvement in faculty recruiting to identify potential needs. There have just been too many other priorities.

But the tide is turning, much of it driven by the researchers and support staff themselves. Researchers and the IT staff who support them are often isolated by school, discipline, or campus geography. Research computing groups within central IT can play a mobilizing role to foster a forum for like-minded people to share practices, skills, ideas, and fellowship. When the University of Iowa conducted an e-research needs assessment, it listed "community" as the first of its five critical research computing issues.

> A strong and consistent message we heard was the desire of better communication, coordination, identity, and leadership. A vibrant research computing community will foster visionary leadership … in the increasingly critical area of competition for research funding. It is clear that the University of Iowa faces a significant challenge in transforming itself into a collaborative enterprise where internal organizational boundaries are far less important than in the past. An active research computing community is a good step in that direction. (University of Iowa, 2006, p. 8)

Building a community and partnership requires small trust-building steps. Ken Klingenstein noted,

> The single phrase we heard most at the NSF workshop is that researchers don't know how to manage security. Or they don't know how to manage files since their grad student left. There are lots of low-cost measures IT organizations can take, like data management courses for faculty, or partnering on grant proposals to provide support services. The trick is to identify places where the enterprise can minimize out-of-pocket expenses and maximize what the researcher gains, such as federated account management, firewall management, or training for researchers in data protection and HIPAA. We can train graduate students on how to do applications programming on local high-end resources. There are virtual organization management services that allow users to manage access and significantly improve collaborations generally. (personal communication, May 5, 2006)

Jim Bottum explained the steps that enabled building a research IT community at Purdue:

> Faculty have a variety of research support needs, and it is always good if you can start with some low-hanging fruit. Find a few good projects where you can add value; don't try to do too much. We found a faculty group that had created some very useful technology that had been supported by grant money that had expired. We offered to support that technology, which freed them up to focus on research and explore new directions. You need to figure out what you can do and then establish yourself as a real partner.… It's the contact, the interaction, and what you can bring to the table, not the money or the facilities. Learn the research environment, what the researchers need, and how they operate. (personal communication, May 5, 2006)

Over time, opportunities arise to rethink how roles and responsibilities are distributed. As we saw in Chapter 8, the University of Iowa campus group that reviewed IT support for research envisioned central IT taking on greater responsibility for physical space, systems administration, data storage, security, Web hosting, data management, and OS and e-mail support. The support related to research, scholarship, and creative works was still seen as being primarily local. Effecting such

redistribution of responsibilities will require, of course, that central IT, local IT units, and researchers develop a common perspective on optimal arrangements and that they must be worked through and negotiated by interested parties from across the institution.

The data from this study showed that one of the strongest indicators of central IT commitment to research was the presence of a unit within central IT dedicated to the support of research. Institutions with such a unit generally had

- higher funding levels for research-related IT,
- higher levels of formal engagement with researchers,
- higher self-reported satisfaction of researchers with central infrastructure and services,
- faster expected growth in central IT research-related staff,
- greater likelihood of long-term planning around research IT needs, and
- a stronger relationship with the institutional research office.

Of course, establishing a unit will not in itself create these outcomes, but it does indicate the benefits that can accrue when, after building trust and relationship, an institution takes the next step and formalizes its commitment to the service of research.

Focusing on the Science

IT in the service of research can, like the research it supports, push the frontiers of knowledge. For some IT professionals, the thought of advancing the technical state of the art is an appealing challenge. But many researchers don't want or need the leading edge. Ken Klingenstein framed the issue for one technology: "Visualization centers are an interesting capability, but there is a feeling that you don't need the high-end stuff like video walls as much as you need convenient stuff. So there is a tension between creating class-A facilities versus the researchers' desire to work within their building" (personal communication, May 8, 2006).

As they continue to partner with researchers and their support units, central IT organizations must focus on the research itself. Linda O'Brien of the University of Melbourne offered a cautionary note by comparing today's state of e-research with the development of e-learning:

> We have a great deal to learn from our past experiences with e-learning that is directly applicable to e-research. New and emerging areas are exciting, but we also must realize it is about developing core capabilities into the future. Allowing the mavericks to run things ended in an unsupportable environment, with many different approaches and poor quality from a student perspective. We had to standardize and move to a more seamless environment. E-research has the same culture and the lessons are very relevant. Our approaches to e-learning, their successes and failures, can give us ideas about how to proceed with e-research and maybe avoid some of the pain we went through. E-learning is much more mature now, and provides lessons that can allow us to get e-research mature more quickly, how to manage the change process better, and how to implement more effectively. (personal communication, January 23, 2006)

In the final analysis, it's all about the science.

Endnote

1. The "Workshop on Effective Approaches to Campus Research Computing Infrastructure" was held April 25–27, 2006, in Arlington, Virginia. The agenda and presentations from the workshop can be found at <http://middleware.internet2.edu/crcc/>.