Roadmap

Tools for Navigating Complex Decisions

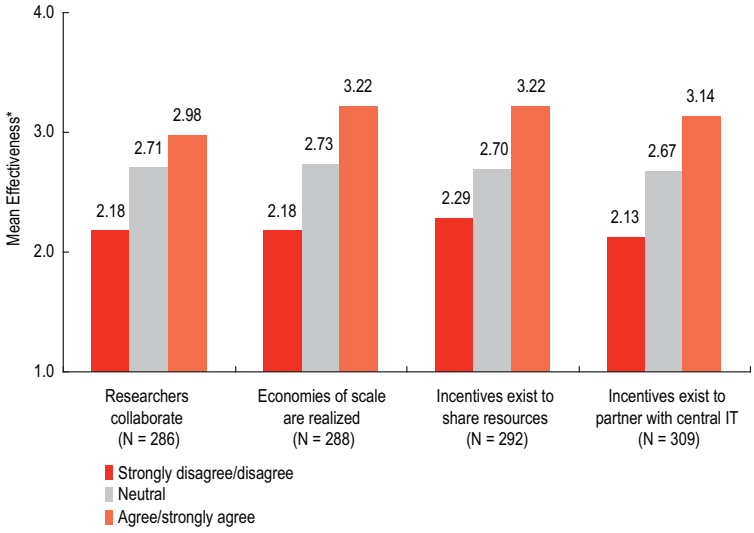
Higher Education IT and Cyberinfrastructure: Integrating Technologies for Scholarship

By Judith A. Pirani and Mark C. Sheehan, Fellows, EDUCAUSE Center for Applied Research

KEY FINDINGS

- About two-thirds of respondents use three of the five core cyberinfrastructure (CI) technologies (CI applications and tools, data storage and management resources, and resources for collaboration within virtual communities); about half use the other two (high-performance computing resources and advanced network infrastructure). Institutions foresee substantial growth in the importance of all CI technologies, both to their research activities and to their teaching and learning activities.
- For all CI technologies, the CIO's knowledge about them increases with the extent to which central IT provides access to and funding for the technology. The CIO's overall knowledge about each CI technology is greater where the institution's documented inventory of that technology is more complete.

Effectiveness at Integrating Cyberinfrastructure Resources, by Agreement about Collaborative Practices and Incentives



** Scale: 1 = not effective, 2 = slightly effective, 3 = moderately effective, 4 = very effective, 5 = extremely effective*

- Respondents identified three items that would most help the central IT organization support more effective research use of each of the five CI technologies at their institutions: more funding for central IT infrastructure, more funding for IT services, and better communication and outreach between researchers and central IT.
- Respondents are reserved in their assessments of the degree to which their institutions' researchers collaborate in the use of CI resources and to which their institutions realize economies of scale. They are also reserved in their agreement that effective institutional incentives exist to encourage the sharing of CI resources among researchers and for researchers to partner with central IT to achieve economies of scale.
- On average, the more collaborative and efficient the research community is judged to be and the stronger the respondent's agreement that incentives for collaboration and efficiency exist, the more effective the central IT organization is at integrating CI technologies to provide seamless support for research.

Cyberinfrastructure is not a term that trips lightly off the tongue or one that reveals its precise meaning without some coaxing, but its importance to research is clear. The National Science Foundation (NSF) defines CI as the layer of hardware, algorithms, software, communications, institutions, and personnel that enables and supports research. This layer should provide an effective and efficient platform empowering specific communities

This ECAR Roadmap summarizes the 2008 ECAR study, Higher Education IT and Cyberinfrastructure: Integrating Technologies for Scholarship, *by Mark C. Sheehan. That study synthesizes 369 responses to a November 2007 online survey of the EDUCAUSE membership and interviews with 12 executives and staff members involved with cyberinfrastructure resources and practices at 11 institutions. To access the full study or to learn about other ECAR publications, visit the ECAR website at http://www.educause.edu/ecar/ or contact us at ecar@educause.edu.*

ECAR EDUCAUSE Center for Applied Research

TERMINOLOGY

Higher Education IT and Cyberinfrastructure: Integrating Technologies for Scholarship studied five sets of technologies that lie at cyberinfrastructure's core:

- *High-performance computing resources:* Supercomputers and clusters of computers or other computational devices integrated in such a way as to provide supercomputer-like performance to individual applications. (Note: The network infrastructure on which high-performance computing relies is addressed below, under the heading "advanced network infrastructure resources.")
- *Cyberinfrastructure applications and tools:* General CI applications and tools that support research but are not specific to a particular discipline. These include software for simulation, parallelization, visualization, job scheduling, data mining, statistical analysis, and so forth but not specific sequencing, chemical analysis, or other disciplinary applications.
- *Data storage and management resources:* Large-scale research data storage systems for real-time use and for archival purposes, as well as facilities, software, and procedures for periodic backup of research data sets.
- *Advanced network infrastructure resources:* The institution's high-performance networks on campus and their connections to off-campus high-performance networks that support such capabilities as massive data transfers to and from clusters, real-time visualization, and use of remote instrumentation. Off-campus networks used for advanced network infrastructure in North America include regional or university consortial networks and such networks as Internet2, National LambdaRail, and CANARIE.
- *Resources for collaboration within virtual communities:* Facilities and support for teleconferencing, for hosting collaborations with off-campus researchers, and for the operation of remotely located research instrumentation and related devices; support for identity management and associated middleware in collaborative research activities.

of researchers to innovate and eventually revolutionize what they do, how they do it, and who participates.[1] Simply put, CI brings to the highly distributed research community many of the strengths of the enterprise IT infrastructure. It aims to integrate critical research technologies, with an eye to encouraging and enabling collaboration among researchers, achieving economies of scale within the institution or the discipline, and developing a seamless fabric of research support infrastructure and services. In short, an optimal CI promotes superior research.

Despite the clarity of the CI vision, many practitioners find it ambiguous when it comes time to put research infrastructure into place or to support it. For CIOs, the dividing line between IT infrastructure for advanced research and scholarship on the one hand, and infrastructure for everyday computing on the other, is blurred by researchers' variable levels of need, by the rapid progress of technology, and by higher education's pervasive, decentralized, cottage-industry style of conducting advanced research.

For these reasons, especially the last, it is often a struggle for CIOs to establish an enterprise-wide foundation for CI. Until recently, CIOs looking for strategic guidance about CI have had relatively few resources to turn to. Information about CI practices at other campuses has been scarce and largely anecdotal. The ECAR study, *Higher Education IT and Cyberinfrastructure: Integrating Technologies for Scholarship*, fills this need by documenting the patterns of use of CI resources and institutional leaders' knowledge about them, assessing the status of integration and collaborative use of those resources, and identifying potential ways in which IT leaders can help institutions enhance their use.

CI's very nature and substantial resource requirements require outreach beyond the central IT organization; *Higher Education IT and Cyberinfrastructure: Integrating Technologies for Scholarship* shows this to be a work in progress. Most respondents report the use of individual CI technologies, but institutions have yet to interweave them in a unified way that realizes CI's ultimate potential. Respondents tend to rate their overall knowledge of CI technologies high but their resources for meeting their research-related responsibilities low. They give low ratings to research-related collaborative practices at their institutions and to their own organizations' effectiveness at integrating CI technologies.

The study emphasizes, too, the vital roles of the CIO and the central IT organization in integrating CI technologies effectively. The CIO is more successful when he/she has a higher level of knowledge about the research uses of high-performance computing and CI applications and tools; central IT is more successful where it plays an engaged, enabling role, by providing and/or funding researchers' access to CI technologies. Thus knowledge and outreach appear to be the first steps to harnessing CI's full potential across the institution.

METHODOLOGY

Higher Education IT and Cyberinfrastructure: Integrating Technologies for Scholarship used the following research approach:

- A literature search of articles and studies from journalistic, academic, and IT practitioner sources, including publications from the NSF's Blue-Ribbon Advisory Panel on Cyberinfrastructure and its Cyberinfrastructure Council.
- An online survey, designed with input from an advisory committee as well as members of the EDUCAUSE Research Mission Support Constituent Group and the members of the EDUCAUSE Net@EDU Campus Cyberinfrastructure Working Group. Invitations sent to 1,688 EDUCAUSE member institutions in the United States and Canada in late November 2007 resulted in 369 qualified responses (a 21.9% response rate).
- Qualitative interviews with 12 individuals, including higher education CIOs and others, involved with cyberinfrastructure practices and resources at 11 U.S. and Canadian institutions, to gain deeper insights into findings from the quantitative analysis and to capture ideas and viewpoints.

Institutional Use of CI Technologies

About two-thirds of respondents report at least some research use of CI applications and tools, data storage and management resources, and resources for collaboration within virtual communities. By contrast, just under half make any use of high-performance computing resources or advanced network infrastructure for research, perhaps because of the relatively high cost or the complexity of these technologies. Unsurprisingly, most institutions that self-identified as having a research-focused mission report that their researchers use all five CI technologies; researchers at institutions with a teaching mission are much more likely to use only a few of them.

Researchers obtain access to most CI technologies most frequently through the use of their own resources or those of their labs, reinforcing the perception that research IT resources are highly distributed. Central IT is the next most frequently used source for these technologies, though at most institutions central IT is the primary source of advanced network infrastructure.

With the exception of resources for collaboration within virtual communities, most institutions rated CI technologies as being most important to research in science and engineering and of substantially less importance to other academic areas.

CIO's Knowledge and CI Integration

Our research suggests that information is an important driver in achieving CI's full potential at an institution. Effectiveness of integration of CI resources is strongly positively associated with the CIO's overall knowledge about all of our CI technologies except advanced network infrastructure. The greater the role that central IT plays in providing and funding a CI technology, the more knowledgeable about it the CIO is. Inventories of CI technology resources are important information sources; the completeness of the inventory for each CI technology is positively associated with the CIO's reported level of knowledge about it.

The level of knowledge of other campus administrators about certain CI technologies seems to influence central IT's effectiveness at integrating CI resources, too. The knowledge level of the chief academic officer (CAO) is most potent: for all five technologies, where the CAO is more knowledgeable, central IT does a better job, on average, at integrating CI resources. Knowledge levels of the chief research officer and of academic deans in science and engineering are of similar, though more technology-specific, influence. Thus it appears that information sharing among the CIO, the chief research officer, and the science and engineering deans has the potential to facilitate more effective CI utilization campus-wide.

Actions That Enable an Enterprise Approach

Higher Education IT and Cyberinfrastructure: Integrating Technologies for Scholarship highlights items that would most help the central IT organization support more effective research use of each of the five CI technologies at their institutions. From a slate of eight items, majorities of respondents report that increased funding for infrastructure and services, as well as increased communication/outreach between researchers and central IT, would be important aids to improving support. Perhaps as one mechanism to improve the funding supply for central IT's support of research, substantial numbers also identify a greater role for central IT in developing budgets for the institution's research grants and contracts.

RECOMMENDATIONS

Based on its findings in *Higher Education IT and Cyberinfrastructure: Integrating Technologies for Scholarship*, ECAR offers a trio of interrelated recommendations to foster a collaborative and seamless institutional CI environment:

1. Engage the university research leaders and community.

The CIO position is increasingly one that requires outreach, diplomacy, and institutional-level leadership. The CI context seems to be no exception if an institution has as its goal the integration of CI technologies to provide seamless support for research. Our respondent population, however, gave lackluster marks to the collaborative tendencies of researchers, to institutional incentives for collaborative practices, and to their own effectiveness in integrating CI resources. It is clear from our data that CIOs need to exert influence far beyond their own organizations to build an institutional CI platform. As CIOs consider which alliances to build in pursuit of this goal, science and engineering would likely be the best disciplines to start in because of CI's importance to this community and their deans' greater knowledge about key CI resources.

2. Encourage open sharing of information about CI technologies.

An important component of the CIO's engagement should be information sharing. Our research suggests that a more open environment, one in which information about the use of CI technologies is readily shared between the research community and the CIO, is one in which those technologies are used more efficiently and are integrated more effectively. An inventory of CI resources facilitates information sharing, and where the inventory is more complete, the CIO's knowledge is significantly better. Respondents indicated that increased communication and outreach between researchers and central IT would be nearly as beneficial as increased funding for central IT infrastructure and services.

3. Increase central IT involvement in research grants and contracts budgeting.

Funding, too, seems to affect many aspects of CI integration. Respondents cited increased funding for central IT infrastructure and services and a greater role for central IT in developing budgets for the institution's research grants and contracts as potentially beneficial in their support for effective research use at their institutions. Increased central IT involvement in grant and contract budget development also promotes the goals of greater engagement and more open information sharing outlined above.

Finding Synergies between Central IT and Researchers

One of the goals of CI is to create a rich research IT environment that is highly available to researchers across the institution. As we have seen, central IT has a role in providing and funding such resources and services, but the universe of need and the pace of technological change are well beyond the capacities of even the best-funded central IT organization. Achieving economies of scale in the use of specialized research IT resources depends in part on researchers themselves offering some of the capacity of those resources to their colleagues under a variety of collaborative scenarios, and it relies on them as well to work with central IT where appropriate to ensure optimal management of shared or shareable resources.

Nevertheless, our respondents, on average, gave their institutions' researchers mediocre scores for collaboration and gave their institutions somewhat poorer scores for achieving economies of scale and for incenting collaborative behavior. We discovered, however, that some institutions are doing better than others. For example, we found respondents more positive about collaborative practices and incentives to collaboration at institutions where central IT provides and funds various CI technologies and where the CIO is more knowledgeable about them.

These same factors—central IT as provider and funding source for CI technologies, level of knowledge of the CIO and other officers about CI technologies, and the status of CI technology inventories—appear in our findings in relation to respondents' self-assessment of central IT's effectiveness at integrating CI resources for support of research, suggesting that what helps collaboration grow is also helpful in central IT's CI integration efforts. Indeed, we found that, on average, institutions giving higher marks to researchers for collaboration and to the institution for realizing economies of scale, and those agreeing more strongly that incentives for those practices are in place, reported that their central IT organizations were significantly more effective at integrating CI technologies to provide seamless support for research.

Endnote

1. Daniel E. Atkins, et al., *Revolutionizing Science and Engineering through Cyberinfrastructure* (Washington, D. C.: NSF, 2003). http://www.nsf.gov/od/oci/reports/atkins.pdf.