Trends in Technology-Enabled Research and Discovery

Gary K. Allen, CIO, University of Missouri-Columbia; Vice President for IT, University of Missouri System

James Davis, Vice Provost for IT and CIO, Iowa State University

It is frequently observed that technology advances at a rapid rate, perhaps following *Moore's Law* by doubling in capacity every 18-24 months. Information technology service providers, including technology manufacturers and higher education institutions, are leveraging these gains to create new types of services that have a direct and supportive impact on research computing. The IT delivery model is also changing. It's becoming more common for universities to collaborate with peers to develop services that provide an economy of scale and the functionality that meets the specific needs of higher education. Also changing is the way that university research is structured, with increased emphasis on large cross-cutting (and often multi-institutional) thrusts. The concurrent changes in technology and the way it's used leads to a dynamic environment. Higher education technology providers must become agile and collaborative to align with the campus and global research communities. In the remainder of the paper, we'll briefly explore how these emerging trends in research and information technology are giving rise to new opportunities to accelerate campus research programs.

The pursuit to create and share knowledge is the essence of the academy. Discovery, innovation, learning, and engagement are core to the mission of the university, and support for those objectives pervades all processes.

An effective information technology infrastructure (or *cyberinfrastructure*) is central to the success of a robust research program. The infrastructure must provide anywhere, anytime access to information, peer collaborators, systems, and services needed to advance the program. When cyberinfrastructure is congruent with the needs of the research program, the ensuing interdependence creates significant synergy that acts as a "force multiplier" to propel research activities forward. Although beyond the scope of this paper, we note that aligning IT with the missions of the university has a similar beneficial effect on technology-enabled learning, the ability to base decisions in institutional information through multiple analytics, and efficient engagement of broad constituent groups. It also contributes to a forward looking environment that assists with recruiting and retaining students and faculty. In a very real way, campus information technology is a strategic asset of the university.

Building an effective cyberinfrastructure to support discovery and innovation is a national priority. Campus technology and research visionaries

have been working with the National Science Foundation (NSF) Office of Cyberinfrastructure the past few years to develop strategies for advancing campus research in a comprehensive and coordinated way. The National Science Foundation Advisory Committee for Cyberinfrastructure Task Force on Campus Bridging offered insightful recommendations for research universities and federal funding agencies in their 2011 final report[1]. One recommendation calls for a "healthy national cyberinfrastructure ecosystem" and a rethinking of some of the barriers to large-scale collaboration. We end this paper with a call for a regional approach to nurturing the campus dialog that brings national cyberinfrastructure efforts to researchers in a consequential way.

Two Research Trends Impacting Cyberinfrastructure

"Big Data Science"

The escalating capacity and reduced cost across almost all components of research cyberinfrastructure is encouraging the expansion of computational models to yield results with improved fidelity. As the problem size scales up, so does the demand for computing resources. Large-scale high performance computing clusters, once used by just a handful of disciplines for specialized problems, are now an essential tool in any area where timely results are important. Large computational problems nearly always consume and produce significant collections of unstructured data that must be stored and transmitted. Additionally, several federal funding agencies now require that project data be retained and made available to

the public so that the reported results can be validated and used for follow-on research (for example, see National Science Foundation data management plan guidelines[2]). Even with the rapidly increasing capacity of contemporary storage, the management of large collections of data is demanding. New storage architectures have been developed to implement hierarchies with integrated networking to balance performance and cost, yet it remains challenging to perform even basic operations such as searching large data stores or archiving collections. The complexity of maintaining large data stores coupled with curation requirements and rapidly expanding security requirements (e.g., FISMA compliance[3]) makes a compelling case for developing an institutional approach to data management.

Another challenging component of processing large research data sets is simply moving them quickly and reliably, for example between a laboratory instrument and a HPC cluster. Even with contemporary 10-gigabit or the emerging 100-gigabit network connections, data transfers can often be measured in hours. This is exacerbated when large data sets are moved between institutions, sometimes at "Internet speed". Additional barriers pertaining to format and data compatibility can arise.

Interinstitutional Research Collaboration

As we recognize the need to substantially improve the capacity of research cyberinfrastructure, we also note that a slow shift in the way scientists collaborate brings new requirements to create "virtual communities" for partici-

pants to gather (virtually) to share ideas, information, systems, and services. While it was more common in the past for grants to have one or two investigators, there has been a steady growth in large multidisciplinary grants with teams that span multiple universities, centers, and corporations. In one approximation for the degree of interinstitutional collaboration, a 2012 National Science Foundation survey of R&D expenditures at universities and colleges[4] noted that the "pass-through" funds represented 7% of the total academic R&D expenditures in FY 2009 as compared to 5% in FY 2000. Project teams expect that technology will mitigate the impact of distance and create a community where participants can interact as if they were located in the same space.

These two aspects taken together have implications for research cyberinfrastructure:

- An increased demand for HPC resources in disciplines that have not in the past been notable consumers of those services.
- A pressing need for significantly larger compute clusters capable of scaling up problems to enable new discoveries.
- The need for specialized one-off computing resources such as GPU-based HPC clusters.
- The ability to transmit, store, and share massive data collections, while addressing cost, security, curation, and backup. New business models must be developed to support the management of data beyond the duration of a research grant, in partnership with institutional libraries.

- Seamless support for collaboration, including video conferencing and shared data, systems, and documents. Services need to be layered on a commonly used identity management paradigm that supports federated identities, such as the Internet2 InCommon suite.

Implicit too is establishing a collaborative relationship between campus technology providers, institutional leadership for research, and research centers and their faculty and staff. As previously noted, the case for providing an enabling research technology infrastructure is clear, but the issues surrounding complexity and cost require an institutional approach to obtain an effective outcome, given the rapid changes in technology and the evolving way that campus research is carried out.

Information Technology Trends Impacting Research Cyberinfrastructure

Market forces have created demand for a new operational model within IT structures. These forces include the economy, price of utilities, personnel costs and IT complexity, among others. Organizations can no longer afford to manage everything on their own. IT is in the largest outsourcing trend in history, and public cloud Infrastructure as a Service (IaaS) will be a key component of outsourcing decisions because it offers commoditized infrastructure and increased agility. This allows IT to focus on core business outcomes. In reaction to these forces, public cloud IaaS providers have developed services that aim to solve many of the business issues IT organizations face[5].

In higher education institutions, these trends are clearly evident in areas of administrative computing and academic technology services supporting teaching and learning. In these areas, migrations toward outsourced or shared solutions are well underway. A different rate of evolution is evident in IT services supporting research activities; these frequently require architectures and applications that are not "commoditized" in the sense that the research goals are sometimes optimally met using non-standard IT toolkits not readily available in the marketplace. Historically, there has not been a sufficient market demand for some of the commonly needed research IT workflows to merit investment by for-profit providers. Additionally, the information security and data management issues at play with research-related IT often requires approaches far different from other types of information systems from research universities. This includes issues such as intellectual property protection (e.g., for investigators, or corporate sponsors), data classification (e.g., for certain types of federally funded research), and data curation (e.g., for mandated public dissemination of research results).

Given these additional concerns and demands for management of research IT infrastructure and data, new models are rapidly developing among communities of use whose members have these requirements in common. Research consortia for IT infrastructure and services are now developing in ways reminiscent of similar approaches used by academic libraries developed and deployed over the past several decades. Significant cost and complexity drivers for these research tools add significant momentum for this evolution, along with the heightening of interdisciplinary research approaches that blend data systems and tools that provide new perspectives into research questions.

Joint efforts, many of which are regionally focused, are also growing in number as institutions work together to develop new approaches to satisfy their research cyberinfrastructure needs. Examples include: the Renaissance Computing Institute (RENCI)[6], Rocky Mountain Supercomputing Center (RMSC)[7], University of Chicago Computing Institute (UCCI)[8], Computational Center for Nanotechnology Innovations (CCNI)[9], and the Victoria Life Sciences Computation Initiative (VLSCI)[10]. In many cases, these efforts include both research universities and corporate partners. They are frequently specifically targeted at not only enhancing research capacity, but also providing nuclei of cyberinfrastructure equipment and expertise designed to foster economic development efforts.

Federal agencies and higher education organizations such as Internet2 are directly supporting these approaches. The National Science Foundation has long been an active supporter of cyberinfrastructure enhancement efforts across the U.S. The Global Environment for Network Innovation (GENI)[11] Project was established in 2007 by the NSF. The goal was to provide an environment for the development of novel networking approaches and applications that obviate the constraints of the current internet and wireless network environment such as might be observed when trying to

move very large datasets at very high speeds. As indicated above, the research computing needs of higher education increasingly demand the ability to move very large volumes of data at very high speeds; such ability is limited on today's established networks. Accordingly, GENI targets promoting networking innovations that utilize existing network infrastructure in ways that are increasingly scalable, flexible (in terms of routing and quality of service), and secure. Significant efforts by GENI program participants relate to software and networking protocols supporting "software-defined networking" as exemplified by OpenFlow. OpenFlow creates virtual network "slices" that operate on existing network infrastructure and share that infrastructure with other networking protocols. Numerous universities are experimenting with the development and deployment of OpenFlow, and both Internet2 and National LambdaRail are implementing OpenFlow in their networks. In 2011, the Open Networking Foundation (ONF)[12] was created as a not-for-profit entity focused on further development and standardization of software-defined networking. ONF is currently supported by more than 70 global companies, and is accelerating the evolution of these new approaches to networking.

The recently announced Internet2 Net+ services and Innovation Platform initiatives are providing both middleware and end-user services targeted specifically at higher education research activities, offered in ways that are congruent with policy frameworks required by universities, and that meet regulatory and compliance concerns. These offerings include extremely high-bandwidth network connectivity, dynamic configuration of networking protocols, data security, and a new generation of monitoring and management tools. For example, Internet2's Advanced Layer 2 Service is a reliable Layer 2 transport environment that provides a flexible end-to-end, high-bandwidth and deeply programmable environment. The Advanced Layer 2 Service builds on Internet2's NDDI/OS3E initiative to provide network connections that can be used to easily create VLANs with a range of characteristics, with reachability through the network, regional networks, campuses, and partners like ESSnet, GEANT and other inter-domain enabled global networks. (see the Internet2 Innovation Platform[13]).

Call to Action

Discovery and innovation are core to the academy mission. Information technology can enable remarkable research when there is alignment between the IT resources available (the cyberinfrastructure) and the needs of the research enterprise. Alignment is challenged by rapidly changing trends in the information technology field, by dramatically increasing computational needs, and also by a shift toward large interinstitutional grants where collaborators work at a distance. In recognition of the need for institutions to collaborate to achieve the scale required to address these issues, several national and international initiatives have formed around the goal of improving campus cyberinfrastructure.

One of the most challenging aspects of determining what would work best locally is engaging campus leadership and key research faculty in explorations of adapting consortia services and experiences to a local context. Given the long standing collaboration among the "four corners" universities at the Merrill workshops, we believe that there's an important role for a regional approach to developing strategies to bridge between the campus and national research cyberinfrastructure initiatives. Additionally, all of the represented states in the Merrill workshops are EPSCoR-eligible; this could represent a significant opportunity to obtain federal funding support (e.g., in the form of NSF EPSCoR Track 2 grants) to begin creating a regional support model for cyberinfrastructure. It is proposed that follow-on meetings with the chief research officers, chief academic officers, and chief information officers of the universities that participate in the Merrill workshop explore this further.

References
1. National Science Committee Advisory Committee for Cyberinfrastructure Task Force on Campus Bridging. Final Report. March 2011. Available from: http://www.nsf.gov/od/oci/taskforces/TaskForceReport_CampusBridging.pdf
2. National Science Foundation data management plans. www.nsf.gov/eng/general/dmp.jsp
3. National Institute of Standards and Technology Federal Information Security Management Act (FISMA). http://csrc.nist.gov/groups/SMA/fisma/index.html
4. Hale, Katherine. "Collaboration in Academic R&D: A Decade of Growth in Pass-Through Funding" NSF 15-325, August 2012. http://www.nsf.gov/statistics/infbrief/nsf12325/
5. Hilgendorf, Kyle. Market Profile: Public Cloud IaaS, Gartner, June 12, 2012
6. Renaissance Computing Institute (RENCI). http://www.renci.org
7. Rocky Mountain Supercomputing Center (RSMC). http://www.rmscinc.org
8. University of Chicago Computing Institute (UCCI). http://www.ci.uchicago.edu
9. Computational Center for Nanotechnology Innovations (CCNI). http://www.rpi.edu/research/ccni
10. Victoria Life Sciences Computation Initiative (VLSCI). http://www.vlsci.org.au
11. Grochow, Jerrold M. "GENI Workshop on Layer 2/SDN Campus Deployment: A Report of the Meeting (July 2011)". Retrieved 9/4/2012 from http://www.educause.edu/library/resources/geni-workshop-layer-2sdn-campus-deployment-report-meeting-july-2011
12. Open Networking Foundation (ONF). https://opennetworking.org
13. Internet2 Innovation Platform FAQ. http://www.internet2.edu/pubs/Internet2-Innovation-Platform-FAQ.pdf