UC SANTA BARBARA

Office of the Chief Information Officer
UCSB Cyberinfrastructure Committee
Santa Barbara, CA 93106
cio.ucsb.edu

UCSB Cross-Campus Perspective on Future Needs for Advanced Cyberinfrastructure to Support Science & Engineering Research

UCSB Cyberinfrastructure Committee, April 2017

Research domain(s), discipline(s)/sub-discipline(s) of the author(s).

Paul Atzberger, Mathematics
Doug Drury, Office of the CIO
Jim Frew, Bren School of Environmental Science & Management
Kirk Grier, Enterprise Technology Services
Matt Hall, Associate Vice Chancellor - IT and CIO
Richard Kip, Letters and Science IT
Miriam Metzger, Communication
Linda Petzold, Computer Science and Mechanical Engineering
Fuzzy Rogers, Materials
Tim Sherwood, Associate Vice Chancellor for Research
David Stuart, Physics
Chris Van de Walle, Materials
Cristina Venegas, Film and Media Studies
Paul Weakliem, California NanoSystems Institute
Rich Wolski, Computer Science

An abstract (200 words) summarizing the response.

Cyberinfrastructure for interdisciplinary science and engineering requires additional, sustainable investment in computing (including both the ability to perform computations at scale with minimal administrative overhead but also in terms of quickly and efficiently performing computation on one's desktop or shared with student researchers), networking (in

securely, efficiently, and economically managing data and computation with collaborators and the public around the globe), and in data management and storage (including viable models for long term data preservation, scalable resources for data analysis, and data sharing/editing with revision control). These imperatives must be achieved while simultaneously shifting our thinking from a facility-oriented model to the service-oriented model to assure broad access by researchers across disciplines and with different computing skill levels and backgrounds.

Question 1 (maximum 1200 words) - Research Challenge(s).

Describe current or emerging science or engineering research challenge(s), providing context in terms of recent research activities and standing questions in the field.

Rather than provide perspective from a specific science or engineering domain, the notes below stem from a cross-campus discussion on future cyberinfrastructure needs with an understanding that scientific advances are increasingly driven by interdisciplinary teams both large and small. Our cyberinfrastructure committee pulls together faculty and staff from across the campus research enterprise (including but not limited to physical sciences, social sciences, humanities, and engineering) to help shape the strategic cyberinfrastructure for the campus. Interdisciplinary teams bring a powerfully diverse set of perspectives on science and engineering problems, and they generate an intense demand for tools enabling collaboration, a broader set of constraints, expectations, and abilities. However, our group's discussion clarified that there are ongoing challenges for computing (including both the ability to perform computations at scale with minimal administrative overhead; but also in terms of quickly and efficiently performing computation on one's desktop or shared with student researchers), networking (in securely, efficiently, and economically managing data and computation with collaborators and the public around the globe), and in data management and storage (including immediate storage from scientific instruments, viable models for long term data preservation, scalable resources for data analysis, and data sharing/editing with revision control). In addition, two very important cross-cutting themes emerged: an important change in thinking from "facility" to "service" and the need to ensure broad access to resources by researchers with backgrounds spanning the entire multi-disciplinary gamut of campus research efforts.

Question 2 (maximum 1200 words) - Cyberinfrastructure Needed to Address the Research Challenge(s).

Describe any limitations or absence of existing cyberinfrastructure, and/or specific technical advancements in cyberinfrastructure (e.g. advanced computing, data infrastructure, software infrastructure, applications, networking, cybersecurity), that must be addressed to accomplish the identified research challenge(s).

Networking: Recent dramatic improvements in large-scale data processing capabilities drive an ever-increasing need among researchers to share large and small data sets. Simply put, because it is possible to process ever-larger data sets, researchers make new discoveries by combining data that previously could not be aggregated. Thus, for large research tasks, it is increasingly the case that data are not local to a single user, department, campus, or even continent. Computation is often shared between campus clusters, cloud services, national science infrastructure, and large international collaborations. The ease with which data of significant scale can be migrated and safely shared, and the trust researchers have in the security of their network attached systems, including the data integrity of those systems, both directly impact not just specific existing research projects but researcher creativity in conceiving new projects. The advantage can sometimes be hard to quantify a priori, as it has a feedback effect on behavior. This holds true for both internal wired and wireless campus networking and the broader connections to the internet or private education, government, or commercial networks.

Computing: While some felt that large-scale computing was not a limiting resource for/in their work given current available resources, other investigators held the opposite feeling. Independent of this perceived need in the total amount of computing resources, there was the consensus that the difficulty of working with computing resources remained high for many and that the demands will only grow. There is a need to support both a hierarchy of computing resources (from a single node to small cluster, to large shared parallel systems, to off-campus supercomputing) and to make those resources readily available and easily accessible to all researchers whether they be faculty, staff, graduate students, or undergraduates. Campus computing clusters, designated spaces for computing and research, and easy access to expertise in the domain help lower the barrier to entry where time, training, and cost are limiting factors. While campus-level shared facilities are critical to the research community, the maintenance, logistics, security, and environmental factors are costly. There is currently a lack of clarity on best practices for funding models that are both sustainable and mutually appreciated by both NSF and campus especially as it relates to the difference between capital expense versus ongoing operational expense.

Storage and Data Management: The bulk of our group discussion happened around the challenges surrounding the needs to store, retrieve, and share research data. The deployment of large numbers of inexpensive sensors (e.g. underwater video), the digitization of traditionally physical-only artifacts (e.g. the cylinder audio collection), the growing data provenance requirements on scientific data (e.g. climate data), all coupled with the powerful new techniques for data analysis enabled by both large data sets and machine learning place storage acquisition, maintenance, and ongoing operations as a major concern for researchers across all disciplines. The need for storage space to hold important research data and to ensure that the data are retrievable in various states or at various points in the lifecycle of the research process is critical to the success of the research. A particular challenge around storage is unique disciplinary requirements. While some require high bandwidth streaming read access (e.g. video archive), others are write-mostly (e.g. sensor capture) or require truly random access (e.g. key-value stores). For some, flat storage is all that is required, while others need a complete file system abstraction, and still, others require extensive curation and metadata. Cyberinfrastructure does not currently exist that allows a researcher to easily and consistently extend the available storage and retrieval capabilities to meet the needs of research data collection as collection techniques and technologies emerge and advance over the course of the research activity. Assuming that sufficient storage/retrieval resources are available, management and curation of the stored data are an additional challenge to research. While digital libraries exist and provide data curation for a subset of the collected research data, significant research data reside outside of curated environments. Research data can, and should, be coupled to service. A repository service may support versioning, multiple access methods (file system, web services, etc.), varying governance models (who can modify/access the repositories, who pays, how often, etc.), and auditing/provenance, but where other services (preservation, working space, etc.) provide complementary functionality. Most important, the complete lifecycle -- from creation, through collaboration, into peer review, followed by dissemination, archiving, and ultimately into the creation of new data and research across different research needs should be considered.

Question 3 (maximum 1200 words, optional) - Other considerations.

Any other aspects, such as organization, process, learning and workforce development, access, and sustainability, that need to be addressed; or any other issues that NSF should consider.

As mentioned in our response to Question 1, two critical cross-cutting themes emerged. 1) The need to shift thinking by both PIs and agencies to a service rather than a facility model of infrastructure (where those services would include a mix of on campus, academic off-campus, and commercial services) and 2) The need to ensure that the power of this infrastructure is broadly available to researchers with a highly diverse set of backgrounds, skills, and attempting to answer as broad a set of research questions as possible.

When we discuss accessibility, we mean not simply by policy, but accessible in practice. Among the current barriers for entry include a lack of clarity among PIs about whom they should contact to discuss research infrastructure, an insufficient understanding of the numerous resources already available on campus, and a dearth of knowledge on the capabilities of modern cyberinfrastructure to address research needs. Building a culture of interdisciplinary collaboration can be facilitated by cyberinfrastructure by providing a point of common interaction across fields if support to help grow that culture existed. Common tools and community building support to help connect researchers with infrastructure and on-campus expertise (where one's natural intellectual home is) would then be a gateway to then further inter-organizational collaboration and resource sharing. The community cannot grow from a single person but rather requires a continued engagement over time from commonly incentivized groups of researchers (students, faculty, and staff).

Finally, two recent technological advances are changing the Cyberinfrastructure needs that UCSB researchers currently experience, and will continue to experience into 2030. The first is a switch from "facility centric" to "service-centric" approaches to the use of computing capabilities. The second is the emergence of "converged" architectures that bundle computing, storage, networking, and security into converged modular units of functionality. Thus researchers now think in terms of services that meet their computational needs (and not the facilities that provide these services) and any computational capability comprises a mixture of computing, storage, networking, and security components indivisibly. A sub challenge remains the consistency of well established best practices for security policies for research services. If we move to shared, service-oriented models, there would be a way for remediation to be implemented (e.g. CVEs and patch sets) while establishing a common language for discussing research data sensitivity (e.g. student data is typically more sensitive than archived simulation runs although not always). Increasing the accessibility, availability, and quantity of diverse computational services paired with local expertise abstracted by cyberinfrastructure will be a necessity to facilitate future ground-breaking research.