Campus Bridging Report
Relation to CI proposals
Jim Pepin
CTO
Clemson University
Campus Bridging ACCI Task Force Report

• ACCI task forces
  – Campus Bridging
  – Cyber learning and Workforce Development
  – Data and Visualization
  – Grand Challenges
  – High Performance Computing
  – Software for Science and Engineering

• 3 workshops for campus bridging

• Link to report
    TaskForceReport_CampusBridging.pdf
The charge of the Task Force on Campus Bridging is to address the broad issues involving improved campus interactions with cyberinfrastructure, broadly construed. It will include a number of different types of bridging:

- Campus grids to national infrastructure (both compute and data-oriented approaches) and international CI;
- Campus networks to state, regional, and national;
- Departmental cluster to campus HPC infrastructure; and
- Campus-to-campus and campus-to state/regional resources.
Finding 1

The diversity in the US cyberinfrastructure environment creates tremendous opportunities for US science and engineering research, but adds new types of complexity and new challenges in campus bridging. The cyberinfrastructure environment in the US is now much more complex and varied than the long-useful Branscomb Pyramid. As regards computational facilities, this is largely due to continued improvements in processing power per unit of money and changes in CPU architecture, continued development of volunteer computing systems, and evolution of commercial Infrastructure/Platform/Software as a Service (cloud) facilities. Data management and access facilities and user communities are also increasingly complex, and not necessarily well described by a pyramid.
Finding 2

The reward system as perceived by individual faculty researchers in science and engineering does not support the development of a coordinated national cyberinfrastructure. It encourages a highly diffuse, uncoordinated cyberinfrastructure that makes sharing and collective investment difficult and does not optimize the effectiveness of cyberinfrastructure support for research and development in science and engineering in the United States. In particular, the current reward structure does not align rewards to faculty with a focus on collaboration in ways that support NSF’s stated views on Virtual Organizations as an essential organizational structure in scientific and engineering research.
Finding 3

The current state of cyberinfrastructure software and current levels of expert support for use of cyberinfrastructure create barriers in use of the many and varied campus and national cyberinfrastructure facilities. These barriers prevent the US open science and engineering research community from using the existing, open US cyberinfrastructure as effectively and efficiently as possible.
Findings 4

The existing, aggregate, national cyberinfrastructure is not adequate to meet current or future needs of the US open science and engineering research community
Finding 5

A healthy national cyberinfrastructure ecosystem is essential to US science and engineering research and to US global competitiveness in science and technology. Federal R&D funding overall is not sufficient to meet those needs, and the NSF share of this funding is not sufficient to meet even the needs of basic research in those disciplines that the NSF supports.
Finding 6

Data volumes produced by most new research instrumentation, including that installed at the campus lab level, cannot be supported by most current campus, regional, and national networking facilities. There is a critical need to restructure and upgrade local campus networks to meet these demands.
Strategic Recommendation to NSF 1

As part of a strategy of coherence between the NSF and campus cyberinfrastructure and reducing reimplementations of multiple authentication systems, the NSF should encourage the use of the InCommon Federation global federated system by using it in the services it deploys and supports, unless there are specific technical or risk management barriers.
The NSF must lead the community in establishing a blueprint for a National Cyberinfrastructure. Components of this leadership should include the following strategic approaches to funding cyberinfrastructure:

• **When funding cyberinfrastructure projects that are intended to function as infrastructure, the NSF should use the review criteria and approaches that are generally used for research infrastructure rather than the criteria used for scientific discovery awards.** Such awards should be made in ways that complement existing infrastructure and align with best practices, appropriate international standards, and the NSF vision and plans for CIF21.

• **The NSF should establish a national cyberinfrastructure software roadmap.** Through the Software Infrastructure for Sustained Innovation (SI2) or other programs, the NSF should seek to systematically fund the creation and ongoing development and support of a suite of critical cyberinfrastructure software that identifies and establishes this roadmap, including cyberinfrastructure software for authentication and access control; computing cluster management; data movement; data sharing; data, metadata, and provenance management; distributed computation / cycle scavenging; parallel computing libraries; network performance analysis / debugging; VO collaboration; and scientific visualization. Funding for personnel should be a strong portion of such a strategy.

• **The NSF should continue to invest in campus cyberinfrastructure through programs such as the Major Research Infrastructure (MRI) program, and do so in ways that achieve goals set in the Cyberinfrastructure Vision for 21st Century Discovery and a national cyberinfrastructure software roadmap.**
Strategic Recommendation to NSF 3

• The NSF should create a new program funding high-speed (currently 10 Gbps) connections from campuses to the nearest landing point for a national network backbone. The design of these connections must include support for dynamic network provisioning services and must be engineered to support rapid movement of large scientific data sets.
Strategic Recommendation to NSF 4

- The NSF should fund national facilities for at least short-term storage and management of data to support collaboration, scientific workflows, and remote visualization; management tools should include support for provenance and metadata. As a complement to these facilities and in coordination with the work in Recommendation #3, the NSF should also fund the development of services for bulk movement of scientific data and for high-speed access to distributed data stores. Additionally, efforts in this area should be closely coordinated with emerging campus-level data management investments.
Strategic Recommendation to NSF 5

• The NSF should continue research, development, and delivery of new networking technologies. Research priorities funded by the NSF should include data intensive networks, sensor nets, networking in support of cyberphysical systems, geographically distributed file systems, and technologies to support long distance and international networking.
Strategic Recommendation to NSF 6

- The NSF should fund activities that support the evolution and maturation of cyberinfrastructure through careful analyses of needs (in advance of creating new cyberinfrastructure facilities) and outcomes (during and after the use of cyberinfrastructure facilities). The NSF should establish and fund processes for collecting disciplinary community requirements and planning long-term cyberinfrastructure software roadmaps to support disciplinary community research objectives. The NSF should likewise fund studies of cyberinfrastructure experiences to identify attributes leading to impact, and recommend a set of metrics for the development, deployment, and operation of cyberinfrastructure, including a set of guidelines for how the community should judge cyberinfrastructure technologies in terms of their technology readiness. All NSF-funded cyberinfrastructure implementations should include analysis of effectiveness including formal user surveys. All studies of cyberinfrastructure needs and outcomes, including ongoing studies of existing cyberinfrastructure facilities, should be published in the open, refereed, scholarly literature.
Tactical Recommendation to NSF

• Tactical Recommendation to the NSF #1: The NSF should fund the TeraGrid eXtreme Digital program, as currently called for in existing solicitations, and should continue to fund and invest in the Open Science Grid.

• Tactical recommendation to the NSF #2: The NSF should commission a study of current reward structures and recommendations about the reward structure – particularly as regards promotion and tenure for faculty – that would better align reward structures as perceived by individual faculty members with the type of large, collaborative virtual organizations that the NSF asserts are required for successful approaches to pressing, large scale scientific problems and transformative research.

• Tactical Recommendation to the NSF #3: The NSF should support joint efforts with organizations such as the Association for Computing Machinery (ACM), the IEEE Computer Society, and/or Computing Research Association (CRA), to develop and maintain curriculum materials for undergraduate education in computer science and computational and data-driven science and engineering.
Strategic Recommendation to Higher ED

• Strategic Recommendation to university leaders and the US higher education community #1: Institutions of higher education should lead efforts to fund and invest in university-specific, state-centric, and regional cyberinfrastructure – including human resources to support use of cyberinfrastructure – in order to create local benefits in research accomplishments and economic development and to aid the global competitiveness of the US and thus the long-term welfare of US citizens.

• Strategic Recommendation to university leaders and the US higher education community #2: Every institution of higher education should have a strategic plan, developed and endorsed at the highest level of its governance, for the establishment of a coherent cyberinfrastructure. Such a plan should have as one of its features a strategy for maximizing effective utilization of the institution’s aggregate research cyberinfrastructure and minimizing impact on the global environment. Such a plan should also include ongoing funding for staff to support implementation and use of cyberinfrastructure hardware and software.

• Strategic Recommendation to university leaders and the US higher education community #3: Institutions of higher education should adopt criteria for tenure and promotion that reward the range of contributions involved in the production of digital artifacts of scholarship. Such artifacts include widely used data sets, scholarly services delivered online, and software (including robust widely useable cyberinfrastructure software and other forms of academic contributions). Such an effort must include creation of new ways to provide peer review of these other, newer types of contributions.
**Tactical Recommendation to Higher ED**

- **Tactical recommendation to university leaders and the US higher education community #1**: Institutions of higher education should continue to press publishers to adopt a strategy of enabling multiple ‘primary authors’ on research papers particularly so that computer, computational, and informatics scholars can contribute to larger collaborative projects while still being rewarded as primary authors.

- **Tactical recommendation to university leaders and the US higher education community #2**: US colleges and universities should systematically consider inclusion of some costs for research cyberinfrastructure in negotiation of facilities and administration rates. When this is done, the best use of facilities and administration income associated with grant awards to universities will be to use it strategically within the context of a campus cyberinfrastructure plan.