Submission in Response to NSF CI 2030 Request for Information

DATE AND TIME: 2017-04-05 14:52:35
REFERENCE NO: 254

Author Names & Affiliations

- Ben Rogers - Director of Research Computing - University of Iowa

Contact Email Address (for NSF use only)

(Hidden)

Research Domain, discipline, and sub-discipline

Academic Scientific Computation/HPC and Research Computing

Title of Submission

A Campus Cyberinfrastructure Provider's Perspective on Future Challenges and Opportunities

Abstract (maximum ~200 words).

The three topics below represent the areas of greatest challenge and opportunity from my perspective as an individual leading research cyberinfrastructure efforts at a comprehensive, research intensive university.

• Developing and ensuring a stable and rewarding career path for the professional workforce that is involved in all facets of cyberinfrastructure is critical. ?
• Developing sustainable funding models is critical as all disciplines increasingly rely on broader and deeper cyberinfrastructure to advance science.
• Support of a computing ecosystem that scales from the individual PI to leadership class facilities is key to continued success.

Question 1 Research Challenge(s) (maximum ~1200 words): Describe current or emerging science or engineering research challenge(s), providing context in terms of recent research activities and standing questions in the field.

Developing and ensuring a stable and rewarding career path for the professional workforce that is involved in all facets of cyberinfrastructure is critical. ?
Research cyberinfrastructure depends heavily on professional staff. The roles of these staff are increasingly diverse and include individuals ranging from research systems and storage administrators, to software developers, and research facilitators/consultants/trainers. A successful cyberinfrastructure support organization needs the capability to run hardware, maintain a full and custom software stack, and educate and support users on how to utilize technology tools requiring expertise across all of these areas. Supporting the continued development of this workforce is paramount to the success of future cyberinfrastructure initiatives. This should include educational and training programs, development of career paths, a strong professional community, and mechanisms to ensure continuous periods of stable

funding and employment for these individuals.

Developing sustainable funding models is critical as all disciplines increasingly rely on broader and deeper cyberinfrastructure to advance science.
In a world where more disciplines require more cyberinfrastructure and IT is transitioning away from traditional capital heavy expenditures the current research funding mechanisms for cyberinfrastructure are being called into question. Continued work is needed to develop ways to sustainably fund cyberinfrastructure that are efficient, effective, and provide broad choice to science communities all while incentivizing best practices in areas such as data and code management.
A national conversation surrounding funding models and barriers may be helpful in addressing challenges in areas such as:
• The preference for equipment purchase over services due to lack of overhead on equipment and that equipment has the potential to run beyond the life of a grant sustaining research programs through dry spells in funding.
• How to better accommodate and incentivize support for professional staff in grants and research contracts.
• Long term maintenance of research software.
• Data management including long-term preservation and distribution.
• Leveraging cloud services to accelerate science while mitigating vendor lock in.
Additionally a national conversation is needed to help determine the baseline level of cyberinfrastructure that research institutions provision versus what should be funded through research grants.

Support of a computing ecosystem that scales from the individual PI to leadership class facilities is key to continued success.
Support for a national cyberinfrastructure of greater capacity than available on campuses remains critical. This national infrastructure should have significant capacity for medium scale computational workloads in addition to meeting the needs of large-scale computational scientists.
Continued engagement from the national facilities with campuses is very valuable and should continue to be encouraged through incentivizing continued community building for programs such as Xsede’s Campus Champions.
Innovative environments and platforms should play a role in national resources. However I have personally not seen strong success yet on our campus in this area with national resources beyond innovations that primarily require scaling. Study of the balance between capacity/production operations and innovation at national level may be a useful area of inquiry.
Collaboration with other federal agencies should be strongly encouraged where partnership opportunities exist. Two possible examples are with DOE for leadership class computing capabilities and with NIH for support of biomedical computing needs.

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

As a cyberinfrastructure provider it is our goal to be responsive to a variety of Science disciplines as opposed to driving specific technologies forward. However, four broad areas are worth noting.
• We are continuing down a path of greater hardware and software diversity. This leads to continued increases in software stack complexity. Tools, training, and methods that help to decrease this complexity both for science users and for those administering and supporting cyberinfrastructure will be broadly helpful in efficiently realizing the potential of new technologies.
• Data is now at least as important as computation. The technology and tools surrounding computational environments are more mature than those related to data and the most frequent and hardest problems in our campus cyberinfrastructure often involve data and storage.
• Machine learning and artificial intelligence techniques are experiencing a renaissance and there is broad interest across disciplines in how to leverage these techniques. Supporting these techniques requires different computational infrastructure than has generally been deployed. Evolving cyberinfrastructure at both the campus and national level to accommodate structured and semi-structured data as well as compatible computational platforms will be important to supporting this change in direction.
• Instruments continue to become more intertwined with research cyberinfrastructure. Instrumentation from sensor networks to microscopes to gene sequencers are generating ever greater amounts of data and are requiring more complex workflows and data integrations to be successful deployed. Campus cyberinfrastructure providers are called upon to support a broad array of such devices and more focus on the development of tools to support such environments would be of great value.

Consent Statement