The future of the TeraGrid and the national advanced cyberinfrastructure: a perspective from the Coalition for Academic Scientific Computation
11 January 2008

The Coalition for Academic Scientific Computation (CASC – http://www.casc.org/) is an educational nonprofit 501(c)(3) organization with more than 50 member institutions. CASC members represent many of the nation's most forward-thinking universities and computing centers. CASC is dedicated to advocating the use of the most advanced computing technology to accelerate scientific discovery for national competitiveness, global security, and economic success, as well as to develop a diverse and well-prepared 21st century workforce. All of the current TeraGrid Resource Partners are also CASC members. CASC is commenting as an organization primarily from the standpoint of the larger number of CASC member sites who are not currently funded by the NSF to provide hardware resources via the TeraGrid. CASC herein presents four recommendations regarding steps that should be implemented so that the TeraGrid and its successors may best serve the NSF mission and US national interests.

CASC presumes that the TeraGrid and its successors will include many of the largest supercomputers available in the US to support the National Science Foundation's goals of "discovery, learning, research infrastructure and stewardship" (http://www.nsf.gov/about/glance.jsp). Dr. Daniel Atkins presented a talk describing the NSF High Performance Computing strategy at the TeraGrid '06 conference. In that talk Dr. Atkins presented an updated and annotated version of the Branscomb Pyramid:
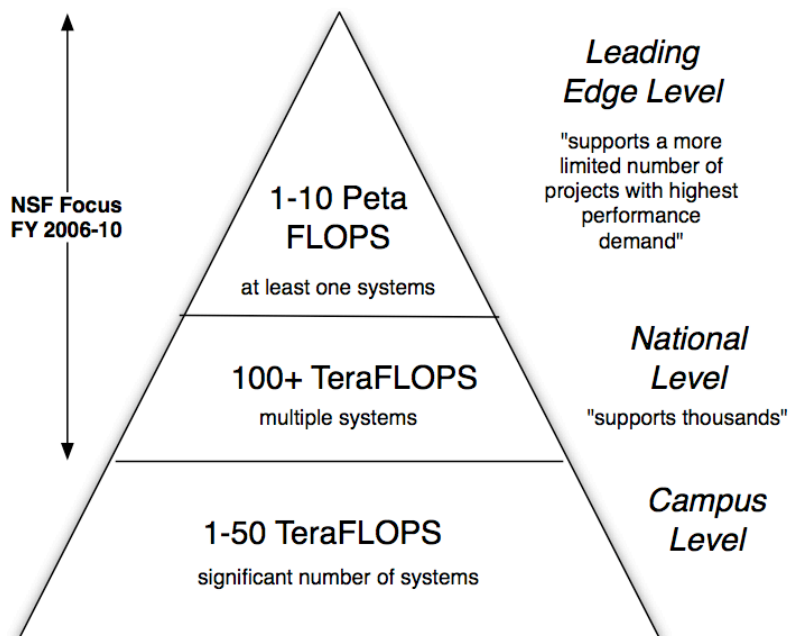

Figure 1. Diagram of distribution of systems, system size, and number of systems from presentation by Dr. Daniel Atkins at the TeraGrid '06 Conference ("NSF and Cyberinfrastructure: Background, Vision, and Activities" presented at the TeraGrid '06 conference, Indianapolis IN 12-15 June 2006. http://www.teragrid.org/library/TeraGrid-06_Atkins.pdf).

The resources provided to the nation by the TeraGrid scale up to the highest levels available supporting unclassified research in the US. At present, the TeraGrid includes an aggregate of more than 750 TFLOPS of processing capability with several thousand individual authorized users. The other resources managed by CASC member sites serve tens of thousands of individual users. CASC members currently manage advanced HPC systems with an aggregate of more than 380 TFLOPS exclusive of those integrated within the TeraGrid, the majority of which fall into the "campus level" of 1-50 TFLOPS.

In our view, the TeraGrid support staff provide excellent technical service (recognizing that this is a statement made by CASC about staff of CASC member sites). However, the nontechnical obstacles to adoption of the TeraGrid are significant. "Quick" turnaround on small requests constitutes at present roughly a month, and once authorization has been granted scientists must first adopt use of a new set of credentials specific to the TeraGrid. Approval of larger requests can easily take a quarter of a year or longer. This is not the pace of 21st century innovation!

Recommendation: The allocation process for the TeraGrid and its successors should be modified to respond much more rapidly to the needs of US academic researchers working in areas related to the NSF mission.

As stated on the TeraGrid Future Web site (http://teragridfuture.org/) the TeraGrid is an important part – but far from all – of the cyberinfrastructure funded by the NSF. Many researchers use HPC resources at the campus or state center level and migrate to the TeraGrid or other national resources when the scale of their needs merit or require. CASC believes that the US national interest, global competitiveness, and the discovery and innovation agenda of the NSF will all be best served in the future if it becomes much simpler and faster for US academic researchers to migrate from campus and state level resources to the TeraGrid and its successors. A straightforward migration path between campus, state, and TeraGrid facilities would ensure that the many tens of thousands of researchers working in areas of science related to the NSF mission make the best possible strategic use of local, state, and national resources.

Recommendation: Authentication and authorization processes should correspond to national standards, perhaps by adoption of InCommon credentials (http://www.incommonfederation.org/).

In ways that parallel the comparison between CASC generally and the TeraGrid particularly, CASC education and training services scale out well beyond the current scope of TeraGrid educational activities. The US Science Technology and Mathematics (STEM) communities generally recognize the pressing need to interest more young Americans in pursuing careers in STEM disciplines, and to educate more US researchers in the use of advanced cyberinfrastructure to accelerate research. The TeraGrid's HPC University program provides training and education for hundreds of individuals annually. CASC members, collectively, educate tens of thousands annually. Most training tends to

happen locally, and a great body of excellent training material exists at CASC member sites and other US colleges and universities.

Recommendation: The NSF should strongly support continuation of the current TeraGrid HPC University program, either as part of the TeraGrid and its successors or as a related but separately funded activity. The HPC University program should strive to integrate with and take advantage of the many other HPC educational activities of US colleges and universities to create a suite of materials that represents the "global best" that may be repeatedly reused locally.

Continued education in high performance computing and advanced cyberinfrastructure will increase the number of US researchers sufficiently knowledgeable to program and use the TeraGrid and its successors but, no matter how aggressive and effective educational efforts are, not all US researchers will become computational science experts. There are many thousands of US scientists whose research would benefit from using advanced cyberinfrastructure tools and high performance computing, but who cannot divert time from their core scientific pursuits to become computational science experts. Many CASC members are developing or have implemented tools such as portals, workbenches, and Science Gateways that allow scientists to use computation at scale via intuitive Web interfaces, without requiring the scientists to first become computational science experts. Support for Web portals, workbenches, and science gateways as modes of access to computational systems at scale can make high performance computing resources more widely accessible and useful. Implementing such access tools in ways that support interactive computing at scale can greatly enhance the pace of US discovery in science and engineering.

Recommendation: The NSF should strongly support the continued development and delivery of Web portals, workbenches, and science gateways as a way to access 'science at scale' via the TeraGrid and its successors. Such modes of access should be implemented in ways that support migration of computational activities as appropriate from campus and state level resources to the TeraGrid, and in ways that support interactive computing, discovery, and innovation.

This white paper has been approved by a majority vote of the current members of CASC (with no votes opposing endorsement of this white paper recorded as of time of submission). Authors: The CASC Executive Committee (Craig Stewart, Chair; Stan Ahalt, Vice Chair; Richard Pritchard, Secretary-Treasurer) and other CASC member representatives. This paper may be cited as:

Stewart, C.A., S. C. Ahalt, R.H. Pritchard. 2008. The future of the TeraGrid and the national advanced cyberinfrastructure: a perspective from the Coalition for Academic Scientific Computation. Coalition of Academic Scientific Computation (www.casc.org).