

SURA EPSCoR Grid Research Infrastructure Development (SEGRID)

An OCI / EPSCoR Proposal Concept Summary

Evolving from discussions at the NSF EPSCoR National Workshop on Cyberinfrastructure held on May 10 -12 in Nashville, Tennessee, SURA (the Southeastern Universities Research Association) proposes a coordinated program of planning, infrastructure deployment, education and application analysis targeted at increasing the use of high performance computing and grid technology among the 10 EPSCoR states that are within the SURA region (Alabama, Arkansas, Delaware, Kentucky, Louisiana, Mississippi, Oklahoma, South Carolina, Tennessee and West Virginia).

The development of cyberinfrastructure for eScience has become a highly visible national priority, expressed through multiple “blue-ribbon” panel reports and federal agency funding initiatives. Due in part to the continuing increase in capacity and capability of High Performance Computing (HPC) systems and pervasive high performance network infrastructure, grid computing has become a cornerstone of the NSF’s cyberinfrastructure program. While there are obvious benefits to making new tools for scientific discovery available to the nation’s researchers, this focus on cyberinfrastructure has the potential to widen the gap between researchers who have access to HPC resources and those who do not. A review of NSF High Performance Computing allocation statistics reveals that *researchers from the 25 EPSCoR states are allocated only 11.5% of total NSF computing center allocations*; on a per capita basis, researchers from non-EPSCoR states are allocated twice as many service units on NSF funded HPC systems as researchers from EPSCoR states. The same 2:1 ratio holds true for the 10 EPSCoR states within the SURA region.

SURA is proposing the creation of a pilot program, the SEGRID Program, targeted at significantly increasing the ability of faculty and students located in the 10 EPSCoR states within the SURA region to gain access to and take advantage of grid computing for teaching, learning and research. The program being considered could be piloted in the SURA region and, when demonstrated successful, could be scaled up for all EPSCoR states. SEGRID will be a comprehensive program that will include the following components:

- Recruiting EPSCoR participation in SURAGrid, adding systems at EPSCoR institutions to the existing SURAGrid environment (http://www.sura.org/programs/sura_grid.html) thereby providing a mechanism for on-going peer support and mentoring through the SURAGrid community (see description of SURAGrid below).
- Applications discovery, porting, optimizing – built on existing funded work, SURA will work with schools in EPSCoR states to identify applications that would be advantaged by access to grid computing resources.
- Community building - SURA will facilitate regular gatherings of the growing EPSCoR grid community, at dedicated meetings or workshops and conferences of common interest, to further motivate and catalyze peer support and collaborative work.
- The specification and development of an EPSCoR TeraGrid Science Gateway providing a venue for specially facilitated access to TeraGrid resources for the EPSCoR community.
- Training – two types:

- User Training - built on SURA Grid Technology Cookbook and NMI train-the-trainer model (http://www.sura.org/programs/nmi_index.html) to develop training specifically targeted at EPSCoR institutions.
- System Admin training – built on relationships with hardware vendors to leverage special training events for schools participating in the program.
- SURA would extend existing and developing partnerships with leading grid computing companies to provide heavily discounted cluster and HPC hardware to institutions in SURA EPSCoR states. Systems would be integrated into the existing SURAgrid environment and purchased with funds from 3 sources:
 1. Funds from participating EPSCoR institutions;
 2. Significant SURA negotiated discounts from SURA’s current and future corporate partners;
 3. Co-funding from NSF OCI and EPSCoR programs.
- Additional matching funds from EPSCoR for network connectivity improvements to allow integration of deployed systems into regional SURAgrid fabric, leveraging SURA’s relationship with AT&T (donated dark fiber) where appropriate.

The SEGRID Program will build upon the considerable foundation of related activities of SURA and SURA member institutions, including the National Science Foundation Middleware Initiative (NMI) Integration Testbed Program (NSF Award #ANI-0123937), the strategic direction of the SURA Information Technology Steering Group (http://www.sura.org/programs/it_strategy.htm), SURA’s Cyberinfrastructure workshop series (http://www.sura.org/programs/it_workshop.htm), an NSF OCI award OCI-0545550 (a grid application development effort), the development of a Grid Technology Cookbook (funded by US Medical Research and Materiel Command) and the evolution of SURAgrid, a multi-institutional collaboration to build distributed, sustainable, secure grid infrastructure for inter-institutional sharing of computing resources. We will also leverage expertise from SURA member developed research application communities such as ChemGrid, an NSF/NMI funded program that is becoming a valuable resource for the national computational chemistry community and SCOOP, a coastal ocean modeling and prediction initiative funded by the National Oceanic & Atmospheric Administration and the Navy Research Laboratory.

SURA intends to leverage our experience in building SURAgrid and the grid technology expertise of the SURAgrid participating institutions to develop the SEGRID Program. In parallel to creating an EPSCoR focused regional grid computing resource, the proposed work will enhance the SURA EPSCoR community’s ability to develop, port, execute and optimize applications in a HPC grid environment. By building regional expertise and capabilities in the use of grid computing and creating operational capabilities within the region, we will improve the ability of researchers in EPSCoR states to win allocations on NSF funded HPC systems. With its unique emphasis on diversity, heterogeneity and inclusion, SURAgrid is well positioned to extend these benefits to the SURA EPSCoR community of researchers, improving their access to high performance and grid-based computing.

The SURAgrid initiative has achieved success in its first stage of development through the involvement of a growing community of SURA member institutions and the voluntary contribution of computing resources – with 27 institutions now participating and contributing a total resource of over 900 processors with combined peak computing capacity in excess of 3TFlops. SURA has just finalized an agreement with IBM that will make IBM’s newest HPC systems available through SURA at very significant discounts creating the potential to more than triple the resources available through SURAgrid over the coming months.