

NSF CI-TEAM AWARD 0537498: MINORITY SERVING INSTITUTIONS CYBERINFRASTRUCTURE INSTITUTE (MSI-CII)

<http://www.educationgrid.org>

Geoffrey Fox, Indiana University, gcf@indiana.edu; Richard A. Alo, University of Houston Downtown, Center for Computational Science, www.uhd.edu/ccsds, ralo@uh.edu; Karl Barnes, NAFEO, www.nafeo.org karl.barnes@gmail.com; Diane Baxter, SDSC, dbaxter@sdsc.edu; Julie Foertsch foertsch@wisc.edu; Al Kuslikis, AIHEC, AKuslikis@aihec.org; or Alex Ramirez, HACU, aramirez@hacu.net

Updated January 13 2008

Research and Education Activities

The Minority Serving Institutions CyberInfrastructure Institute (MSI-CII) is a coalition of leading Cyberinfrastructure experts, projects and resources with the American Indian Higher Education Consortium (AIHEC), the Hispanic Association of Colleges and Universities (HACU), the National Association for Equal Opportunity in Higher Education (NAFEO), and the more than 330 MSIs they represent. MSI-CII is led by the Alliance for Equity in Higher Education with Indiana University, NCSA, SDSC and the University of Houston Downtown helping the three national MSI organizations. Julie Foertsch performed the evaluation function and substantial funding was provided to several MSI faculty and staff for the cost of participation in Cyberinfrastructure activities. MSI-CII is a demonstration CI-Team project that is followed by the implementation project MSI-CIEC or the Minority Serving Institutions CyberInfrastructure Empowerment Coalition. MSI-CIEC has been responsible for following on MSI-CII meeting and outreach activities during 2007. This includes CI days which were proposed on the basis of MSI-CII findings. During the “no cost extension” for MSI-CII, work focused on the portal whose design emerged from the first year’s work. This portal has two major components: a Research and Resource Web 2.0 matching site and a collection of carefully edited video nuggets describing key Cyberinfrastructure concepts.

Major Meetings (details on <http://www.educationgrid.org>)

1. October 3-6 2005: MSI attendance at Global Grid Forum GGF15 at Boston
2. October 18-19 2005: Attendance at EPIC All Hands Meeting Albuquerque
3. November 12-18 2005: MSI attendance at SC05 Seattle
4. January 30-31 2006: Planning and Cyberinfrastructure Education (train the trainer) meeting at SDSC
5. April 11-12 2006: NCSA MSI Summit
6. June 26-30 2006: General Cyberinfrastructure Summer School at SDSC
7. August 7-10, 2006: Enabling Minority Serving Institutes to be Cyberinfrastructure Resource Providers Workshop at NCSA
8. October 27 2006: SACNAS (Society for Advancement of Chicanos and Native Americans in Science) National Conference Session organized by Chaitan Baru and Ann Gates, Tampa Florida.

Meetings 2 and 8 were aimed at planning and we did not fund MSI participants. Meetings 1 and 3 just involved a few participants. Meetings 4-7 were focused on MSI participants

with Meeting 4 and 5 the largest with total attendance of 40 and 50 respectively. These latter two were also formally evaluated as reported in appendices to our findings.

Activities at Meetings

1. Training in Cyberinfrastructure
2. Substantial Planning and discussion leading to new projects including:
 - a. MSI-CIEC concept and implementation plan |
 - b. Elizabeth City State University leadership in use of Cyberinfrastructure
 - c. Planning study for the “Internet to the Hogan and Dine Grid” project led by Navajo Tech IT faculty and academic dean, at Crownpoint, NM
 - d. Cyber-ShARE Center of Excellence: A Center for the Sharing of Cyber-Resources to Advance Science and Education from UTEP led by Ann Gates. This was recently funded by NSF CREST program and Fox is on board of advisors.
3. SDSC meetings were Access Grid-enabled and plenary presentations were videotaped with CI Channel and available on Web at <http://www.educationgrid.org>. During 2007, selected videos were edited to capture “nuggets” in key areas of Cyberinfrastructure that could be used for MSI faculty and administration.
4. Web 2.0 MSI Portal research matching site supporting exchange of interests and tagged documents including federal grant opportunities, was designed and was developed during 2007

Key Objectives of Project

1. Identify potential early CI adopters among MSI faculty and students
2. Pilot an approach to expansion of CI use in research and education among MSI faculty and students, starting with those identified above
3. Analyze scalability of the approach for eventually reaching the over 330 MSI institutions represented by the coalition of AIHEC HACU and NAFEO
4. Provide access to physical infrastructure needed to support use of CI in research and teaching
5. Support curriculum development, research, mentoring, and teaching teams
6. Exploit key Cyberinfrastructure (Grid) resources
7. Develop prototype portal (mashup) supporting broad participation in CI with low barrier to entry
8. Use project evaluation to strengthen the approach and processes supporting its wider implementation.

Key Ideas behind project

1. Cyberinfrastructure is critical to all involved in Research and Education
2. Cyberinfrastructure is intrinsically democratic, supporting broad participation
3. MSI's should lead MSI integration with Cyberinfrastructure
4. Experts in Cyberinfrastructure should act as advisors in efforts to broaden its use
5. From the outset, one key criterion for success should be the scalability (systemic infusion) of the approach.

6. Mentoring is seen as just an initial step toward the goal of peer collaborations involving all institutions of higher education.
7. Collaboration, coordination, and trust-building must occur across institutional, cultural, and geographical barriers
8. Activities should be professionally evaluated

Advisory Team

1. Malcolm Atkinson, NESC (UK National e-Science Center), ICEAGE (EU Grid Education)
2. Fran Berman, SDSC
3. Jay Boisseau, TACC
4. Charles Catlett, Argonne National Laboratory
5. Kelvin Droegemeier, Oklahoma, LEAD
6. Tom Dunning, NCSA
7. Mark Ellisman, SDSC, BIRN
8. Ian Foster, Chicago, Open Science Grid Globus etc
9. Juan Meza, LBL
10. Dan Reed, UNC, Renaissance Computing
11. Richard Tapia, Rice
12. Larry Smarr, UCSD, Cal(IT)2

Implementing National Cyberinfrastructure for MSI's

Our findings included a study of the National Cyberinfrastructure (Open Science Grid, TeraGrid, Internet2, and National LambdaRail etc.) and the need for integrating MSI's with it.

- Fund local infrastructure and local infrastructure support
- Establish Tiger teams (focused groups mobilized to solve particular IT problems) to visit MSIs and establish plans for institution CI-Enablement
- Identify a “simple” and “robust” CI-lite software stack (could be an existing stack)
 - Security issues for MSIs
 - Role of Portals/Science Gateways versus “power user” access
- Define, Implement and Support an education and training model including “distance learning” (including “Access Grid” infrastructure) and institution curriculum integration
- Fund internships and other research opportunities (e.g. REU) for MSI faculty/students
- Provide (remote) MSI/Community CI Operations Center to provide support to MSI users of CI
- Empower (=fund) Centers of Excellence to provide institutional support
- Research use of VM technology and shared desktops to allow remote hardware and remote support

- Involve “all organizations” including Internet2 to support network access and community organizations to support scalability
- Establish Partnerships between MSI’s and experienced National CI institutions for smooth CI-Enablement (Internet2 Strategy)
- All aspects should be Systemic (scale to all 335 MSIs) and aim at peer collaborations and not elite to non-elite relationships

Broadening MSI communities involved

Our findings identified a problem in expanding from “early adopters” to the wider MSI community. The challenge is identifying MSI faculty and institutions that are open and ready for training to use CI resources. Often the different NSF activities in this area all work with the same rather small group of people. We suggested proactive projects like CI-days as one approach to this issue. Another more speculative idea is to develop Web 2.0 style portals for community building and scientific discovery as Web 2.0 allows a broad range of people to participate with low barriers. Thus in the second year of MSI-CII we developed a MSI portal built with Web 2.0 described below to foster communities. Further we are working with Navajo Tech on use of Web 2.0 gadgets as an attractive technology for “easy to use” science gateways to TeraGrid aimed at education.

MSI-CIEC Portal Web 2.0 Research Matching Site

Problem Statement: One of the MSI-CII/MSI-CIEC project’s primary activities is to identify new MSI faculty and researchers to participate in cyberinfrastructure-related research activities such as the NSF TeraGrid. A key problem we identified in MSI-CII was the currently difficult process for identifying new participants: many outreach projects worked with the same relatively small number of MSI researchers.

Solution: In addition to conventional outreach activities, MSI-CIEC has investigated the use of Web technologies for outreach and community building. These include an informational Wiki (http://www.msi-ciec.org/eduwiki/index.php/Main_Page) and a prototype Web 2.0-style social networking Web Research Matching Site.

Social Networks such as Facebook, LinkedIn, and MySpace, with millions of registered users, demonstrate the power of online communities. In our analysis, we decided to build our social network prototype around the central concept of *shared bookmarking*. In this, we were inspired by bookmarking sites such as del.icio.us (a general purpose site), Connotea (specialized for bookmarking journal articles and other scholarly URLs), and CiteULike (also for journals). Although these sites provide some social networking capabilities (one can form groups), they generally do not provide detailed user profiles and one cannot tag easily key information such as TeraGrid allocations.

We identified social bookmarking as one way for likeminded individuals to discover one another. We further thus chose to add value to these social bookmarking by providing more detailed, customizable user profile pages that can be decorated by the user with personal RSS feeds, profile tags, and contact information. This has been implemented

in the MSI-CIEC portal (currently at <http://www.msisiec.org>) effort led by Indiana University. Two software engineers implemented the portal.

In addition to the functional capabilities of the MSI-CIEC Portal, we also imported information on previous funding activities from the NSF and also project information (such as allocations) from the NSF TeraGrid. This information was converted into tags.

The following summarized capabilities of the MSI-CIEC Portal:

- Can be used to bookmark and tag arbitrary URLs.
- Can be used to search for NSF projects, principal investigators, and collaborators by a rich set of imported tags (NSF division, project funding size, project funding year, TeraGrid allocation size, etc). Tags are displayed as tag clouds, indicating for example NSF principle investigators with the most collaborators.
- Can be used to “click tag” NSF RSS feeds for new funding announcements as “interesting” or “uninteresting”. Users can view their own “interesting” tags and also “interesting” from all users.

The MSI-CIEC Portal is implemented as a fully open source project, and code is available from SourceForge. The prototype was completed for SC07 but we identified issues and we will encourage broad use in January 2008.

MSI-CII Video Resource

We got excellent feedback on talks given at MSI-CII events and so we undertook to build a library of Cyberinfrastructure video nuggets gotten be breaking up approximately 12 hours of video taken with the SDSC Cchannel into few minute segments on particular topics. MSI-CIEC will use these at meetings with MSI administration and faculty and use them to support other online resources. We will later in 2008 integrate them into the portal described above generating suitable tags. We will also post to Web 2.0 sites YouTube and SciVee. This work was performed by SDSC and the selected talks were

- Cyberinfrastructure for Humanities, Arts and Social Science ~Fran Berman
- Bioinformatics ~ Mark Ellisman
- Grid Technology ~ Karan Bhatia
- Virtual Infrastructure ~ Kate Keahey
- Linked Environments for Atmospheric Discovery (LEAD) ~ Jay Alameda
- Security and Grids ~ Victor Hazelwood
- Data Grids and Data Management ~ Reagan Moore
- Future of the Internet - kc Claffy, SDSC
- Linking the Back Country - Hans-Werner Braun
- The International Connection with PRAGMA Grid - Peter Arzberger
- Grids and the Home Institution (Campus Grids) – Access and Resource Providers - Laura McGinnis
- Grid Technology : Software Engineering for Grids ~ Shava Smallen
- What can we learn about unusual earthquake hazards from large scale numerical simulations on the Grid? ~ Bernard Minster

- Grids in Biology (Bioinformatics and Medicine) ~ Nicholas Schork
- Computational Science and Drug Discovery ~ Andy McCammon
- Computational Astrobiology ~ Kim Binsted