CI-TEAM Implementation Project: Collaborative Research:
"CI-TEAM-Work"

Cyberinfrastructure promises profound benefits for scientific discovery in all disciplines by providing scientists, educators, students and administrators with unprecedented access to computing and communications facilities of vastly enhanced scale, capability, and usability.
CI-TEAM-Work's vision is to realize this promise with a sustainable, nationally scalable process to **create and sustain undergraduate curricula that integrate cyberinfrastructure within science, technology, engineering and mathematics (STEM) disciplines.**

The promise of cyberinfrastructure (CI) and the vision of this project will be realized by developing resources and strategies among diverse campuses, faculties, administrators and student populations for integrating the national investment in cyberinfrastructure into the practices of learners, educators and researchers. The project focuses on working with and empowering faculty, administrators and students beginning with fourteen two- and four-year colleges and universities to create a campus environment for sustained integration and use of cyberinfrastructure in the preparation of the workforce across STEM disciplines. A key focus is to leverage and utilize the national Grid infrastructure of TeraGrid and the Open Science Grid. The project is organized as a virtual organization that blends–and amplifies through teamwork–the expertise of high performance computing specialists, computational scientists, educators, and science informatics experts from TeraGrid, Open Science Grid, the National Computational Science Institute, the annual SC Conference, and other nationally renowned HPC science and education organizations. These experts work in close collaboration with diverse populations among diverse academic institutions to support their growth through the use of CI.

Intellectual Merit

The project will develop effective strategies and resources for 1) training and motivating faculty to effectively apply cyberinfrastructure; 2) assisting faculty with integrating cyberinfrastructure methods into the curriculum; 3) nurturing the advancement of faculty and students, including women and minorities, into leadership positions; 4) mentoring of faculty and students to increase the recruitment and retention of significantly more under-represented students into professional careers in STEM fields; and 5) assisting the participating institutions with enhancing their own infrastructure in support of these strategies. The strategies will be designed, implemented, tested, and evaluated for their ability to be sustained well into the future by a growing movement of organizations and people. Strategies for scaling-up to other institutions beyond those directly supported by this project will be designed, implemented, tested, and evaluated.

Broader Impact

> "*No first-world nation can maintain the health of its economy or society when such a large part of its population remains outside all scientific and technological endeavors.*"
> – Dr. Richard Tapia, University Professor, Rice University

CI-TEAM-Work addresses this challenge directly, by engaging under-represented researchers, faculty and students from minority serving institutions, community colleges and schools in EPSCoR states, to significantly increase the diversity of people pursuing and flourishing in scientific careers. The project will engage and scale-up these vital practices among other academic institutions as well. Materials and strategies developed will be shared with the community through the National Science Digital Library, professional societies, publications and conferences. The project will collaborate with other CI-TEAM funded projects, NSF funded science and Grid projects, and other projects with similar goals funded by other agencies, foundations and companies. Beyond education, CI-Teamwork activities will empower the *research* efforts of STEM faculty and students through the use of cyberinfrastructure, and support advancement and mentoring of faculty and students as vital components of the process.

1. Vision and Goals

CI-TEAM-Work's vision is a sustainable, nationally scalable process to create and support undergraduate curricula that integrate cyberinfrastructure within science, technology, engineering and mathematics (STEM) disciplines. The project focuses on working with faculty, administrators and students to create campus commitments to sustaining courses that teach the application of the national cyberinfrastructure for scientific discovery using the computing resources provided by TeraGrid and the Open Science Grid. CI-TEAM-Work's training and education focus will have broad impact, empowering the research efforts of STEM faculty and students through the use of cyberinfrastructure. The project supports advancement and mentoring of faculty and students as vital components of the process.

The goal of CI-TEAM-Work is to facilitate the education of a larger and more diverse workforce of people that are well prepared to apply cyberinfrastructure across all STEM fields. It will create a sustainable continuum of workforce development and support in three complementary dimensions:

- *Disciplinary Breadth*: moving along a continuum from a specific discipline to an interdisciplinary approach to discovery,
- *Computing Spectrum*: gaining the skills to move from using desktop resources to effectively utilizing HPC and the national Grid computing cyberinfrastructure, and
- *Scientific Preparation*: moving from learning about cyberinfrastructure to applying cyberinfrastructure within scientific practices, for discovery.

In each of these dimensions we seek to advance the skills of faculty, students and ultimately practitioners to apply cyberinfrastructure in appropriate and meaningful ways within their areas of specialization.

CI-TEAM-Work will **implement** this vision by **creating an integrated process** comprising:

1) hands-on cyberinfrastructure training workshops focusing on applications to scientific inquiry and the undergraduate science curriculum; 2) grants for curriculum development by faculty; 3) regular campus visits by cyberinfrastructure experts; 4) on-line mentoring of participants; and 5) a variety of services to support the deployment, promotion, evaluation, and enhancement of courses that teach the application of cyberinfrastructure in STEM disciplines. This three-year project engages fourteen two-and four-year colleges and universities within seven regions across the country, linking preeminent research intuitions and underserved populations that are underrepresented in STEM disciplines. CI-TEAM-Work will support the development of training materials, curricula, and collaboration environments that can, with follow-on investment, be scaled to reach ever-larger audiences. External evaluators will measure both the effectiveness and sustainability of the approach, and the strategies for scaling to a broader national level.

CI-TEAM-Work is executed by a virtual organization [FKT01] which combines the expertise of high performance computing specialists, computational scientists, educators, and science informatics experts from TeraGrid, Open Science Grid, the National Computational Science Institute (NCSI), the annual SC Conference, and other nationally renowned HPC science and education organizations working in close collaboration with researchers, faculty, administrators and students at diverse academic institutions.

This proposal is organized as follows. In section 2 we describe the motivation for this work. Section 3 describes our prior work in this area under NSF support, and section 4 briefly surveys other related work. Section 5 describes our work plan in detail, organized by the integrated set of activities that make up our proposed process. Sections 6-10 describe our management structure, collaboration, risk management, evaluation and dissemination plans. We conclude with a brief summary of merit and impact.

2. Background and Motivation

> "*Computational science – the use of advanced computing capabilities to understand and solve complex problems – is now critical to scientific leadership, economic competitiveness, and national security*." –J. H. Marburger, Science Advisor to the President and Director, OSTP. [HPCWire]

We describe in this section three principles that derive from this mandate and motivate our work:

1. The need to enhance the size and skill level of our national workforce in STEM fields is critical to the future prosperity and vibrancy of our nation.
2. Undergraduate students are a critical population segment to reach in order to build this workforce.
3. The national Grid infrastructure – TeraGrid and Open Science Grid – is *the* critical and durable *computing* cyberinfrastructure that we must train students to master and apply across all STEM fields.

Why STEM education? The 1991 report *"Computational Science and Education: Workshop on the Role of HPCC Centers in Education"* [Smarr, 1991] urged NSF to establish a vigorous, coordinated program to integrate high performance computing and communications technology with general science and mathematics education to provide an effective and wide ranging mechanism to increase the number of people engaged professionally in science, mathematics, and engineering and to improve general scientific literacy. The 2003 NSF Blue Ribbon Advisory Panel on Cyberinfrastructure [Atkins, 2003] acknowledges that there are significant education challenges to meeting the demands of the new STEM research agenda that require blended expertise in disciplinary science and engineering, mathematical and computational modeling, numerical methods, and visualization.

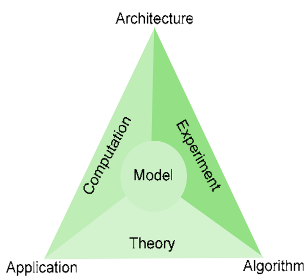

Figure 1: Third facet of science

"The PITAC calls on Federal research and development (R&D) agencies and universities to make coordinated, fundamental changes to their research and education structures to promote and reward collaborative approaches essential to computational science." – Report of The President's Information Technology Advisory Committee. [PITAC]

Academic, industrial and government researchers are making extensive use of computational methods in all STEM fields. Computational science has been recognized as the third leg of science for more than 30 years, as depicted in Figure 1 [Shodor, 2006]. The ability to build conceptual links between content knowledge and computational models and visualization techniques is the most significant learning challenge students face [Helland, 1997], [Honey, 1995-2001]. To be prepared for a future of complex knowledge manipulation, reasoning, and problem solving with multidimensional data and sophisticated representations, students must acquire different kinds of knowledge and thinking skills than those emphasized previously [Bransford, 2001], [NCTM, 2000], [NRC, 1996].

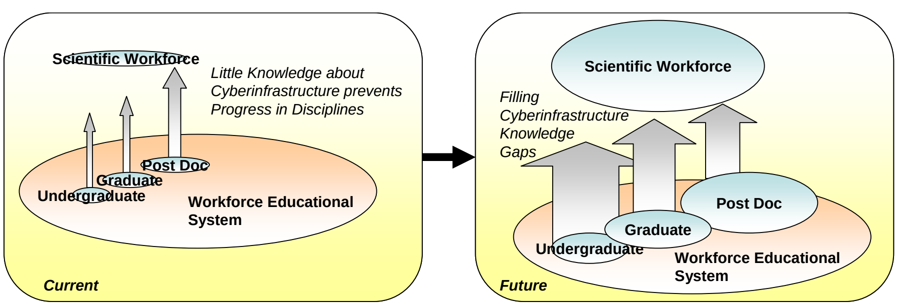

Figure 2: CI-TEAM-Work's impact on the scientific workforce

The CI-TEAM-Work project seeks to transform the nation's education system, from one in which the number of people advancing through the educational system to become active members of scientific teams is currently small and not well diversified, to an environment in which many people, across diverse backgrounds, are well prepared in cyberinfrastructure and computational science skills (Figure 2). This project will not only influence such a major paradigm shift in education during its three-year span: it will create the foundation for the sustained adoption of tools and methods needed for national scale change.

"Although we feel NCSI has been successful in establishing a true computational science education community, it is fairly vulnerable at this time: vulnerable to slow curricular change processes, vulnerable to departments and institutions with different values and priorities, vulnerable to overcrowded instructor schedules, vulnerable to lack of resources, vulnerable to demands made by

creating intense learning environments, etc. All these things can create barriers that impede progress and could possibly even stop it altogether. There is a need for continued work in this community and it really needs to be addressed on a consistent basis. The community still needs support to sustain what has already been started and to help it continue to grow." –***From the 2005 NCSI Evaluation Report***

Why focus on undergraduate Faculty? Undergraduate faculties are teaching the future K-12 teachers, the future graduate students, and the future work-force. If we want to have the biggest impact on educational reform and improvement, undergraduate education is the **pressure point**. In addition, research universities continue to advance in their use of computational science as a new methodology in scientific discovery and in their use of HPCC technologies, with less emphasis on adaptations for the undergraduate educational experience. Yet the gap between the predominately undergraduate institutions and the nation's research institutions is widening, even though across the country most graduate students still start in predominately undergraduate institutions, and a growing number of the four-year institution graduates start in two-year institutions. We seek to increase the number of qualified students who apply to graduate school or who can transfer to four-year institutions, with a special emphasis on students who are underrepresented in the areas of science, technology, engineering, and mathematics. CI-TEAM-Work reaches this student group by helping their professors develop relevant, engaging, focused courses that teach the value of using cyberinfrastructure and the excitement of mastering it.

Why focus on the national-scale Grids? The National Science Foundation (NSF) through the Office of Cyberinfrastructure is deploying the cyberinfrastructure needed to support significant advances in all fields of STEM research and education. The TeraGrid and the Open Science Grid are exemplars of the cyberinfrastructure designed to address these research and education needs. Wikipedia (http://en.wikipedia.org/wiki/Cyberinfrastructure) defines cyberinfrastructure as:

"...the new research environments that support advanced data acquisition, data storage, data management, data integration, data mining, data visualization and other computing and information processing services over the Internet. In scientific usage, cyberinfrastructure is a technological solution to the problem of efficiently connecting data, computers, and people with the goal of enabling derivation of novel scientific theories and knowledge."

TeraGrid (www.teragrid.org) is an open cyberinfrastructure for scientific discovery. It combines leadership-class resources at nine partner sites to create a powerful, integrated, persistent computational resource. The TeraGrid provides researchers with access to more than 102 teraflops of computing capability, more than 15 petabytes of online and archival data storage, and over 100 discipline-specific databases, with rapid access and retrieval over high-performance networks. The TeraGrid is the world's largest, most comprehensive distributed cyberinfrastructure for open scientific research supporting over 1,500 researchers, educators and students across nine Resource Provider sites. The communities of practice that benefit from TeraGrid are projected to grow from the current 1,500 users to well over 6,000 users within five years by engaging researchers in using the computational resources, scientific data collections, unique scientific instruments, and science gateways within various scientific domains. National efforts are also underway to significantly expand the computing capacity available to the community with the deployment of petaflop computing systems by 2008.

The Open Science Grid (http://www.opensciencegrid.org/index.php) is a national production-quality Grid computing cyberinfrastructure for large-scale science, built and operated by a consortium of universities and national laboratories. The Open Science Grid (OSG) Consortium enables diverse communities of scientists to access a common grid infrastructure and shared resources. Consortium members contribute effort and resources to the common infrastructure. The OSG provides a stable, supported environment for sustained applications. It unifies a vast array of computing resources across over ten huge experiments in the physical sciences, and is aggressively making its resources available to other scientific communities that are still developing the capabilities to compute at this scale. OSG is projected, over the next five years, to provide over 100,000 CPUs and ten petabytes of storage to its user communities, all accessible through uniform Grid services and protocols. The OSG community has created the governance and management procedures necessary to operate as a consortium where members can contribute computing facilities and expertise in exchange for shared access to a greater and more dynamic ensemble of resources than any single stakeholder could afford or manage.

Together, TeraGrid and the Open Science Grid represent the largest collection of organized, national-scale openly shared computing resources, and form a key pillar for a significant portion of the computing resources that will power the advancement of science for the next decade. They represent a flexible mechanism for scaling up computational power in an affordable manner and will continue to exist and evolve even as the power of the individual computing, storage and communications resources that form these grids continues to scale up. It is critical to integrate these facilities into the national science education process and to teach students how to apply these resources to the practice of scientific research and discovery. From 2000 to 2010 the US will have committed over $300M to the construction and operation of these two Grids. The work proposed here is a small investment with great potential for enabling us to leverage the maximum benefit of this vital cyberinfrastructure.

3. Work to Date: Results from Prior and Current NSF Support

DUE-0127488 CCLI-ND Grant to Shodor. The National Computational Science Institute (NCSI), funded by NSF's Division of Undergraduate Education, has assisted over 1,000 undergraduate faculty to integrate computational science methods into the curriculum over the last six years. NCSI enables faculty to explore computational science in the undergraduate classroom, lab, and research experience.

NCSI has conducted 59 weeklong workshops through the end of summer 2006 and nearly 1000 faculty have attended with many attending more than one workshop several years in a row. More than a third of the participants were minority group mambers, and about the same percentage female. NCSI offered workshops for about a dozen Sigma Xi chapters, who in turn offered workshops for area high school teachers. Overall more than 80 faculty have received funding for a variety of proposals to implement lessons learned, including five schools that based winning HHMI undergraduate science education reform grants on computational science projects. NCSI grew into a national program from a much smaller initial experiment, the Shodor Computational Science Institute, focused primarily on the Carolinas.

One anecdotal piece of evidence of NCSI's effectiveness, as communicated by the Office of Science of the Department of Energy, is that students in computational courses, minors or degrees initiated by NCSI participants now dominate many of the internships and fellowships offered by national labs in the computational sciences. A second data point, also from the DOE, is that both winners of the Krell Institute's Undergraduate Computational Engineering and Science (UCES) awards for 2005 were NCSI participants, and each has gone on to secure substantial NSF funding for their efforts to implement computational science curricular reform at their institutions. This faculty recognition is important for faculty advancement, and we are pleased to have the Krell Institute as a collaborating partner.

The CI-TEAM-Work project builds on the success of NCSI by adding three new components: 1) three years of continuous interaction with the faculty, students and administrators at the participating institutions to achieve sustained change within those institutions, 2) attention to scaling-up to more institutions to impact a larger workforce, and 3) a broader perspective of applying cyberinfrastructure to do science, including Grid computing and collaboration environments for learning and mentoring.

DUE-0527943 Provoking a National Discussion on Computational Science Education. As a precursor to this proposal, NSF asked Shodor and SRI to propose a mechanism to stimulate a national discussion–reaching new voices other than the "usual suspects"–on how to implement computational science across the curriculum at all education levels. The strongest and most striking result of the first few months of this discussion has been the near-unanimous sentiment that for this to happen, NCSI or a program grow out of it would be needed to provide vision, support, and coordination of various efforts. Given the short-term nature of federal funding of such projects, interactions with the NCSI advisory board set a firm direction that we should develop a comprehensive plan for NCSI that would transition to long-term, independent funding. This project will pursue such a comprehensive plan.

ITR-0086044 GriPhyN – The Grid Physics Network - pioneered techniques for harnessing distributed resources for data analysis within scientific communities. Building on security, resource discovery, data management, and resource management methods, this work has developed new techniques for representing and managing virtual data, scalable replica location, scheduling, and resource discovery, explored applications, and provided new insights into community data sharing. The GriPhyN outreach, program yielded two deliverables that will benefit CI-TEAM-Work. First, GriPhyN virtual data workflows have been applied to hands-on high-school physics education as part of the **QuarkNet Grid** project which led to the NSF-supported pilot program **PHY-0538356 I2U2 – Interactions in Understanding the Universe** on which co-PI Wilde also serves. This program provides a model that can be readily applied to the hands-on teaching of Grid computing techniques for scientific data analysis. I2U2 [I2U205] is a complementary program to CI-TEAM-Work. I2U2's main goal is to excite high school students, through hands-on learning, into considering a science career. CI-TEAM-Work then takes students that I2U2 has attracted into a STEM major or minor, and retains and educates these students with the skills they will need to participate as (or with) practicing scientists. Second, for the past three years (2004-2006) GriPhyN and **iVDGL (International Virtual Data Grid Laboratory–NSF Grant 0122557)** have been major contributors to the US Grid Summer Workshop – a one-week workshop to train advanced undergraduate and early graduate students in the hands-on application of Grid techniques to scientific research. This workshop (http://cgwa.phys.utb.edu/Outreach/SummerSchool.php) hosts approximately 40 students per year from across a range of backgrounds, geographical regions, and disciplines in a highly-engaged, interactive and practically-focused educational experience. These workshops will be a core element of the OSG education and training effort, and will be adapted as Grid training modules for the CI-TEAM-Work workshop and online-tutorial activities described below.

SC Education Programs at the annual SC Conferences. The SC Conference is the annual premier international conference on high performance computing, networking, storage and analysis. The SC Conference includes an Education Program designed to engage educators from around the country in learning about the impact of these technologies to enhance science and education. The PI for CI-TEAM-Work led the SC|02 and SC|03 Education Programs in partnership with NCSI to bring together over 240 faculty and teachers from around the country to focus on integrating computational science into their courses. During the SC|05 conference, the PI was asked to develop a long-term Education Program that would offer greater continuity and greater sustained impact in return for the investment in the SC Education Program. A plan for the SC07-09 Education Programs was developed in collaboration with about 50 scientists, technologists and educators from around the country. The resulting Education Program plan, along with a Pathways to SC07 plan to prepare materials in advance of this revised program were submitted to the SC Steering Committee. Both plans are included in the appendix. In January, the SC Steering Committee approved a substantial match of $197K that directly complements the CI-TEAM-Work project. The SC|07 and SC|08 Conference chairs have enthusiastically endorsed the Education Program plans for each of those years. Once a chair is selected for the SC|09 conference, their support will also be sought to sustain the program for a full three years. The SC plan directly complements CI-TEAM-Work and includes these components described in detail in the appendix:

- Pathways to SC07 – part 1 – currently funded at $197K (letter of support is in the appendix)
- Pathways to SC07 – part 2 – recommended for funding of an additional $197K based on the financial success of the SC|06 conference
- SC07-SC09 Education Programs – a proposal requesting $315K for the SC|07 Education Program, with modest increases for SC|08 and SC|09 to engage participants from five more institutions each year, has been submitted and will be decided upon after the SC|06 conference.

4. Comparison to Previous Work and Similar Programs

A CI-TEAM award in 2005 to Capital University, OSC, and OLN has similar motivation and elements.

The Ohio work has a statewide focus, and focuses on traditional HPC approaches rather than Grid-based approaches. We know that the Ohio team is planning an implementation project to expand upon this two-year effort. These efforts highly complement CI-TEAM-Work, and can leverage each other's deliverables and expertise. The science education materials developed in Ohio will be utilized in our program. Our Grid-based material and techniques for making the national Grids accessible at the undergraduate level will benefit the Ohio program. Dr. Ignatios Vakalis (Capital University) has agreed to serve on our Advisory Committee, which will help facilitate communications and collaborations with efforts in Ohio.

A CI-TEAM project is led by Indiana University (Geoffrey Fox) and the Alliance for Equity in Higher Education, called the "Minority Serving Institutions Cyberinfrastructure Institute". We are working with them to raise awareness of Grid resources with training workshops planned at SDSC and NCSA this summer. We will work with their proposed CI-TEAM implementation project called: Minority-Serving Institutions Cyberinfrastructure Empowerment Coalition (MSI-CIEC). Alex Ramirez (Hispanic Association of Colleges and Universities) states, "We mention your proposal as being synergistic with ours and an example of activity that can help sustain the efforts of the coalition." He adds, "Your project does look at an important aspect that we also look at, education and curriculum at MSIs, but not to the level of specificity that you do. Three other pieces of connective tissue are: you as a collaborator within the coalition, the TeraGrid itself, e.g., we expect to utilize TeraGrid 07 and 08 for some of our training activity, and Tom Davis as a highly regarded leader in the MSI community who along with you can provide effective practices and polices as well as lessons learned from your project to MSI-CIEC."

We have involved more than 50 scientists, researchers, and educators from multiple organizations with expertise in computational science, scientific computing, HPC and grid computing in developing the plans for the SC Education Program and the CI-TEAM-Work project. The objective is to identify the best resources and approaches, leverage existing programs, and avoid duplication of effort. We are aware of no other programs, other than those cited in this proposal, that have a similarly broad national scope, diverse range of institution types, and a specific program to engage the use of the national Grid infrastructure in undergraduate STEM education, research and faculty advancement.

5. Implementation Plan

The CI-TEAM-Work program of curricular development will be implemented through the activities summarized below and described in detail in the rest of this section. Each activity embodies one or more of the *components* of the CI-TEAM-Work name and program foci: *cyberinfrastructure (CI), training, education, advancement, mentoring, teamwork, and workforce impact*. Figure 3 shows how these activities form a coherent process. We will work with campus personnel, in a manner convenient and cost-effective for them, via live and virtual cyberinfrastructure-mediated interactions to allow the project staff to better understand and address the needs of the participants. The primary activities are:

- *Create an engaged, nationwide community of educators and specialists dedicated to the goals of this program.* We have engaged faculty and administrators across the STEM fields at 14 two- and four-year institutions, including minority serving institutions and EPSCoR schools (listed in Table 1 below). We have also engaged specialists among 10 *Collaborating Organizations* (*ital*. in Table 1). This process of inclusive engagement will continue to actively recruit more practitioners among these and other institutions and organizations throughout the course of the project.
- Adapt and utilize grid-based cyberinfrastructure to teach and learn effective applications of cyberinfrastructure in research and education. This will create techniques and a support structure – a so-called "on ramp" – to make it easy and affordable to use the national Grids as an effective laboratory tool in undergraduate education.
- *Conduct week-long summer workshops* at seven regional sites each year to provide participants with training to utilize cyberinfrastructure tools and resources such as those offered by TeraGrid and OSG (e.g. Grid computing, visualization, data analysis, etc.) – a total of 21 intensive sessions.
- *Provide curriculum development support* to enable at least 45 faculty members over three years to

integrate CI tools and methods into new and existing courses across the STEM spectrum.

- *Conduct regular campus visits* to seven regional sites each fall and each spring to support the participants in their integration of CI into their research and courses; and meet with administrators to address sustaining CI integration and scaling to other institutions;
- *Support over 300 faculty and students to attend the SC conferences over 3 years* to experience the benefits of HPC technologies on scientific discovery, and participate in conference leadership;
- *Provide on-line tutorials, discussion forum and collaboration spaces* for continuing education and mentoring of faculty, administrators and students during their adoption of CI;
- *Disseminate the resulting training and curricular materials* via the National Science Digital Library, conferences, workshops, newsletters and the web; and
- *Evaluate the effectiveness* of these efforts towards achieving the project's vision via surveys, interviews, and requirements analysis with faculty and administrators.

A critical aspect of the project is sustaining change within the participating institutions and scaling-up successful practices to additional institutions. The first two years of the project will focus on effecting change within the 14 participating institutions while working with administrators and faculty to design and implement campus and national approaches to sustaining these changes. Years two and three will design and implement mechanisms for scaling up the activities to other institutions. The third year of the project will place an emphasis on evaluating the approaches for sustaining national-scale change.

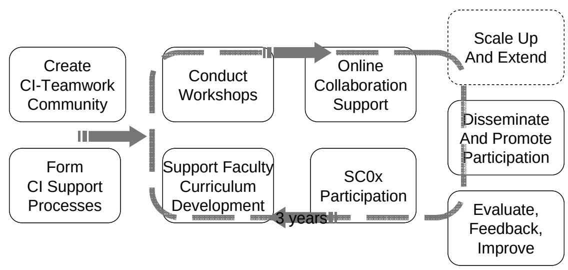

Figure 3: CI-TEAM-Work Implementation Activities

The process for addressing sustained change and scaling-up of programs will include discussions among campus administrators and faculty, OSG and TeraGrid science and technology leaders, HPC center directors, the Advisory Committee to this project, and the SC Steering Committee. The project will work with current and future CI-TEAM projects to regularly compare strategies and approaches. Approaches to be explored in depth will include adding new courses using cyberinfrastructure to the course catalog, helping faculty engage their peers in teaching similar course content, urging administrators to fund the campus cyberinfrastructure needed for sustained human and technological change, creating financial models for sustained training and mentoring, and engaging faculty and administrators as active practitioners in professional societies that advance the use of cyberinfrastructure including discipline specific societies and the SC Conference. These and other approaches will be discussed, prototyped, and evaluated for their potential for sustaining change. The project staff and participating institutions will produce a report with recommendations, evaluations and findings of their collaborative efforts to create sustained change.

Activity: Engage Diverse Communities of Practice. *(Teamwork and Workforce Components)* The project has selected 14 *Participating Institutions* (**bold** in Table 1 below) from seven geographically distributed regions around the country to engage multiple institutions in each region to further broaden the project's national impact (depicted as red circles indicating the anticipated "driving" range of regional impact in Figure 4). These institutions researchers, faculty and administrators have committed to participate in this project for three years. Their letters of commitment are included in the Appendix. These institutions form a diverse mix of two- and four-year colleges, minority serving institutions (MSI), and Research-1 universities. They will serve as workshop hosts for faculty and students from their campus and neighboring institutions to

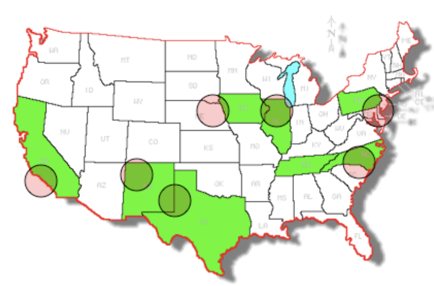

Figure 4: Regions of Participation

broaden the impact within each region. Institutions were chosen that will facilitate collaboration among all participants. *Collaborating Organizations* (ital. in Table 1) have agreed to offer substantial resources, cyberinfrastructure assistance, and STEM content as instructors and mentors working with the *Participating Institutions'* faculty, administrators and students.

Table 1: Participants and ***Collaborators***

San Diego, California Region
National University
San Diego State University (m)
San Diego Supercomputer Center, University of California San Diego/CalIT²
Washington, DC Region
Howard University (m)
Pittsburgh Supercomputing Center
Waukegan, Nebraska Region
Little Priest Tribal College (m, e)
Chicago, Illinois Region
Chicago State University (m)
Governors State University
University of Chicago (r)
Argonne National Laboratory, Krell Institute, National Center for Supercomputing App., Oak Ridge Nat. Lab., Purdue University
Durham, North Carolina Region
North Carolina Central University (m)
North Carolina Agricultural and Technical State University (m),
Wake Tech Community College (2)
Shodor Education Foundation, Inc.
Crownpoint, New Mexico Region
Crownpoint Institute of Technology (m)
El Paso, Texas Region
University of Texas at El Paso (m)
New Mexico State (m)
University of New Mexico (r, m, e)
Texas Advanced Computing Center
m=MSI, r=research-1, 2=2-year, e=EPSCoR

Activity: Use of Cyberinfrastructure. *(CI Component)* The project will utilize collaborative cyberinfrastructure to support the learning and use of cyberinfrastructure within research and education. The focal point will be the national shared Grid infrastructure – OSG and TeraGrid – and the software to facilitate easier access between the user's desktops and these Grids. Included will be the use of on-line collaboration systems such as discussion spaces (wiki forum - http://www.wiki.org/), learning environments (Moodle, http://moodle.org/), and videoconferencing (Access Grid - http://www.accessgrid.org/) to support collaboration, mentoring and on-line learning to complement and extend the interactions from the workshops, seminars and campus visits.

The collaboration environment teams from Boston Univ. (Jennifer Teig von Hoffman) and NCSA (Edee Wiziecki) have been instrumental in leading the Access Grid in Education Virtual Institute for EPIC (www.eotepic.org) and have been national leaders in working with institutions across the country to install, support and effectively utilize video conferencing technologies to support collaboration, teaching and learning. They will coordinate CI-TEAM-Work's efforts with the Partnering Institutions to ensure that they have ready access to a range of collaboration technologies.

The first year of the project will focus on working with the Participating Institutions to support the adoption and use of this infrastructure on the campuses. Training materials on cyberinfrastructure tools and resources will be developed and expanded for use in workshops and via on-line learning modules. Years two and three will utilize this infrastructure extensively to support on-line training events, seminars, consulting and mentoring relationships. The training materials developed for use in the live workshops will be adapted for on-line self-paced learning using the Moodle learning environment system, and other emerging technologies found to be most effective for self-paced learning and collaboration.

The TeraGrid software development team is working on the infrastructure needed to provide any faculty or student, at a school with Shibboleth [Shib] installed, with a Grid credential that could be used to access TeraGrid using their local campus login. Shibboleth will enable faculty and students to use TeraGrid for research and/or classes requiring any resource provided through TeraGrid. The institutions that can immediately benefit from this capability are those included on the following list: http://inqueue.internet2.edu/who/. Similar mechanisms will be explored for OSG access.

As the Grid infrastructure matures, a significant amount of change will occur over the next few years. To address the evolution of cyberinfrastructure this project will provide the community with a conceptual overview, with a set of course materials that allows the workforce development of more technical experts, and a set of teaching materials that address the use of the Grid cyberinfrastructure within all disciplines.

This development will involve experts in the appropriate discipline domains to address the unique domain needs, and to avoid redundancy of development across the nation. A wide set of materials will become available that address different levels of expertise and educational goals from the novice to the expert.

Activity: Conduct Week-long Summer Workshops. *(Training Component)* Week-long hands-on workshops will be held at an institution in each of the 7 regions each summer. Faculty, students and administrators from neighboring institutions will be encouraged to attend so that all 14 participating institutions are involved. These workshops are intended to introduce faculty, administrators, and students to the tools, resources, and methodologies for effectively using Grid cyberinfrastructure for research and education. The workshops will be led by faculty, scientists, educators and technologists from the collaborating institutions identified in Table 1. The content will tailored to adapt to the changing needs and requirements of the local participants. The workshops will be modeled on the hands-on approach of those offered by NCSI ([NCSI], OSG Summer Grid workshops [OSG], TeraGrid workshops [TeraGrid], and the astronomy Cyberinfrastructure workshop [Astro], and the International Grid Summer School. There will be plenary sessions and parallel tracks to address specific participant interests. These workshops collectively offer comprehensive end-user-focused training on the broad spectrum of what's needed to use the national shared grid infrastructure. These topics include networking prerequisites, grid concepts, service-oriented architecture, grid security infrastructure, grid organizational structures, job management and scheduling, data transport and cataloging, and workflow management.

Ever expanding collections of training material (summarized below) are available to help the participants understand closely related cyberinfrastructure, and how it can enhance scientific discovery. The workshops will blend a variety of these topics to best meet the needs of the participants. There will be general plenary sessions as well as parallel tracks for participants to delve into topics of greatest benefit to their needs and interests. The topics below have been covered in TeraGrid, OSG and NCSI workshops and found to be very popular with the community. Additional details about the topics are in the appendix.

Applications of technology to scientific practice:

- Wireless Sensor Networks characteristics, signal processing, and deployment
- Visualization - overview of hardware and software resources and services
- Introduction to the TeraGrid resources, services, and MPI performance
- The Virtual Data System for scientific workflows in grid environments
- Science Gateways using grid-based resources from multiple science domains
- Hierarchical Data Format (HDF5) for beginners and advanced HDF5 users
- Data Collections and their access, management, publication and usage
- Developing Web Services provides an introduction to programming Web services
- Collaboration technologies (discussion spaces, learning environments, videoconferencing)

Applications for learning science:

- First Look Interdisciplinary introduction to computational science methods
- Curriculum Development for significant curriculum development
- Computational Biology for Biology Educators for modeling and bioinformatics
- Computational Chemistry for Chemistry Educators in undergraduate courses
- Parallel, Cluster and Grid Computing for scientific computing
- Computational Science Concept Map of computational science skills

Sustaining and scaling-up cyberinfrastructure:

- Administrator Engagement to address why and how to support CI on campuses
- Process of identifying good practices for adoption in research and education
- Models for sustainability at the course level, department level, and the campus level
- Models for scaling-up good practices to other institutions

Additional topics will be developed based on the needs of the community from the gap analysis being conducted through the "Pathways to SC07" project during 2006, described in the leveraged efforts later in this document. Topics to be added will include:

- The role of Grid computing in scientific computing and computational science
- The creation of advanced portals for inter-disciplinary scientific discovery
- Examples of Grid computing applications in multiple scientific domains, such as watershed models to allow students to understand environmental complexity or genomics databases to explore metabolic pathway functions across organisms.

Activity. Provide curriculum development support. *(Education and Teamwork Components)* During the workshops, participants will be exposed to materials developed by and in use by other faculty in their disciplines. Experience from previous workshops show that faculty are interested in one or more of the following paths for integrating new technologies and methods into their own courses: 1) using other people's models and materials, with minor changes; 2) modifying others people's models and materials to be more appropriate for their students and their learning goals; and 3) developing new materials when they can't find any existing materials that address their classroom needs. Following the summer workshops each year, participating faculty will be provided the opportunity to request support (e.g., release time, software, materials) to explore one or more of these three paths to replicate, adapt or develop curriculum materials for their courses, fifteen grants will be awarded. Teams of faculty that jointly pursue interdisciplinary curriculum development efforts will be highly encouraged. Participants will be encouraged to involve students in the development effort. All faculty will be provided access to OSG and TeraGrid computing resources to augment their own campus resources. The faculty will be asked to test the materials in their own classroom, refine the materials based on that classroom experience, and share their materials and findings with the community.

Activity: Conduct regular campus visits. *(Advancement and Mentoring Components)* Follow-up visits will be made to seven regional sites in the fall and in the spring of each year to support the faculty working on curricular materials with on-going support and mentoring. Faculty, students and administrators from neighboring institutions will be encouraged to attend so that all 14 participating institutions are involved. During the visits, seminars and short workshops will be offered to engage faculty and students from the campus and the local region. Additionally, the campus visits will include meetings with campus administrators to help them address the long-term support issues required for sustained change within their institutions.

Activity: Support travel of faculty and students to annual SC conference. *(Training and Advancement Components)* Through the resources provided by the leveraged support of the SC Steering Committee and the SC Conference (described in more detail later in this document), teams of faculty and students from each participating institution will have their expenses covered to attend the annual SC Conference. The SC Education Program will structure a series of activities to broaden their exposure to the applications of state-of-the-art high performance computing and communications technologies for research and education. The faculty and students will also be encouraged to participate in the technical sessions, to visit the exhibit floor, and to submit their own papers and posters for presentation at the conference. Further, the faculty and students will be encouraged to join various conference committees to help shape and support future conferences.

Activity: Provide on-line tutorials and discussion forum. *(Training and Mentoring Components)* The project will publicize the existing on-line tutorials that have been developed among the collaborating institutions and many other organizations. For example, NCSA, SDSC, TACC, Purdue and Shodor have an extensive array of self-paced learning materials. Additional on-line learning materials will be developed for broad, on-demand access. The current support from the SC Steering Committee is supporting the development of those materials during the next year.

Between visits to the campuses, and throughout the three years of the project, there will be a technology consultant (at U. Chicago) and a computational science consultant (at Shodor) available to assist people adapting the technologies and methods into their research and educational practice. Researchers and educators will be encouraged to utilize the TeraGrid user support services as they use the Grid resources. OSG support for the CI-TEAM-Work virtual organization will be performed by the technology consultant and can hopefully be supplemented by OSG EOT resources in the currently pending OSG grant.

These support people will interact with the faculty, students and administrators using the wiki, moodle, and Access Grid collaboration cyberinfrastructure environments described above. The faculty, students and administrators will be empowered to share with one another and to interact with the project staff and instructors. The project staff will regularly reach out to the people at the participating institutions to share new information, offer assistance, and sustain regular communications. We will also encourage the use of existing mentoring networks, such as MentorNet, to provide under-served people with the support to encourage the continued pursuit of their studies and retain them in STEM fields.

Activity: Disseminate project materials. *(CI, Advancement and Teamwork Components)* All materials developed will be submitted to the Computational Science Education Reference Desk (CSERD), a Pathways project of the National Science Digital Library (NSDL) to foster broad dissemination of the materials. CSERD provides a Verification, Validation, and Accreditation (VV&A) quality review process on materials submitted to CSERD. Educators have indicated their preference for using materials reviewed by their peers as quality resources. The faculty from the participating institutions will be encouraged to become reviewers and to participate in the NSDL "ask the experts" program for their own advancement as leaders and contributors in their field.

In years two and three, the curricular materials developed in the previous years will be shared with faculty in the workshops. Experience has shown that many faculty will "use other people's models" in their own courses, or adapt other people's models to be appropriate for their own courses. People that adapt and modify these materials will be encouraged to submit their changes to CSERD for broad dissemination.

Evaluate project effectiveness. *(Teamwork, Advancement, and Workforce Components)* A key aspect of this project is regular and persistent attention to the needs of the diverse communities participating in the project. The needs will vary considerably among each campus, each department, and each individual. Beginning with the start of the project, the project team will initiate regular surveys, interviews and discussions with administrators, faculty and students among the participating institutions. This information will be used to determine how the project should evolve over time to best serve the needs of the communities.

For example, Little Priest Tribal College has provided us with a description of the needs and interests that led them to join this project. The availability of cyberinfrastructure tools and capabilities can have a powerful effect on the ability of tribal communities to eventually become part of the solution to the crisis concerning the STEM-educated work force in this country. Tribal colleges and universities are particularly interested in curriculum aspects of cyberinfrastructure as it relates to the potential for enhancing tribal college and university education. A model that seems to have the most promise in this regard for this institution is included in the appendix. National University envisions its role in the CI TEAM Work project as a partner for broadening the applicant pool of women and underrepresented minority populations in the STEM fields through a Cyber Infrastructure Education, Recruitment, Resource, and Access (CIERRA) component, described further in the appendix.

As evidenced through NCSI, we expect that faculty trained early in the project will become workshop instructors in subsequent years. This is a key factor in scaling-up the effort and providing a sustainable approach to broadening the adoption of cyberinfrastructure, computational science, and Grid computing resources and methods. The number of faculty becoming instructors is a metric of project effectiveness.

We will assess the effectiveness of the project to help the participating institutions with

- identifying tools and materials useful for research, teaching and learning,
- providing support to help faculty find and incorporate those materials,
- developing proposals to expand human capacity and physical cyberinfrastructure

An annual face-to-face meeting will be held with the Advisory Committee (described in the Management Plan in this document) to seek their feedback on the year's efforts, and their advice for planning for subsequent years. They will receive feedback from the external evaluator to assist in their review.

The University of Michigan (Ann Zimmerman) has been selected to conduct an external evaluation. Their evaluation plan is included later in this document.

6. Timeline of Major Project Milestones

Year 1 ***Q1**:* conduct user requirements and formative evaluation; develop workshop topics and schedule workshops and campus visits; develop strategies for sustainability
***Q2**:* conduct seven 1-week summer workshops, one per region; curriculum development projects selected and supported; faculty and students selected to attend SC|07; refine strategies for sustainability
Q3: conduct 7 campus visits – one per region; 100 faculty and students attend SC|07; implement practices that will support sustained change; define strategies for scaling-up to other institutions
***Q4**:* conduct 7 campus visits – one per region; implement procedures to engage other institutions in year 2; review progress towards sustainability and scaling-up; external evaluation report; annual meeting of Advisory Committee; report to NSF; submit materials to NSDL for dissemination

Year 2 ***Q1**:* conduct user requirements and formative evaluation; develop workshop topics and schedule workshops and campus visits; refine and implement additional strategies for sustainability and scaling-up
***Q2**:* conduct week-long summer workshops – 7 workshops – one per region; curriculum development projects selected and supported; curriculum development projects selected and supported; faculty and students selected to attend SC|08; continue to address sustainability and scaling-up issues and practices;
***Q3**:* conduct visits to each region – 7 campus visits – one per region; 100 faculty and students attend SC|08; identify benefits and challenges of sustainability and scaling-up;
***Q4**:* conduct visits to each region – 7 campus visits – one per region; engage more institutions to participate in year 3; external evaluation report; annual meeting of Advisory Committee; report to NSF; submit materials to NSDL for dissemination

Year 3 ***Q1**:* conduct user requirements and external evaluation; develop workshop topics and schedule workshops and campus visits; continue to address sustainability and scaling-up practices
***Q2**:* conduct week-long summer workshops - 7 workshops – one per region; curriculum development projects selected and supported; faculty and students selected to attend SC|09; develop recommendations for sustaining and scaling the program
***Q3**:* conduct visits to each region – 7 campus visits – one per region; 100 faculty and students attend SC|09 conference; publish recommendations and findings from efforts on sustainability and scaling
***Q4**:* conduct visits to each region – 7 campus visits – one per region; submit materials to NSDL for dissemination; final external evaluation report; annual meeting of Advisory Committee; final NSF report

Outcomes. The project will result in the following measurable impacts.

- Approximately 800 researchers, faculty and students engaged at 20 institutions to address the creation of a workforce capable of effectively using cyberinfrastructure in STEM fields
- New training materials developed and disseminated to introduce cyberinfrastructure and its applications to diverse research and education communities
- New curricular materials developed by at least 30 faculty, tested and revised through their classroom use, and quality review of the materials and disseminated through the National Science Digital Library and outreach by TeraGrid and OSG institutions
- Integration of collaborative technologies in use by the participating colleges and universities
- Attendance by a total of over 300 faculty and students at the annual SC|07-09 conferences

- New communities of practice among faculty and students established among multiple academic institutions – measurable through the use of on-line collaboration systems, surveys and interviews
- Recommendations and findings of efforts for sustained change and scaling-up of good practices
- Documented user requirements for TeraGrid, Open Science Grid, and other cyberinfrastructure providers among diverse communities for research and education

7. Project Management Team Structure and Budget Summary

The CI-TEAM-Work project is a virtual organization led by a team of nationally recognized experts from diverse organizations committed to meeting and exceeding the project goals.

Executive Committee. The project PI (Lathrop) will be the project director. He will take responsibility for the overall management of the project, management of finances, and coordinate interactions with NSF. The PI and co-PIs (Davis, Jacobs, Teller, Wilde, and von Laszewski) will serve as the Executive Committee for the project, making key decisions on all aspects of the project. Additional members of the Executive Committee will include Dr. Caesar R. Jackson (Dean of the College of Arts and Sciences at North Carolina Central University), Dr. Marcus Alfred (assistant professor of physics at Howard University), and Jerry Sheehan ($CalIT^2$ at the University of California San Diego).

The University of Chicago will lead the training development and coordinate technology applications content for the workshops (Wilde), and will serve as a facilitator to REU projects and participate in the outreach with the rest of the team to establish interest groups for the use of cyberinfrastructure in various scientific disciplines (von Laszewski). The Shodor Education Foundation (Panoff) will lead the curriculum development and coordinate the classroom learning content for the workshops. The PI (Lathrop) will lead the sustainability and scaling content for the workshops. A Shodor staff member will coordinate the project logistics, including scheduling of events, tracking expenses, posting materials to NSDL and the project web site. Staff at the U of Chicago will track and manage the project finances.

Advisory Committee. An Advisory Committee will provide regular and on-going advice to the Executive Committee on the goals and the plan to achieve those goals. The Advisory Group will include an administrator from each Partnering Institution to address sustained transformation and change within each institution. In addition, other experts from the community have agreed to serve including: Ruth Pordes (OSG), Dr. Geoffrey Fox (Indiana University), Dr. Eric Jakobsson (NCSA/University of Illinois at Urbana-Champaign), Dr. John Ziebarth (Krell Institute), Dr. Holly Hirst (Appalachian State University), Dr. Marilyn “Marty” McClelland (North Carolina Central University), Dr. Alex Ramirez (Hispanic Association of Colleges and Universities), and Dr. Ignatios Vakalis (Capital University).

CI-TEAM-Work Project Personnel. The CI-TEAM-Work project builds on a strong leadership team with many years of experience and success in conducting national education and outreach projects. The team has the experience in all of the dimensions required for this effort to deliver and succeed: visionary leadership, experienced, innovative, building communities of practice, experience managing collaborative projects of national scale, and considerable talent in helping faculty and students making effective use of computing in STEM fields. Below are brief bios for the project staff. Complete CVs are in the appendix.

Scott Lathrop *(PI)* is the Director of Education, Outreach and Training for the TeraGrid Project, a member of the SC Steering Committee, and co-PI on the NSF funded Computational Science Education Reference Desk Pathways project of the National Science Digital Library. His heart, his passion, and over 30 years of HPC outreach experience are embodied in this project. He commits his time and energy to engaging and broadening the community that participates in and benefits from this project.

Tom Davis *(co-PI)* is the new president of Little Priest Tribal College in Winnebago, Nebraska. President Davis has over 30 years of working with tribal colleges and other Indian organizations. He has held leadership in the American Indian Higher Education Consortium (AIHEC), World Indigenous Nations Higher Education Consortium, and Advanced Networking for Minority-Serving Institutions. He currently serves of the Executive Board and is chair of the STEM Committee for AIHEC.

Mike Wilde (co-PI) is co-PI of the NSF pilot project "Interactions in understanding the Universe", an outgrowth of the project QuarkNet and has pioneered with collaborator and PI Marjorie Bardeen of Fermilab the concept of "e-Labs" for classroom and "i-Labs" for informal (museum) settings. Wilde served as GriPhyN project coordinator and organizes the US summer workshops in Grid computing for undergraduate/graduate students. He serves as OSG EOT coordinator, and his research in scientific workflows and provenance is directly applicable to the scientific education and discovery process.

Patricia Jacobs *(co-PI)* is the Project Manager for Computational Science Education Reference Desk (CSERD), a Pathways project of the National Science Digital Library (NSDL) program. She has developed Computational Science curriculum and trained workshop instructors for the Pathways to CyberInfrastructure project funded by CI-TEAM. She is the Women and Mathematics (WAM) Program Mentor of Durham County, in which she participates in encouraging young women interested in mathematics, science, and technology to take more challenging courses and pursue careers in these fields.

Patricia J. Teller (co-PI) has been involved in high performance computing since 1980, and is actively involved in local, regional, and national research and education efforts. She was awarded a Center for Effective Teaching and Learning Fellowship (2001) in recognition of her excellent teaching. She was the elected chair (2004-2005) of the Coalition to Diversity Computing (CDC). Her current research projects are funded by the DoD, DoE Office of Science, IBM Corporation, and NSF. She is a member of the SC Conference Steering Committee, and elected General Chair of SC|08.

Gregor von Laszewski (co-PI*)* served as Co-PI of the NSF REU Site for Grid Computing and Bio Informatics in which over the last years selected undergraduate and graduate seniors had the opportunity to participate in hands on interdisciplinary research projects focused on Grid Computing and BioInformatics. In addition, he has been active for over ten years in the education of the workforce for Grid related projects. Tools developed by his group to utilize advanced Grid cyberinfrastructure are used by many scientific and educational users and lower the entry barrier to Grid computing.

Robert Panoff is founder of Shodor Education Foundation, Inc. a non-profit research and education corporation with nearly $2M annually in federal, state, and private funding to undertake interdisciplinary research in the appropriate uses of technology in education. Director of NCSI, a comprehensive, interdisciplinary training program for undergraduate faculty funded by NSFs Division of Undergraduate Education. Director of the Computational Science Education Reference Desk, a Pathway project of the National Science Digital Library. Co-founder of the NSF Corporate and Foundation Alliance, bringing together leaders from NSF, industry and foundations to address critical issues in undergraduate education.

Ann Zimmerman (external evaluator) is a Research Investigator with the Collaboratory for Research on Electronic Work (CREW), of the University of Michigan's School of Information. She is conducting the evaluation of TeraGrid, and is a member of the research team of a 5-year NIH-funded project to develop conceptual models, computational infrastructure, and knowledge bases for complex biomedical processes.

Jennifer Teig von Hoffman is the Assistant Director of Scientific Computing and Visualization at Boston University. She is a co-PI of the NSF Broadening Participation in Computing project, "New Voices and New Visions for Engaging Native American Students in Computer Science." She organized one of the first Access Grid (AG) conferences in 1999, chaired SC Global component of the SC|03 Conference, and co-chaired the Access Grid performance and presentation space at SIGGRAPH 2005.

Edee Norman Wiziecki, Coordinator of Education Programs at NCSA, is the co-Lead of the Virtual Institute for using the AG in Education for the NSF Engaging People in Cyberinfrastructure Program. She is co-PI of the NSF REVITALISE Program, creating a virtual learning environment using AG technology and Moodle learning environments for rural K-12 science teachers. She brings 25 years of experience as a teacher, curriculum developer, and professional development leader for K-20 communities.

Dr. Caesar R. Jackson is the Dean of the College of Arts and Sciences at North Carolina Central University (NCCU). He joined NCCU in 2005 after 13 years at NCA&T as Interim Dean of the College of Arts and Sciences 2002-2005, Associate Dean for Research and Graduate Programs in the College

1998-2000, Chair of the Department of Physics 1994-1998, and Assistant Professor 1992-1994. He has been engaged in nuclear physics research at Triangle Universities Nuclear Laboratory in Durham, NC and at Thomas Jefferson National Laboratory in Newport News, VA. He was recently involved in geophysical science research at NCA&T and was the PI and Director of the NSF funded project TALENT-21.

Dr. Marcus Alfred is an assistant professor of physics at Howard University. His research areas include high energy physics, physics education research, biophysics and computational electromagnetics (CEM). Alfred's work has funded several students who subsequently graduated; two physics undergraduates and two graduate students were also supported, all from under-represented populations.

Project Budget Summary. CI-TEAM-Work's use of funds leverages existing resources to the maximum extent by collaborating with and engaging related projects. All co-PIs draw only token support from this program, and contribute as part of their roles through closely related projects, which will also benefit greatly from this investment. The NSF budget components include: Management (PIs, materials, advisory meetings, and staff travel) – 22%, support of cyberinfrastructure among sites (wiki, moodle, AG) – 11%, technology person (U Chicago) and curriculum person (Shodor) for advice and support to campuses – 19%, participant costs (workshops, campus visits, curriculum development) – 38%, and external evaluation – 10%. The expected matching contribution from the SC Education Program budget (est. $985K over three years) includes 43% to support participants at three SC conferences, 42% to support workshops and visits at the campuses, and 15% to support logistics and evaluation.

8. Collaboration Plan: Leveraging Partnerships for Success

This project builds on the significant experience of its collaborators in research and educational efforts that link large-scale experiments and grid computing to hands-on training and education programs. The work is based on the collective experience and talents of the collaborating institutions and organizations including TeraGrid (and its Resource Provider partners), Open Science Grid, the National Computational Science Institute, GriPhyN, iVDGL, the SC|xx Education Programs, QuarkNet, I2U2, the TeraGrid Science Gateways, the Krell Institute, and CalIT^2. Many of the institutions participating in this proposal are already members of the OSG and TeraGrid infrastructures.

Science Drivers. The project will work with scientific teams that are pushing the envelope on the applications of cyberinfrastructure in support of research. These science drivers are each inherently interdisciplinary in nature. They serve as exemplars for setting the requirements for the range of training and education needed to prepare people to contribute to advancing research in each of these disciplines. This project will work with the science drivers to identify emerging tools and methods that should be incorporated into the training and education activities. In addition, we will explore opportunities for internships within the science drivers for students and faculty to both practice what they have learned and expand their skills and knowledge. We are in contact with the following science drivers:

- Math and Physical Sciences – GryPhyN, iVDGL, ATLAS, Mathematical Sciences Pathway, Materials Science Pathway and Physics and Astronomy Pathway of NSDL
- Geosciences – GEON, DLESE of NSDL
- Biosciences – NEON, BEN Biological Sciences Pathway of NSDL
- Engineering – Neutron Science, Engineering Pathway of NSDL

We will leverage the significant resources which we have developed in the joint efforts of the collaborating partners to date: Virtual data mechanisms from GriPhyN that enable productive leverage of Grid systems by cataloging datasets and applications and processing them on the Grid with automated workflow tools; Commodity Grid (CoG) Kits; Grid education workshops with hands-on exercise material that we have developed and applied in three annual summer workshops for undergraduates and graduate students in the sciences and computer science; mentorship programs with undergraduate students through the REU program; and educational workshops that we have organized in the annual Supercomputing conference, the flagship international conference in high performance computing.

Collaboration with other Proposed CI-TEAM Projects. The CI-TEAM-Work project will proactively engage current and future C-TEAM projects to explore information exchange so that we can learn from one another, avoid duplication, and share good practices. These exchanges will occur through conferences, workshops, AG sessions, conference calls, and on-line discussions among the projects.

We will work with "The PetaScale Training Network" proposal led by the Pittsburgh Supercomputing Center, which involves the San Diego Supercomputing Center, the National Center for Supercomputing Applications, and the Texas Advanced Computing Center. That proposal complements this project by focusing on training on petascale computing by assisting researchers with code optimization, scaling issues, performance monitoring, and other related high performance aspects of scientific discovery. CI-TEAM-Work will be a feeder into the activities of the PetaScale project providing further support along the continuum of computing and scientific practice, as referenced in the vision and goals section above.

International Collaboration. The project will collaborate with international organizations to foster dialogue and the exchange of ideas and resources on projects with similar goals and interests. For example, TeraGrid is an active member of the Global Grid Forum (GGF) which has launched an international education, outreach and training effort. The GGF Education and Training Community Group, led by Malcolm Atkinson, met at the meeting in Greece to develop stronger collaborations among US and international organizations. The group is pursuing collaborative activities for future GGF meetings, including the September meeting in Washington, DC. Multiple members of the TeraGrid and OSG EOT working groups will be involved in GGF activities. Discussions are also underway with the Latin American Grid (LA Grid) to foster the interchange of information and collaborative opportunities.

Industrial Partnerships. A major benefit from being associated with the SC Conference is the ability to forge partnerships with the industries that participate in the SC Conference series. We will work with the Industrial Advisory Committee of the SC Conference to share the Education Program and CI-TEAM-Work plans to secure their support in the programs. To date the SC Education Programs have been very successful in raising cash contributions, volunteer staff expertise, and other in-kind contributions to augment the impact of the program. We will also leverage the partnerships that TeraGrid and OSG have with major vendors to secure additional support to enhance the program and broaden the impact.

9. Project Risk Analysis and Management Plan

The project team has identified the following potential risks associated with achieving the project goals, and the strategies designed to mitigate these risks.

Risk: Contingency of SC matching funds. This project has secured support from the SC Steering Committee to begin to collect and prepare training materials and curricular materials to be used in the workshops supported by the SC07-09 Education Programs and the CI-TEAM-Work workshops . The initial support includes $197K for the Pathways to SC07 project – the proposal for that effort is included in the appendix. The proposed plans for the SC07-09 Education Programs are also in the appendix.

The SC Steering Committee has recommended funding a second year of the Pathways project, pending sufficient finances following the conclusion of the SC|06 conference in November, 2006. We will know if the SC conference will support Pathways year 2 and the SC07 Education Program by January, 2007. The level of financial support for the SC08 and SC09 Education Programs depend on the financial success of SC|07 and SC|08. The conference has been on solid financial ground with attendance steadily increasing, and the requested level of support for the SC07-09 Education Program is on par with the level of support provided over the last three years. The SC|07 and SC|08 conference chairs have endorsed these plans. The SC|09 Education Program direction depends on the selection of the SC|09 chair.

The CI-TEAM-Work project is not dependent on the SC support to meet its goals, except for covering travel and housing for the participants to attend the SC conference. Without SC support, this project would reduce the number of workshops and campus visits.

Risk: Sufficient resources for engaging experts and mentors. This project builds on the model developed by the National Computational Science Institute (NCSI), which has developed a cadre of faculty around the country who are now the instructors for many of the sessions. These faculty agree to lead the instruction at multiple workshops each summer. They are compensated for their time, travel, and expenses in return for their efforts. The CI-TEAM-Work project will engage these and many other faculty to be instructors at workshops, to visit the campuses, and to provide on-line mentoring and support. The list of faculty that have been NCSI instructors are included in the Appendix.

Risk: Sufficient support of faculty for classroom change. The project will fund about half of the time, travel and expenses of a technology expert and a science curriculum expert to develop materials, conduct training, visit the campuses, and provide on-line mentoring and support. These two people will serve as facilitators for connecting faculty from the participating institutions to the instructors with the most relevant background and experience as needed. These people will also actively engage the faculty at the institutions on a regular basis to ensure they are receiving the support they need throughout the year.

10. External Evaluation and Dissemination Plan

External Evaluation Plan. The University of Michigan (Ann Zimmerman) will conduct an external evaluation for the project. This group is familiar with evaluation within interdisciplinary settings using cyberinfrastructure and they are conducting an external evaluation of TeraGrid.

In order to measure CI-TEAM-Work's progress in meeting its goals, the external evaluator will conduct an evaluation organized into formative and summative components. The formative component will generate feedback to the project team and to key stakeholders about what works and what does not, in order to help modify the implementation plan or redesign activities to increase the likelihood of meeting the goals. The summative component will be an objective analysis of outcomes, as captured using qualitative and quantitative measures to analyze changes in baseline over time. The evaluation will consist of three elements that evaluate the following: 1) progress in meeting education and training needs, 2) progress in integrating cyberinfrastructure into education and research, and 3) participant satisfaction. The following section provides more detail of each component.

Evaluating progress in meeting education and training needs. The evaluator will identify broad needs and priorities for education and training compared against CI-TEAM-Work plans and priorities. The needs analysis will include 1) interviews to guide development of survey questions and to provide detail about user education and training needs that cannot be captured through questionnaire responses, and 2) responses to an Internet-based survey to identify needs and priorities of the participants. The survey will be given to a statistical sample stratified across participating institutions in years 1 and 2 of the project.

Progress will also be evaluated by participants attending education and training events through surveys of the following issues: 1) quality of instruction and materials; 2) appropriateness of level of presentation; 3) likelihood of participation in future events; and 4) level of enthusiasm for cyberinfrastructure before and after the event. The evaluator will develop the survey and analyze all results. The survey will be administered by project staff during education and training activities. Representative evaluation questions for this portion of the evaluation include: Do project plans and activities match user needs and requirements? How do participants judge the effectiveness and quality of the activities and materials? Are milestones related to the number of training workshops and campus visits being met?

Evaluating progress in integrating cyberinfrastructure into education and research. It is not possible to transform the research and education system within the 3-year timeframe of this proposal. It is possible, however, to assess progress toward this goal by evaluating the impact of CI-TEAM-Work activities on institutional commitment, curriculum and research practice. Representative evaluation questions include: How does attendance at sessions impact the utilization of cyberinfrastructure? Does participation in curriculum development activities impact development of curricular materials? Do the activities impact administrator commitment to high performance computing on their campuses?

These questions will be addressed through interviews, surveys, and analysis of project data. Interviews will be conducted with a sample of faculty attending education and training events to evaluate how (or if) CI-TEAM-Work changes the content of courses. Quantitative measures will track the number of new courses developed, implemented, and added to an institution's course catalog; the number of new training materials developed and used; the number of students enrolled in the courses; and the number of requests for access to Grid resources. To evaluate the influence of CI-TEAM-Work activities on institutional commitment to cyberinfrastructure, administrators at participating institutions will be surveyed before the program begins, mid-way through the program, and at the end of the project. Interviews will be conducted with a subset of administrators to guide development of survey questions and provide fine-grained detail.

Evaluating participant Satisfaction. The evaluation will focus on the degree of satisfaction among key participants in the project: participating institutions, project team, and collaborating institutions. Participant satisfaction data will be collected through a survey, interviews, and through analysis of communications and interactions between project participants, such as through the on-line collaboration system. Survey questions will collect data on the level of satisfaction and interaction among the groups. We will gather Participating Institution opinions about Collaborating Institutions and project personnel. Survey items will ask about the nature of the relationship, partner responsiveness, degree of perceived coordination between partners, and types and frequency of interaction. Representative evaluation questions for this portion of the evaluation include:

- To what extent does participation in the project contribute to communication and/or collaboration among participating institutions?
- How do individuals at participating institutions judge the effectiveness of the support provided by CI-TEAM-Work (on-line, face-to-face, and telephone)?

The evaluator will prepare an annual report. The report will contain a summative evaluation focusing on changes to baseline over time using the data from the formative evaluation activity.

Dissemination Plan. The CI-TEAM-Work project will host a website for news, schedule of events, multi-media resources, and other information of interest to scientists, educators, students, and the public, as well as all training and curricular materials. In addition, the material will be submitted to the Computational Science Education Reference Desk (CSERD), a Pathways project of the National Science Digital Library (NSDL) to foster broad dissemination of the materials.

Project personnel will present project activities at various conferences and meetings. The project team will utilize the Access Grid to make workshops accessible to other interested organizations. We intend to distribute information about our activities in the newsletter “CyberInfoBeat” that was launched by The Engaging People in Cyberinfrastructure (EPIC) project. It is an electronic newsletter that focuses on education, outreach and training news for the community. The newsletter sponsors have included EPIC, TeraGrid, the National Science Digital Library, OSG/Science Grid This Week, and the National Computational Science Institute. This project will ensure the continued publication of this newsletter and seek to expand the readership and engage additional contributors to the newsletter content.

We have initiated discussions to publish the project’s computational science resources and methods. The discussions to date have received positive support for creating a series of texts on “Computational X”, e.g. Computational Chemistry, Computational Biology, and so on for as many disciplines as the team is able to develop materials for during the three years of the project. Some of the most useful material we hope to publish in book form. Several positive discussions with well known publishers have taken place.

11. Conclusion

CI-TEAM-Work will achieve sustained, systemic change in scientific teaching, learning and practice in US colleges and universities and contribute to the growth in size and skill of our STEM workforce. Its **intellectual merit** derives from a highly qualified and diverse staff with strong technical and scientific backgrounds, significant experience in CI use and training, and in *achieving* diversity in participation. The project efficiently leverages resources and contributions from related collaborations. Its likelihood of success derives from empowering faculty through collaboration and community to use CI in an engaged, non-prescriptive manner. The project is simple and scalable, leverages power of *teamwork*, and thus provides a network for support and sharing. Its significant **broader impacts** include engagement of diverse communities across multiple dimensions, facilitation of research and thus faculty advancement in addition to education, and valuable feedback and improvement on extending the usability and *use* of national Grid cyberinfrastructure at the undergraduate level.