OFFICE OF CYBERINFRASTRUCTURE

$219,000,000
+$19,720,000 / 9.9%

Office of Cyberinfrastructure Funding
(Dollars in Millions)

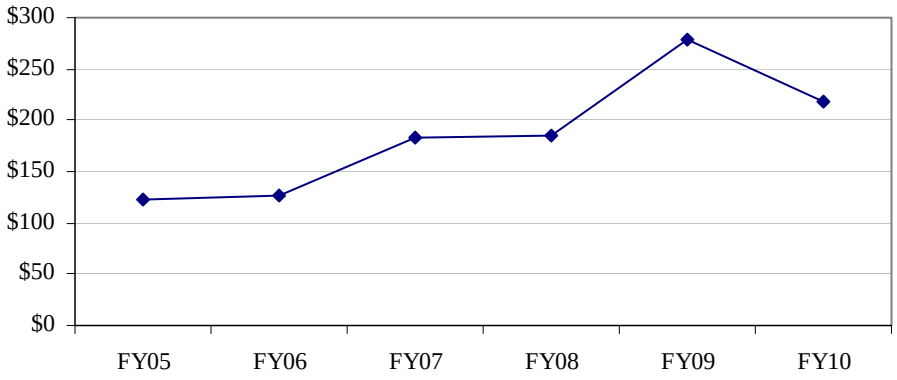
	FY 2008 Actual	FY 2009 Current Plan	FY 2009 ARRA Estimate	FY 2010 Request	Change Over FY 2009 Plan Amount	Change Over FY 2009 Plan Percent
Total, OCI	**$185.15**	**$199.28**	**$80.00**	**$219.00**	**$19.72**	**9.9%**
Major Components:						
Research and Education Grants	31.48	38.63	63.00	58.03	19.40	50.2%
Track-1 (Blue Waters)	13.65	45.23	-	97.00	51.77	114.4%
Track-2	14.19	68.73	-	26.95	-41.78	-60.8%

The Track-1 activity is designed to provide researchers with a sequence of leadership-class systems at intervals of four to five years. The Track-2 activity is designed to maintain a diverse portfolio of national-scale supercomputing and storage infrastructure for the research and education community.

The Office of Cyberinfrastructure (OCI) supports research, development, acquisition, and operation of advanced shared and connecting cyberinfrastructure (CI) that enables otherwise unrealizable advances in 21st century science and engineering research and education. It increasingly supports the use of advanced CI to attack frontier science problems through the growing discipline of computational science and engineering, as well as the computational scientists who develop and use it. OCI capitalizes on a broad range of fundamental scientific and engineering research as well as education and research in other directorates to create and expand the next generation of CI. This CI is critical to converting data to knowledge, understanding complexity through simulation and prediction, and creating more systematic knowledge about the social and technical issues of large-scale, multidisciplinary collaborative communities, known as virtual organizations, needed to address complex problems and grand challenges facing science and society.

OCI-supported CI includes the comprehensive set of deployable hardware, software, and algorithmic tools that support research, education, and increasingly collaboration across and among all research disciplines, whether they are experimental, theoretical, and/or computational. CI consists of computing systems, data storage systems, data repositories and advanced instruments, visualization environments, people, and the necessary intellectual capital, all linked together by software and advanced networks to sustain and improve scholarly productivity and enable breakthroughs in complex problem solving. OCI supports socio-technical research on the use of CI and on ways of improving its effectiveness. It supports training in the development and use of advanced CI as well as research on its use to enhance learning. OCI also supports the scientific and engineering professionals who create and maintain these leading-edge resources and systems, and who provide the Nation's researchers and educators with essential CI services. OCI makes investments that improve CI for science and engineering research, funding the deployment of current CI and innovative developments in future CI. In doing so, it both leverages and complements investments made by other federal agencies. For example, some of NSF's high-end computing investments take advantage of expertise at laboratories funded by the Department of Energy (DOE) and hardware and software developments funded by the Department of Defense's Defense Advanced Research Projects Agency (DARPA). In addition, OCI investments in petascale applications and tools complement those of DOE's Scientific Discovery through Advanced Computing (SciDAC) program, and OCI's TeraGrid infrastructure is used by researchers funded by the National Institutes of Health (NIH), DOE, and other agencies.

OCI Funding
(Dollars in Millions)



OCI in Context

OCI supports the development and deployment of cyberinfrastructure (CI) that is shared by all scientific and engineering disciplines, making possible potentially transformative basic research in areas such as nanotechnology, physics, chemistry, materials science, sustainable energy, climate/weather, and engineering. It also promotes interoperability between components of CI both here in the U.S. and abroad. About two-thirds of NSF's investments in CI are made by the directorates and offices responsible for fundamental domain specific research and education in science and engineering, with the remaining one-third provided by OCI. Through coordinated planning and investments facilitated by NSF's Cyberinfrastructure Council, OCI works to support integrated applications and teams that use advanced CI to solve complex multidisciplinary science and engineering problems, providing economies of scale and scope to ensure that NSF's CI portfolio delivers the highest returns on the Nation's investment.

OCI's investments are guided by NSF's *Cyberinfrastructure Vision for 21st Century Discovery* (www.nsf.gov/dir/index.jsp?org=OCI), a comprehensive CI strategic plan for the Foundation; by many blue-ribbon panel and advisory reports, such as the 2005 Presidential Information Technology Advisory Committee (PITAC) (www.nitrd.gov/Pitac/reports/20050609_computational/computational.pdf) report; by the America COMPETES Act; and by the opportunities identified by the academic and industrial research community through workshops and white-papers. The America COMPETES Act calls for the Foundation to conduct long-term basic and applied research on high-performance computing and networking, and OCI's investments are central to advancing that goal.

OCI activities are key components in the federal government's Networking and Information Technology Research and Development (NITRD) program. The technologies developed and the systems deployed by OCI facilitate discovery and innovation and bolster national competitiveness. PITAC specifically recommended in 2005 that federal agencies reorganize to support more effectively multidisciplinary computational science, which it called the "third pillar" of science and engineering of the 21st century. OCI was created in 2005, and is now specifically taking on a role as steward of computational science activities in coordination with directorates and offices across the Foundation.

The FY 2010 Request for OCI includes $2.0 million to leverage activities aimed at increasing support for transformative research. Examples of potential foci for these investments include the Strategic

Technologies for Cyberinfrastructure (STCI) Program and the EArly-concept Grants for Exploratory Research (EAGER). The primary purpose of the STCI Program is to support work leading to the development and/or demonstration of innovative cyberinfrastructure services for science and engineering research and education that fill gaps left by more targeted funding opportunities. In addition, STCI considers highly innovative cyberinfrastructure education, outreach and training proposals that lie outside the scope of targeted solicitations. The EAGER mechanism supports high-risk, exploratory and potentially transformative research.

Office-wide Changes and Priorities

Learning (+$8.45 million to a total of $13.10 million).

Consistent with NSF's Cyberinfrastructure Vision for 21st Century Discovery, the President's Council of Advisors on Science and Technology (PCAST) 2007 report and in response to OCI's 2008 Committee of Visitors (COV) recommendation, OCI will significantly expand its participation in the area of learning and workforce development. This will be accomplished by co-funding OCI mission consistent Graduate Research Fellowships (GRF), and expanding participation in the REU program by initiating support for REU sites in addition to providing ongoing support for REU supplements. These are all new areas for OCI investments. In addition, the OCI 2008 COV identified the lack of an "extensive workforce initiative" in the context of Learning activities. During FY 2010, OCI shall introduce an education and training program exploiting the dual nature of CI, i.e. both as a resource for teaching and as a resource for learning. Such an activity is important to development of a 21st century workforce.

OCI Disciplinary & Interdisciplinary Research (+$7.04 million to a total of $40.70 million).

The 2008 OCI COV report states "... in order for [OCI] to create cutting-edge research infrastructure, OCI must actively and deeply engage in the research itself." Addressing that recommendation, OCI will augment its Cyber-enabled Discovery and Innovation (CDI) funding and complement it with investments in several research areas within OCI, in partnership with directorates across the Foundation. Of particular interest will be programs in software, large data, and scientific applications that require advanced CI. In conjunction with NSF's other research directorates, OCI will participate in Science and Engineering Beyond Moore's Law (SEBML), and in Climate Research. In addition, OCI will increase its investments in supporting grand challenge communities and virtual organizations as well as in programs seeking innovative technologies in CI. As OCI begins to play a stronger role in computational science, it will actively increase its participation in NSF's research programs in partnerships with other units across the Foundation, and expects to catalyze new multidisciplinary programs in computational sciences in the future.

CAREER (+$3.88 million to a total of $3.88 million).

The CAREER program remains the primary mechanism for jump-starting junior faculty toward independent careers in research and education. The program has been very successful in supporting traditional areas of science, but less so in computational science, an area so important to the future of science and engineering that it has been called the "third pillar" of scientific inquiry. In order to address this, OCI will support CAREER awards in computational science, across all disciplines for outstanding young faculty who prototype and develop the next generation cyberinfrastructure, and/or apply it to advance their basic science disciplines.

Stewardship (+$67,000 to a total of $4.32 million).

A number of activities are funded directly from NSF's programs to advance NSF's Stewardship goal. These include Intergovernmental Personnel Act appointments, NSF-wide studies and evaluations, and mission-related information technology investments.

Networking and Computational Resources Infrastructure and Services (-$353,000 to a total of $156.65 million)

OCI FY 2010 infrastructure activities are highly dynamic, principally as a direct result of the NSF Track-2 awards acquisition phase winding down, the transition to operations and maintenance in these awards, and the beginning of significant milestone payments to the petascale Track-1 award coming due. As OCI expands activities in computational science research, OCI expects to increase investments in software both at the infrastructure level and to support complex application development in and among the various science domains. OCI will strengthen its broadband investments that cover research and development in high-performance network deployment, tools, and scientific applications. The goal is to "raise the bar" in national and international end-to-end network connectivity and availability in support of scientific research and education. Activities funded in this program are expected to result in an array of new integrated CI capabilities ranging from major advances in network performance, domestic connectivity for institutions in need, new global partnerships in international networking, and leadership class operational awareness in network enabled scientific applications and collaboration. FY 2010 represents a transition year between current TeraGrid award activities and the new eXtreme Digital (XD), also known as TeraGrid III. There are TeraGrid resources that will transition off while new, less costly, resources will be made available through the most recent Track-2 acquisition. In addition support for some programs that were initiated prior to the creation of OCI (by the division of Shared Cyberinfrastructure in the Directorate for Computer and Information Science and Engineering) ends in FY 2009.

- *Track-1 and Track-2 HPC Activities (+$9.98 million to a total of $123.95 million).* Driven by the needs of the science and engineering communities, two complementary "tracks" for provisioning high-performance computing, storage and personnel resources are currently in progress. The Track-1 activity is designed to provide researchers with a sequence of leadership-class systems at intervals of four to five years. The first Track-1 award will support the deployment of a system capable of sustaining performance exceeding a petaflop per second on very difficult problems in a range of research areas, permitting ground-breaking research in areas as diverse as astrophysics, climate, engineering and life sciences. This award, of approximately $208 million, was made in 2007 to the University of Illinois at Urbana-Champaign (UIUC) in a partnership with IBM to develop a system to be named Blue Waters. The first set of deliverables from IBM, relating mostly to software, were received by UIUC in Fall 2008 and were subsequently reviewed by the project team, an external review panel, and NSF. The IBM deliverables were accepted by UIUC in April 2009. The schedule for the commencement of full operations remains mid-2011. UIUC is constructing a specially designed facility to house Blue Waters and its support staff. Begun in October 2008, it is on schedule, due to be completed in 2010.

  The Track-2 activity is designed to maintain a diverse portfolio of national-scale supercomputing and storage infrastructure for the research and education community. The first phase of this activity included four annual solicitations beginning in FY 2005 and ending in FY 2009. No solicitation is planned for FY 2010 and the next phase of high-end computing competitions will begin in FY 2011. The first Track-2 award was made in FY 2006 to the Texas Advanced Computing Center for a total of approximately $59 million for acquisition and operation of a system, named Ranger, built by Sun Microsystems Inc. This entered full service in February 2008. The second award, of $64.4 million for the acquisition and operation of a sequence of systems built by Cray Inc., was made in FY 2007 to the University of Tennessee at Knoxville. The first full production phase, a Cray XT4, entered

service in August 2008. The next phase, a Cray XT5, entered production use in February 2009. A final upgrade is scheduled for the second half of 2009. The third Track-2 award is currently being negotiated with the Pittsburgh Supercomputing Center (PSC) and is expected to enter production service in 2010. The final round of the first phase of Track-2 competitions anticipated making several awards intended to provide the research community with access to innovative and experimental production computing architectures. Two awards resulting from this round are currently being negotiated with the potential awardees. One is for a data-intensive computing system, scheduled to enter full production in 2011. The other is for an experimental high-performance computing system with a parallel architecture that includes many-core accelerators. At least one more award from the fourth Track-2 competition is anticipated.

The Office of Cyberinfrastructure is working with NSF's Advisory Committee for Cyberinfrastructure (ACCI) to gather input from the researchers and educators who use computing and the technologists who develop high-performance computing on future requirements and opportunities for the national CI.

Program Evaluation and Performance Improvement

The Performance Information chapter describes the Foundation's performance evaluation framework, which is built upon the four strategic outcome goals in NSF's Strategic Plan: *Discovery, Learning, Research Infrastructure*, and *Stewardship*. Performance evaluation is conducted at all levels within the Foundation, using both qualitative and quantitative measures – including an agency-wide annual review of research and education outcomes by an external expert committee and periodic reviews of programs and portfolios of programs by external Committees of Visitors and directorate Advisory Committees. Other performance indicators, such as funding rates, award size and duration, and numbers of people supported on research and education grants, are also factored into the performance assessment process.

To ensure the highest quality in processing and recommending proposals for awards, OCI convenes a Committee of Visitors (COV), composed of qualified external evaluators, to review each program every three years. These experts assess the integrity and efficiency of the processes for proposal review and provide a retrospective assessment of the quality of results of NSF's investments. The most recent OCI Committee of Visitors met in FY 2008. The COV focused on two specific areas, in the context of OCI's four focus areas of High Performance Computing, Data, Virtual Organizations, and Learning and Workforce Development: (1) assessments of the quality and integrity of program operations and program-level technical and managerial matters pertaining to proposal decisions; and (2) comments on how the outputs and outcomes generated by awardees have contributed to the attainment of NSF's mission and strategic outcome goals. The COV made a number of recommendations that OCI is currently working to address.

Number of People Involved in OCI Activities

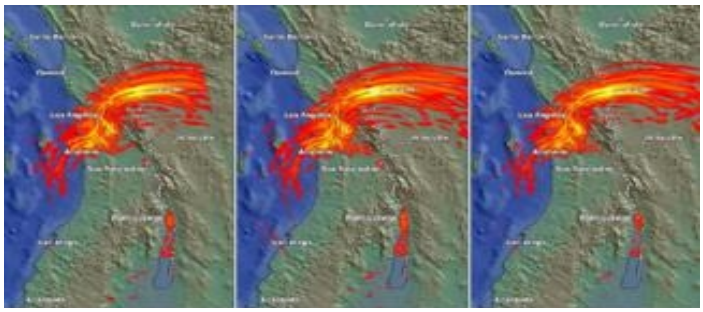
	FY 2008 Estimate	FY 2009 Estimate	FY 2009 ARRA Estimate	FY 2010 Estimate
Senior Researchers	535	395	158	448
Other Professionals	596	628	251	712
Postdoctorates	51	71	29	80
Graduate Students	252	196	78	222
Undergraduate Students	67	60	24	68
Total Number of People	1,501	1,350	540	1,530

OCI Funding Profile

	FY 2008 Estimate	FY 2009 Estimate	FY 2010 Estimate
Statistics for Competitive Awards:			
Number of Proposals	501	555	580
Number of New Awards	98	147	119
Regular Appropriation	98	105	119
ARRA	-	42	-
Funding Rate	20%	26%	21%
Statistics for Research Grants:			
Number of Research Grant Proposals	482	518	558
Number of Research Grants	80	124	98
Regular Appropriation	80	86	98
ARRA	-	38	-
Funding Rate	17%	24%	18%
Median Annualized Award Size	179,398	192,797	212,188
Average Annualized Award Size	221,237	237,761	261,674
Average Award Duration, in years	2.3	2.7	2.5

Recent Research Highlights

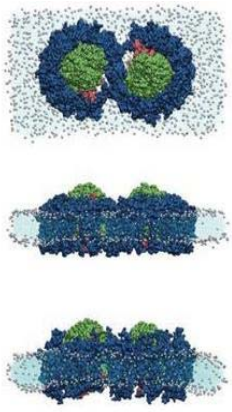
► **ShakeOut Exercise Prepares Californians for Earthquakes:** The Great Southern California ShakeOut, held in November 2008, was the largest earthquake drill in U.S. history. The exercise was designed to inspire Southern Californians to get ready for major earthquakes and, through public preparedness, to prevent disasters from becoming catastrophes.



Comparison of ShakeOut Simulations performed at different sites using different codes. The good agreement between the results helped to build confidence in the ground motions projected by the simulations for the ShakeOut scenario event. TeraGrid resources at TACC and PSC were used to run these 1Hz simulations. *Credit: SCEC.*

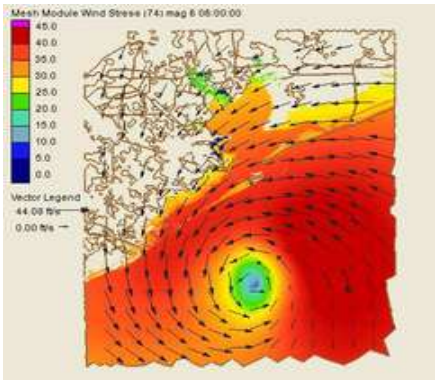
The ShakeOut Scenario defines a plausible Mw7.8 earthquake on the southern San Andreas Fault. Scientists and computer scientists from the Southern California Earthquake Center have performed several large-scale simulations using TeraGrid computer resources at the Texas Advanced Computer Center, San Diego Supercomputer Center, and Pittsburgh Supercomputer Center to ensure that the scenario was as realistic as possible. Simulations performed in support of the scenario extend from basic research to earthquake rupture processes, to collaborative simulations to verify results, and to innovative visualizations that help the general public understand the exercise and its implications. This exercise involved several million people in California.

► **The Computational Microscope:** Simulating cellular molecules from the atomic level up is essential to comprehending the human body and to designing effective medicines and treatments. The functions of biomolecules like proteins and DNA are well known but certain aspects of these molecules still elude researchers even with the aid of powerful microscopes. A researcher at the University of Texas at Austin collects and analyzes information from electron microscopy, X-ray crystallography, quantum chemistry, and multi-scale molecular dynamics, and uses TeraGrid's Ranger supercomputer as a "computational microscope" to integrate the data on thousands of parallel processors. The researcher is using Ranger to model the largest and most complex biomolecular machinery to date: the 100 million atom chromatophores of purple bacteria. His experience using his highly scalable code on Ranger is pointing the way to future algorithms and codes capable of modeling larger biomolecular systems on next-generation high-performance computing systems.



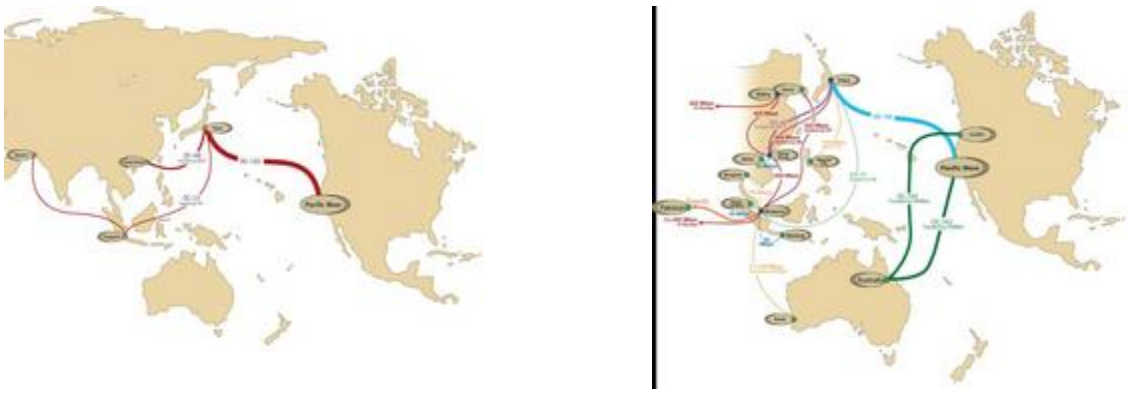
Simulation of the LH1-RC-PufX dimer. LH1 is colored in blue, the RCs in green, and PufX in red. a) Snapshots at the beginning of the simulation, viewed from the cytoplasm (top), and in the plane of the membrane (bottom). b) Snapshot of the simulated system at the end of the 20 ns equilibration. The LH1 protein exhibits a slight bending towards the periplasmic side. The membrane adapts to the LH1's change in shape. *Credit: Theoretical and Computational Biophysics Group, University of Illinois at Urbana-*

► **New Computer Model Helps Coastlines Prepare for Storm Surges:** Recent events in the Gulf of Mexico and around the world have demonstrated the vulnerability of coastal populations and infrastructure to storm surges. In response, researchers from the University of Texas at Austin, using the TeraGrid Ranger computing system, are developing the Advanced Circulation Storm Surge Model. The model has been used in design and planning mode prior to the hurricane season, in forecasting mode as storms approach land, and in hindcasting mode after the event. The researchers are developing a new model which includes highly refined details of the Texas coast, and which will be much more computationally intensive. This model will be used to perform hundreds of hurricane scenarios for the state of Texas, develop flood risk maps, and study potential inundation in high risk areas, such as the Houston-Galveston corridor and the Brownsville region.



The eye of the hurricane with wind vectors and wind speed just before landfall. *Credit: University of Texas at Austin.*

► **Extending Research and Education Connectivity to Pakistan:** Indiana University is leading an NSF-funded effort to upgrade Pakistan's internet service to gigabit level and extend its connectivity to public institutions, including libraries, health centers, and schools. The new Pakistan and Education Research Network also will provide researchers in Pakistan with connectivity to their international colleagues and to the global internet. The goal of this project is to increase cooperation between American and Pakistani scientists; the establishment of high-performance network connectivity is only a first step. Three to six months after the network connection is active, the project will host a science collaboration seminar in the United States, which will bring together researchers from different scientific disciplines from the United States and Pakistan to discuss how the network connection can enhance their research.



Map of the Pakistan and Education Research Network (PERN). *Credit: Indiana University.*

► **Remote Sensing of Ice Sheets Provides Ideal Training Opportunity:** The vision for the NSF project "Cyberinfrastructure for Remote Sensing of Ice Sheets" is to equip the current and next generation of traditionally, under-represented minority scientists, engineers, and educators with the knowledge and skills necessary to conduct interdisciplinary research in areas, including cyberinfrastructure, remote sensing, engineering, and modeling related to glaciers and ice sheets. The project is a significant step forward in providing the necessary virtual training environment and grid computing power that the students and scientists associated with the NSF Science and Technology Center for Remote Sensing of Ice Sheets require. To date, the project has deployed students to Antarctica, sponsored virtual field work at Elizabeth City State University in North Carolina, and funded travel for 16 students to attend a 2008 symposium in Virginia Beach on computing issues relevant to minority students, education, and institutions.

ECSU graduate student Je'aime Powell is shown in Ilulissat, Greenland where he provides technical support for the Cyberinfrastructure project which provides a field and base camp server for storage and visualization of CReSIS Polar Science SAR data sets. *Credit: ESCU.*

► **Development of a Regional Climate Model for U.S. Midwest Applications:** A team of scientists at the Illinois State Water Survey and the University of Illinois at Urbana-Champaign used high-performance computers to examine ways in which global climate change and human-driven emissions will affect future air quality. The team applied a new modeling system that integrates a regional climate model, an emissions model, and an air quality model nested within global climate and chemistry models. Among the issues the team examined was the impact of "transboundary emissions" from Mexico and Canada on ground-level ozone concentrations. They found that these emissions had a major impact on U.S. air quality, indicating that efforts to reduce emissions must be transnational. The TeraGrid Resources provided the critical computation and human support needed to evaluate the simulations.

Relative contributions percent) of the projected emissions (EMS) and climate (MET) changes to total surface ozone concentration trends between 2050 and 1998: A1Fi scenario and B1 scenario. Green denotes the dominance of the EMS effect (contribution greater than 70 percent), yellow denotes the MET effect (contribution greater than 70 percent), and red denotes where both effects are deemed to be important. *Credit: Xin-Zhong Liang, University of Illinois at Urbana-Champaign.*