OFFICE OF CYBERINFRASTRUCTURE $200,000,000

The FY 2008 Budget Request for the Office of Cyberinfrastructure (OCI) is $200.0 million, an increase of $17.58 million, or 9.6 percent, over the FY 2007 Request of $182.42 million.

Office of Cyberinfrastructure Funding
(Dollars in Millions)

	FY 2006 Actual	FY 2007 Request	FY 2008 Request	Change over FY 2007 Request Amount	Percent
Cyberinfrastructure	$127.14	$182.42	$200.00	$17.58	9.6%

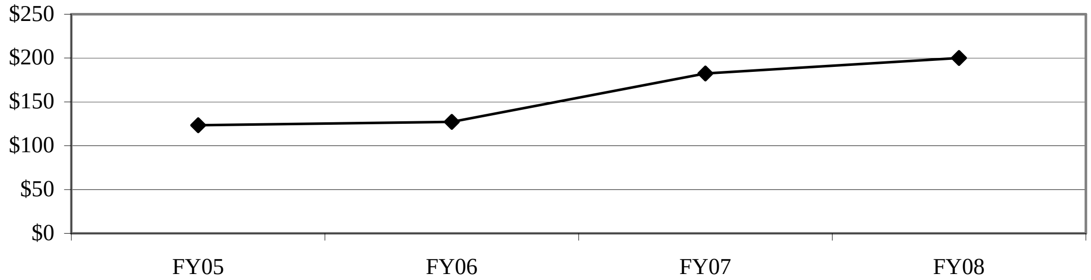
The Office of Cyberinfrastructure supports the development, acquisition, and operation of state-of-the-art cyberinfrastructure resources, providing cyberinfrastructure services that promote otherwise unrealizable advances in 21st century science and engineering research and education. OCI was created in July 2005 in an organizational realignment that moved the CISE Division of Shared Cyberinfrastructure (SCI) into the Office of the Director. At the same time, a Cyberinfrastructure Committee (CIC), composed of members of NSF's senior management, was created. The CIC provides integration and strategic vision across NSF's portfolio of cyberinfrastructure activities. In FY 2007, funds were added to the OCI budget to begin the acquisition of a leadership-class high-performance computing (HPC) system optimally configured to enable *petascale* performance (computing at rates on the order of 10^{15} floating point operations per second (petaflops) or working with very large datasets on the order of 10^{15} bytes (petabytes)) on important science and engineering problems.

OCI-supported cyberinfrastructure includes information technology resources and tools such as supercomputers, high-capacity mass-storage systems, system software suites and programming environments, scalable interactive visualization tools, productivity software libraries and tools, large-scale data repositories and information management systems, networks of various reach and granularity, an array of software tools and services that enhance the usability and accessibility of computational, observational and experimental infrastructure, and digital collaboratories. OCI also supports the scientific and engineering professionals who create and maintain these IT-based resources and systems, and who provide the Nation's researchers and educators with essential cyberinfrastructure services.

OCI activities directly respond to the President's advanced networking, high-end computing and cyberinfrastructure priorities, and are key components in the interagency Networking and Information Technology Research and Development (NITRD) priority. The technologies developed and the systems deployed by OCI facilitate discovery and innovation and bolster national competitiveness. The American Competitiveness Initiative (ACI) describes the goal of providing world-leading, high-end computing capability, coupled with advanced networking, to enable scientific advances through modeling and simulation at unprecedented scale and complexity across a broad range of scientific and engineering disciplines. OCI investments in high-performance computing for research and education, the TeraGrid infrastructure, and international network connections directly contribute to this goal.

OCI will participate in the new NSF-wide investment area of "Cyber-enabled Discovery and Innovation," enabling cyberinfrastructure in collaboration with several of NSF's directorates.

OCI Subactivity Funding
(Dollars in Millions)



RELEVANCE

How does a protein fold? What happens to space-time when two black holes collide? What impact does species gene flow have on an ecological community? What are the key factors that drive climate change? Did one of the trillions of collisions at the Large Hadron Collider produce a Higgs boson, the dark matter particle or a black hole? Can we create an individualized model of each human being for personalized healthcare delivery? How does major technological change affect human behavior and structure complex social relationships? What answers will we find – to questions we have yet to ask – in the very large datasets that are being produced by telescopes, sensor networks, and other experimental facilities?

These and other questions are only now coming within our ability to answer because of advances in computing and related information technology. Once used by a handful of elite researchers in a few research communities on select problems, advanced computing has become essential to future progress across the frontier of science and engineering. Coupled with continuing improvements in microprocessor capabilities, converging advances in networking, software, visualization, data systems and collaboration platforms are changing the way research and education is accomplished.

Recognizing that cyberinfrastructure capabilities are essential to advances in all science and engineering fields, NSF has developed a comprehensive cyberinfrastructure strategic plan entitled, *NSF's Cyberinfrastructure Vision for 21st Century Discovery* (www.nsf.gov/dir/index.jsp?org=OCI). This plan describes the agency's commitment to:

- Develop a stable, human-centered cyberinfrastructure (CI) that is driven by science and engineering research and education opportunities;
- Provide the science and engineering communities with access to world-class CI tools and services, including those focused on: high-performance computing and advanced networking; data, data analysis and visualization; virtual organizations; and learning and workforce development;
- Promote a CI that serves as an agent for broadening participation and strengthening the Nation's workforce in all areas of science and engineering; and
- Provide a sustainable CI that is secure, efficient, reliable, accessible, and usable, and which evolves as an essential national infrastructure for conducting science and engineering research and education.

OCI supports the development and deployment of cyberinfrastructure that is shared by all scientific and engineering disciplines, making possible potentially transformative basic research in areas such as nanotechnology, physics, chemistry, materials science, and engineering, as called for in the ACI. It also promotes interoperability between components of cyberinfrastructure both here in the U.S. and abroad.

About two thirds of NSF's investments in cyberinfrastructure are made by the directorates and offices responsible for fundamental domain specific research and education in science and engineering, while the remaining third, which is shared across all of NSF, is provided by OCI. Through coordinated planning and investments facilitated by NSF's Cyberinfrastructure Council, OCI provides economies of both scale and scope, ensuring that NSF's cyberinfrastructure portfolio delivers the highest returns on the Nation's investment.

Summary of Major Changes in Office-wide Investments ***(Dollars in Millions)***

FY 2007 Request, OCI…………………………………………………………………………….$182.42

<u>Discovery Research for Innovation</u> +$10.58

- *Software and Services for Complex Science and Engineering (+$5.08 million).* OCI will support the development and provision of software and services that facilitate complex science and engineering research. The emphasis will be on software and services that enhance the utility and impact of NSF's parallel investments in high-performance computing and advanced network control and transport mechanisms. These include innovative approaches to the management of data collections; software and practices that enhance the semantic interoperability of data and tools; robust middleware that supports distributed applications, distributed collaboration, interactive remote observation, and the tele-operation of instruments and experimental facilities; as well as advanced data analysis and visualization tools. Such advances in the analysis and management of data from experiments and computational models are critical to advancing ACI goals in data-intensive areas such as nanotechnology, materials science, weather and climate prediction, and the prediction of hazards from events such as earthquakes and hurricanes.

- *Petascale application software development (+$3.50 million).* Supporting the development of numerical models, data analysis tools, new algorithms, and new programming paradigms that take full advantage of the very large-scale HPC systems becoming available over the next few years is a key element of NSF's HPC strategy. With advanced petascale applications, researchers will be able to determine the three-dimensional structures of proteins and study how structure influences function; examine the patterns of emergent behavior that occur in models of very large societies; study nucleosynthesis in supernovae; understand what sort of abrupt transitions can occur in Earth's climate and ecosystem structure and why these happen; pursue the capability to design catalysts atom-by-atom, potentially transforming industrial synthesis; find strategies that optimize the management of complex infrastructure systems; gain a better understanding of language processing in large assemblages of neurons; and facilitate the planning and response to natural and man-made disasters that prevent or minimize the loss of life and property. OCI support for petascale software development will focus on preparing computer codes, in strategic science and engineering research areas, to run effectively on petascale computing systems. This type of development is critical for the ACI goal of using world-leading computing capability to advance a broad range of science and engineering through modeling and simulation at unprecedented scale and complexity. OCI will collaborate with NSF's research directorates to identify the appropriate science and engineering research areas. OCI support of petascale application software development will leverage the support for fundamental research provided by these divisions and include co-investment opportunities.

- *Strategic Technologies for Cyberinfrastructure (+$2.0 million).* The current level of innovation in cyberinfrastructure is very high. In addition to providing robust cyberinfrastructure for science and engineering research, OCI will provide "venture capital" to researchers who wish to capitalize on new ideas emerging from computer science and elsewhere and to explore whether these have the potential to be the next revolutionary strategic technologies in cyberinfrastructure. As such technologies mature, they will contribute to strengthening the capabilities of computing systems and advanced networks, highlighted in the ACI, and to the provision of new tools for basic research.

<u>Transformational Facilities and Infrastructure</u> +$14.48

- *High Performance Computing: Operations and Maintenance (+$20.0 million).* Increased operations and maintenance funding will support the sustained operation of high-performance computing (HPC) systems in university supercomputing centers. Such centers provide access to HPC resources, coupled with sophisticated user support and training, to a diverse mix of researchers and educators in the academic community. These resources are used in innovative research in areas ranging from biology to social science. They also provide a bridge to the larger-scale computing systems being brought on-line in NSF's national supercomputing grid, the TeraGrid or Extensible Terascale Facility. The latter is a high-end computing capability that is coupled with advanced networking and is designed to enable scientific advances across a broad range of disciplines, as called for in the ACI.

- *Other Infrastructure and Tools (-$5.52 million).* Adjustments will continue to be made in the current OCI portfolio to accommodate the strategic priorities described herein and in more detail in the document, *NSF's Cyberinfrastructure Vision for 21st Century Discovery*. These include a transition from a centralized model for the provision of cyberinfrastructure development and services to a broader and more diverse portfolio of cyberinfrastructure developers and resource providers.

<u>Preparing the Workforce of the 21st Century</u> -$7.48

Integration of research and education through cyberinfrastructure. In collaboration with partners across the Foundation, OCI will support creative explorations and demonstrations of the use of cyberinfrastructure to integrate research with education, the development of innovative technologies that will facilitate the integration of research and education, and research on how educators and students interact with cyberinfrastructure. One aim of this support is to connect students and educators with the types of complex science and engineering that are becoming increasingly prominent in contemporary research, that are themselves being facilitated by cyberinfrastructure, and that are difficult to reproduce in a school laboratory or informal education setting using traditional methods.

Subtotal, Changes +$17.58

FY 2008 Request, OCI…………………………………………………………..……………………$200.00

QUALITY

OCI maximizes the quality of the projects it supports through the use of a competitive, merit-based review process. The percent of funds that were allocated to projects that undergo external merit review was 99 percent in FY 2006, the last year for which complete data exist.

To ensure the highest quality in processing and recommending proposals for awards, a Committee of Visitors (COV) for the Division of Shared Cyberinfrastructure that preceded OCI was held in June 2005. The COV was composed of qualified external evaluators. These experts assessed the integrity and efficiency of the processes for proposal review and provided a retrospective assessment of the quality of results of NSF's investments. The COV found that: (i) the "merit review process is run with high integrity with appropriate care on criteria, consideration and judgment;" (ii) "reviews and summaries did an excellent job of addressing both quality and impact criteria;" and (iii) "the overall quality of accepted projects seems gratifyingly high."

The COV recommended:

- The development of a "long-term strategic vision for the integration of complementary activities across NSF" that included more attention to data-intensive applications, the use of service-oriented architectures to advance interoperability, and networking. NSF responded by developing the document, "NSF's Cyberinfrastructure Vision for 21st Century Discovery," inviting public comment, and incorporating this wider view in its plans for FY 2007 and FY 2008 investments;
- Remaining engaged in activities with promise for longer-term impact. The subsequent organizational realignment has created a clearer division of responsibilities with OCI focusing primarily on near-term development, deployment and sustaining infrastructure and CISE focusing on longer-term research activities; and
- The establishment of an external panel to provide advice on strategic directions in cyberinfrastructure. NSF has subsequently chartered such a committee, the Advisory Committee for Cyberinfrastructure (ACCI).

In partnership with NSF's directorates and offices, the ACCI provides guidance on issues such as: the mission, programs, and goals that can best serve the science and engineering community; how OCI can promote quality graduate and undergraduate education in the computational sciences and engineering; and priority investment areas in cyberinfrastructure. The ACCI meets twice a year. Members from both academe and industry represent a cross section of the science and engineering field, with representatives from many different disciplines. The ACCI includes a balanced representation of women, underrepresented minorities, and individuals from a range of geographic regions and institutions.

PERFORMANCE

The FY 2008 Budget Request is aligned to reflect funding levels associated with the Foundation's four strategic outcome goals highlighted in the FY 2006-2011 Strategic Plan. These goals were designed as a mechanism to better enable assessment of program performance and to facilitate budget and performance integration.

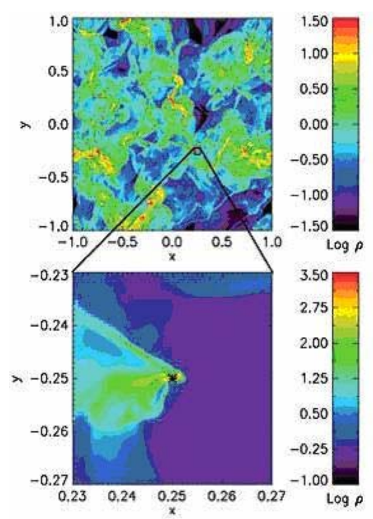
Office of Cyberinfrastructure
By Strategic Outcome Goal
(Dollars in Millions)

	FY 2006 Actual	FY 2007 Request	FY 2008 Request	Change over FY 2007 Request Amount	Change over FY 2007 Request Percent
Discovery	$9.85	$4.17	$14.75	$10.58	253.7%
Learning	11.48	11.48	4.00	-7.48	-65.2%
Research Infrastructure	103.76	164.72	179.20	14.48	8.8%
Stewardship	2.05	2.05	2.05	-	-
Total, OCI	$127.14	$182.42	$200.00	$17.58	9.6%

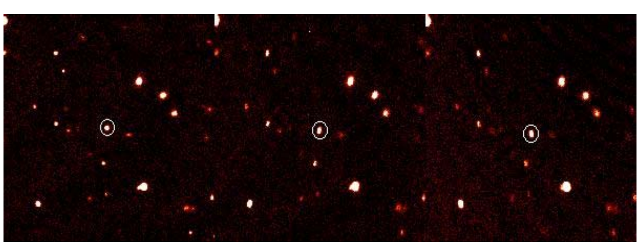
Totals may not add due to rounding.

Recent Research Highlights

► **A Star is Born:** Astrophysicist Richard Klein of the University of California, Berkeley, and his colleagues have carried out a series of massive computer simulations that greatly clarify our understanding of how stars form inside immense clouds of interstellar gas. Astronomers have long agreed on the basics. Stars form in the clouds because gravity pulls the gas into clumps, which eventually become so hot and so dense they ignite by thermonuclear fusion. But there has been much less agreement on the details. Do stars form all at once, as in the "gravitational collapse" model? Or do they start small and then grow over time, as in the "competitive accretion" model? To find the answers, Klein and his colleagues have developed the first computer simulation that can fully take into account the complex motions within a collapsing cloud. The researchers had to run their simulation for nearly two weeks on one of the most powerful machines in the world: the San Diego Supercomputer Center's DataStar system, which has a total memory of 7.3 trillion bytes, and is capable of 15.6 trillion arithmetic operations per second. The result was a victory for the gravitational collapse model. Competitive accretion cannot account for what's observed, either in the simulation or in the observations.



Top figure shows a slice through a star formation region, with densest areas shown in red. The key finding, shown in the zoom-in (bottom), is that once a protostar forms, creating a dense wake behind it (left), the turbulent wake prevents the protostar from gaining much additional gas from the surrounding clump, as required by the competitive accretion theory. *Credit: Mark Krumholz, Princeton U.*

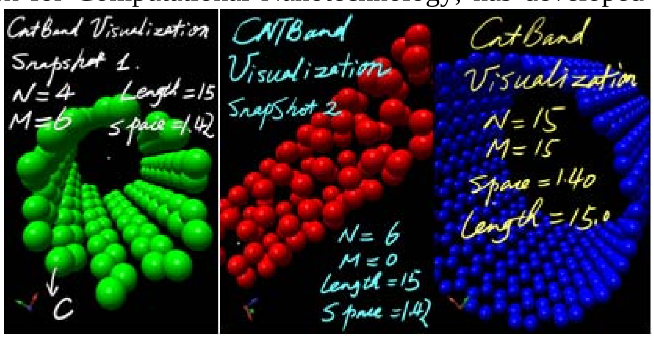


These time-lapse images of a new found planet in our solar system were originally taken in October 2003, using the Samuel Oschin Telescope on Mt. Palomar. The planet, circled in white, is seen moving across a field of stars. The three images were taken about 90 minutes apart. *Credit: Hans-Werner Braun.*

► **What do the discovery of a new planet and fighting large-scale wildfires have in common?:** The NSF-funded High Performance Wireless Research and Education Network (HPWREN) is a prototype system now operating in California's San Diego and Riverside counties. HPWREN is partly intended as a testbed for several of NSF's large-scale sensor network initiatives. These include EarthScope, the Ocean Observatories Initiative, the National Ecological Observatory Network, and the Network for Earthquake

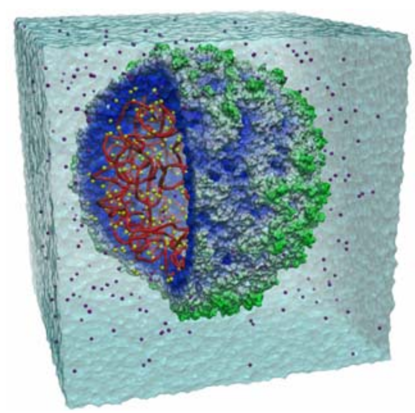
Engineering Simulation. At the same time, however, HPWREN is a working system, with multiple remote sites that are providing high-speed Internet access to field scientists in a variety of disciplines. Recently, astronomers from around the world used HPWREN to analyze the flood of data produced by a 161-megapixel camera at the Palomar Observatory — and in the process, discovered another "planet" in our solar system. Other remote HPWREN nodes include seismometers and ecological sensors. HPWREN also serves the first-responder community. For example, the California Department of Forestry and Fire Protection routinely accesses HPWREN's mountaintop cameras and sensors to monitor the notoriously fire-prone region. Firefighters at the scene of a blaze can rapidly deploy a wireless HPWREN node to access maps, aerial imagery, and telemetry data. Finally, HPWREN provides educational opportunities for rural Native American learning centers and schools in the area. Many other applications are described on the HPWREN Web site: http://hpwren.ucsd.edu.

► **Cyberinfrastructure education for future nanotechnologists:** The nanoHUB, operated by an NSF-funded research consortium known as the Network for Computational Nanotechnology, has developed next generation cybertools that encourage students to do collaborative simulations of nanoscale systems on tablet PCs and mobile devices. The nanoHUB has also deployed video and audio podcasts that facilitate anytime, anywhere learning. The video podcasting service on the nanoHUB has attracted over 1,000 downloads in just 2 months, suggesting that there is indeed a demand for such content. In order to measure the impact of such tools on learning, the nanoHUB team is integrating the assessment engine of the open-source Sakai Collaboration and Learning Environment using web services, a technology at the core of the nanoHUB middleware.

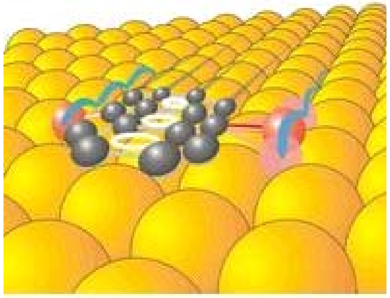


Ink-based interfaces to simulations that facilitate better student collaboration. *Credit: Krishna Madhavan and Sebastien Goasguen, Purdue.*

► **The first atom-by-atom simulation of a life-form:** For the first time, researchers have visualized the changing atomic structure of a virus by calculating how each of the virus' one million atoms interact with each other every femtosecond--or one-millionth-of-a-billionth of a second. The researchers' hope is that the insights gained from such simulations may one day help us combat viral infections in plants, animals and even humans. Using a supercomputer at the NSF-funded National Center for Supercomputing Applications (NCSA), the team ran the simulation about 100 days to generate just 50 nanoseconds of virus activity. For comparison, notes team leader Klaus Schulten of the University of Illinois at Urbana-Champaign, it would have taken the average desktop computer 35 years to come up with the results. The simulation revealed key physical properties of satellite tobacco mosaic virus, a very simple, plant-infecting virus. Ultimately, using the next generation of "petascale" supercomputers, scientists hope to generate longer simulations from bigger biological entities. The National Institutes of Health also provided support for this study.



The first all-atom simulation of a satellite tobacco mosaic virus. *Credit: University of Illinois at Urbana-Champaign's Theoretical and Computational Biophysics Group.*



Propelled by two sulfur (red) atoms as feet, DTA 'walks' across the surface setting step in front of step and never veering off course. *Credit: Ludwig Bartelis, UC – Riverside.*

► **Walking Molecule Provides a Key to "Molecule Memory:"** UC Riverside professor Ludwig Bartels and his team have designed and simulated a molecule that can "walk" across a flat surface in a straight line. Indeed, 9,10-dithioanthracene (DTA), as the molecule is known, can walk for more than 10,000 steps on molecular appendages that act as feet. Such a DTA "nano-walker" could form the basis of a molecular memory 1,000 times more compact than current computer memory devices. That, in turn, could make it important to the nascent field of "molecular computing." The new concept of molecular propulsion may also have far reaching benefits for the development of surface nano-robots, with applications ranging from information storage to the control of surface chemical reactions. The molecule design and simulations were done using one of the TeraGrid's supercomputers located at the San Diego Supercomputing Center.

► **Industrial Partnerships through the National Science Foundation's Supercomputing Resources:** Forty companies participated in a study involving NSF supercomputing resources. The study concluded that the partnership between the NSF Centers and the U.S. businesses "… clearly has been successful." The HPC resources were utilized at the centers to advance the industrial research and development efforts, advance strategic work, develop new products, solve specific problems and to get products to market more rapidly. Benefits were characterized as both financial and from a competitive perspective. The results obtained were identified to be beneficial in the areas of increased revenue growth, increased market share and the ability to respond to actual competitive threats. Several of the industrial users were able to assign a dollar value to their relationships, ranging from $100,000 to $57 million. Further, more than half of the participants reported that their partnerships had resulted in "… a breakthrough or discovered something totally new." This is a significant finding in today's global market-place where innovation can often provide a competitive advantage. Identified as a principal reason for entering into the partnership was access to the scientific and HPC expertise available at the centers, with actual access to the HPC resources themselves coming in as next in significance. Finally, 15 of the industrial users installed HPC systems after their experience with the centers.

Other Performance Indicators

The table below shows an estimate of the number of people benefiting from OCI funding based on the types and number of awards.

Number of People Involved in OCI Activities

	FY 2006 Estimate	FY 2007 Estimate	FY 2008 Estimate
Senior Researchers	339	350	360
Other Professionals	439	450	515
Postdoctorates	16	20	25
Graduate Students	136	145	175
Undergraduate Students	85	90	90
Total Number of People	1,015	1,055	1,165

However, OCI investments directly impact a much larger number of researchers and educators within the U.S. and around the world who use OCI-supported cyberinfrastructure facilities, resources and tools. For example, OCI-funded cyberinfrastructure enables the work of an estimated 150,000 senior researchers, graduate students, undergraduate students, and K-12 teachers annually.

The OCI funding profile is provided below.

OCI Funding Profile

	FY 2006 Estimate	FY 2007 Estimate	FY 2008 Estimate
Statistics for Competitive Awards:			
Number	42	60	50
Funding Rate	32%	20%	18%
Statistics for Research Grants:			
Number of Research Grants	34	55	50
Funding Rate	28%	18%	18%
Median Annualized Award Size	$253,000	$255,000	$190,000
Average Annualized Award Size	$287,000	$270,000	$210,000
Average Award Duration, in years	2.6	2.7	2.8