COMPUTER AND INFORMATION SCIENCE AND ENGINEERING

COMPUTER AND INFORMATION SCIENCE AND ENGINEERING $618,050,000

The FY 2005 Request for the Computer and Information Science and Engineering Activity is $618.05 million, an increase of $13.40 million, or 2.2 percent, above the FY 2004 Estimate of $604.65 million.

Computer and Information Science and Engineering Funding
(Dollars in Millions)

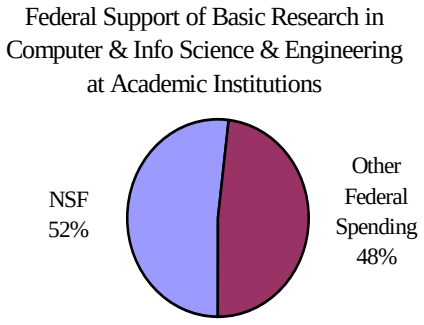
	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Change over FY 2004 Percent
Computer and Networked Systems	117.15	114.93	132.39	17.46	15.2%
Computing and Communication Foundations	81.15	78.93	91.41	12.48	15.8%
Information and Intelligent Systems	82.15	80.05	92.54	12.49	15.6%
Shared Cyberinfrastructure	95.07	112.63	123.60	10.97	9.7%
Information Technology Research	213.77	218.11	178.11	-40.00	-18.3%
Total, CISE	$589.29	$604.65	$618.05	$13.40	2.2%

Totals may not add due to rounding.

The Computer and Information Science and Engineering (CISE) Activity supports research, infrastructure, and education in the computer science, computer engineering, information science, networking, and computational science disciplines. It also supports shared cyberinfrastructure that enables cyber-science across the full range of NSF-supported science and engineering disciplines.

RELEVANCE

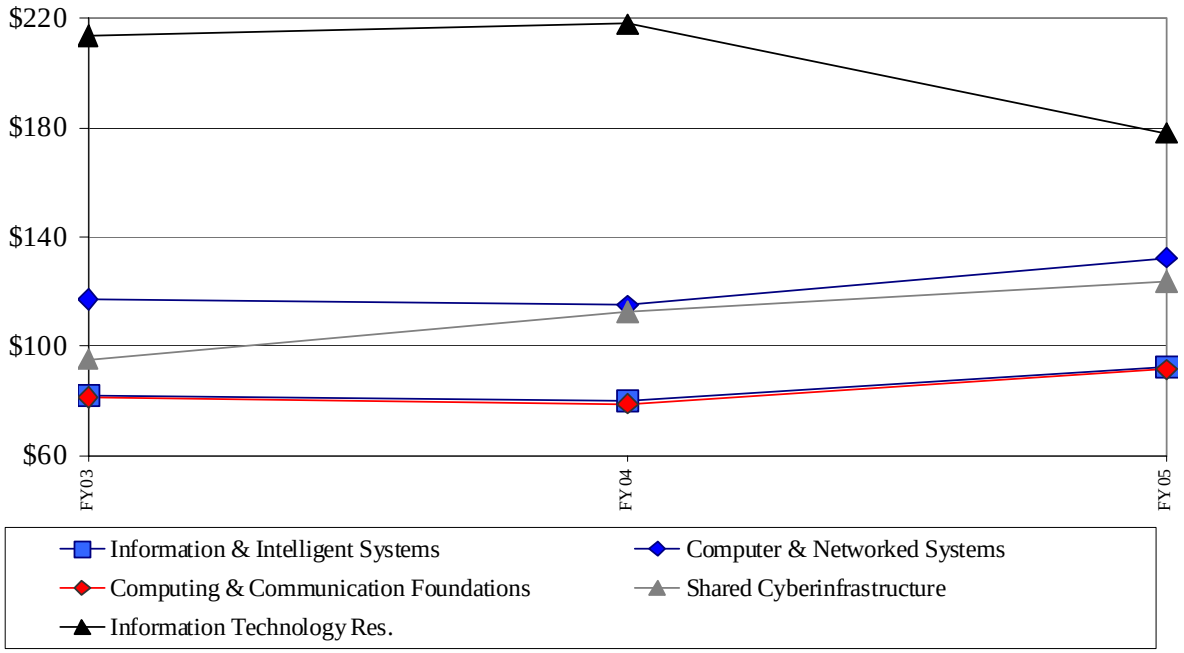
CISE is the principal source of federal funding for university-based basic research in the computer science, computer engineering, information science, networking, and computational science disciplines, providing over half of the total federal support in this area. The CISE Activity exerts a lead role in the multi-agency Networking Information Technology Research and Development program by providing 36% of total NITRD funding in FY 2003 and by chairing many of the working groups that promote interagency coordination. Building on past accomplishments, such as developing the Internet and supporting fundamental advances in numerical methods, digital libraries, data mining, computer languages, and computer systems, CISE is positioning its activities for the future with new efforts to address the most prominent challenges and opportunities of information technology:



- Cyber Trust: Growing concerns about the vulnerability and trustworthiness of computers, networks and information systems have led to increased CISE investments in Cyber Trust (also called cybersecurity) research, education and training.
- Cyberinfrastructure: CISE will manage and support the creation of a widely shared cyberinfrastructure that will revolutionize the conduct of research and education across the science

and engineering enterprise, and will invest in research to develop new generations of cyberinfrastructure CISE will provide broadly accessible and well-supported high-end computing, communications, storage, and analysis resources. CISE will also support the provision of services to support the effective use of these resources by domain scientists and engineers; and will support the education, outreach and training to take full advantage of or to support this new infrastructure-Workforce. Continuing needs for a U.S. workforce with the world's leading Information Technology (IT) skills drive CISE efforts to broaden participation of all people, regions, and institutions in IT education and career paths.

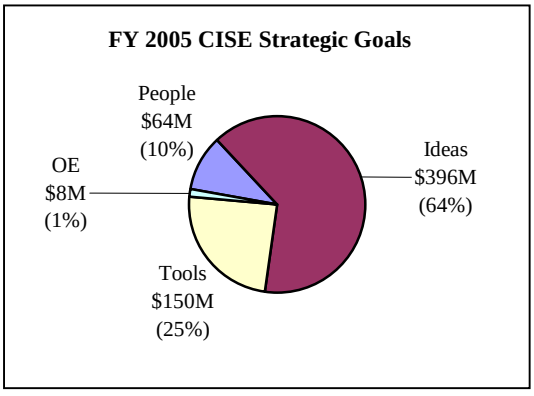
CISE Subactivity Funding
(Dollars in Millions)



Note: CISE subactivities have been reorganized; crosswalk data prior to FY 2003 do not exist.

STRATEGIC GOALS

NSF's four strategic outcome goals guide CISE activities.



- PEOPLE: CISE advances education and training for current computer and information scientists and engineers, increases the diversity of these communities, facilitates education of future generations of computer and information scientists and engineers, and enhances the public's knowledge of IT-related disciplines.

- IDEAS: CISE supports advances in knowledge across the computer science and engineering spectrum, providing core support for all IT fields and identifying opportunities where focused investments can play a catalytic role in advancing scientific progress.

- TOOLS: Provision of advanced tools to support CISE research and the development and support of an integrated cyberinfrastructure to support all areas of NSF science and engineering research and education, are CISE priorities in FY 2005. To provide tools for computer and information science and engineering research, CISE supports infrastructure at the small to mid-scale level that enables research in such areas as computer systems, information systems, robotics, and networking. In support of the full range of NSF-supported research and education, CISE will identify, develop, and support a shared cyberinfrastructure. Cyberinfrastructure will integrate sensors and instruments, data archives, digital libraries, high-end computing platforms, and visualization facilities, to enable completely new ways to advance science and engineering in the long-term.

- Organizational Excellence (OE): Organizational Excellence provides for administrative activities necessary to enable NSF to achieve its mission and goals. These investments include support for Intergovernmental Personnel Act appointments and for contractors performing administrative functions.

CISE's support for ongoing core and new activities contributes to NSF's efforts to achieve its strategic goals, and to the organizational excellence activities necessary to achieve these goals.

Funding by Strategic Goal: Summary
(Dollars in Millions)

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Change over FY 2004 Percent
People	54.64	56.67	63.51	6.84	12.1%
Ideas	385.70	408.37	396.24	-12.13	-3.0%
Tools	142.00	132.50	150.34	17.84	13.5%
OE	6.95	7.11	7.96	0.85	12.0%
Total, CISE	$589.29	$604.65	$618.05	$13.40	2.2%

PEOPLE (+ $6.84 million, for a total of $63.51 million)

CISE Investments in People
(Dollars in Millions)

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Change over FY 2004 Percent
Individuals	42.34	42.74	44.91	2.17	5.1%
Institutions	9.99	9.77	14.44	4.67	47.8%
Collaborations	2.31	4.16	4.16	0.00	0.0%
Total, CISE People	$54.64	$56.67	$63.51	$6.84	12.1%

Special emphasis in FY 2005 will be on broadening participation of U.S. citizens, including women and minorities, in the CISE enterprise and on reaching a wider range of institutions more effectively.

INDIVIDUALS

- An increase of $2.0 million to $7.39 million for IGERT and will support about 40 additional graduate students. Graduate Research Fellowships funding is maintained at $1.66 million.
- Support for Research Experiences for Undergraduates will increase by $170,000 to a total of $2.92 million for supplements. This will support participation of approximately 70 additional students in CISE-funded projects.
- Support for education and training for cyberinfrastructure totals $5.37 million. This will prepare individuals to effectively use these new integrated facilities.

INSTITUTIONS

- An additional $3.99 million to a total of $7.49 million will support demonstration projects that effectively link research and education and use best practices to attract more women and minorities to CISE fields; these projects will support recruiting and retaining students in computing science and engineering tracks along with improved outcomes for all students. Efforts will build on prior CISE awards that create new understanding on the reasons for the low participation of women and minorities in computer and information science and engineering activities.

COLLABORATIONS

- Funding for GK-12 will increase by $80,000 to a total of $240,000.

IDEAS (-$12.13 million, for a total of $396.24 million)

CISE Investments in Ideas
(Dollars in Millions)

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Change over FY 2004 Percent
Fundamental Science and Engineering	305.10	329.77	316.74	-13.03	-4.0%
Centers Programs	80.16	78.00	79.00	1.00	1.3%
Capability Enhancement	0.44	0.60	0.50	-0.10	-16.7%
Total, CISE Ideas	$385.70	$408.37	$396.24	-$12.13	-3.0%

Particular cross-cutting CISE emphases for FY 2005 are Cyber Trust, research to create new IT technologies for cyberinfrastructure, Science of Design, and Information Integration.

FUNDAMENTAL SCIENCE AND ENGINEERING

Funding for fundamental science and engineering will decrease by $13.03 million to a total of $316.74 million. In FY 2004, CISE reorganized its subactivities to better mirror the research communities supported and to position itself to take advantage of new research opportunities. In FY 2004 and 2005, the CISE directorate will focus its investments in eight clusters that position the Activity to manage its activities more strategically. To support research priorities in the eight clusters, funding is redirected from the ITR subactivity to other subactivities. This also enables the clusters to increase award size and duration, and to support larger-scale projects. Within the clusters, FY 2005 emphases will be:

- Systems in Context: research to improve the security of data intensive applications.
- Data Inference and Understanding: a new effort on shared data resources such as archives of annotated speech, videos and web logs (blogs) that will leverage existing research on digital libraries and accelerate research on human language and communication; and a new cross-cutting thrust on Information Integration that will create capabilities for the meaningful fusion of information from disparate sources.
- Science and Engineering Informatics: a new research focus on the collection, annotation, archiving, access, and analysis of all types of scientific data.
- Formal and Mathematical Foundations: support new efforts on parallel computing architectures and computation, algorithms for computational science, integrated sensing, and signal processing.
- Foundations of Computing Processes and Artifacts: new efforts on software design, parallel methods for computing, and graphics and visualization. Science of Design will also be emphasized.
- Emerging Models and Technologies for Computation: new efforts on computational methods for nano-scale design and computational neuroscience.
- Computing Systems: focus on scalable systems that are representative of the challenges of large-scale, modern systems of the future. Cyber Trust will also be emphasized.
- Network Systems: multi-institutional projects and projects that provide results that will scale to future networks.

CENTERS PROGRAMS

CISE-supported centers include the Information Technology Centers and the Science and Technology Centers (STCs).

CISE Centers
(Dollars in Millions)

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Change over FY 2004 Percent
Information Technology Centers	76.16	74.00	75.00	1.00	1.4%
Science and Technology Centers	4.00	4.00	4.00	0.00	0.0%
Total, CISE Centers	$80.16	$78.00	$79.00	$1.00	1.3%

CISE will continue its support of the Information Technology Centers. These center-scale awards have allowed CISE to support projects with ambitious goals and complex efforts that enrich both research and the training of students. New centers started in FY 2003 include "Sensitive Information in a Wired World" led by Stanford University researchers and Linked Environments for Atmospheric Discovery led by University of Oklahoma researchers. These projects are focusing on important IT issues and drawing on both traditional disciplinary as well as interdisciplinary expertise.

In FY 2005, CISE will continue support of the Science and Technology Center for Embedded and Networked Sensing (CENS) at UCLA. CENS is developing Embedded Networked Sensing Systems and applying this revolutionary technology to critical scientific and social applications. Embedded networked sensing systems will form a critical infrastructure resource for society - they will monitor and collect information on such diverse subjects as plankton colonies, endangered species, soil and air contaminants, medical patients, and the health of buildings, bridges and other man-made structures. Across this wide

range of applications, embedded networked sensing systems promise to reveal previously unobservable phenomena.

During FY 2004, CISE expects to make new center-scale awards in the area of Cyber Trust. Support for these awards will continue in FY 2005.

CAPABILITY ENHANCEMENT

- CISE will maintain participation in special programs such as RUI in FY 2005. Additionally, CISE-based special programs such as digital government, and research infrastructure will focus on developing new capabilities and fields to support innovations in IT research and to support IT research applications.

TOOLS (+$17.84 million, for a total of $150.34 million)

CISE Investments in Tools
(Dollars in Millions)

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Change over FY 2004 Percent
Facilities	95.38	109.44	127.44	18.00	16.4%
Infrastructure and Instrumentation	46.62	23.06	22.90	-0.16	-0.7%
Total, Tools Support	$142.00	$132.50	$150.34	$17.84	13.5%

Totals may not add due to rounding.

FACILITIES

In FY 2005, the NSF's shared Cyberinfrastructure efforts will build on the successes of prior CISE-supported programs. The emerging cyberinfrastructure will incorporate data archives, instruments and sensors, visualization, and enabling software, and will be managed in the new Shared Cyberinfrastructure subactivity.

CISE Facilities
(Dollars in Millions)

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Change over FY 2004 Percent
PACI / Shared Cyberinfrastructure Facilities	73.24	87.00	90.00	3.00	3.4%
Terascale Computing Systems	11.17	10.00	25.00	15.00	150.0%
Other CISE Facilities[1]	10.97	12.44	12.44	0.00	0.0%
Total, CISE Facilities	$95.38	$109.44	$127.44	$18.00	16.4%

[1]Other CISE facilities include equipment support programs for CISE disciplinary research providing equipment such as experimental cluster computers, Networking testbeds, and visualization equipment.

Cyberinfrastructure activities build on the successes of:

- PACI. The Partnerships for Advanced Computational Infrastructure program will end in FY 2004. The support for high-end computing will be provided through the new integrated cyberinfrastructure investments.
- Terascale Computing. Construction of the Extensible Terascale Facility (ETF), funded by the MREFC account, will be completed in FY 2004. The ETF will also become part of the coordinated cyberinfrastructure effort. Cyberinfrastructure, funded at $20.0 million in FY 2004, will be incorporated in FY 2005 as part of this coordinated effort.
- Advanced Networking Infrastructure, described below in Infrastructure and Instrumentation will also become part of the enabling cyberinfrastructure.

The Directorate also manages other CISE Facilities supporting CISE-focused research.

- The Computing Research Infrastructure cluster will support mid-scale instrumentation needed by all CISE research and education programs. The program supports many types of institutions and projects. Funding will remain at $12.44 million.

INFRASTRUCTURE AND INSTRUMENTATION

CISE Investments in Infrastructure and Instrumentation
(Dollars in Millions)

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Change over FY 2004 Percent
Advanced Networking Infrastructure	46.62	23.06	22.90	-0.16	-0.7%
Total, Infrastructure & Instrumentation	$46.62	$23.06	$22.90	-$0.16	-0.7%

Advanced Networking Infrastructure, which formerly included an applied research component, has been split into research (IDEAS) and advanced development and infrastructure (TOOLS) components. The development and infrastructure activities will be coordinated in the cyberinfrastructure efforts. The research component will be carried out in the Computer and Network Systems subactivity.

ORGANIZATIONAL EXCELLENCE (+$850,000, for a total of $7.96 million)

Organizational Excellence supports Intergovernmental Personnel Act appointments, IPA's travel, and the administrative contracts necessary to conduct the level of program activity at the Request Level.

PRIORITY AREAS

In FY 2005, CISE will support research and education efforts related to broad, Foundation-wide priority areas in Biocomplexity in the Environment, Nanoscale Science and Engineering, Mathematical Sciences, Human and Social Dynamics, and Workforce for the 21st Century.

CISE Investments in NSF Priority Areas
(Dollars in Millions)

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Change over FY 2004 Percent
Biocomplexity in the Environment	7.36	8.00	8.00	0.00	0.0%
Nanoscale Science and Engineering	11.14	15.79	19.40	3.61	22.9%
Mathematical Sciences	2.29	2.29	2.29	0.00	0.0%
Human and Social Dynamics	N/A	3.00	3.00	0.00	0.0%
Workforce for the 21st Century	N/A	N/A	2.56	2.56	N/A

- **Biocomplexity in the Environment**: In FY 2005, CISE will maintain investment in the Biocomplexity in the Environment priority area at $8.0 million. These funds will contribute to NSF's coordinated central competition and will support focused environmental informatics activities such as multi-scale modeling and simulation, dynamic data analysis and interpretation, synthesis studies, and data mining and data management.
- **Nanoscale Science and Engineering:** CISE support totals $19.40 million in FY 2005, an increase of $3.61 million over the FY 2004 Estimate of $15.79 million, for research on quantum computing, simulation of atomic and molecular scale systems, self-assembly of bio-molecular computer components, nano-robotics, and design automation to support a new approach to molecular architectures.
- **Mathematical Sciences:** CISE support totals $2.29 million in FY 2005, continuing the same level of support as FY 2004. CISE support will emphasize interdisciplinary research bridging IT and mathematical disciplines with focus on algebraic and geometric algorithms, algorithms for scalable scientific computation, algorithms for visualization, and statistical learning algorithms
- **Human and Social Dynamics:** CISE support totals $3.0 million in FY 2005, unchanged from FY 2004. Research will focus on improving use of IT systems including visualization, human-computer interaction, and language interfaces; modeling uncertainty, representing uncertainty of data objects, reasoning with uncertain objects, and semantics of distributed reasoning on uncertain objects; and mechanisms for how humans and groups interact with them. The Vulnerabilities Analysis, Consequence Management, and Threat Reduction program will initiate research that creates and manages integrated multidisciplinary and multi-sector resources for response to extreme events, including: data repositories, tools for planning and decision-making, communications infrastructures, sensor infrastructures, and real-time data-driven simulations.
- **Workforce for the 21st Century:** CISE support totals $2.56 million in FY 2005 with emphasis on increasing the capacity and quality of the nation's IT workforce.

QUALITY

CISE maximizes the quality of the R&D it supports through the use of a competitive, merit-based review process. The percent of basic and applied research funds that were allocated to projects that undergo merit review was 97 percent in FY 2003, the last year for which complete data exist.

To ensure the highest quality in processing and recommending proposals for awards, CISE convenes Committees of Visitors, 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.

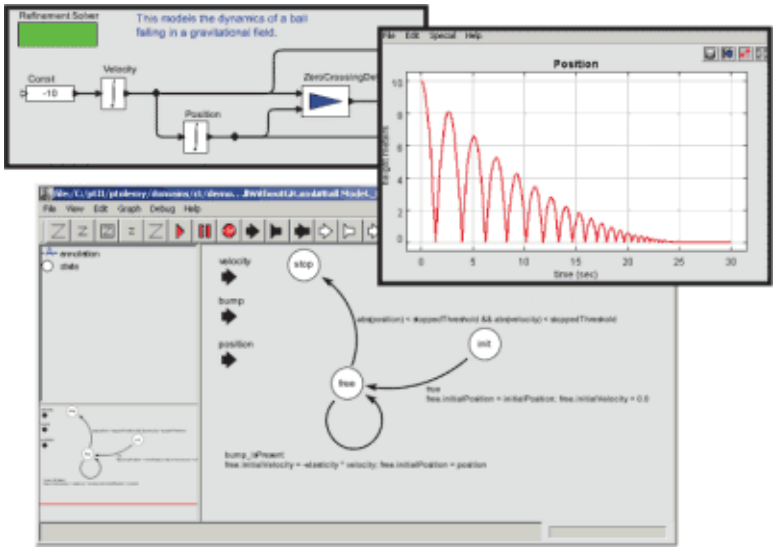
The Directorate also receives advice from the Advisory Committee for Computer and Information Science and Engineering (CISEAC) on such issues as: the mission, programs, and goals that can best serve the scientific community; how CISE can promote quality graduate and undergraduate education in the computer and information science and engineering disciplines; and priority investment areas in CISE research. The CISEAC meets twice a year and members represent a cross section of computer and information science and engineering with representatives from many different sub-disciplines within the field; a cross section of institutions including industry; broad geographic representation; and balanced representation of women and under-represented minorities. The CISE directorate also received advice from an advisory committee on Cyberinfrastructure, which in early 2003 issued the report, *Revolutionizing Science and Engineering through Cyberinfrastructure: Report of the National Science Foundation Advisory Panel on Cyberinfrastructure* (available from the NSF).

PERFORMANCE

Recent Research Highlights

A New System Science: Bridging Physical and Computational Systems

Teams from the University of California at Berkeley and Vanderbilt University have united to build new foundations for computational, physical, and engineered systems science. The Center for Hybrid and Embedded Software and Systems (CHESS) was established as the focal point for this NSF Large ITR project. This center is aimed at developing model-based and tool-supported design methodologies for real-time fault tolerant software that must execute on heterogeneous distributed platforms, and must control or otherwise interact with physical and engineered systems. The research seeks to bridge the gap between computer science and systems science by developing the foundations of a modern systems science that is simultaneously computational and physical. This represents a major departure from the current, separated structure of computer science (CS), computer engineering (CE), and electrical engineering (EE): it reintegrates information and physical sciences.



The center has convened a "curriculum council," which is an advisory board consisting of deans and department chairs of several California universities and community colleges to develop a strategy for propagating curriculum reform in systems science. The objective is to develop and disseminate courses that bridge the gap between physical system modeling and design and computational system modeling and design.

This project aims to revolutionize system science, integrating physical and computational aspects of systems on a new foundation of hybrid systems theory. The educational framework seeks also to revolutionize engineering education. This is done through an educational consortium for outreach to regional and minority-serving institutions.

Graph-Based Data Mining

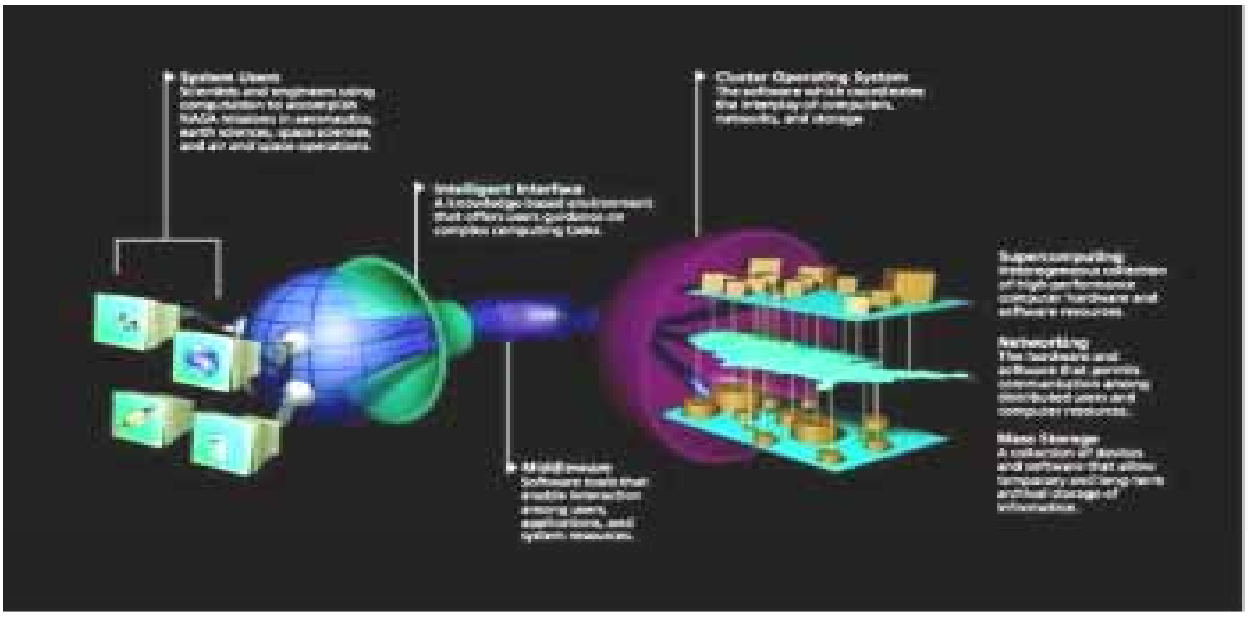
Current methods for extracting knowledge from databases are deficient in handling the growing amount of structural data expressing relationships among data objects. A CISE-supported research project, based at the University of Texas – Arlington, investigates a method for discovering knowledge in structural data. The main objective of this project is to design, implement and evaluate new methods for performing pattern learning on structured data represented as graphs and evaluate their application to structural, relational databases. The research team implemented the SUBDUE substructure discovery system that discovers interesting and repetitive subgraphs in a labeled graph representation using the minimum description length principle. The team is investigating approaches for using probabilistic graphs instead of deterministic graphs as the pattern language used by SUBDUE.

Experiments show SUBDUE's applicability to several domains, such as molecular biology, discovery of patterns in protein secondary structure, DNA gene transcription sites, carcinogenic chemical compounds, toxicology application, aviation incident reports, seismic events, image analysis, computer-aided design and program source code. The SUBDUE system is available to the research community at: http://cygnus.uta.edu/subdue/index.html.

NSF Middleware Initiative's GRIDS Center

Advances in science and engineering are driven increasingly by collaborations that focus on sharing data, computing, code, and access to experimental facilities. Network-driven computers, storage, data collections and scientific instruments are now central to the day-to-day practice of many research disciplines and are emerging as a model for cyberinfrastructure.

For example, NSF's Grid Physics Network (GriPhyN) project uses an international network of computational systems and data collections to address next-generation particle physics experiments at the Large Hadron Collider (LHC), while the NSF-funded Network for Earthquake Engineering Simulation (NEES) is revolutionizing seismology via network enabled access to experimental facilities, data, and simulations. GriPhyN and NEES represent not today's standard practice, but five- to ten-year strides for these disciplines. GriPhyN is preparing for a torrent of data from LHC experiments to begin in 2006 with a 15-year duration, while NEES is expected to be in place until 2014. Though these communities are ready now to develop new modes of research, scientists and engineers are frustrated by the scarcity of network-enabled services to suit their applications.



The NSF Middleware Initiative has begun to lead a path toward next-generation infrastructure for flexible resource sharing on national and international scales. The Grid Research Integration Development and Support (GRIDS) Center will define, develop, deploy, and support an integrated national middleware infrastructure for 21st century science and engineering applications. GRIDS involves the University of Southern California's Information Sciences Institute, the National Center for Supercomputing Applications at the University of Illinois, the University of Chicago, the San Diego Supercomputer Center at the University of California-San Diego and the University of Wisconsin-Madison.

In 2002, two NMI software releases were made. These software releases were built on widely used middleware such as the Globus Toolkit (the de facto standard for G rid environments), Condor-G and Network Weather Service, heavily leveraging open protocols based on IETF and W3C standards. In addition to GRIDS, the releases also included tools from a second NMI team called EDIT (for "Enterprise and Desktop Integration Technologies"), led by the University Corporation for Advanced Internet Development (UCAID), EDUCAUSE and the Southeastern Universities Research Association. The EDIT project focuses on security and directory service middleware for applications involving inter-campus collaborations. The GRIDS Center and EDIT team are also working closely with the NSF Partnerships for Advanced Computational Infrastructure and private industry to define and create an open, extensible architecture that integrates extant middleware.

Other Performance Indicators

The tables below show the number of people benefiting from CISE funding, and trends in growth of award size, duration, and number of awards.

Number of People Involved in CISE Activities

	FY 2003 Estimate	FY 2004 Estimate	FY 2005 Estimate
Senior Researchers	4,160	4,200	4,400
Other Professionals	1,330	1,400	1,200
Postdoctorates	501	500	500
Graduate Students	4,602	4,700	4,800
Undergraduate Students	774	810	1,000
Total Number of People	11,367	11,610	11,900

CISE Funding Profile

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request
Statistics for Competitive Awards:			
Number	872	880	900
Funding Rate	22%	22%	26%
Statistics for Research Grants:			
Number of Research Grants	1,030	1,030	1,040
Funding Rate	20%	22%	23%
Median Annualized Award Size	$116,193	$116,000	$116,000
Average Annualized Award Size	$160,174	$163,604	$165,000
Average Award Duration, in years	3.0	3.0	3.1

COMPUTER AND NETWORK SYSTEMS $132,390,000

The FY 2005 Budget Request for the Computer and Network Systems (CNS) Subactivity is $132.39 million, an increase of $17.46 million, or 15.2 percent, above the FY 2004 Estimate of $114.93 million.

Computer and Network Systems Funding
(Dollars in Millions)

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Change over FY 2004 Percent
Computer and Network Systems	117.15	114.93	132.39	$17.46	15.2%
Total, CNS	$117.15	$114.93	$132.39	$17.46	15.2%

The CNS subactivity supports research and education activities that invent new computing and networking technologies and that explore new ways to make use of existing technologies. The Division seeks to develop a better understanding of the fundamental properties of computer and network systems and to create better abstractions and tools for designing, building, analyzing, and measuring future systems. The Division also supports the computing infrastructure that is required for experimental computer science, and it coordinates cross-divisional activities that foster the integration of research, education, and workforce development.

The CNS Division is organized into four clusters, each of which is responsible for a related set of activities.

- **Computing Systems:** Future computing systems will be required to control a greater variety of computing, communication, storage, and external devices; to support a broader range of increasingly demanding applications; and to manage hundreds of asynchronous activities correctly, securely, and reliably. This cluster supports research and education activities that address these requirements in a variety of systems, including distributed, mobile, and embedded systems; sensing and control systems; dynamically configured, multiple-component systems; parallel systems; and trusted systems.
- **Network Systems:** Future networks are likely to exhibit unpredictable and complex behavior and dynamics; to span a broad range of technologies and bandwidths, from wireless sensors to a high-performance core; and to carry increasingly large amounts of increasingly demanding traffic. This cluster supports a range of research and education activities in network systems, including networking research, new technologies, and networking research test beds.
- **Computing Research Infrastructure:** An important component of experimental computing is building prototypes and test beds, and this requires having an experimental infrastructure. This cluster provides support for the acquisition, enhancement, and operation of experimental facilities for all CISE research and education areas. Supported facilities range from instrumentation needed by a few projects to major experimental facilities for an entire department. Support is also provided to enhance the computational and human infrastructure in minority-serving institutions and to support the equipment needs of collaborative, distributed research projects. An emphasis is to expand support to include a wider range of infrastructure needs, research projects, and institutions.
- **Education and Workforce Cluster:** Rapid advances in computing technology lead to the need to transfer research results into the classroom. Developing and making effective use of new research results requires a well-educated and diverse workforce that is representative of and can interact with the entire populace. This cluster supports projects that integrate research and education across CISE, study the causes of the current lack of diversity in the information technology workforce, and lead to a broadening of participation by all under-represented groups. The cluster works closely with all

CISE divisions to achieve these goals. It also coordinates the participation by CISE in a portfolio of NSF-wide education and workforce programs.

The following are examples of major research efforts supported by CNS.

Computing Systems Cluster
Computer systems are subject to the ever-increasing demands of the information revolution. Applications are rising in complexity, requiring continuous improvement of software, computational resources, storage, and other devices. Massive quantities of data must be managed accurately, reliably, and securely across an array of disparate systems. The computation required for visualization of very large data sets can be as complex and expensive as the physical experiment or computer simulation that produced the data in the first place. Researchers at the University of California, Davis are exploring ways to optimize the mechanisms and methodologies that generate visualization data using parallel supercomputers. This project is system-level research, examining all stages of the visualization process, from preprocessing and rendering algorithms, to compression for transport and application control for storage.

Computing Research Infrastructure Cluster
Prototypes and testbeds are essential to experimenting with new computing systems. NSF supports the acquisition, enhancement, and operation of infrastructure in order to provide experimental facilities and advance computer and information science and engineering research and education. Research and education at Tuskegee University, an historically black university in Alabama, has been enhanced by an NSF grant that helped establish laboratories and facilities including a multimedia lab, a 4-server cluster, and High Performance Computer lab. These facilities have helped the university attract and retain African American students in computer science and engineering. The university has subsequently undertaken an outreach program that has been effective in creating a continuing interest in computation and computer science in a large number of pre-college African Americans.

Network Systems Cluster
Networks in the future are predicted to become more dynamic, complex, and unpredictable. An increasing amount of traffic with new demands will travel over a diverse set of technologies and bandwidths. Research seeks to improve the ability of network technologies to scale, adapt, and protect data to meet emerging demands. The development of robust and stable ultrascale networking, at gigabit per second (gbps) speeds in the wide area, is critical to support the new generation of high-end computing and Petabyte to Exabyte datasets that promise to drive discoveries in fundamental and applied sciences of the next decade. Researchers at CalTech are developing the theories and the algorithms for such networks of the future. Tests have been successful in transmitting data the equivalent of a full length DVD movie in approximately 7 seconds.

COMPUTING AND COMMUNICATIONS FOUNDATIONS $91,41,000

The FY 2005 Budget Request for the Computing and Communication Foundations (CCF) Subactivity is $91.41 million, an increase of $12.48 million or 15.8 percent, over the FY 2004 Estimate of $78.93 million.

Computing and Communications Foundations Funding
(Dollars in Millions)

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Change over FY 2004 Percent
Computing & Communications Foundations	81.15	78.93	91.41	$12.48	15.8%
Total, CCF	$81.15	$78.93	$91.41	$12.48	15.8%

The CCF subactivity supports research and education activities that explore the foundations of computing and communication devices and their usage. The Division seeks advances in computing and communication theory, algorithms for computer and computational sciences, and architecture and design of computers and software. CCF-supported projects also investigate revolutionary computing paradigms based on emerging scientific ideas and integrate research and education activities to prepare future generations of computer science and engineering workers. The Division is particularly active in CISE' Science of Design and Nanoscale Science and Engineering research efforts.

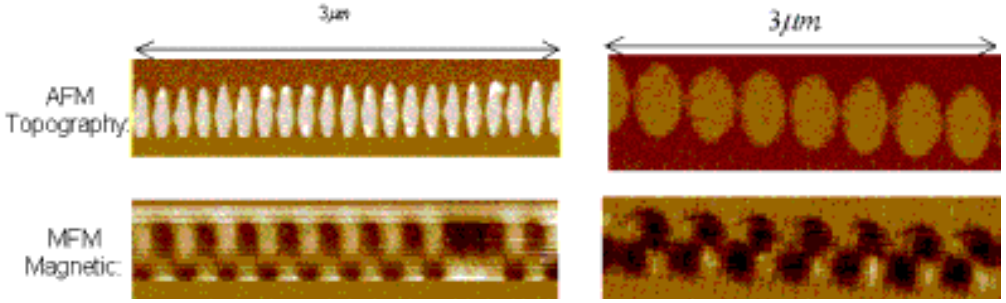
Because of the breadth of research it supports, CCF has identified three clusters of programmatic efforts.

- **Emerging Models and Technologies for Computation:** This cluster seeks to explore computational models, techniques, and systems based on emerging and future technologies. Research and education projects are supported in computing systems based on nanotechnology, quantum computing and communication, and computational devices and architectures inspired by the processing of information in living matter. The portfolio of awards examines concepts in new computing architecture, quantum and biologically inspired computing, as well as micro- and nano-systems. Topical areas include: computational algorithms and simulation techniques for nanoscale systems; design and architecture of systems based on molecular scale devices; quantum algorithms for computation, communication, and coding; realization of quantum computing; algorithms and computational modeling of biological processes; and computing models and systems for future technologies.
- **Formal and Mathematical Foundations:** This cluster seeks to determine inherent limits of computation and communication, and to obtain optimal solutions within those limits. Research and education projects supported examine information representation methods and computational techniques for advancing information technology and all scientific and engineering disciplines. Topical areas include: models of computation; computational complexity; parallel and distributed computation; random and approximate algorithms; algorithmic algebra, geometry, topology, and logic; computational optimization; computational algorithms for high-end scientific and engineering applications; techniques for representing, coding and transmitting information; mobile communication; optical communication; signal processing systems; and analysis of images, video, and multimedia information.
- **Foundations of Computing Processes and Artifacts:** This cluster seeks to advance the science, formalisms, and methodologies for building computing and communication systems. Research and education projects in software engineering, programming language design and implementation, graphics and visualization systems, computer architecture, and design automation are supported. Topical areas include: software design methodologies; tools for software testing, analysis, and

verification; semantics, design, and implementation of programming languages; micro-architectures; memory and I/O subsystems; application-specific architectures; performance metrics; VLSI electronic design; analysis, synthesis and simulation algorithms; system-on-a-chip; and architecture and design for mixed or future media (e.g., nanotechnology).

Some examples of the research promoted by CCF are:

Emerging Models and Technologies for Computation Cluster: The next breakthrough in computation capability may well be very far removed from the known paradigms. Research in areas such as biology, nanotechnology, and quantum physics provides fundamentally different models and inspiration that could lead to faster, more robust computer software, hardware, and architectures. Researchers at Notre Dame University, conducting an NSF-funded project entitled, "Computing Architectures for Coupled Nanomagnets," have discovered that magnetic interactions can be used to communicate complex information between nanoscale elements, in much the way that silicon is used in microprocessors. This finding opens the door to new computing architectures based on nanoscale elements.



Magnetic interactions to communicate information between neighboring quantum dots

Formal and Mathematical Foundations Cluster: The inherent limits of computation and communication are not well understood. Research at the foundational level is attempting to define the limits and optimize the solutions that can be produced in computer science, scientific computing, communication theory, signal processing theory, and mathematics to bring understanding across all science and engineering domains. NSF is sponsoring research at Carnegie-Mellon University that may help to make the difficult task of integrating multiple databases easier. Identifying duplicate entries of the same data from separate data sources is a vexing problem for scientists, engineers, and other data consumers. Using a natural machine-learning algorithmic approach, researchers have automatically identified hand-labeled data duplicates, providing theoretical insight into large-scale data integration.

Foundations of Computing Processes and Artifacts Cluster: NSF is seeking to advance the science, formalisms, and methodologies for building computer and communications systems. From the theoretical frameworks to the technical implementations, consideration is given to the artifacts and processes as they are involved in specifying, designing, and building complex systems. Researchers at the University of Minnesota – Twin Cities are focused on improving parallel computation methods in order to solve large-scale engineering and scientific problems. Advances are being made for three important components of parallel systems: effective and scalable algorithms, effective computer science tools and data structures, and testing and validation. Parallel systems rely on commodity hardware and can greatly reduce the costs and time involved in solving complex science and engineering problems.

INFORMATION AND INTELLIGENT SYSTEMS $92,540,000

The FY 2005 Budget Request for the Information and Intelligent Systems (IIS) Subactivity is $92.54 million, an increase of $12.49 million, or 15.6 percent, above the FY 2004 Estimate of $80.05 million.

Information and Intelligent Systems Funding
(Dollars in Millions)

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Percent
Information and Intelligent Systems	82.15	80.05	92.54	$12.49	15.6%
Total, IIS	$82.15	$80.05	$92.54	$12.49	15.6%

The Division of Information & Intelligent Systems supports research and education that increases the capabilities of human beings and machines to create, discover and reason with knowledge by advancing the ability to represent, collect, store, organize, locate, visualize and communicate information. The Division contributes to interdisciplinary research on how observational data leads to discovery in the sciences and engineering.

The IIS subactivity is organized into three clusters, each of which is responsible for a coordinated strategy across a set of research and education areas.

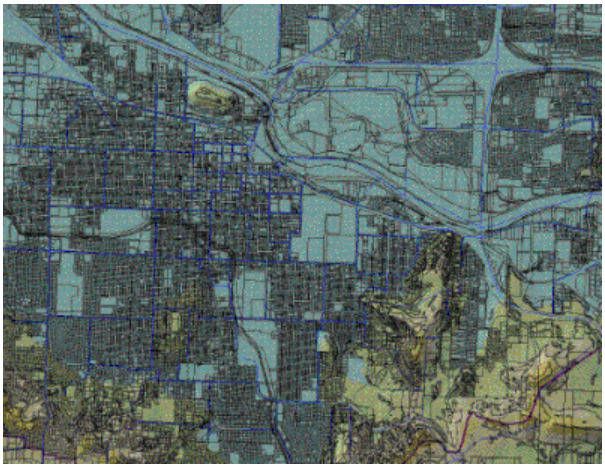
- **Data Inference and Understanding:** This cluster supports basic computer science research and education with the goal of creating general-purpose systems for representing, storing and accessing data, information and knowledge. It also supports research and education in automated methods of drawing conclusions from data and knowledge. Topical areas include: artificial intelligence and cognitive science; information and data management; computer vision; and human language and communication.
- **Science and Engineering Informatics:** This cluster supports research and education focused on advances in information technology that address problems in specific sciences and engineering domains (e.g., biology, geology, chemistry). Traditionally, scientists develop hypotheses, design experiments to test hypotheses, collect observational data, and publish results based on experiments. Data were often published to allow others to build upon or verify the results. In nearly every field of 21st century science, including all of the disciplines funded by NSF, new knowledge is generated by teams of researchers and educators analyzing data sets that are far too large to publish in journals and often collected independently by other scientists with different goals in mind. Characteristics of the research and education activities within this cluster are: integrative; focused on tools and analysis; and supportive of the data infrastructure for science and engineering. As an example, this cluster will be responsible for the Joint NSF/NIH Initiative to Support Collaborative Research in Computational Neuroscience. This project seeks to enhance our understanding of nervous system function by providing analytical and modeling tools that describe, traverse and integrate different levels of information. The Science and Engineering Informatics cluster will support similar projects across all fields of science and engineering.
- **Systems in Context:** This cluster supports research and education on the interaction between information, computation and communication systems and users, organizations, government agencies, the scientific community and the external environment. Research results provide requirements for the design and construction of future systems so that more system deployments are successful by design. The result of the integration of research and education ensures that future generations of researchers

and educators are well prepared to support new discovery over the long run. Topical areas include: human-computer interaction; digital society and technologies; data and applications security; digital government; and robotics.

Among the successes resulting from prior IIS funding are as follows.

Systems in Context Cluster: The International Children's Digital Library. Research by the University of Maryland and the Internet Archive focuses on developing interface technologies for children (ages 3-13) to access an international collection of 10,000 digital children's books (http://www.icdlbooks.org). On November 20, 2002, the first software demonstration, which included 200 books from 27 cultures in 20 different languages, was launched on the Internet with a celebration at the U.S. Library of Congress. A unique aspect of this research is the collaboration and partnership that has been established (http://www.cs.umd.edu/hcil/kiddesign/introduction.shtml). Interdisciplinary researchers from computer science, library studies, education, art, and psychology are working together with children (ages 7-11) to design this new library. Children's ideas are heard throughout the entire technology design process. Therefore, children work in the labs as researchers twice a week during the school year, and for two intensive weeks over the summer. Together this interdisciplinary and intergenerational team brainstorms, sets project directions, tests new ideas, and implements new technologies. The research has advanced understanding of how children access written materials.

Software Architectures for Microsimulation of Urban Development Transportation and Environmental Impact. The UrbanSim project at the University of Washington has developed a software-based simulation model for integrated planning and analysis of urban development, incorporating the interactions between land use, transportation, and public policy. It is intended for use by Metropolitan Planning Organizations and others needing to interface existing travel models with new land use forecasting and analysis capabilities. Based on the successes thus far with the UrbanSim project, researchers Alan Borning and colleagues are researching ways of building a complete, flexible and scalable microsimulation of urban growth. One of the tests applied to UrbanSim was a historical validation, launching the model with 1980 data for Eugene/Springfield, Oregon, and running it through 1994, comparing the simulated results with what actually happened. Results were very encouraging, with a correlation of better than 0.9.



Eugene/Springfield Oregon Input Data: Parcels

SHARED CYBERINFRASTRUCTURE $123,599,000

The FY 2005 Budget Request for the Shared Cyberinfrastructure (SCI) Subactivity is $123.60 million, an increase of $10.97 million, or 9.7 percent, over the FY 2004 Estimate of $112.63 million.

Shared Cyberinfrastructure Funding
(Dollars in Millions)

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Percent
Shared Cyberinfrastructure	95.07	112.63	123.60	$10.97	9.7%
Total, SCI	$95.07	$112.63	$123.60	$10.97	9.7%

The Shared Cyberinfrastructure (SCI) subactivity supports design, development and deployment of a coherent set of interconnected computational engines, data repositories, digital libraries, sensors and field-specific instruments known as cyberinfrastructure. Such resources are widely shared across multiple scientific and engineering domains and enable shared digital knowledge environments in which researchers and educators create and promulgate new knowledge across distance, time and fields of expertise.

The Shared Cyberinfrastructure subactivity has a single cluster with several areas that together provide the foundation for the shared elements of Cyberinfrastructure.

- **Infrastructure Planning, Construction & Operations Cluster:** SCI supports acquisition, operation and upgrading of national infrastructure in support of high-end computation for the academic research and education community. These resources may include: supercomputers; high-capacity mass-storage systems; system software suites and programming environments; productivity software libraries and tools; large-scale data repositories; and the experts and support staff that create and maintain the facilities.

Advanced Networking Technologies and Infrastructure: SCI supports networks of various reach and granularity from high-speed backbone networks that connect high-performance computational resources and high-end instrumentation sites, to wireless networks that connect embedded sensor nodes in remote scientific field sites. SCI fosters deployment of networks as well as development and fielding of networking technologies that enhance cyberinfrastructure. Some of the key areas include end-to-end networking protocols; performance monitoring tools and measurement infrastructure; wireless networks; strategic international collaborations; and testbeds to support trial deployment.

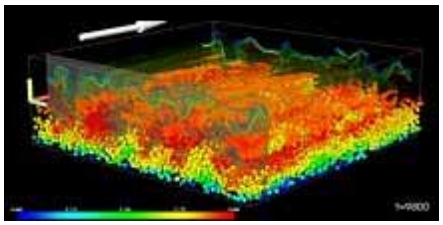
Advanced Services and Cybertools: There is a need for development and support of an array of software tools and services that hide the implementation complexities and heterogeneity while offering clean logical interfaces to users. These tools and services include information management systems and data services, scalable interactive visualization tools, middleware service building blocks for high-end computational resources as well as for networked instrumentations and sensors.

In each of above areas, the subactivity stresses fielding, testing, and ongoing support of advanced technologies beyond basic proof-of-concept demonstrations. The Division collaborates closely with other NSF Directorates and Offices to ensure the advancement of cyberinfrastructure will meet the demands of tomorrow's science and engineering communities.

Examples of SCI supported efforts include:

Tiny Bubbles: SCI-funded computational facilities and innovative application codes are enabling fundamental research on drag reduction for ocean-going vessels that could potentially save substantial fuel costs.

It has long been known that a surface layer of small bubbles on the hull of a large ocean-going vessel reduces the frictional drag as a ship moves through the ocean. However, the physics of this phenomenon has never been well understood. A team with participants from 14 universities led by George Karniadakis of Brown University is advancing understanding of these bubble phenomena. The 20,000 or so simulated bubbles in the Brown team's studies move in a representation of a three-dimensional channel. To produce these models, the Brown team relies on a computational fluid dynamics code called NekTar, created over the years by Karniadakis using NCSA computers and other machines supported by the NSF's Partnerships for Advanced Computational Infrastructure program.



Flow streamline velocity contours for a system of 21,600 microbubbles.

NSF Middleware Initiative's Grid Research Integration Development and Support (**GRIDS) Center**. The GRIDS Center has been created to define, develop, deploy, and support an integrated national middleware infrastructure in support of 21st century science and engineering applications. GRIDS is a partnership of the University of Southern California's Information Sciences, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the University of Chicago, the San Diego Supercomputer Center at the University of California-San Diego and the University of Wisconsin-Madison. Middleware supports the needs of scientists and engineers to collaborate by sharing data, computing, code, and access to experimental facilities. Network-driven computers, storage, data collections and scientific instruments are made available to the day-to-day practice of many research disciplines.

Several innovative cyberinfrastructure projects are already using products of GRIDS. For example, NSF's GriPhyN project uses an international network of computational systems and data collections to address next-generation particle physics experiments at the Large Hadron Collider (LHC), while the NSF-funded Network for Earthquake Engineering Simulation (NEES) is revolutionizing seismology via network-enabled access to experimental facilities, data, and simulations. GriPhyN and NEES represent not today's standard practice, but five- to ten-year strides for these disciplines. GriPhyN is preparing for a torrent of data from LHC experiments to begin in 2006 with a 15-year duration, while NEES is expected to be in place until 2014.

INFORMATION TECHNOLOGY RESEARCH $178,110,000

The FY 2005 Budget Request for the Information Technology Research (ITR) Subactivity is $178.11 million, a decrease of $40.0 million, or 18.3 percent, from the FY 2004 Estimate of $218.11 million.

Information Technology Research Funding
(Dollars in Millions)

	FY 2003 Actual	FY 2004 Estimate	FY 2005 Request	Change over FY 2004 Amount	Change over FY 2004 Percent
Information Technology Research	213.77	218.11	178.11	-40.00	-18.3%
Total, ITR	$213.77	$218.11	$178.11	-40.00	-18.3%

FY 2004 is the last year of Information Technology Research as an NSF priority area. Thus, the CISE activity in FY 2005 focuses on institutionalizing the gains made as a result of ITR investments. Gains of particular note include:

- Enhanced support for cutting edge IT research and related education activities.
- Enhanced support for more focused research in areas of national importance, such as cybersecurity, homeland security, and cyberinfrastructure.
- Enhanced support for larger, more complex projects, including those that are multidisciplinary in nature.

In FY 2005, ITR in CISE will consolidate these gains and continue its transformation of the CISE activity through support of new theme areas that cross the CISE disciplinary areas, require larger and longer duration awards, and that are responsive to national and disciplinary needs. These themes will be multi-year investments and will build on those established in FY 2004 as described below:

Cyber Trust: Networked computers reside at the heart of systems on which people now rely, both in critical national infrastructures and in their homes, cars, and offices. Today, many of these systems are far too vulnerable to cyber attacks that can inhibit their function, corrupt important data, or expose private information.

Cyber Trust promotes a vision of a society in which these systems are:

- more predictable, more accountable, and less vulnerable to attack and abuse;
- developed, configured, operated and evaluated by a well-trained and diverse workforce; and
- used by a public educated in their secure and ethical operation.

To improve national cyber security and achieve the Cyber Trust vision, NSF will support a collection of projects that together: advance the relevant knowledge base; creatively integrate research and education for the benefit of technical specialists and the general populace; and integrate the study of technology with the policy, economic, institutional and usability factors that often determine its deployment and use.

Science of Design: This effort will support science and engineering research and education that develops the foundations of making the design of IT systems a *science*, leading to more effective development, evolution and understanding of systems of large scale, scope and complexity. The emphasis of this program is on software-intensive computing, information and communication systems, (i.e., systems for which software is the principal means to conceptualize, define, model, analyze, develop, integrate,

operate, control, and manage such systems). Other disciplines with a longer history than computing and software have scientifically discovered and validated facts, volumes of codified experience, and formalized, teachable principles. Analogous foundations are needed for a Science of Design for software-intensive systems.

Information Integration: The Information Integration theme will focus on advancing the state of the art in the application of advanced information technology to science and engineering problems in specific domains, such as astronomy, biology, the geosciences, public health and health care delivery. Since many scientific problems have common needs for information management and data analysis, the advancement of these technologies is central. Similarly, within computer science, the study of complex distributed computer and network systems requires the collection and analysis of timely, accurate and reliable information. Within this effort, the NSF intends to support a group of projects that will advance the understanding of technology to enable scientific discovery, and that will creatively integrate research and education for the benefit of technical specialists and the general population.

In FY 2005, CISE will reallocate $40.0 million in ITR funds to the other four CISE sub-activities for the purpose of supporting these themes and other emerging IT priorities, as well as to continue support for larger scale, interdisciplinary projects.