INFORMATION TECHNOLOGY FOR THE TWENTY-FIRST CENTURY: A BOLD INVESTMENT IN AMERICA'S FUTURE

Proposed in the President's FY 2000 Budget

IMPLEMENTATION PLAN

June 1999

National Science and Technology Council
IT[2] Working Group

Table of Contents

1. Overview

With this Information Technology for the Twenty First Century (IT^2) initiative, the Federal Government is making an important re-commitment to fundamental research in information technology. The IT^2 initiative proposes $366 million in increased investments in computing, information, and communications research and development (R&D) to help expand the knowledge base in fundamental information science, advance the Nation's capabilities in cutting edge research, and train the next generation of researchers who will sustain the Information Revolution well into the 21st Century.

Leading-edge information technology, due to its enormous and profound socioeconomic benefits, has, seemingly in a few short years, become critical to our Nation's continued well-being and prosperity. Information technology is also changing the way we live, work, learn, and communicate with each other. For example, advances in information technology can improve the way we educate our children, allow people with disabilities to lead more independent lives, and improve the quality of healthcare for rural Americans through technologies such as telemedicine.

Information technology advances in supercomputers, simulations, and networks are creating a new window into the natural world, making high end computational experimentation an essential tool for pathbreaking scientific discoveries. Advances in weather and climate forecasting are providing early warnings of severe weather, which saves lives, lessens property damage, and helps business be more efficient. U.S. leadership in information technology is also vital to our national security. Our military strategy now relies on information superiority to gain advantage over our adversaries and to keep our troops out of harm's way. The technologies being developed by the Department of Energy's (DOE) Accelerated Strategic Computing Initiative (ASCI) are a key element in assuring the safety, reliability, and performance of our nuclear arsenal.

The technologies resulting from past Federally-sponsored high end research (for example, the Internet, the first graphical Web browser, advanced microprocessors) have helped strengthen American leadership in the information technology industry. Information technology now accounts for one third of U.S. economic growth and employs 7.4 million Americans at wages that are more than 60 percent higher than the private sector average. All sectors of the U.S. economy are using information technology to compete and win in global markets, and business-to-business electronic commerce in the U.S. alone is projected to grow to $1.3 trillion by 2003.

The Nation needs significant new investments in information technology research to help ensure the underpinnings of future economic growth and to address important national problems in defense, education, the environment, health care, and transportation.

IT^2 builds on the Government's previous accomplishments and existing investments in High Performance Computing and Communications (HPCC), including the Next Generation Internet (NGI) and the DOE's Accelerated Strategic Computing Initiative. The IT^2 research agenda responds directly to the findings and recommendations of the President's Congressionally-chartered Information Technology Advisory Committee (PITAC), which concluded in a report released in February 1999 that the Federal investment in Information Technology R&D is inadequate relative to its importance to the Nation. The investments planned by the IT^2 initiative will augment the base HPCC programs to fund critically-needed extensions of some ongoing HPCC research agendas and expansions into entirely new research areas, as recommended by the PITAC. The IT^2 focuses explicitly on long-term, fundamental research to address the underinvestments noted by the PITAC in its report. When added to existing HPCC investments, new funding through IT^2 will provide a necessary first step in restoring the imbalance between fundamental research and development and shorter-term, mission oriented research and development in the current Federal portfolio.

The initiative will extend some existing research and development and provide opportunities to address new, complementary research and development topics in three key areas:

- Long term information technology research and development leading to fundamental advances in computing and communications
- Advanced computing infrastructure to facilitate scientific and engineering discoveries of national interest
- Research on the economic and social implications of the Information Revolution, and the training of additional Information Technology workers at our universities

Six agencies — all highly dependent on advances in Information Technology to carry out their increasingly complex missions — are participating in the initiative:

- Department of Defense (DoD) (including the Defense Advanced Research Projects Agency [DARPA])
- Department of Energy (DOE)
- National Aeronautics and Space Administration (NASA)
- National Institutes of Health (NIH)
- National Oceanic and Atmospheric Administration (NOAA)
- National Science Foundation (NSF)

Each agency will bring its own unique mix of capabilities and expertise to bear on an appropriate subset of the R&D objectives of the initiative. Planned activities will be closely coordinated to connect complementary efforts across research disciplines and funding agencies. All partners in IT^2 are committed to maintaining fundamental research as the core of the initiative. Leadership by NSF, the only agency with a clear mission to support fundamental research, and annual reviews by the PITAC are mechanisms that will help ensure a properly balanced Federal research portfolio.

The IT2 initiative builds on base funding in HPCC, for which the proposed FY 2000 budget is $1.462 billion (including $543 million for ASCI). The proposed IT^2 FY 2000 budget is as follows:

Agency	Fundamental Information Technology Research and Development	Advanced Computing for Science, Engineering, and the Nation	Social, Economic, and Workforce Implications of Information Technology	Total
DoD	$100M	—	—	$100M
DOE	$ 6M	$62M	$ 2M	$ 70M
NASA	$18M	$19M	$ 1M	$ 38M
NIH	$ 2M	$ 2M	$ 2M	$ 6M
NOAA	$ 2M	$ 4M	—	$ 6M
NSF	$100M	$36M	$10M	$146M
Total	$228M	$123M	$ 15M	$366M

The IT^2 initiative assumes that the activities reported as the HPCC programs will continue at the current funding levels. Continued funding of the HPCC programs is critical to the success of the IT^2 initiative since many of the IT^2 activities build on the HPCC base programs.

Should the initiative be funded, it will be managed together with the HPCC programs to ensure the best leverage of Federal investments in complementary information technology research and development programs and to avoid duplication of efforts. Integration of the management structures for IT^2 and HPCC should be completed by September 1999. Beginning in FY 2001, the HPCC programs and the IT^2 initiative will be reported in a single budget crosscut based on the integrated HPCC and IT^2 programs.

Fundamental Information Technology Research and Development

The information technology underpinning modern society is in large measure the result of past advances in the field of computer science and engineering — the theoretical and experimental science base on which computing applications build. Fundamental research in computer science and engineering generates the knowledge and concepts that will become the information technology of the future. The importance of this fundamental research and development in information technology can only increase in the future. As the complexity of computing grows, so will the need for well-understood concepts and theories to manage this complexity. Entirely new research problems and opportunities are created each day by rapid technological advances in information technology.

The fundamental information technology research and development component of the IT^2 initiative will address long-term, high risk investigations in issues that confront computer science and engineering. Four research focal points address the PITAC's recommendations:

- Software
  - Software engineering
  - End-use programming

  - Component-based software development
  - Active software
  - Autonomous software
- Human computer interfaces and information management
  - Computers that speak, listen, and understand human language
  - Computer sensors and actuators that enhance human physical and mental capabilities
  - An "electronic information exchange
  - Information visualization
- Scalable information infrastructure
  - Deeply networked systems
  - Anytime, anywhere connectivity
  - Network modeling and simulation
- High end computing
  - Improving the performance and efficiency of high end computers
  - Creating a computational grid
  - Revolutionary computing

These research areas constitute a diversified program for conducting long-term research on making computing and information systems easier to use, more reliable and secure, more effective, and more productive.

This multiagency effort will be led by NSF, with participation by DoD, DOE, NASA, NIH, and NOAA.

Advanced Computing for Science, Engineering, and the Nation

IT^2 will enable a wide range of scientific and technological discoveries by allowing complex simulations to run on the highest capability computing systems accessible to researchers around the country. Recent scientific and technical advances make possible a quantum leap in computational and data management capabilities and the application of advanced computing to nationally important strategic problems. We can realistically project that by the year 2005 end-to-end computations 1,000 times more powerful than those of today will be achieved. Challenging problems not otherwise solvable with existing computational and experimental resources will require and foster fundamental new developments in computer science and engineering as well as parallel progress in algorithm development, database management, networking, and related areas. In order to reach these goals by 2005, this component of the IT^2 initiative will establish a staged program of research and development commencing in FY 2000.

IT^2 will procure and deploy the world's most powerful computers to tackle Grand Challenge-class computing problems. This component of the initiative will:

- Obtain computers that are 100 to 1,000 times more powerful than those now available to the research community and make them available to researchers on a competitive basis
- Develop scientific and engineering simulation software and other tools needed to make these new machines useful, including new mathematical algorithms and parallel programming environments and tools for collaboration, visualization, and data management
- Build multidisciplinary teams that allow researchers working in the most challenging science and engineering research areas to benefit from advances in fundamental information science resulting from other IT^2 work, and computer scientists to explore challenging new information technology problems

This multiagency effort will be led by DOE and NSF, with participation by NASA, NIH, and NOAA.

Social, Economic, and Workforce Implications of Information Technology and Information Technology Workforce Development

Social, economic, and workforce implications of information technology

Information technology is important to the nation not only because of its value in advancing science and technology and U.S. competitiveness, but because of the ubiquity of the technology and its effect on all aspects of citizens' lives. The impacts of information technology on our society, economy, and workforce include massive changes in the nature of work, commerce, education and training, entertainment, financial management, and quality of life. Understanding these changes will enable rational decisions to be made in Government and the private sector in allocating resources and planning policies to take advantage of the opportunities that information technology brings to society.

IT^2 will support focused research on the social, economic, and workforce implications of information technology as an integral part of its activities. Projects will include research to:

- Enhance the usefulness of information technology, for example, use of information technology in educational settings
- Limit potential misuse of information technology, for example, issues related to data privacy and security
- Understand how the knowledge, values, and systems of society influence the spread of information technology and the acceptability of information technology in various aspects of our lives

Research will encompass activities within and among disciplines. It will include both focused activities and social, economic, and workforce components of fundamental information technology research and computational infrastructure development efforts.

This multiagency effort will be led by NSF, with participation by DOE, NASA, and NIH.

Information technology workforce development

Science agencies must do more to address the rapidly growing demand for workers with high end information technology skills at the undergraduate, graduate, and post-graduate levels. Today, colleges and universities are reporting increased undergraduate enrollments in information technology while applications for graduate study are actually declining. Faculty in two-year colleges, four-year colleges, and research universities need access to modern curricula and instructional materials that provide a strong foundation for future work or further study. Likewise, students need graduate traineeships, research assistantships, and postdoctoral positions that enable them to pursue advanced degrees or additional training.

IT^2 will support a number of activities that address workforce development issues, including:

- Activities to improve education in information technology
- Support — through a competitive solicitation — for a national center or group of regional centers to assure that an appropriate range of training opportunities, including opportunities to review and develop modern curricula, are available to undergraduate faculty
- Development of model curricula in computational sciences and computer literacy to help raise the level of understanding, employ computing in other science and engineering disciplines, and help college students become more computer literate
- Opportunities for graduate and postdoctoral participation on research awards, to be provided by all agencies participating in IT^2
- Workforce Retraining and Access to Distributed Learning, including:
  - Enhanced and expanded outreach programs operated through existing programs and centers, with emphasis on newly funded centers and new information technology-related projects funded at existing centers.
  - Development of the technology and research base for advanced distributed learning, including software development and distribution (for software that implements distributed learning); research on using technology in education (delivery, assessment, collaboration) and assistance to new sites in its adoption; and establishment of a clearinghouse for information on distance learning in information technology
  - A new digital library focusing on coursework in information technology and training for information technology professionals that builds on existing digital library research for K-12 through undergraduate levels

This multiagency effort will be led by NSF with participation by DOE, NASA, and NIH.

Management and Coordination

The IT^2 initiative will be coordinated within the framework of the National Science and Technology Council (NSTC) in conjunction with existing base HPCC programs and the NGI Initiative.

A Senior Principals Group from NSTC, chaired by the Assistant to the President for Science and Technology, has been formed to set policy and coordinate the work of the initiative. This Group initially consists of the principals from the IT^2 funding agencies:

- Director of the National Science Foundation
- Administrator of the National Aeronautics and Space Administration
- Under Secretary of Energy
- Under Secretary of Commerce for Oceans and Atmosphere and Administrator, National Oceanic and Atmospheric Administration
- Director of the National Institutes of Health
- Under Secretary of Defense (Acquisition and Technology)

It also includes senior officials from the Office of Management and Budget (OMB) and the National Economic Council (NEC). The Senior Principals Group assists the Assistant to the President for Science and Technology in establishing program goals and monitoring the progress of the initiative, balancing requirements for meeting IT^2 goals with agency missions and capabilities, ensuring tight coordination of Federal efforts, and making sure that the research program allocates funds in an open, competitive process geared toward supporting the best ideas.

The Senior Principals Group is supported by an IT^2 Working Group chaired by the Assistant Director of NSF for Computer and Information Science and Engineering. The IT^2 Working Group consists of agency representatives with operational authority over information technology research and/or information technology infrastructure within their agencies. The National Coordination Office (NCO) for Computing, Information, and Communications provides coordination support.

The IT^2 Working Group is charged with preparing research plans, budgets, and the detailed implementation for the initiative. It is also developing a transition plan to provide for coordination of IT^2 with ongoing base research programs, including HPCC and NGI, by September 1999. The integrated coordination structure will use existing groups and teams to the maximum extent possible, facilitating coordination among agencies in projects that require close and continued partnerships to ensure that collective research activities provide a sound and balanced national research portfolio. Some tasks may require new subgroups. One such new subgroup, chaired by NSF and DOE, is preparing a plan to develop and operate the new advanced infrastructure funded by the two agencies through IT^2. This will ensure that the infrastructure is purchased, sited, and made available on an open, competitive basis. The subgroup will assure that the systems are made available to the research teams with the most compelling research concepts and the best ideas for building partnerships between experts in state-of-the-art information and computational science, and groups familiar with research challenges in application areas such as biochemistry or climate modeling that can benefit from access to ultra-fast machines.

Within the integrated management and coordination structure, agencies will retain control over their own budgets and will support the coordinated effort only in areas where they have authority. Each agency will use its own method for inviting proposals. Agencies will coordinate their activities through program managers and experts from other agencies in review processes and in interagency IT^2 activities.

Time Line Summary

The following table lists IT^2 deliverables. Outyear deliverables depend on regular increases in funding for this initiative.

Deliverables	**First Achieved**
Research and development in software, including high confidence systems, human computer interaction, large scale networking and scalable infrastructures, and high end computing, begins in response to open competitive solicitations	FY 2000
Teams for pursuing mission agency problems sets begin work	FY 2000
Combined 10 teraflops computing power available to the science and engineering research community for both open competitive research and mission-directed research	FY 2001
Prototypes and demonstrations of software, including high confidence systems, human computer interaction, large scale networking and scalable infrastructures, and high end computing, begins in response to open competitive solicitations	FY 2002
Combined 80 teraflops computing power available to the science and engineering research community for both open competitive research and mission-directed research	FY 2004*
High-resolution three-dimensional simulations of five or more different complete complex systems (such as an airplane's flight, combustion devices, and the human body) for agency mission applications and general science and engineering are demonstrated	FY 2004

*IT^2 anticipates availability of machines with this power in FY 2004, but acquisition of this capability for users will depend on price and IT^2 funding levels.

2. Fundamental Information Technology Research and Development

Introduction and Strategy

Research and development in this area will investigate long-term, high risk issues confronting computer science and engineering. This IT^2 component has four focal points: software, human computer interfaces and information management, scalable information infrastructure, and high end computing. Research and development in these areas will encourage a diversified, long term strategy to make computing and information systems easier to use, more reliable and secure, more effective, and more productive.

NSF will lead this multiagency effort, and DoD will make significant contributions in software and scalable information infrastructure. DOE, NASA, NIH, and NOAA will also fund activities in this area. Each agency will encourage participation through its normal mechanisms. Selection of research proposals will be coordinated through cross agency review panels, and agency program managers will help ensure that the proposals do not unnecessarily duplicate other efforts. The intent is to ensure a range of complementary research efforts including single and multiple principal-investigator (PI) research, multiple-PI/multiple-field collaborations, intramural research in institutes and Federal laboratories, and joint industry-government-academia experiments or proofs-of-concept.

Details of each agency's plans, including research programs, funding mechanisms, and milestones, are described in the Agency Specifics at the end of this section.

Software

From the desktop computer to the phone system to the stock market, our economy and society have become increasingly dependent on software. The PITAC concluded that not only is the demand for software exceeding our ability to produce it, but that the software produced today is fragile, unreliable, and difficult to design, test, maintain, and upgrade. The small software failures that shut down large parts of the Nation's phone systems and the "Year 2000" (Y2K) problem are but two significant examples of what can go wrong.

Software — unlike bridges, airplanes, and nuclear power plants — is constructed without the benefit of standard practice by technologists who lack thorough training in the engineering sciences of reliability, maintainability, and cost-effectiveness.

The research and development proposed in this area will significantly improve the concepts, techniques, and tools to engineer the software infrastructure, and will help educate software developers in well-founded and more productive engineering methods to develop techniques and tools to create productive, reliable, and useful software.

- *Software engineering:* Currently, we do not understand how to design and test complex software systems with millions of lines of code in the same way that we can verify that a bridge is safe. Fundamental software research will increase software productivity, make software more reliable and easier to maintain, and automatically discover errors. These efforts will reduce the cost and time penalties currently required to make systems safe and reliable. Progress on safety and reliability is important in critical systems such as the telecommunications network, medical devices, the electric power grid, and the air traffic control system.

- *End-user programming:* One way to address our shortage of programmers is empower end-users through the creation of domain-specific development tools — the spreadsheet being perhaps the most common example. Just as telephone companies, when faced with a scalability barrier with switchboard operators, solved the problem by putting dialing technology in the hands of the end-users, so can software developers push portions of development out to end-users through various kinds of "smartware," domain-specific and user-friendly templates for filling a variety of needs. Developing tools and techniques to make end-user programming more widespread will require advances in intelligent templates, domain-specific languages, and programming-by-example.

- *Component-based software development:* Today, most programs are written from scratch, one line of code at a time. The software industry lacks the equivalent of the interchangeable parts used in manufacturing. Component-based research will make it easier to find the right software component, to accurately predict the behavior of a software system assembled from smaller components, and to support an electronic marketplace in software components.

- *Active software:* Active software participates in its own development and deployment. We see the first steps towards active software with "applets" that can be downloaded from the Internet. Research in active software will lead to software that can update itself, monitor its progress toward a particular goal, discover a new capability needed for the task at hand, and safely and securely download additional software needed to perform that task.

- *Autonomous software:* Research in this area will lead to more intelligent software and robots, such as unmanned vehicles that keep our troops from harm's way; intelligent agents or "knowbots" that search the Internet on our behalf; robots and knowbots that plan, react appropriately to unpredicted changes, and cooperate with humans and other robots; cars that can drive themselves and automatically avoid collisions; and robots that can explore planets or places on Earth that are unsafe for human travel.

Human-Computer interaction and information management

The amount of information available today is outrunning the combined attention span of all the people in the world put together. The number of processors in use worldwide will soon be larger than our planet's population. To use this information for societal and economic benefit, we need better ways for computers to manage information and interact with people. The research proposed in this area will significantly improve the concepts, techniques, and tools to engineer the software infrastructure and educate software developers in well-founded and more productive engineering methods. Research and development in this area will develop techniques and tools to create productive, reliable, and useful software. Such efforts will include:

- *Computers that speak, listen, and understand human language:* Today, more than 40 percent of American households own computers. However, most Americans find computers too hard to use, and surveys show that computer users waste over 12 percent of their time because they can't understand what their computers are doing. Research in this area will result in improved human-computer interfaces that make computers easier and more enjoyable to use, resulting in an increase in productivity.

  Ideally, people could converse with a computer, as opposed to being limited to today's WIMP (windows, icons, mouse, pointer) interface. Researchers will attempt to make computer speech more intelligible, increase the accuracy of speech recognition, and allow computers to ask questions to confirm or clarify what a user has said. This ability will be particularly useful for people who lack access to a keyboard, such as doctors and mobile professionals, or for people with disabilities. Coupling this technology with robotic extenders will enhance productivity even more. Additional research will enable simultaneous translation from one language to another. The knowledge and techniques gained in this area will make possible a realtime translating telephone, and allow people to accurately search foreign language databases in their native tongue.

- *Computer sensors and actuators that enhance human physical and mental capabilities:* Multiplying individual physical and mental capabilities via computer sensors and actuators to control devices will aid people at work and at home. For example, robotic systems could be developed to allow surgeons to manipulate and repair blood vessels.

- *An "electronic information exchange":* Today's digital libraries represent a key step in this direction, but are limited both in content and accessibility. Building an "electronic information exchange" will allow every American to retrieve text, data, visual, and other information in a wide range of subject areas.

- *Information visualization:* Computer users are attempting to understand increasingly complicated phenomena. Scientists trying to make long-term climate predictions, for example, must analyze data on hundreds of phenomena, such as changes in jet streams, snow and cloud cover, atmospheric carbon dioxide, and ocean circulation. The goal of visualization research is to significantly improve our ability to see, understand, and manipulate huge quantities of data.

Scalable information infrastructure

The growth of the Internet has been phenomenal. In 1985, the Internet connected 2,000 computers. Today's Internet connects 37 million computers and an estimated 153 million users. As the Internet becomes more pervasive, the networks of tomorrow will need to support a billion or more users sending and receiving voice, video, and high-speed data at home and while traveling; transmitting tens of millions of simultaneous requests for information from popular Web sites; and relying on the Internet to run a business, deliver government services, or respond to a medical emergency. The Internet of the future will connect billions or trillions of devices. Computers will be combined with sensors, wireless modems, global positioning system (GPS) locators, and devices that can interact with the real world-all shrunk to the size of a single chip and transparently embedded in everyday objects.

However, today's Internet technology was not designed to support this explosion in the number of users and devices. The goal of Scalable Information Infrastructure (SII) research and development is to develop tools and techniques to enable the Internet to grow ("scale") to support these new demands. Research goals are:

- *Deeply networked systems:* Researchers will develop software and expertise enabling a dramatic increase in the number of devices that can be attached to the Internet. This will encourage compelling new applications, such as low-cost wireless sensors to provide realtime information on air and water pollution, will improve our ability to monitor the environment and respond to manmade disasters; "guardian angels" that monitor the health and safety of fire fighters, law enforcement officials, soldiers, and home health care patients; and crisis management centers that use sensors carried by response teams and airplanes to improve responses to forest fires, floods, and hurricanes. Research will focus on naming, addressing, and network configuration, and on developing less expensive network interfaces.
- *Anytime, anywhere connectivity:* Research will focus on wireless technology to bring high-speed "anytime, anywhere" connectivity to all U.S. citizens. Such wireless networks will expand access and reduce the cost penalty associated with deploying advanced telecommunications in rural areas — critical to rural economic development since companies are increasingly basing site selection on the quality of the telecommunications infrastructure. Wireless networks can extend services such as distance learning and telemedicine to remote rural areas in the U.S. and to markets in developing countries.
- *Network modeling and simulation:* The goal of this research is to develop tools to model the behavior of networks and to provide the capability for "faster than realtime" simulation to allow network operators to automatically prevent congestion or collapse.

High end computing

Advances in high end computing benefit the country, the economy, and the lives of all U.S. citizens. Research in high end computing can have far-reaching benefits to society including new medical insights, weather forecasts with greater accuracy, more accurate climate change predictions, and advanced weapon systems design. Over the next decades, computation will become even more essential to the Nation, as we rely more heavily on high end computers for science and engineering endeavors, both civilian and military, and explore new applications of high end computing such as data mining and crises management.

As a result of investments by the Federal HPCC programs and DOE's ASCI, DOE national laboratories operate computing systems capable of sustaining a trillion calculations per second on key applications code to ensure the integrity of our nuclear stockpile. However, achieving this computational rate on a broad range of applications and making efficient use of massively parallel machines requires an aggressive research program — particularly in systems software — that includes long-term research to develop computers capable of a thousand trillion (10^{15}) calculations per second. In addition, there is evidence that current massively parallel computers may not be well suited for some applications — especially applications with highly irregular or where large quantities of data need to be transferred from memory to perform the calculation. To address these challenges research and development is needed on software to improve the performance of high end computers and in innovative computing technologies and architectures.

The Government has a particularly important role to play in high end computing. Since the market for this kind of computing is small, national security and other Government missions frequently require machines that are faster and more complex than those in use by the private sector. Research topics include:

- *Improving the performance and efficiency of high end:* This research will focus on developing programs that can be easily moved from one high end computer to another, and on creating new algorithms, compilers, and problem-solving environments that are designed with massively parallel machines in mind. U.S. companies can now make high end computers that can perform trillions of calculations per second. However, they are "massively parallel" machines that contain thousands of individual microprocessors. It has proven difficult to configure and program these machines to perform at their theoretical capacity. The goal of this research is to make these machines easier to program and improve the efficiency of typical programs.
- *Creating a computational grid:* This research will develop software to translate the application requirements into requirements for computers, networks, and storage; security mechanisms permitting resources to be accessed only by authorized users; and computers and operating systems that are more tightly integrated with high speed networks. The goal is to develop a computing grid that works the same way that the electric power grid that provides universal access to electrical power. A computational grid could provide widespread access to computational power, allowing users to request additional computer resources on demand, construct a supercomputer from many smaller computers connected to the Internet, take advantage of computers that are idle, interact with simulations and very large databases in real time, and collaborate with colleagues — who may be halfway around the world — in three-dimensional (3-D) virtual environments.
- *Revolutionary computing:* This research will explore fundamentally new computer technologies such as biological computing, single electron transistors, quantum logic computers, and devices based on carbon nanotubes, possibly leading to computers more than a thousand times faster than today's fastest supercomputers.

Agency Specifics

DoD

The DoD research and development efforts will enable revolutionary capabilities of importance to the DoD and the Nation. DoD's priorities in this regard are aligned with Joint Vision 2010, the Chairman of the Joint Chiefs' conceptual template for achieving new levels of warfighting effectiveness. This vision forecasts dynamic change in the nature of potential adversaries, emphasizes the increasingly critical nature of technological advances and their implications, and outlines the emerging importance of information superiority.

DoD participation in IT^2 will be coordinated by the Deputy Under Secretary of Defense (Science and Technology) DUSD(S&T), whose primary responsibility is to maintain US technological superiority over potential adversaries. DoD IT^2 participation includes DARPA ($70 million), DUSD(S&T) University Research Initiatives ($10 million), and the Advanced Research and Development Activity in Information Technology for the Intelligence and Information Security Communities ($20 million). DoD's activities will involve research and development to provide the foundations for orders-of-magnitude increases in information processing capabilities, through programs managed by the NSA, DARPA, and DoD research offices.

Fundamental Information Technology Research and Development

The "fundamental" descriptor used to describe this component is not the DoD formal definition which would cover only basic research. Much of the DoD IT^2 program is advanced research or advanced technology development.

DARPA

SOFTWARE RESEARCH

This research and development will develop software to enable reliable, safe, and cooperative operation of free ranging autonomous systems, including software for mobile robots (air, land, or maritime unmanned vehicles) performing tasks in dynamic unstructured (physical) environments without the need for synchronous operator control inputs or high quality communications links. Similarly, this effort includes the development of software agents (knowbots) that can range over cyberspace performing information services, and negotiating for and assigning select resources. These autonomous systems should be able to learn and adapt to change and uncertainty while improving with experience.

Autonomous Systems will enable revolutionary, asymmetric military capabilities, such as the ability to autonomously convey military payloads (both lethal and nonlethal) to any portion of the battlefield without requiring human operators, and the ability to autonomously retrieve, process, and deliver information.

The Common Software for Robotics component will develop a software platform that can be reused across a range of experimental autonomous systems.

The Software Enabled Control component will leverage increased processor and memory capacity to vastly increase our ability to maintain control over mobile devices by developing novel techniques such as predictive mode changes, dynamic control scheduling, and composable control and dynamic sensor and actuator allocation.

The Negotiation component will enable autonomous operation of large collections of agents negotiating resource allocation issues, such as those encountered in logistics and countermeasures.

SCALABLE NETWORKS

Extending DoD's ability to monitor and shape the physical environment will require a much deeper approach to information systems — one that manages the vast quantities of physical information that can be accessed by sensors and actuators in direct contact with real world processes. In order to do this, the network infrastructure must be extended to deal with a wide diversity of embedded devices accessing physical world information; vast increases in the numbers of nodes with realtime transmission requirements; and operating regimes in which network-based nodes must host services on behalf of embedded clients.

The large scale networking of embedded and autonomous devices creates new requirements for multimode network interface technologies that can achieve drastic reductions in costs while being compatible with a wide range of network media and flexible mechanisms for naming, addressing, configuring, and administering that will make the deployment and operation of a hundred billion part infrastructure feasible. (These challenges are addressed in the Network Interface component of this project.)

Future defense uses of the network will increasingly emphasize the direct exchange of realtime sensor-derived information among autonomous embedded devices. This reflects a significant change in network traffic from the present environment that is dominated by the exchange of symbolic information among human users. The new traffic models, architectures, and protocols needed to effect this transition will be investigated in the Near Realtime Networking component of this project.

Many applications of deeply networked systems will perform data dissemination and fusion operations that could most efficiently be performed at nodes within the network. The Agile Network Services component of this project will leverage the capabilities of a programmable network substrate to deploy nomadic middleware that can go where network connectivity permits, allowing network elements to host services on behalf of embedded and autonomous devices.

High end Computing

DoD's high end research will focus on component technologies that, while suitable for use in next generation supercomputers, also have the potential to revolutionize the design of systems that acquire and process information directly acquired from the physical world. These include:

- Configurable Processors to allow high performance applications to reconfigure the processor logic on a cycle-by-cycle basis
- Data-Intensive Systems that use logic-in-memory fabrication techniques and programmable caches to achieve dramatic reductions in memory system latency
- Bio-Digital Interfaces and Processing Techniques that will bridge the gap between silicon and organic systems, extending our ability to sense and control a range of chemical processes

University Research Initiative (URI)

The DoD University Research Initiative IT^2 activities will support university research to fundamentally advance the collection, storage, processing, communication, networking, dissemination, retrieval, and display of large quantities of multimedia information such as speech, data, graphics, and video. Research projects will be conducted by multidisciplinary teams of investigators in related science and engineering fields. Projects will be selected through open competition. Research includes, but is not limited to, the following:

- Quantitative information processing and data fusion in large arrays of microsensors
- Computational methods and decision making in command and control involving information uncertainty
- Fundamental algorithmic principles for mobile computation, visualization, and augmented reality
- Foundations of adaptive, realtime, fault-tolerant protocols for communication and computer networks
- Information processing by optical computing and solitons
- Quantum computing, quantum communication, and quantum memory for information processing and transmission

- Adaptive mobile wireless networks technology for distributed and highly dynamic computing and communication environments

DoD will also provide funding to universities to purchase equipment supporting defense related research in:

- Software engineering — end-user programming, component-based software development, active software, and autonomous software
- High confidence systems — assurance foundations and technologies, information security, survivability technology, software control systems, and public key infrastructures
- Networking — reliability, security robustness, quality of service, bandwidth allocation, collaboration technologies, network modeling and simulation, digital libraries, distributed computing, and anytime/anywhere connectivity
- Human centered systems — collaborative workspace, information storage, information agents, distributed cognition, knowledge acquisition, multilingual technologies, virtual reality environments, and visualization systems
- High end computing — advanced computer hardware, cooperative computing, system software technologies, future-generation computing, advanced computing technologies, high end simulation software, national computer grid, and high end computing architectures

Advanced Research and Development Activity (ARDA)

The Advanced Research and Development Activity will fund high-risk/high-payoff research in information technology to support the Intelligence and Information Security Communities. ARDA will also implement a resource enhancement program to attract talented scientists and researchers to supplement the Intelligence Community's R&D workforce, transfer technology developed or discovered as a result of the initiative, and form strategic partnerships with world-class researchers in academia and industry.

ARDA's IT^2 activities will focus long-term research on a small number of problems relevant to the Information Security and Intelligence Communities and to support enabling technologies relevant to both. In FY 2000, the majority of the funds will be directed toward three research areas:

- Information exploitation investments to develop technology to automate time-consuming functions currently performed by scarce intelligence analysts. These include advanced question and answering capability to allow community analysts to pose complex queries in a natural form; data mining to look for new, unusual or unexpected information within very large sources of multimedia, multilingual data; and information analysis and synthesis tools to organize, analyze, and represent information appropriately.
- High performance computing R&D to ensure that the intelligence community has the computing and data storage technology needed in the future. This area includes initiatives in supercomputer architecture design, very high speed photonic logic, quantum computing, and high capacity (terabyte) data storage.

- Scalability Information Infrastructure R&D to address issues related to the vulnerabilities of modern networks. This area includes modeling networks to assess their vulnerability, developing high speed networks, and reverse engineering hardware and software.

Milestones

DARPA

FY 2000

- Multi-Mode Network Interfaces ($9.0 million)
  - Develop laboratory prototypes of efficient multi-mode interface to networks employing diverse link protocols, symbol rates, channel codings, and signaling technologies
  - Investigate protocols for dynamic configuration of client interfaces
  - Specify a heterogeneous network architecture that integrates wireless, wireline, and satellite communication; investigate addressing schemes and routing protocols to support vertical hand-offs across network boundaries
- Near Realtime Networking ($11.0 million)
  - Experiment with time stamping of network payloads in support of performance monitoring
  - Analyze network and client nodes to identify key contributors to network latency
  - Develop new traffic models to analyze the impact of sensor generated information on network performance; investigate protocol interactions and cross talk
- Agile Network Services ($5.0 million)
  - Develop a framework for automated migration of client specified proxy services to internal network nodes
  - Investigate techniques for load balancing and service placement optimization
- Common Software for Robotics ($13.0 million)
  - Develop an architecture for integrating deliberative, reactive, and learning behavior
  - Define a strategy to account for and integrate emergent behaviors
  - Identify alternative approaches to software for distributed robotics
- Software Enabled Control ($9.0 million)
  - Specify an architecture for a hybrid control system that synthesizes the control law approach with computationally-enabled mode logic scalable to very large state spaces of 100K+ states

- Develop active transition control and joint mode logic/control law designs
- Implement tools for active model creation, augmentation, and query

• Agent Based Negotiation ($5.0 million)

- Develop a framework for bottom-up organization of autonomous software
- Analyze autonomous software ability to predict, negotiate, and track resource requirements under changing environments
- Implement software toolkit for knowbot development, generation, and deployment

• Common Operating Environment for Embedded Systems ($8.0 million)

- Develop methods for distributing and operating embedded software to allow the C^3I capabilities of embedded devices to be upgraded or altered on a near-realtime basis
- Initiate rapid prototyping activities using functional equivalent to target operating environment based on augmenting COTS technology to originate and host mobile code
- Investigate suitability of aspect-oriented approaches to specification and generation of embedded systems software

• Large Scale Networks of Sensors ($7.0 million)

- Specify a gradient-based approach to automated aggregation and distribution of information from large numbers of multi-taskable sensor nodes
- Determine the mapping of gradient primitives to multicast and incast network protocols

• Declarative Tasking and Querying of Embedded Systems ($3.0 million)

- Investigate the use of declarative interfaces for tasking and querying of networked embedded systems; develop alpha level prototype based on relational database query technology

FY 2001

• Multi-Mode Network Interfaces

- Prototype implementation of network software and application interfaces for multi-mode interfaces
- Demonstrate high quality-of-service connectivity among wireless, wireline, and satellite networks
- Prototype implementation of a network architecture for transparent vertical handoffs between different access interfaces

• Near Realtime Networking

  - Specify architectural requirements to achieve order of magnitude reduction in worst case latency induced by packet buffering
  - Prototype demonstration of sensor/actuator feedback loop operating over near realtime Internet

- Agile Network Services
  - Demonstrate dynamic programmability of select network components
  - Develop capability to support the migration of continuously operating client proxy services
  - Demonstrate agile network services in support of distributed robotics applications

- Common Operating Environment for Embedded Systems
  - Alpha level prototype of embedded systems operating environment incorporating native support for mobile code and ability to perform in-situ upgrades of all runtime modules
  - Specify interfaces supporting common runtime services required by signal processing and generation applications
  - Demonstrate dynamic generation and loading of signal processing applets

- Large Scale Networks of Sensors
  - Implement experimental prototype supporting automated aggregation and distribution of sensor derived information involving at least 50 nodes and 100 sensors
  - Demonstrate dissemination of sensor tasking, data collection, integration, and analysis over a multi-node sensor network spanning tens of square kilometers

- Declarative Tasking and Querying of Embedded Systems
  - Prototype demonstration using declarative interfaces for tasking and querying of multitaskable sensor networks
  - Prototype implementation of middleware services that bridge the gap between logical and physical infrastructure — for example, by providing and using geo-location information

- Common Software for Robotics
  - Prototype implementation demonstrating integration of deliberative, reactive, and learning behavior
  - Laboratory demonstration of compatible knowledge representations for reprogrammable behavior-based control
  - Prototype demonstration and experimental evaluation of software for distributed robotics capable of coordinating the operation of 10+ vehicles

- Software Enabled Control

- Alpha-level prototype implementation of multimode control architecture and framework

- Develop parametric predictive and adaptive control frameworks

- Complete multilevel multimodal advanced design tools

• Agent Based Negotiation

- Prototype demonstration of autonomous software ability to use negotiation in logistics scenario

- Demonstrate and evaluate software agent's ability to approximate behavior tradeoffs and use negotiation in advanced logistics scenario with a three second response requirement

- Prototype implementation of negotiation technology in realtime with a 500 millisecond response requirement

DOE

DOE's IT^2 activities are part of its Scientific Simulation Initiative (SSI). SSI is an integrated project to develop a national scientific simulation computational infrastructure to solve complex scientific problems that cannot be addressed by other means. To do so requires (1) the creation of a high end computing system 10 to 100 times more powerful than any now available to civilian science, (2) the development of computer science, applied mathematics, and system software environments necessary to exploit these computers, and (3) the simultaneous development of scientific models that use the full capability of the computing system. SSI will build on advances in high end computing, high speed networking, and scientific model development realized in the HPCC programs and ASCI.

While DOE's Scientific Simulation Initiative (SSI) focuses on revolutionary uses of computing as tools for science, many of its investments in computer science and enabling technologies will contribute to and benefit from the broader Fundamental Information Technology Research component of the IT^2.

Fundamental Information Technology R&D

The goal of the SSI Computer Science and Enabling Technology (CSET) program is to develop and deploy innovative computing technologies that will make possible new scientific simulations. Such simulations will require breakthrough software tools and strategies that can handle highly complex problem domains on the most advanced ultra-scale computing and communications environments. The strategy in these areas is to take advantage of the most recent research in software components and extend it to function at very high performance levels. In addition, hardware available at SSI performance levels requires that its individual components be fault tolerant and that these components can be assembled into reliable fault tolerant systems.

In response to the PITAC recommendations on software development, DOE is positioning SSI as a significant fundamental research program in software development and component technologies, and will begin to develop a national repository of software components. DoE activities in the Fundamental Information Technology R&D component of the IT^2 initiative include efforts in each of the four CSET components:

• Algorithms, models, methods and libraries — building blocks for scalable scientific applications, with a focus on robustness and novel interfaces that will enable programmers to combine software from different libraries.

- Problem-solving environments and tools — software components for developing and managing software, with a focus on portability, not only for moving programs from one machine to another but from one hardware generation to the next.
- Distributed computing and collaboration technology — research to enable collaborative simulation, telepresence, and remote resource sharing; research topics will include new resource management, security, and communication protocols.
- Visualization and data management — significant new research in human computer interfaces and interaction to enable researchers to navigate and extract scientific knowledge from petabyte-scale data archives from SSI scale simulations and the coming generation of large experimental facilities; SSI planning workshops and ASCI Visualization Corridors workshops identified these areas as candidates for an "Expedition to the 21st Century."

Management Plan

DOE will rely on its existing management structure and is instituting additional measures to ensure corporate direction, oversight, and effective management of SSI, including the readiness, feasibility, and added value of future applications for terascale systems. The following management plan covers all three DOE elements of the IT^2 initiative:

The Existing DOE R&D Council, chaired by the Under Secretary and vice-chaired by the Director of the Office of Science (SC), is comprised of Assistant Secretaries from each of DOE's science and technology programs. The Council is the governing body for corporate Science and Technology (S&T) direction and management, and will establish a new Large-scale Simulation Subcommittee to oversee Department-wide large-scale simulation activities including SSI.

The Large-Scale Scientific Simulation Subcommittee will be co-chaired by the Director of SSI and the Head of ASCI and will consist of representatives from each DOE program area. This Subcommittee will establish Department-wide program direction and policy and will coordinate the scientific simulation activities across the Defense Programs (DP), Office of Science (SC), and other technology programs.

ASCI/SSI Technical Working Groups will be formed to coordinate the detailed implementation plans for the computer science and enabling technologies (for example, large scale data visualization and management) and other areas that promise mutual benefit. These groups of technical researchers will report to the Large-Scale Scientific Simulation Subcommittee.

An SSI Director will be responsible for day-to-day SSI management and coordination of activities across all SC programs. Working closely with SC program mangers, the SSI Director will oversee development of scientific applications, guide the establishment and operations of the hardware facilities, coordinate the activities of the S&T crosscutting technology and application teams, and provide implementation direction. The SSI Director will report to the Director of the Office of Science, co-chair the Large-scale Simulation Subcommittee, and act as a liaison to the IT^2 Interagency Working Group. Working closely with the IT^2 agency partners, the SSI Director will manage the DOE program to ensure coordination and collaboration in interagency planning, policy, and implementation activities and, where appropriate, joint solicitations. The SSI Director will direct the coordination and dissemination of SSI information to the broader scientific and engineering community.

Management of DOE's HPCC and NGI programs will not change, but will be coordinated with SSI.

Milestones

FY 2000 ($6.0 million)

To prepare for the large terascale systems anticipated in FY 2003 and FY 2004, DOE will initiate projects in the following areas:

- Algorithms, models, methods, and libraries ($1 million)
- Problem solving environments and tools ($1 million)
- Distributed computing and collaboration technologies ($1 million)
- Visualization and data management ($3 million)

NASA

NASA's IT^2 activities are part of its Intelligent Systems (IS) program. IS is designed to begin a national mid- to long-term research program to fulfill or exceed the NASA Administrator's vision for next generation information technology. The IS program will develop revolutionary, state-of-the-art IS technologies by leveraging Government and university research, and moving maturing technologies to ongoing NASA missions and activities, industry activities, and other Government agencies.

The overall objective of the IS program is to conduct research, develop, and transfer the information technologies required to support NASA's 21st century missions. Our increased dependency on pervasive information technology coupled with the leveraging nature of IS will enable a wide range of applications and missions, some of which we can only dimly glimpse today. NASA will support the following mission-critical application areas, which are listed with representative five-year IS program milestone goals that IT^2 will enable:

- *Autonomous Spacecraft and Rovers:* Spacecraft monitor, diagnose, and recover from faults. They respond to high-level goals, rather than low-level action sequences, accomplishing more per mission. Smaller ground teams are needed, dramatically cutting operations costs.
- *Automated Software Engineering:* Domain experts build high-assurance, domain-specific, cost effective software using program synthesis systems.
- *Aviation Operations:* Airport ground operations streamlined with automated planning and scheduling. Ground-based airline operations teams can work with remote onboard automation and aircrews for fleet routing and management.
- *Support for Intelligent Synthesis Environment:* Teams in remote sites can work together over a distributed computation and communication infrastructure. Design work is focused on models of the artifacts being built, not on details of the tools used for analysis, including sophisticated virtual reality visualization tools.

- *Human Exploration of Space:* Mission-specific mobile computing and communicating devices aid astronauts. Smaller ground teams are needed, substantially cutting operations costs.

Fundamental Information Technology R&D

The IS program is a comprehensive strategy to integrate high risk research, concept and prototype development, and the transfer of mature technologies to program customers. Research areas include:

- Automated reasoning that takes the programmer out of the loop. Scientists can solve problems in their own terms rather than having to program solutions, and spacecraft can operate autonomously rather than responding to ground commands.
- Human centered computing that focuses on system designs that look at how humans and machines interact, taking into account basic human perceptual, cognitive, and social abilities
- Revolutionary computing resulting in breakthrough technologies that can change the way we think of computation.

Execution Plan

The IS program will be developed and implemented consistent with NASA Procedures and Guidelines (NPG 7120.5A) governing the planning, development, implementation, and review of major NASA programs. The IS concept has been under development within NASA for several years and an Integrated Product Development Team (IPDT) with an assigned lead NASA center has been established. The IPDT will integrate the IS program elements.

Milestones

The following milestones focus on implementing self-tasking, self-repairing robotic networks. These technologies will contribute to future decisions on developing permanent robotic outposts at key points of scientific interest throughout the solar system, and will be developed in close coordination with science objectives. Near-term demonstrations of these technologies will take place within micromissions in NASA's Space Science programs.

FY 2000 ($18.0 million)
- Mature model-based programming process and tools (Intelligent Systems)

FY 2001
- Mature model-based programming process and tools

FY 2002
- Techniques for remote collaboration with distributed environments

FY 2003
- Coordinated planning and execution for fleets

FY 2004
- Demonstrate miniaturization of processors, memory devices and sensors based on biochips and biometric architectures
- High-bandwidth sharing of distributed human/machine resources

- Cognitive task analysis for human-computer systems with variable autonomy

NIH

One of the world's foremost leaders in medical research, NIH funds almost 30 percent of the Nation's health-related research and development. The NIH mission is to uncover new knowledge leading to better health for everyone. NIH conducts supporting research in its own laboratories; supports research of non-Federal scientists in universities, medical schools, hospitals, and research institutions throughout the country and abroad; helps to train research investigators; and fosters communication of biomedical information.

To achieve its mission, NIH relies on bioinformatics, high end computing systems, high-speed networks, and other advanced information technologies to provide the performance needed in today's biomedical computing and information rich environment. The information generated by fundamental science, such as genetic sequencing and mapping, and the information and computation required for optimal practice of medicine, have exceeded the capacity of earlier methods of data storage and manipulation. Further advances in biology, and our ability to diagnose and treat many of our most devastating illnesses and disabilities, are dependent upon information technology.

Recent advances and current NIH projects demonstrate how the medical sciences have relied on information technology. These projects include:

- The full blueprint for a multi-celled organism *C. elegans* or roundworm. Announced in December, 1998, this marked an historic step for the Human Genome Project by being the first time scientists deciphered the genetic instruction of a complete organism that shares many basic life functions with humans. *C. elegans* shares 70 percent of its genes with humans and therefore serves as a model for human genetic research and provides entrée into the discovery of cures to human diseases.

- In the area of information technology workforce, NIH is working with science teachers in the Nation's K-12 classes by providing to them state-of the-art curricula developed by NIH scientists and an expert curriculum development group. NIH will provide the science curricula to teachers over the Internet, and will provide video presentations by NIH scientists using future improved Internet services such as interactive Web sites and faster and crisper video imaging.

Fundamental Information Technology R&D

NIH's support of IT^2 will further the Nation's biomedical research progress, and improve health, through a number of initiatives:

SOFTWARE AND ALGORITHM RESEARCH AND DEVELOPMENT

The increased complexity and interdependency of biomedical research and computation has created a tremendous need for improved and robust software in bioinformatics, imaging, statistical analysis, ontologies (knowledge bases), and three-dimensional image labeling and searching. NIH proposes to:

- Establish programs to develop software and algorithms to meet the needs of biomedical researchers. These programs, in part driven by the rapid expansion of sequencing the human genome, will also create transferable software that NIH-funded researchers can use at national centers and at their own institutions.

- Develop simulation methods to implement realistic models of molecular, cellular, organ, and epidemiological systems on parallel architectures and create component libraries that will provide a framework to allow researchers to assemble these parts into applications that serve their specific needs.

- Further the work of the Brain Molecular Anatomy Project (BMAP), a multi-institute initiative that supports research on the genomics of the nervous system, with initial efforts focusing on the discovery of new genes and the study of gene expression to quantify and track the expression of tens of thousands of genes in space and time, and to generate substantial amounts of useful data.

Execution Plan

The Fundamental Information Technology Research and Development program will be implemented through NIH extramural and intramural mechanisms, and will be administered as a part of the larger NIH portfolio. Overall coordination will be integrated through a trans-institute management group as an expansion of the current HPCC management.

Milestones

FY 2000 ($2.0 million)

- Develop software and algorithms designed to meet the needs of biomedical researchers

- Develop simulation methods for realistic models of molecular, cellular, organ, and epidemiological systems on parallel architectures

- Further the work of the Brain Molecular Anatomy Project

FY 2001 - FY 2004

- Future NIH activities will be guided by the report of the Advisory Committee to the NIH Director (ACD) Working Group on Biomedical Computing, released in June, 1999.

NOAA

NOAA relies heavily on high performance computing, storage, and communications to carry out its mission. Large-scale computer modeling of the Earth's atmosphere and oceans has been a fundamental tool for both weather prediction and climate research in NOAA since the early days of computing, and realtime collection and storage of environmental data and the rapid dissemination of forecasts and research results has always been key to NOAA's mission. Today, NOAA develops and uses comprehensive computer models of the atmosphere and oceans for research and operational forecasting; maintains geophysical and satellite databases approaching a thousand trillion bytes (one petabyte) in size; routinely collects environmental data in real time in response to disasters; and disseminates information products to emergency managers, business, and the public to protect life and property.

Fundamental Information Technology R&D

NOAA's key interest in fundamental information technology research is to develop improved software capabilities to support climate and weather research and forecasting. NOAA research will:

- Explore the application of component-based software to complex coupled simulations of the oceans, atmosphere and land surface. Key objectives are to identify and develop software components that support rapid scientific development. The behavior of the combined components must be stable and predictable. Such software must be flexible to permit rapid prototyping as new hypotheses are generated and advanced geophysical understanding is modeled.

- Improve the performance of cache-based systems for high priority atmosphere, ocean, and climate problems. The current mix of hardware and software systems is capable of achieving only a fraction of the potential of modern cache-based commodity processors. Research is needed to develop more efficient tools, algorithms, and problem-solving environments to take full advantage of the capabilities these processors offer.
- One of the key impediments to widespread collaboration in climate research is the lack of widely-used adaptable data formats that can be used efficiently on modern scalable architectures. In cooperation with its university partners, NOAA will develop and use self-describing standards-based data formats to facilitate sharing data and analysis tools among different research groups, and implement them in advanced model diagnostics and analysis applications.

Execution Plan

NOAA has centered the responsibility for IT^2 in the Office of High Performance Computing and Communications, which identifies opportunities for NOAA participation in interagency activities such as IT^2 and manages the coordination of activities with other agencies. NOAA research has identified clear needs for improved component-based methods, better approaches for use of commodity cache-based processors for geophysical research, and improved standards-based data formats that are optimized for use in highly parallel systems.

Building on the fundamental research of NSF, DARPA, and others in these areas, NOAA will partner with key universities to explore and develop these new technologies to augment existing knowledge and skills in the three areas identified above.

Milestones

FY 2000 ($1.5 million)

- Initiate the NOAA Geophysical Fluid Dynamics Laboratory (GFDL)-university software collaboration in high performance computing to address scalability, portability, and optimization of fluid dynamics codes on highly parallel cache-based systems

FY 2001

- In cooperation with university partners, NOAA will refine and use advanced programming and cache management techniques to develop and to optimize a more advanced version of the GFDL Hurricane Prediction System for execution on different scalable architecture systems within the agency

FY 2002

- Develop component-based design methods to simplify the exchange of model physics packages, maximize the usability of the modeling system, and facilitate model maintenance. These will be implemented in specific codes including the next-generation GFDL coupled atmosphere-ocean research model

FY 2003

- In cooperation with the university community, GFDL will develop and use self-describing, standards-based data formats to facilitate the sharing of data and analysis tools among different research groups and implement them in advanced model diagnostics and analysis applications

NSF

NSF's mission is to promote the progress of science and engineering in the national interest. NSF supports fundamental research and education activities, including support for user facilities in many fields. Over the past decade, such activities in computer and information science and engineering have received high priority in NSF planning and budget development. As an active participant in both the HPCC Program and the NGI Initiative, NSF has helped develop the foundations on which the IT^2 will flourish.

NSF's role in fundamental information technology research is to mobilize the broad base of academic researchers to provide a solid foundation for computer and information science and engineering. Over the past decade, many elements of computer and information science and engineering have seen rapid growth in the knowledge base, much of it in the context of attempting to solve specific problems. Much of the activity has been ad hoc in character, minimizing the cross-problem flow of information and the usefulness of the work in development of theoretical underpinnings. The ubiquity of information technology is a clear argument for more systematic development of tools for its use. Thus, an underlying theme of NSF's approach to fundamental research in IT^2 will be to build a strong science and engineering base for future development of information technology that is applicable across levels of scale and types of activity

Several changes in the scope and scale of NSF programs will occur as NSF implements IT^2, including:

- Increased emphasis on long-term, high-risk research going beyond current paradigms
- Continuing discussion with the science and engineering communities to identify emerging opportunities and adapt NSF emphases to take advantage of them
- Larger awards for longer durations enabling investigators to tackle more ambitious problems and provide appropriate levels of support for students and technicians
- Opportunities to develop research groups and centers that bring multiple perspectives, sources of data, and realtime requirements for communication to bear on difficult fundamental information technology problems

Such changes will enable NSF-supported researchers to take on the challenges of today's and tomorrow's information technology systems with their mix of local and distributed processors of different types, rapid communication capabilities, and orders-of-magnitude changes in speed, size, and storage and retrieval capability.

Fundamental Information Technology R&D

SOFTWARE

Basic information technology research is concerned with the foundations underlying the design and use of computing systems. It covers foundational research on computing, algorithms, development of systems, and tools and techniques.

Planned IT^2 activities build on the base of FY 1999 activities in:

- The software development process, including languages and tools that support and improve that process
- Theoretical foundations of algorithm design, particularly aspects key to solving advanced complex scientific and engineering problems

- High performance processor architecture
- Communications systems and signal processing software and hardware
- Design methods to facilitate the design of billion-transistor chips
- High-capacity data storage systems and information retrieval
- Digital government

The IT^2 initiative will permit expansion and linkage of these activities through efforts of greater scope and scale that take advantage of the wide range of software. Program announcement language will be designed to elicit proposals for innovative approaches with the potential to contribute to:

- Building "no-surprise," performance-engineered software systems to deliver scalability, functionality, predictability, and security. There will be a strong emphasis on research to extract design principles from large, successful software projects and to validate promising approaches such as component-based software design.
- Developing hardware/software co-design to help build vertically integrated electronic systems that may be contained on a single chip. In the future, such systems will be the heart of many critical systems and technologies.
- Building high-confidence systems by addressing a range of significant software challenges such as predictability, reliability, and security.

HUMAN COMPUTER INTERACTION AND INFORMATION MANAGEMENT

Researchers in human computer interaction and information management seek to improve technologies enabling humans to work with computing systems more effectively. Scientists, engineers, educators, students, the workforce, and the general public all benefit from increased accessibility and usability of computing systems and communications networks. Planned IT^2 activities build on the base of FY 1999 activities in the following areas:

- Knowledge repositories, information agents, and digital libraries
- Multimodal interfaces between humans and computers, including communication by sight, sound, touch, and gesture
- Visualization, virtual reality, and 3-D imagery tools
- Collaboratories that facilitate knowledge sharing, group decision-making, control of remote instruments, and datasharing across large distributed systems
- Cognitive and perceptual abilities as they affect computer design
- Multilingual technologies, speech recognition tools, and audio interfaces
- Sensor-based databases and systems

While continuing to pursue these areas, NSF will encourage integrative activities through IT^2 that can lead to:

- *Multiplying individual physical and mental capabilities* via computer sensors and actuators to control devices that aid people at work and at home, and research on high performance communications links, high end computational engines, and the worldwide information resources to manage them
- *Meeting, working, and collaborating in cyberspace,* to allow interactions over networks in a realistic 3-D virtual environment
- *Building an ubiquitous content infrastructure* that enables seamless retrieval of text, data, visual, and other available information in all subject areas by all citizens

SCALABLE INFORMATION INFRASTRUCTURE

The networking infrastructure of the nation has been growing at an astounding rate. Many reports indicate that the traffic on the internet is doubling every hundred days. Commerce is adopting the internet as a preferred avenue of doing business. The development of the underlying infrastructure is a major success story. However, if the nation is to maintain a leadership position in this burgeoning industry, major new initiatives in fundamental research in scalable information infrastructure must be undertaken. The focus of NSF research activity is in the following three theme areas:

SCALABLE MOBILE INFRASTRUCTURE

The objective of this theme is to provide wideband access to the information infrastructure anywhere, anytime. A major research challenge is to deliver tetherless multi-megabits/s by 2010. Global seamless roaming, once achieved, will enable user equipment to operate everywhere with the same human interface at tens of megabits per second, thus permitting real-time multimedia access to the Internet regardless of mobility with mixing of traffic types (i.e., voice, video, and data).

To get there, we need systems-level research to enable optimization and adaptation across layers (physical through applications). Areas of research needed to reach this goal include:

- Space-time processing
- Joint source and channel coding
- Wideband multichannel characterization
- Network protocols adaptively matched to channels
- Practical means to implement multiuser theory
- Dynamic resource allocation and management
- Wideband signal processing (software radios)

We will encourage broad approaches to solutions initially along the following lines.

- The huge explosion of the cellular communications industry has prompted communications companies to explore the use of antenna arrays to improve both the channel capacity and the quality of wireless communications by exploiting the enhanced interference rejection capability they provide. However, several basic theoretical issues must be resolved before widespread adoption is possible. The most fundamental of these involves system-level deployment — the development of so-called space-time codes, in which coding is performed concurrently across both the spatial and temporal dimensions to achieve superior redundancy.

- Practical sources, such as speech and images, contain a significant amount of redundancy, which is generally removed through compression. The compressed information is highly affected by channel errors, however. A solution is to impose error-correction coding on the compressed information, which adds redundancy, but in a manner that can be used for error correction at the receiver. This serial process of first removing redundancy by compression and then adding it back for error correction can in practical; situations be made more efficient if done jointly, or as a one-step process.
- Systems-level integration of communications and signal processing requires hardware/software codesign, for example techniques that optimize the end-to-end performance of the system, subject to many local/systems-level constraints, and which are applicable to the highly complex information processing systems envisioned 5-20 years in the future, must make maximum utilization of software radio techniques.
- Major system performance gains may be possible through development of mobile networking architectures and protocols to handle the particular constraints of mobile and fixed wireless systems and to smooth the interface between wireless and fixed networking environments. Fundamental wireless network system limitations and constraints need to be formulated.

Complex Systems Modeling, Simulation and Emulation

Modeling, simulation and emulation, separately or together, provide a substrate for many activities in on computing and communications. In general they allow us as scientists and engineers to predict behaviors of algorithms and designs, where real testing may not yet be possible. The predictions may be used to correct errors in design prior to any deployment or construction. They may be used to configure networks, collections of servers, deploy web caches in networks, etc. They may also be used in real time to tune the execution of functions to changing demands and constraints of users and complex systems in networks.

Network Modeling

Algorithmic models define behaviors. To the extent that we can embody the behaviors of our designs in algorithms, we can use the models to determine the correctness of and bounds on behaviors. The most exciting discovery in distributed computational models in the 90's is that network traffic when aggregated exhibits self-similar behavior after suitable normalization at different time scales. This behavior has now been discovered not only in network traffic, but in long-scene video transmission, and user behaviors as well. Prior to this, these behaviors were assumed to be Poisson or Markovian. The discovery of the behavior ranging from packet level traffic to users has enormous impact on our understanding of the effectiveness of network protocols and distributed behaviors at higher levels of abstraction. Thus some questions that might be addressed within the scope of this theme are:

- Why is self-similarity applicable in these cases, and in general?
- Is there a generalization we can make about the applicability of the long range dependencies that lead to the self-similarity behaviors discovered?
- Is it the case that there are only these two alternatives of assumptions of independence of events or the long range dependencies?

Network Simulation

Although models alone may tell us a great deal, much of what needs to be learned cannot be modeled directly. A next step is simulation, some of the inputs to which will be models to the extent we have them. Simulation allows us to explore more complexity, by considering the interactions of a variety of mechanisms, injecting models for behaviors where they are known, and often injecting real traffic measurements to drive the simulations where possible. The variety of ongoing problems in this area is enormous.

- As our systems become more complex, how do our simulations expand their capabilities to handle this? Certain kinds of problems are amenable to various speed up techniques such as parallelization of the simulator, but it is not clear yet whether these techniques are particularly well tuned to communications and networking problems.
- How can we incorporate our strong interest and need to understand behaviors at various different levels of abstraction simultaneously? The problem at hand might be how a new web-based application may behave from the perspective of users. But at the same time, it will be important to understand the impact it will have on the core trees in the network used for the distribution of multicast traffic, or the ability of the routers at the IP layer to provide the quality of service reserved for the application. In fact, it may be the case that one only wants detailed low level information at certain points in the simulated network, allowing for boring down through the layers only at certain points. But a simulation of this sort needs to have behavioral models of users, perhaps end-nodes (the application or the operating system), routers, and so forth. At the same time, it must enable the incorporation of existing protocol implementations as they are, rather than requiring re-implementation. The reason for this is that experience teaches us that the details of particular implementations are often the issue, as seen when TCP implementations caused the whole network to self-synchronize and exhibit cyclic behavior.
- The bottom line is how do we enable the ability to deal with scaling in our simulations?

Network Emulation

Emulation is the third component of allowing for experimenting with behavior without full deployment. Emulation allows for the deployment of not only the protocols themselves, but presumably more of the complete system. An emulated environment would allow real users to perform real activities in real time. The bottom line is that total effectiveness can only be evaluated by the final effectiveness of the complete complex system. Emulation is often the first step in such real deployment.

Not only are these tools used prior to deployment in order to understand problems prior to that deployment, but the tools, especially modeling and simulation, are necessary components to continuing management and adaptability. Examples of questions one might ask are:

- In order to allow for reservations of specific qualities of service from the network, how might the network be enabled to predict the behavior of the reservations already in place?
- In order to quickly deploy web caches as traffic patterns change, how should the system not only measure network and user behaviors (looking back at the past) but also using that information to predict future behavior in order to act on it?
- What sorts of algorithms and other issues must be addressed in order that these sorts of measurements, models, and simulations be able run in real time in the execution environment with the dual constraints of being effective (fast and correct enough) while not being too invasive?

This theme of modeling, simulation, and emulation is both interdisciplinary in that it spans all the layers of computer science and is an area in which our community has a great deal to learn from other large scale scientific simulation communities (meteorology, astronomy, etc.). It is also key to our community's ability to succeed at deploying the global, pervasive, computational infrastructure we are heading toward.

Pervasive Environments: Interfaces, Computing, and Communication

With the dramatic fall in both size and price of processors, approximately 90-95% of processors manufactured today are being embedded in devices, not used as monolithic computers. As such the processors do not necessarily have all the other elements available in a general purpose computer, such as large amounts of memory or caches, persistent storage or programmable I/O devices. They are often intended for limited functionality, but they are becoming pervasive.

At the same time, such computationally capable devices are also joining networks. Some of these are extremely local and statically organized, such as in a car. Others may be more mobile, with respect to their peers, such as devices that support Bluetooth and other local wireless technologies. Yet others may simply move on and off the Internet, such as personal digital assistants (PDAs) and their descendants.

The pervasive environments theme will cross the boundaries among all the traditional Computer Science disciplines. Research within the scope of this theme could include, but is not restricted to Human Computer Interfaces, Scalable and Pervasive Middleware Infrastructure, Network Architectural issues, and new Computational Paradigms. Although this precis does not address more theoretical issues, there are also significant possibilities addressing such issues as modeling, verification, theorem proving about correctness, and so forth, as related to this theme as well.

Human Computer/Network Interfaces

At the Human Computer Interface, it is not clear where the balance falls between cyberspace moving into real space vs. humans moving into cyberspace. There are communities that believe that the right approach is avatars for humans and capabilities that cannot exist in real space. There are communities that believe that computing should become "the telephone" or "the car," meeting the needs of humans, but on human terms, not cyberspace terms. For example, if the human needs the use of some device, several questions we might ask are:

- Should the human be able to point to a device and identify it to "the system" as a candidate for use, or should the human always be in a virtual reality in which a device is requisitioned for the human? Whichever model might be chosen, as a system becomes more complex and pervasive in our society, it will need to become more sophisticated, we must have fewer expectations of humans in control of the details.
- How do we make the environment as a whole be smarter, and better able to adapt to the needs and requirements of humans? For example, rather than providing an interface to low level data, we must move up the scale to information and thence to knowledge, in the human sense.
- How do we as a society architect networked knowledge bases that better meet our needs as a whole society rather than databases that can only be utilized by the cognoscenti or in isolation?

Scalable and Pervasive Middleware

At a more mundane level, examples of what one might ask are:

- How do we build network environments that become smarter by learning from humans or from the human interfaces at least?
- How does the human cause an environment to be modified or enhanced without having to become a system or network manager?

This leads directly into a variety of middleware and networking issues a subset of which are:

- How do we build a system that incorporates many orders of magnitude more elements?
- How do we deal with the fact that the elements may not fall into some limited set of categories, such as hosts, web sites, etc., but may include not only so many more devices, but also fragments of code, robots, agents or knowbots, avatars, etc. Some may be active, others passive, and perhaps those distinctions are no longer important.
- How do we figure out what we are looking for in such a huge space?
- How do we locate things, once we have figured out what we are looking for?
- How do we organize such "a system"?

Network Architecture

At lower levels of communication, we probably have to imagine a universe in which a variety of network architectures exist. Bluetooth devices do not speak TCP/IP, so there is a need for a proxy of some sort. In other cases, small devices, which have much more computational power than an early PC, might very well speak TCP/IP. But to make the problem more complex, all optical networks, at least as they are being designed to date, do not provide active elements within the network. Some questions one might address here are:

- How one might provide quality of service or other policy mechanism within the network without leaving the optical world, in which case, one may have lost the advantages of an all-optical system?
- In a more general form, how might one provide end-to-end functionality that needs activity inside the network, when the network actually is a network of networks that only provide very different operating models within themselves? It is not clear that our current network architecture models provide the right abstractions in the right places to achieve what is needed. For example, routing which is now done at a very low level, may well be needed at a higher layer, and it isn't clear what a multi-tiered architecture ought to look like, or how we design an architecture that can evolve into the future and allow for partial deployment, since we are likely always to be in a state of transition.

A New Computational Paradigm?

If we are to build such a pervasive environment, what will the computing paradigm be? There are researchers now proposing that "an application" should no longer be simply a pile of code, but rather an active, self-reconfiguring, adaptive entity that monitors situations and either comes into existence only when and as needed, or perhaps reorganizes itself to be most effective at its tasks. This will lead to a need to revisit a great many of the software engineering and operating systems issues that have been addressed, often only more or less successfully in the past. Some questions that might be addressed in this component of this theme are:

- If software is reorganizing itself, how can we have any assurance of its correctness and trustworthiness?
- More specifically, for example, if "an application" is redefined during execution, how do we provide the continuity that might be needed, despite changing supporting infrastructure. We don't have a good model of partial computations in the face of optimizing structure.
- What does it mean to do semantic checkpointing?

There are several significant cross-cutting issues, such as information infrastructure management and security and privacy, which crosscut all of the above research areas.

HIGH END COMPUTING

High end computing research is the core of the original HPCC programs, and includes research needed to advance the state of the art in computing and bring advanced computational capabilities to bear on fundamental science and engineering problems.

NSF's broad range of ongoing activity in high end computing includes:

- Program environments and tools
- Visualization and data handling
- Scalable numerical algorithms
- Multidisciplinary analysis and design
- Handling complex multi-level memory hierarchies
- Remote collaboration on high performance computing applications
- Cutting-edge disciplinary applications such as simulations of complex biological systems, design of new chemical compounds, materials simulations, and comprehensive modeling of coastal systems

Advances in computer hardware technology create the possibility of a three-order-of-magnitude increase in speed with comparable increases in memory capabilities, opening an enormous new arena for scientific discovery and a concurrent drive to solve computer science-related problems to accelerate the effective use of parallel and distributed systems. To empower computational discovery across science and engineering, NSF will use IT^2 to:

- Initiate a multidisciplinary research program, Terascale Opportunities for Promoting Science (TOPS), to accelerate the wide introduction of recent developments in simulation and modeling and the underlying computer science, mathematics, and technology that make them possible into cutting edge research across the full spectrum of science and engineering disciplines
- Extend the interdisciplinary Large Scientific and Software Data Set Visualization program to study the presentation of large data sets, including those from software state and from large scientific simulations and experiments. This program will energize the field of visualization and data handling, and will enable scientific progress in other areas by the study of large, important data sets.

NSF will also extend research activities in related areas such as:

- Visualization or, more generally, presentation of large data sets
- Storage of information on persistent media, for example, intelligent, multilevel, adaptive control of file systems and storage devices for complex access patterns such as subsetting
- Algorithms related to computational complexity, particularly algorithms for solving the partial differential equations that arise from modeling physical phenomena

Execution Plan

NSF has centered responsibility for IT^2 in the Directorate for Computer and Information Science and Engineering (CISE), which also will integrate IT^2 with ongoing efforts. This element of IT^2 will add $100 million for fundamental information technology research.

NSF will issue a unified program announcement that provides opportunities to submit competitive proposals simultaneously for existing programs and the IT^2 activities that integrate them and permit attention to enhanced scale and scope. The announcement will highlight the areas of emphasis listed above and encourage innovative, creative, and potentially risky ideas. It will also point proposers to activities implementing the other components of IT^2 and to joint activities with other agencies and other parts of NSF that may have free-standing announcements appearing in the unified announcement. Proposals will be due beginning in November 1999.

Many of the highlighted research areas have been subjects of workshops involving the science and engineering community in developing priority areas for emphasis in the competitions. Others will be the subjects of upcoming workshops. In addition to helping NSF better understand the cutting edge for research in these areas, the workshops draw opportunities to the attention of potential proposers. Major outreach activities are anticipated in conjunction with issuing announcements of opportunities, providing potential proposers with the chance to meet with program officers and discuss how they might best become involved.

NSF will implement its fundamental research component of IT^2 using three closely coupled modes of support:

- Individual investigator awards, usually including support for students, equipment, and other resources used to conduct the project
- Team awards that may require specialized instrumentation
- Large projects, where the talents and perspectives of several researchers are brought to bear on issues requiring a larger scale of resources due to their larger scope

Individual investigator awards will comprise approximately $30 million of the $100 million of IT^2 funds. Award size and duration are expected to increase significantly over current norms, to an average of $150,000 per year. Most will be made as standard grants to minimize the burden on the community.

Team awards (usually three to five senior scientists) will range up to about $1 million per year for three to five years, with interim project reviews for those of longer duration. NSF plans to devote about $40 million to such awards. Team awards provide the opportunity for a single proposal for research and instrumentation elements of projects requiring special instrumentation. Such awards are particularly relevant for experimental testbeds and will be used in part to expand the community knowledgeable in under-funded research areas.

Large project awards will average $3 million per year for five years. NSF plans to make a few such awards using the cooperative agreement mechanism. The initial award is expected to be for five years, with a promise for up to ten years with appropriate progress. Progress and decision reviews will occur at regular intervals. Both localized and virtual distributed centers are possible. This activity will broaden and strengthen the institutional base by encouraging collaborative partnerships to improve future research and the personnel base.

Proposals for all three categories will be reviewed together in panels that follow due dates in timely ways. The most highly ranked center proposals will undergo a second stage of review designed specifically to compare center proposals across research areas.

Milestones

March to May 1999

- Work with other agencies and other parts of NSF to determine where joint announcements of opportunity would be most valuable

August 1999

- Issue unified announcement of opportunities

November 1999 - January 2000

- Proposals due at NSF

January through March 2000

- Review process

April through July 2000

- Make awards. Individual investigator and team awards early in the period; center awards will require National Science Board (NSB) approval.

March through July 2000

- Determine program structure for FY 2001 and issue announcements

3. Advanced Computing for Science, Engineering, and the Nation

Introduction and Strategy

During the past decade, fueled by the exponential increases in computing power provided by information technology, the U.S. has made dramatic progress in its ability to model the fundamental physical, chemical, and biological processes of nature. For example, the 1998 Nobel Prize in Chemistry was awarded for the development of very sophisticated, yet widely used theory and software for simulating the behavior of small molecules. IT^2 will support a series of ambitious efforts to obtain and use the world's most powerful computers to attack problems of critical national interest. The Advanced Computing for Science, Engineering, and the Nation component of IT^2 will focus on creating and making available to the science and engineering community a previously unavailable jump in computational speed and power. This will enable researchers to attack important new classes of problems than cannot be studied at the present time and enhance the capabilities of current computational models. To accomplish this IT^2 will:

- Obtain computers that are 100 to 1,000 times more powerful than those now available to the research community and make them available on a competitive basis
- Develop the scientific and engineering simulation software and other tools (such as new mathematical algorithms and parallel programming environments as well as tools for collaboration, visualization, and data management systems) needed for these new machines to be useful for research
- Build multidisciplinary teams where scientists working in the most challenging research areas can benefit from advances in basic information science made in other parts of IT^2, and where computer scientists can explore difficult new problems

The Advanced Computing for Science, Engineering, and the Nation component of IT^2 provides for balanced, integrated coordinated access across the entire spectrum of research community needs, consistent with the specific missions of the agencies taking part. There is a very broad and overlapping research community engaged in high end computation in which DOE, NIH, NOAA, NASA, and NSF play important parts. In this the first four agencies focus primarily on their mission responsibilities, and as a result must be selective in the research areas they emphasize. NSF is charged with providing resources for the fundamental research and education community across all science and engineering disciplines. IT^2 is structured to meet both mission-specific and general disciplinary requirements. The agencies' programs are complementary and, taken together and integrated, effectively address the computing needs of the entire national scientific community.

Advanced infrastructure

The fastest computers now available to the civilian research community on a competitive basis are capable of about 100 billion computations (such as a simple addition) a second. This initiative will make it possible for this community to use machines capable of five trillion (a thousand billion) computations per second by the end of FY 2000 and 40 trillion by the year 2003. These machines will be equipped with software and operating systems, data storage, internal memory, and communication links to support a broad spectrum of potential applications for teams located throughout the country. Development of these parts of the high end infrastructure will benefit from the fundamental information technology research and development conducted in IT^2.

The terascale computing facilities made available as part of the IT^2 will be funded by DOE and NSF. Both of these agencies have a long history of making high end computing facilities available to the civilian research community, although the range of disciplines that use the facilities supported by NSF is broader than typical at DOE-supported computing facilities. DOE and NSF will work closely together to site these computers, and to ensure cooperation in order to meet IT^2 goals and objectives. Both NSF and DOE are committed to competition among researchers or research teams for access to IT^2 terascale computing facilities. DOE and NSF will engage in a well-coordinated competitive process (for example, through the issuance of a joint solicitation), open to all research performers for the development, operation, and use of the facilities.

Advanced science and engineering computation

IT^2 will build multidisciplinary teams that allow researchers working in the most challenging science and engineering research areas to benefit from advances in fundamental information science resulting from other IT^2 work, and computer scientists to explore challenging new information technology problems. The HPCC Grand Challenge problems demonstrated that strong interaction between challenging applications and new architectures and software systems, and cooperation between computer and computational scientists can lead to advances in both computing and the particular scientific domain involved. Indeed, the best basic research is often stimulated by attempts to solve tough practical problems. The search for a good substitute for a mechanical switch resulted in the transistor, and this search in turn resulted in extraordinary advances in solid state physics. Information science will be no exception. IT^2 is designed to build a strong working relationship between two types of teams:

- Teams working on scientific and engineering simulations, whose primary interest is attacking problems such as climate change or research physics
- Teams interested in the broad field of information technology research, whether it is in visualization or basic aspects of software design

Both groups should benefit from this relationship. Researchers in IT^2 will take vastly different approaches from those currently in wide use. Individually or in groups, they will need to have deep understanding of the underlying science and engineering, a working knowledge of current computational processes, the ability to modify current processes to accommodate massive parallelism (with all that entails, from software design through message-passing protocols), and analytical skills to take full advantage of the computational output. Knowledge gained about scientific and engineering computation in a highly parallel environment will be put to use by other scientists and engineers and by researchers in information technology to help establish the foundations for a new generation of advanced computation.

Information Technology for the Twenty-First Century:

The following are examples of potential science and engineering problems that will become tractable with the computational power proposed for IT^2:

- *Predicting the climate:* Significant improvements are needed in both the accuracy of the forecasts and their ability to make predictions for each part of the country. This will require higher resolution and will include additional features such as ocean and ecological effects.

- *Predicting severe weather:* Many life-threatening weather conditions such as flash floods and tornadoes are still beyond our ability to model and forecast with adequate skill. To be successful, we need to increase model detail, model complexity and physics, data capacity, and event detection in a reliable, robust computational system. Weather simulations are central to today's weather forecasts but to reach their full potential we will need vastly increased computing and communications capabilities.

- *Understanding genetic function:* Many of the functions of the human body are carried out by proteins — very large, intricately structured molecules. Although it is known that protein function is intimately connected to its 3-D structure, the connection is poorly understood. A detailed understanding of this structure-function relationship would have enormous implications ranging from more effective drugs to more efficient cleanup of waste sites.

- *Computational seismology:* Advanced simulation can improve predictions of the impact of earthquakes on buildings and other structures. IT^2 will result in better understanding of the ground motion of large sedimentary basins during earthquakes by allowing scientists to study these phenomena with much greater accuracy than is possible now.

- *Simulating combustion:* Predictive simulations of the performance of automobile engines hold the promise of decreasing the Nation's dependence on foreign oil and improving environmental quality (for example, by reducing unwanted emissions such as of carbon dioxide). Predictive simulation of diesel engine performance will help the transportation industry meet proposed emission standards in 2004.

- *Materials simulation:* The Nation has a continuing need for new materials — light-weight yet strong materials to reduce the weight of automobiles, flexible plastic batteries for use in portable electronic devices, and new magnetic materials for the computing industry. Computational development of new materials — one of the most challenging and computation-intensive problems in materials science — at greatly reduced cost and effort, would provide for the needs of a modern society and enhance the competitive edge of U.S. industry.

- *Modeling the evolution of the universe:* New space- and ground-based instruments, such as the Hubble Space Telescope, the Keck Telescope, the Sloan Digital Library Survey, and the Cosmic Background Explorer, are creating a revolution in cosmology by constructing an increasingly accurate picture of the universe. IT^2 would allow definitive tests of current cosmological models that are essential to capitalize on the major investments made in the new observational instruments.

Several scientific and engineering community workshops over the last two years have identified many other areas where computation on the scale envisioned for the IT^2 initiative would make revolutionary contributions. These include, but are not limited to:

Atmospheric modeling
Black holes and general relativity
Chemical reaction studies
Cryptography
Drugs by design
Ecological analysis of dynamic systems
Electron transport in complex media
Fusion energy
Genomics
Geophysical and astrophysical turbulence
Global systems simulation and modeling
High energy and nuclear physics
Microbial interactions in soil, water, and/or vegetation
Modeling metabolic systems
Prime number mathematics
Protein folding
Quantum chromodynamics
Sub-surface transport
Supernova science
Theoretical biology

Projects funded under IT^2 will be selected because they are important to science and engineering, are able to take advantage of enormously more powerful computational tools, and involve the necessary collaboration between information scientists and experts in the research areas. Access to high end capabilities will be important for both the development needed to make use of these new tools and for the large-scale modeling efforts envisioned. Projects from all areas of science and technology will be considered, with each agency supporting projects within the scope of its mission. Ties with the information technology research community will be the key to success.

Computer science and enabling technology

These efforts put the use of massively parallel computing on an aggressive trajectory similar to that of DOE's ASCI program. The agencies deploying the infrastructure for computation will partner with ASCI to focus on developing the computer science and applied mathematics technologies that enable the advances noted above in the scientific applications. Critical to the success of the applications is the development and deployment of advanced technologies in computational algorithms and methods, and software libraries; problem solving and code development environments and tools; distributed computing and collaborative environments; visualization and data management systems; and computer systems architecture and hardware strategies. This is a strong synergy between the fundamental information technology research and development and advanced science and engineering computation elements of IT^2 as it is currently structured.

Computation and simulation for large National applications

Many areas of societal interest will benefit enormously from the improved simulation studies and modeling capabilities to be developed by the IT^2 initiative. These include large-scale economic modeling and market forecasting; testing of highly complex engineered systems such as aircraft, bridges, and buildings that are expected to withstand extremes of weather or seismic activity; and complex traffic flow patterns in large cities or within the air traffic control system. Such studies will not necessarily require the most powerful computers, but they will often need the kinds of connectivity, bandwidth, visualization, and software tools foreseen in this initiative.

The capacities of this robust scalable information infrastructure will also make possible many kinds of interconnects, including dramatic new methods of learning, in which one or more students can have direct access to state-of-the-art scientific and technical instruments that are thousands of miles from home, but that are being used as if the students were in the control room. Other kinds of networks can be used to connect vast arrays of weather sensors, computers, and science teams to analyze and to model the complexities of a national weather system or a shock wave emanating from an earthquake's epicenter.

A partial list of activities that can be included in the civilian scalable information infrastructure includes:

- Computation to support Federal, state, and local policy development
- Shortening of technical development cycles
- Dynamics of energy supply and use
- Disaster prediction and mitigation
- Telemedicine
- Understanding the complexity of National systems for:
  - Financial markets and banking
  - Health care delivery
  - The electric power grid
  - The transportation system
  - Air traffic control
  - Agricultural commodities

Many other examples appear in the agency descriptions elsewhere in this document.

Agency Specifics

DOE

The DOE Scientific Simulation Initiative (SSI) is an integral part of the IT^2 initiative. SSI is an integrated project to develop a national scientific simulation computational infrastructure to solve complex scientific problems that cannot be addressed by other means. To do so requires (1) the creation of a high end computing system 10 to 100 times more powerful than any now available to civilian science, (2) the development of computer science, applied mathematics, and system software environments necessary to exploit these computers, and (3) the simultaneous development of scientific models that use the full capability of the computing system. SSI will build on advances in high end computing, high speed networking, and scientific model development realized in the HPCC programs and ASCI.

Work in defining, testing, and integrating the computing facilities and underlying computer science and applied mathematics technologies into tools for scientific discovery can support both the SSI and ASCI programs. Efforts will be coordinated to ensure maximum leverage for both programs. The SSI will also be closely coordinated with NSF to prevent duplication of effort.

The SSI supports the goals of IT^2 as well as the mission objectives of the Department. DOE must exploit the capabilities offered by modern multi-teraflop computing systems to address specific mission-critical scientific issues of national importance — such as those in global systems and combustion — and to advance knowledge across scientific research.

Advanced computing for science, engineering, and the Nation

The principal objective of the DOE is to provide a revolutionary tool for the conduct of research on large, inherently complex, scientific problems of national importance through application of a new generation of teraflops-based simulation capabilities by FY 2003, bringing simulation in parity with theory and experimentation. This requires a centrally managed and integrated project with three major components.

1. *Create a national terascale distributed scientific computing infrastructure.* This network of terascale computers, ultra-high speed communications, and support centers will provide the advanced computing systems required to achieve the principal objective.

   - In addition to procurement of terascale computers, SSI requires dramatic advances in communications. Planning estimates indicate that aggregate network data flow into and out of SSI facilities must be near 100 gigabits/second. This network capability is being demonstrated by the prototype NGI fast test beds, but it is almost 200 times faster than the fastest links on today's Energy Sciences Network (ESnet). Remote sites with major participation in SSI science and CSET require significantly greater network capabilities than ESnet can provide within its current funding profile. The SSI networking activity represents the incremental cost to build on the tools and network technologies developed in the NGI program to enable effective remote access to the SSI terascale computing infrastructure and the millions of gigabytes of data that these computers will produce.

2. *Development and deployment of basic computer science and applied mathematics technologies.* The practical and intellectual challenges to making effective use of terascale computers require the development of a teraflops-scale computer science and applied mathematics technology base in software, networking, data management and visualization, communications, and operating environments to enable scientists to make effective use of the simulation infrastructure. This technology base is essential to make teraflops-scale computers usable for present and future applications. It includes:

   - *Computer Science and Enabling Technologies, Algorithms, Models, Methods, and Libraries*

     - New classes of algorithms must be developed to effectively use SSI systems. SSI computers will be characterized by large-scale distributed, shared memory, with deep memory hierarchies. To achieve maximum performance, new hierarchical algorithms and methods are needed to take advantage of these architecture features. Mathematical techniques that are already under development may yield fundamentally new ways to conduct individual simulations and ensemble calculations, improving predictability and our understanding of errors in computations. New strategies for building mathematical software based on components are also likely to make a large contribution to programmer effectiveness for SSI applications.

   - *Problem Solving Environments and Tools (PSET)*

     - A major challenge for CSET is to develop an integrated problem solving environment to support the rapid construction and testing of applications codes. These environments need to support the most advanced parallel tools and libraries and enable users to collaborate on the development of codes, planning runs, debugging, and performance analysis. PSET systems will allow users to monitor jobs visually, steer computations from their desktop, and share results with others.

- *Distributed Computing and Collaboration Technologies*

  - The SSI will require distributed software development environments as well as support for effective, collaborative access to terascale compute facilities. Much of this software will be based computational grids, but significant SSI research will be required to transform research results into tools that applications scientists and CSET researchers can rely on. In addition, fundamental research is needed in distributed software development, configuration control, and component specification.

- *Visualization and Data Management Systems*

  - Simulations run on 10 or 40 teraflop class computers will generate tens or hundreds of terabytes of output per job. Advanced scientific visualization and analysis systems will be needed to translate this output for human understanding. Some data will best be analyzed visually while other data will be post-processed to extract features and statistics. Both forms of analysis will need sophisticated data management systems to enable rapid and facile manipulation of datasets. Large-scale databases will be used to manage metadata from simulations and advanced data organization techniques will support rapid and hierarchical traversals of data, so a user can move through data in real time. Associated capital requirements for large-scale data management and visualization testbeds will also be supported.

3. *Development of models capable of operating in a terascale computing environment.* DOE has selected global systems and combustion as the agency's initial applications for model development through a competitive process based on the need for high end computing, the readiness of the research community to proceed, and traditional review criteria, such as merit and mission relevance. To realize direct mission benefits in the next 5-10 years, the interdependencies between the development of these models and development of the terascale infrastructure will be closely coordinated. Additional applications, including but not limited to materials sciences, structural genomics, high energy and nuclear physics, subsurface flow, and fusion energy research, could benefit from this capability in the future. Current plans call for the selection of two additional basic science application efforts for inclusion in SSI in FY 2000. These will also be selected through a competitive process using the traditional review criteria outlined above.

   - *Global Systems: The objective of global systems simulation is to understand, model, and predict the effects on the Earth's global environment of atmospheric greenhouse gas emissions, with an emphasis on carbon dioxide.* By using teraflops-scale computers, DOE plans to accelerate climate model development and application to substantially reduce uncertainties in decade-to-century, model-based projections of global environmental change and increase the utility of such projections to the broad research community. Current models of global systems cannot achieve regional specification in global environmental change projections with the requisite accuracy and reliability needed to support national and international energy and environmental policies. To develop such regional specificity, the DOE has a project with the following components:

- *Regional Variability Associated with Global Changes*
  - Comprehensive global coupled air and water model simulation output must be interpreted and down-scaled to be useful for regional studies of environmental changes. Methods for producing derivative information, such as predicting stream discharge from precipitation rates, will be developed and implemented in rule-based systems for automatic extraction. A hierarchy of applied and engineering models will be assembled to interpret the raw simulation output and put it into a familiar and useful form for different user communities. Datamining technologies will be developed and employed to work through the large distributed archive of observational and modeling data to assemble customized data sets for further research on regional changes and their impacts. Two or three dedicated centers connected to the core SSI infrastructure will access the primary data archive and computational facilities.
- *Coupled Model Development, Testing and Application*
  - Together, with its partners in the multi-agency U.S. Global Change Research Program, DOE will establish two or more multi-institutional model development teams to develop and test fully coupled comprehensive models that accurately simulating multi-decade to multi-century changes in the global environment at regional resolution. Additional funding is required to implement state-of-the-art collaboratory technologies at the major sites involved in collaborations to accelerate progress and promote interaction and resource sharing.
- *Component Model Research and Development*
  - DOE will initiate a research program to improve the fidelity and computational performance of component models, particularly atmospheric and ocean general circulation models. Topics for further research include improved predictability, better numerical methods, and more accurate parameterizations.

- <u>Combustion Modeling:</u> *The goal of the combustion modeling effort is to understand, model, and predict the behavior and properties of combustion processes and devices.* With teraflops-scale computing resources and a concerted, multidisciplinary research program in combustion modeling, DOE will develop a new generation of combustion modeling tools for accelerated design of combustion devices to meet National goals for emission reduction and energy conservation. By providing a comprehensive understanding of the details of combustion, design engineers will have the computational tools to predict the chemical outcome of combustion processes with practical reliability. This new capability will avoid the current time-consuming trial and error experimental approach to the design of combustion devices such as gas turbines and internal combustion engines.

- This research effort, part of the overall DOE SSI program, will develop, through simulation and modeling, a detailed understanding of combustion processes to accelerate the development, characterization, and validation of design tools for advanced combustion devices. To address the complexity of combustion chemistry, the problem will be broken into computational regimes: chemistry models, subgrid, and device scale. The chemistry models will use modern quantum mechanical computational methods to calculate the rates of component combustion reactions that cannot be measured. The subgrid models will incorporate the detailed interaction of chemistry with fluid dynamics at a scale independent of the combustion device and provide data on material and energy balance required by the device scale models. The device scale model will complete the simulation, incorporating the results and parameters supplied by the other two regimes into a description of the overall device performance. Models for each of the three regimes will be developed concurrently and interactively. Computational codes that can take advantage of parallel computer architectures are already under development for each of the three regimes. Provision of sufficient fidelity in the prediction of combustion device performance, particularly the emission characteristics of combustion devices, will require computational resources of the scale being planned for the overall SSI. The knowledge gained will be used by the combustion engineering community to develop, design, and optimize codes for new combustion devices.

Execution Plan

The execution plan for this element is the same as the DOE execution plan described in the Fundamental Information Technology Research and Development section of this document.

Milestones

Late FY 1999 - FY 2000

- Competitively select two or three basic science applications for further development
- Competitively select one or two site(s) for upgrade of existing computer facilities to approximately one teraflop system(s)
- Establish multi-institutional model development teams focusing on comprehensive global systems modeling, on integrated combustion modeling, and on computer science and enabling technologies for terascale systems
- Competitively select sites for research, development, and deployment of persistent computational testbeds and related computer science and enabling software
- Competitively select site(s) for science centers (such as Regional Climate Collaboration Centers)
- Competitively select sites for hardware center(s)
- Begin improvement of global and combustion systems component models
- Develop project plan for additional basic science applications

FY 2001 - FY 2002

- Install approximately a 5 teraflops system with internal memory, data storage, and communications links required by the scientific applications
- Deploy the terascale system software developed by CSET research teams
- Begin to couple global and combustion systems models on the interim computer system
- Competitively select two or three basic science applications for further development

FY 2003-2004

- Install 40 teraflop system
- Deploy CSET tools for the 40 teraflop system
- Integrate, test, and initiate simulations on full resolution global systems and combustion models
- Competitively select two or three basic science applications for further development

NASA

NASA's efforts in the Advanced Computing for Science, Engineering, and the Nation component of the IT^2 initiative will be through NASA's Intelligent Synthesis Environment (ISE) Initiative. The ISE Initiative is a national program aimed at making substantial progress toward fulfilling the NASA Administrator's vision for next generation science and engineering capabilities. This involves developing revolutionary ISE-related technologies and coordinating related ongoing NASA activities, industry activities, other Government agency initiatives, and university research.

The ISE Initiative is intended to conduct research, develop, and implement the tools and processes that contribute to a revolution in engineering practice and science integration in the design, development, and execution of NASA's missions. When fully deployed, it will function as an advanced networked collaboration of all of the geographically dispersed entities involved in mission definition, design, execution, and operation. This collaboration will be in a controllable, immersive virtual environment in which humans and analytical models can interact using full sensory perception in a computationally rich mission life-cycle simulation. A diverse set of life-cycle engineering tools that can be integrated seamlessly to provide unprecedented computational speed and fidelity will be researched, developed, and validated by focusing on select NASA applications. A fundamental goal is to combine computationally-based intelligence and distributed smart agents emerging from advanced information technology research with engineering and science team decision-making.

Advanced computing for science, engineering, and the Nation

The ISE Initiative has two functional elements that provide an integrated strategy spanning research, proof of concept, user capability development, and insertion into practice:

- Rapid Synthesis and Simulation Tools (RSST), including research and development of enabling revolutionary methods and tools including cost and risk management tools
- Collaborative Engineering Environment (CEE), including accelerated introduction into practice and acceptance by users of these advanced tools

The long-term goals of the ISE Initiative are to:

- Reduce design and mission development time to 18 months
- Reduce design cycle testing by 75 percent
- Develop the ability to predict mission life-cycle cost to within 10 percent
- Develop the ability to predict quantified mission life-cycle risks to within a 95 percent confidence interval

While significant progress will be made over the next five years, objectives will be to:

- Reduce mission design and development time to 30 months through distributed collaborative teams of engineers and scientists
- Develop very rapid, high fidelity life cycle simulation methods incorporating virtual prototyping to significantly:
  - Reduce testing requirements and costs related to redesign and rework
  - Improve the ability to predict life cycle costs and minimize risk
  - Reduce technology insertion time

And, within 15 years:

- Teams of globally distributed scientists and engineers will work together in controllable, fully immersive, full sensory, environments — including 3-D vision, sound, and touch — with computers functioning as intelligent partners, not simply crunching numbers
- End-to-end mission scenarios will be developed in days and weeks, not years, providing technical feasibility analysis and reducing risk, while bounding life cycle costs and uncertainty to less than 10 percent

The ISE Initiative will be closely coordinated with related efforts in DoD, DOE, NSF, industry, and universities through partnerships and consortia.

Execution Plan

The ISE Initiative will be developed and implemented consistent with NASA Procedures and Guidelines (NPG 7120.5A) that govern the planning, development, implementation, and review of major NASA programs. The ISE concept has been under development within NASA for over two years and an Integrated Product Development Team (IPDT) with an assigned lead NASA center has been established. The IPDT will integrate the RSST and CEE functional elements, with the following product based requirements:

- Progressively improving integrated mission life-cycle analysis tools and tool sets to enable rapid, high fidelity, end-to-end analysis and simulation of all aspects of mission systems, configurations, development processes, and operations by geographically distributed participants operating in a heterogeneous computational environment
- Research and development focus applications for each enterprise to enable the integration and validation of promising new tools and methods and the

investigation of advanced methods for using computer-intelligence to augment life-cycle performance, and risk and cost trade decisions

- A state-of-the-art operational infrastructure in practice on focused NASA applications for collaboration in life-cycle analyses by geographically distributed teams of engineers and scientists. This infrastructure will include all NASA centers as well as industry and academic participants and will evolve in capability for bandwidth/throughput, accessibility, availability, and ease of use as improved tools emerge from research and development and from commercial suppliers

Milestones

FY 2000 ($19.0 million)

- Reusable launch vehicle design testbed on line
- Collaborative assembly simulations of space station
- Collaborative design investigations of reusable launch vehicle
- Deep space mission design testbed online and applied to robotic mission simulation

FY 2001

- Earth observation testbed online for engineering and science observing instruments simulations
- Immersive 3-D design collaborative capability demonstrated

FY 2002

- Virtual prototyping of reusable launch system design

FY 2003

- Virtual co-location capability on line and demonstrated

FY 2004

- Fully immersive 3-D collaborative capability of multiple engineering and science teams trained with smart tools and models, distributed intelligent knowledge bases, and simulation-based design methods

NIH

NIH has participated in HPCC initiatives since its inception. Dr. Donald Lindberg, Director of the National Library of Medicine, served as the founding Director of the National Coordination Office for HPCC. Currently NIH invests in high end computing, NGI, human computer interface research, and research and support in the interagency digital library program.

NIH has built a well rounded portfolio advancing HPCC, distributing those advances to the larger community that includes academia and rural hospitals and clinics, and enabling others to build on the research that NIH has funded. Current NIH challenges include the following:

- The large size of the Visible Human image set and other medical images challenges storage and network transmission technologies. Since the full set of Visible Human images require more than 100 CD-ROMs, NIH is investigating advanced compression and networking techniques to minimize storage and improve transmission speed over the Internet. The images, techniques, and efficient rendering algorithms will provide new tools for researchers, healthcare providers, students, and the general public. NIH is working with academia and industry to encourage the development of interoperable methods for representing and communicating such electronic images.

- Bioinformatics is an essential component of genome research, protein engineering, and drug design through its use of analytical and predictive methods to identify key molecular patterns associated with health and disease. The NIH/NLM National Center for Biotechnology Information (NCBI) focuses on automated systems to store and analyze the vast and growing volume of molecular biology, biochemistry, and genetics data. Within a distributed database architecture, NCBI collects sequence data from researchers worldwide and incorporates them into GenBank, NIH's DNA sequence data bank — a key data resource of the Human Genome Project — to produce an integrated database system consisting of GenBank, the genetic literature of Medline, taxonomy, and 3-D molecular structures. These databases are accessed daily over the Internet from more than 90,000 sites and account for over 4 million hits per day. Basic research on efficient data analysis techniques and large-scale genome analysis conducted within NCBI's Computational Biology Branch has been a key factor in gene discovery.

- NIH is providing parallel computing applications in biochemistry, molecular biology, and cellular biology with access to supercomputing systems, large databases, and other resources through Web-based browsers. Thus, complex simulations related to receptor sites and other drug design R&D can be conducted over the Internet. NIH will integrate 3-D graphics software, software tools for magnetic resonance spectroscopy data analysis and molecular structure determination, and other software tools, to provide new capabilities for structure-based drug design; improve methods for predicting protein-drug binding energies; and develop high performance computing methods for biomedical applications.

- In positron emission tomography (PET), images are reconstructed from a set of projected measurements in what is known as the inverse problem. While advances in detector technology have enabled researchers to use these techniques to study small animals, the data exhibit poor counting statistics. The quality of the reconstructed image can be improved by incorporating a statistical model of the scanning process into the reconstruction algorithm. Reconstruction based on the maximum likelihood (ML) criterion show both reduced noise and improved resolution. These methods are being used to study the phenotypical consequences of genomic manipulation in mice and rats. The Nuclear Medicine Department at the NIH Clinical Center is developing a high-resolution small animal PET imaging system using new scintillation crystal technology. Image sensitivity is being improved both through acquiring 3-D data and using ML-based reconstruction.

- The NLM Medical Connections program provides "jump start" funding to academic medical centers, community hospitals, and other healthcare organizations to connect to the Internet. The goal is to provide Internet connectivity to the top 3,000 healthcare institutions in the U.S. More than 250 U.S. medical schools and healthcare facilities have been connected during the past five years.

Advanced computing for science, engineering, and the Nation

HIGH END COMPUTING

Supercomputing technologies are used across the Government for applications such as imaging, and for processing the volumes of data such as those generated by the Human Genome Project. Activities in this area include:

- Developing an initiative to permit small laboratories to establish their own implementations of the NASA-developed Beowulf technology of clustered low-cost processors. The strategy is to build a supercomputer using many commercial low-cost off-the-shelf (COTS) processors instead of expensive supercomputer processors. This funding will enable development, software conversion, and a cookbook approach for other laboratories implementing these systems.
- Developing methods and algorithms to apply parallel systems to problems of biomedical significance
- Developing an internet accessible database, which will include all federally funded and some privately funded clinical trials where drugs for serious or life threatening diseases and conditions were used. Initially, all NIH Institutes and the Clinical Center will provide data to be loaded into the Clinical Trials database to capture information required by the FDA Modernization Act.
- Developing the information technology to support the genomics for a prototype instrument that will run 384 sequencing lanes on a microfabricated device.

Execution Plan

The NIH activities in this component of IT^2 will be implemented via the NIH extramural and intramural mechanisms, and administered as a part of the larger NIH portfolio. Coordination of these initiatives will be integrated though a trans-institute management group that expands the current HPCC management.

Milestones

FY 2000 ($2.0 million)

- Enable small laboratories to establish clustered low-cost processors (implementing NASA's Beowulf technology)
- Develop methods and algorithms for applying parallel systems to problems of biomedical significance
- Develop an internet accessible database, that will include all federally funded and some privately funded clinical trials where drugs for serious or life threatening diseases and conditions were used

FY 2001 - FY 2004

- Future NIH activities will be guided by the report of the Advisory Committee to the NIH Director (ACD) Working Group on Biomedical Computing, released in June, 1999.

NOAA

NOAA proposes to build on the HPCC base support that has positioned it to take full advantage of the scalable computing advances nurtured by that Federal Program. Since NOAA scientists developed the first coupled ocean-atmosphere climate model in 1969, these simulations have become the primary modeling tool for predicting climate behavior on time scales from months to centuries. Concern about climate change and the 1997-1998 El Niño event have spurred interest in developing National programs to improve climate prediction capabilities. The success of these efforts will depend heavily on physics-based climate modeling and fundamental climate science. A sharp increase in computing and data archival capabilities is the remaining ingredient required to allow NOAA scientists to attack the difficult problems confronting the climate research community and to support ongoing and developing research collaborations with other Government agencies, academic institutions, and research centers around the world.

Advanced computing for science, engineering, and the Nation

- *Acquisition of Large-scale Modeling and Prediction Research System.* Proposed plans include the competitive procurement of a comprehensive scalable computing system, including very large computing and storage capacity and balanced analysis and visualization capabilities as well as support for software development to make most effective use of the state-of-the-art, scalable technologies being developed by the American computer industry. Research support will focus on the design and maintenance of component-based applications, improvements to the portability and usability of climate and weather applications, and optimization of models for specific scalable architectures. Activities will also support the design and implementation of advanced parallel algorithms and the development of techniques for analysis and optimization of system resources in a production research environment.

- *Attacking Major Computational Challenges in Climate and Weather Research.* The enhanced computational capability provided by the acquisition of the large-scale simulation system and the associated software development activities will allow NOAA to attack some of the most difficult but critical problems in decadal-centennial climate change projections, seasonal-interannual climate predictions, and the next-generation hurricane prediction system. These problems include:

  - Cloud-Radiative Feedback, the dominant source of uncertainty in today's climate change projections, which demands a very large increase in computing capability to develop more realistic cloud representations for global climate models

  - Model Climate "Drift" and Improved Initialization of the Model Ocean, one of the most vexing problems in modeling the behavior of the climate system and a major barrier to providing the credible regional climate projections that are a high priority of coming U.S. and Intergovernmental Panel on Climate Change (IPCC) climate assessments

  - Physically Consistent Seasonal-Interannual Climate Prediction, the key to developing improved physically-based model forecasts of El Niño events and gaining a better understanding of the processes controlling these events

  - Development of a More Advanced GFDL Hurricane Prediction System, a critical tool for providing improved prediction of hurricanes, including changes in storm intensity, ocean response, and storm surge

Execution Plan

NOAA plans to acquire a balanced scalable computing system for its Geophysical Fluid Dynamics Laboratory (GFDL) through a competitive procurement in FY 2000. GFDL has extensive experience with similar procurements and has already begun internal planning.

Execution of this NOAA IT^2 activity will be coordinated with the overall HPCC programs, with GFDL's HPCC funding, and other NOAA base funded programs. Interagency coordination will be provided through the IT^2 and HPCC structures described in section 5. Internally, NOAA coordinates major research computing needs of the NOAA HPCC program, through the NOAA-wide HPCC Council comprised of all major organizational elements. The Council provides advice to the NOAA HPCC Office to ensure the best and highest priority research is conducted.

Milestones

July 1999

- Approval of project agreement. Under the newly reinvented Department of Commerce "CONOPS" procedures, full and open competition using a reengineered, efficient selection process open to all qualified entities is assured. This process brings all interested parties to the team including management, procurement, and legal to ensure fairness in the process and full value to the Government.

FY 2000 ($4.2 million)

- Initiate competitive procurement of a comprehensive scalable computing system including very large computing and storage capacity and balanced analysis and visualization capabilities
- Issue request for proposals announcement (October 1999). This will be initiated in line with Congressional approval as soon as funding becomes available.
- Contract award (May 2000)
- Initial system installation (July 2000)

While this time table is aggressive, past experience indicates that it can be met with a focused effort by all parties.

FY 2001

- Evaluate the capabilities of a more advanced GFDL Hurricane Prediction System for providing improved track forecasts as well as predicting other storm features, such as wind and precipitation fields and changes in storm intensity. This system will be designed and optimized for scalable architecture systems.

FY 2002

- Demonstrate progress in improving the capabilities of the next-generation GFDL coupled research model for predicting seasonal-interannual climate and for elucidating some of the processes that control El Niño-Southern-Oscillation events

FY 2003

- Isolate some sources of climate drift and define a strategy for reducing its effect on long-running, higher resolution coupled climate models. This investigation will require very large computational resources and the analysis of very large model datasets.

FY 2004

- Provide higher resolution projections of climate change with improved representations of clouds and ocean circulation to the impacts research community as part of the 2005 IPCC climate change assessment

NSF

NSF maintains a significant infrastructure for advanced computation in science and engineering through its Partnerships for Advanced Computational Infrastructure (PACI) program. Two sets of linked PACI partnerships exist, including academic institutions, other research institutions, and a few Federal laboratories. They provide access to high performance computing systems through merit-based, competitive processes throughout the science and engineering community. Continuous evolution in the capabilities of these partnerships is planned.

Advanced computing for science, engineering, and the Nation

Terascale computing systems provide resources that go beyond evolution in existing systems. Access to terascale systems is critical in order to keep the advanced computation infrastructure at the cutting edge. A workshop in May 1998 at NSF identified numerous important computational applications that could take advantage of a significantly increased computing capability.

Some examples of applications requiring terascale computing capabilities are:

- The simulation of one microsecond of protein folding, The longest undertaken to date, required two months on a 256 node Cray T3E. However, the shortest protein folding time is twenty microseconds.
- The quantum representation of chemical reactions involving three or four atoms requires 100 gigaflops to teraflops sustained speed on machines with very large memories. More complex molecules require correspondingly greater resources.
- An increase in computing power by a factor of 100 would allow definitive tests of current cosmological models through comparison with the wealth of data being accumulated from satellite and ground-based observations.
- Computing systems in the multi-teraflops/terabytes range are needed to compute the gravitational waveforms, which are essential for extracting physical information from gravity wave detectors, such as the Laser Interferometer Gravity-Wave Observatory (LIGO).
- Extending seismological simulations from the San Fernando Valley to the entire Los Angeles Basin, and improving resolution to the level desired by structural engineers for earthquake hazard abatement, would require moving from gigaflops to teraflop/petaflop computers.
- Dynamic modeling of the biosphere, incorporating data gathered in biocomplexity projects to develop predictive models

Thus, NSF will use IT^2 funding to accelerate the development of terascale systems available to the science and engineering community.

Execution Plan

NSF has centered responsibility for the terascale system initiative in the Directorate for Computer and Information Science and Engineering (CISE), which houses the Partnerships for Advanced Computational Infrastructure program (PACI).

Terascale computing hardware will be acquired to enable US researchers in all science and engineering disciplines to gain access to leading edge computing capabilities. This will be coordinated with ongoing NSF efforts, other agencies' activities, and high end computing research to leverage the software, tools, and technology investments. In FY 2000, a five-teraflop capability will be put in place at one site, with the system available to the PACI user community. Allocations of computing resources on this system will be fully merit-based and competitive.

NSF will issue a program solicitation to submit competitive proposals for acquisition, operation, and user support. This solicitation will be developed in coordination with DOE during the spring/summer of 1999. After internal NSF approvals, it will be issued as soon in FY 2000 as funding is assured.

Responses will be evaluated using the following criteria listed in priority order:

- The overall scientific milieu in which the facility will be sited
- The proven ability of institutions to realize large, complex facilities
- Plans for contributing to the education and training of the future information technology and science and engineering workforce
- Resources at the site that can contribute to addressing both scientific and IT^2 objectives, including mission objectives
- The cost-effectiveness of the proposed facility

Milestones

March to June 1999
- Work with other agencies and other parts of NSF to generate a coordinated program solicitation and gain necessary approvals for its release

July 1999
- Issue program solicitation

November 1999
- Proposals received at NSF

December 1999 through February 2000
- Review process involving panelists from all NSF Directorates and other agencies

March 2000
- NSB review and approval of award

May 2000
- Begin system installation

July 2000

- Begin "friendly user" period

September 2000

- System available for competitive allocation

4. Social, Economic, and Workforce Implications of Information Technology and Information Technology Workforce Development

Introduction and Strategy

Information technology is important to the nation not only because of its value in advancing science and technology and U.S. competitiveness, but because of the ubiquity of the technology and its effect on all aspects of citizens' lives. The impacts of information technology on our society, economy, and workforce include massive changes in the nature of work, commerce, education and training, entertainment, financial management, and quality of life. Understanding these changes will enable rational decisions to be made in Government and the private sector in allocating resources and planning policies to take advantage of the opportunities that information technology brings to society.

Various advisory groups have noted the costs associated with the lack of research investment in these areas. One example is the report of the President's Council of Advisors on Science and Technology's Report to the President on the Use of Technology to Strengthen K-12 Education in the United States. The report noted the severe underinvestment in education research and strongly recommended that the Federal government provide that investment. The report states, "The panel believes that a large-scale program of rigorous, systematic research on education in general and educational technology in particular will ultimately prove necessary to ensure both the efficacy and cost-effectiveness of technology use within our Nation's schools." Other areas of social and economic research suffer similarly from underinvestment in the research enterprise.

Our ability to address potential misuse of information technology would also benefit from research investment in these areas. For example, many citizens are concerned about the possible misuse of personal information, such as health records and spending patterns, that can be aggregated and mined at low cost and used in ways not anticipated nor authorized by the affected individuals. Data privacy and security are both technical issues in the design of protocols to implement privacy and security policy and social and political issues of privacy rights vs. governmental needs. Interdisciplinary socioeconomic and technical research can better inform policy decisions and provide potential technical solutions.

Finally, the knowledge and values of society have an impact on the spread of information technologies and the acceptability of their use in various aspects of our lives. Legal and regulatory constraints, tax treatment of investment in information technologies, and the existence of a workforce skilled in the use of information technologies are examples of factors in our society that can facilitate or inhibit the development and deployment of information technologies. Research into social, economic, and workforce issues related to information technology will allow us to better understand where there are barriers and how they might be overcome to ensure that the nation benefits to the greatest extent possible from the information technologies currently being developed.

Focused research on the Social, Economic, and Workforce implications of information technology (SEW)

Objectives

IT[2] will support focused research on the social, economic, and workforce implications of information technology (SEW) as an integral part of fundamental research and computational infrastructure development efforts. SEW research will encompass activities within and among disciplines, and will, in many instances, have immediate practical applications. SEW research conducted in collaboration with other IT[2] research in enabling technology and expedition centers will stimulate development of new disciplines and will enable national dissemination and development in areas judged by peer review to have both research promise and societal and economic importance. It will be carried out in multiple contexts, including: embedding SEW research within existing and new information technology centers, specific centers for national focus on SEW research areas in information technology, and interdisciplinary research grants involving SEW implications of information technology

There are three primary objectives for IT[2]-funded focused SEW research:

1. Provide sound empirical research to complement fundamental information technology research, thus improving our ability to develop and deploy technology and applications. For example, research on the co-evolution of human organization and complex information technology can play a critical role in informing the design of new information systems and enabling those systems to eventually “work” in specific settings and contexts.
2. Develop intellectually sound and useful research areas at the interface of information technology and SEW that can support the planning and policy development necessary for society to realize fully the potential of information technology. Research leading toward a computational grid, for example, could consider issues of access for underserved populations, privacy issues, proper pricing for computational resources, and the role of a future computational grid in providing service in competition with the private sector.
3. Facilitate the education and training future generations of scientists and engineers. Scientists and engineers who have developed interdisciplinary expertise in computer science linked to other scientific and/or social science fields will help to drive future advances in information technology.

Topics for focused SEW research

Possible topics for focused SEW research are listed below:

- *Transformations of Social Institutions.* Examples include:
  - Transformation of work (economics of labor markets; skills base and professional development, quality of worklife, telecommuting, distributed work, and computer-supported cooperative work)
  - Transformations of households ("smart houses"; quality of homelife; risks and benefits in accessing information; changes in socialization as it impacts physical and mental health)
  - Transformations of medicine (telemedicine; privacy and security of healthcare information; quality of healthcare; economics of healthcare)
  - Transformations of education (education standards; distance education; pedagogical changes; learning and cognitive ability; measurement, evaluation, assessment and certification; validation and accreditation)
- *Governance.* Examples include public access to information versus privacy; digital voting; the role of legislation and regulation; standard setting; tax and trade issues; political leadership; and values (such as privacy, civility, diversity, and inclusivity).
- *Electronic Commerce.* Economic examples include shifts in market segments, business models, distribution channels, consumer-supplier networks, and their larger impact on the economy and society. Legal dimensions include patents; intellectual property rights; and private versus public ownership of e-commerce processes and business models. Other examples include consumer trust in e-commerce and transaction security; micro-commerce; legislative and taxation impact on electronic commerce; electronic auctions; and social trust, including security, confidentiality, access, and standards.
- *SEW Simulation and Modeling.* Examples include research on social informatics, simulation of economic models, mathematical modeling of social interactions, simulation and modeling of cultural systems and interactions.
- *Sustainable Use of Large Information Infrastructures.* Examples include the evolution of software over long time frames (such as a Year 2000 type problem); socio-technical risks in sustaining learning networks such as distance learning and virtual universities; software safety; management of large networked centers.
- *Electronic Groups and Communities.* Examples include the use of multi-media in electronic interaction and impression management; fragmentation vs. inclusion in electronic communities; anonymous communications.
- *Barriers to Information Technology Diffusion.* Examples include economic models of information technology diffusion; successes and failures in information technology diffusion; and new theories of information technology diffusion.
- *Human Interaction and Communication Laboratories.* Examples include experimental studies of on-line election systems, auctions, social networks, and communication flows, including testing formal mathematical models and comparing electronic interaction with what is known about interactions in more traditional contexts.

- *Development of Innovative Scientific Methodologies.* Examples include methods for calibrating on-line surveys to compensate for sampling biases, new techniques for conducting distributed research with large numbers of subjects at many locations, and tools for analyzing naturally-occurring streams of communication on the Internet and the Web.
- *Development of a scalable information infrastructure to support research in SEW.* This infrastructure would consist of survey data, trace data, etc. needed for research in SEW that could be shared by multiple users.
- *Digital Government.* Examples on the effect of government actions include studies of the effects of Government regulations and taxation and international implications of information technology regulation. Examples of information technology benefiting the process of government include studies of distributed information such as multi-agency statistical information, means for universal citizen access to government, and the use of information technology for the management of emergencies such as natural disasters and large scale industrial accidents.

Implementation

Focused SEW research activities will be implemented largely through research projects and centers.

- Research in this highly interdisciplinary area requires an expanded proactive, peer reviewed, research program that draws on the social and behavioral sciences and the computer and information sciences. The program will support interdisciplinary projects that are larger than typical SEW research projects, but not yet at the level of a center. Such a program provides for agility in the rapidly moving information technology arena and in the rapid evolution of social phenomena surrounding the deployment of new Information Technology. Current SEW studies are frequently about information technologies that have been superceded before the research can be completed. To be of greater value, the program must allow for realtime modification of research to account for the pace of technological change.
- Centers define and legitimize research areas; they provide the scale of research activity necessary for experimental research in SEW areas; they allow for research teams of sufficient size for multidisciplinary research; and they allow sufficient time for the maturation of a research area. They can also provide training for the next generation of the information technology workforce in undergraduate, graduate, and postdoctoral education. Finally, centers can create a necessary network in which results from research can be adequately discussed and disseminated to others — including researchers, educational institutions, policymakers, and the public.

Assessment

In assessing the effectiveness of these focused SEW research activities, reviewers will consider indicators such as the following:

- The quality and quantity of research in SEW effects of information technology measured against 1999 baseline research activity.
- The quantity, quality, and influence of SEW research arising from collaborative activities.

- The development of new research areas in the SEW effects of information technology as evidenced by new communities of researchers, new interest groups within professional organizations, new journals, and new degree programs.
- Evidence of the systematic application of research results in social, economic, and workforce effects of information technology to practical problems arising in the development and deployment of information technology.
- The number of technology systems developed with SEW input to system design.
- The development of an information technology SEW discipline as evidenced by new interest groups within both SEW and information technology disciplines reflecting multidisciplinary research..

Information technology workforce development

Workforce development is a problem distinctly different from examining how information technology transforms the nature of work and the workforce. We need to ensure that the American workforce:

- Has members with the specialized skills needed to drive development of new information technologies
- Is sufficiently information technology literate to reap the benefits promised by information technology

The PITAC report has identified workforce development issues at many levels ranging from expectations for improved performance in the K-12 system and greater competency and literacy in information technology at the college level, to ensuring an adequate supply of highly trained scientists and engineers to achieve information technology research objectives. Increased research support may help alleviate the information technology workforce shortage at the Ph.D. and Masters degree levels through increased funding for graduate assistants. This will not be adequate to address other aspects of the information technology workforce shortage.

IT^2 will support three sets of activities for addressing workforce development issues. The first addresses undergraduate, graduate, and postdoctoral education; the second, workforce retraining and access to information technology educational materials; and the third, the problem of under representation in the information technology workforce.

Another important element in building a workforce with appropriate expertise in information technology is use of relevant capabilities throughout the educational process. An interagency Education Research Initiative is addressing, among other issues, the use of technology in education. Because this issue was covered under another initiative, it is not discussed further in this IT^2 implementation plan.

Ensure an adequate supply of highly trained scientists and engineers

Recent reports of a shortage in skilled information technology workers have led to unusual pressures on educational institutions as well as industry. Colleges and universities are reporting increased undergraduate enrollments at the same time as applications for graduate study are decreasing, except at the most elite research universities. Graduate students leave before completing doctoral programs, in part because of attractive industry salaries. Universities have reported problems with both recruiting and retaining graduate and teaching faculty. This "pipeline" issue must be addressed at both undergraduate and graduate levels, not just through graduate study and research.

IT^2 takes a two-pronged approach to developing highly trained scientists and engineers: improving education in information technology and providing opportunities for graduate and postdoctoral training.

Improving education in information technology

Two specific activities are proposed in this area:

- An Information Technology Educational Center: Through a competitive solicitation, support will be provided for a national center or group of regional centers that will assure that an appropriate range of training opportunities, including opportunities to review and develop modern curricula, are available to undergraduate faculty. The center(s) would also serve as a focus for building a strong community of undergraduate computer science, computer engineering, and information technology educators. Prior experience with smaller projects indicates that teaching faculty are eager to participate both to improve their competencies and to develop curricula and materials. The center(s) would initiate development, production, and distribution of new curricula and materials, where appropriate.

- Model curricula in computational sciences and computer literacy: This developmental activity would help raise the level of understanding and use of computing in other science and engineering disciplines and help college students be more computer literate. Specific projects could range from computational chemistry to new models of electronic commerce. The aim is to improve the skills of people who use information technology in their professions. Many colleges have broad competency and literacy requirements that could extend to recognize computer based methods such modeling, simulation, use of communication tools in collaborative environments, and statistical calculations as important universal skills.

The payoffs from expanding training opportunities include continuous improvement for the faculty who have most of the contact with undergraduates, uniformly higher quality of undergraduate education in information technology majors across all U.S. educational institutions, and improved transitions for students who move from two-year to four-year institutions or from colleges to graduate schools. Developing and distributing model curricula and course materials will promote information technology literacy as part of the broader goal of a citizenry literate in science and mathematics. Majors outside Information Technology fields will gain more skill in using Information Technology methods across the range of jobs and careers they choose, and more students will have the preparation needed to attract them to information technology careers or to have strong Information Technology roles in other careers. Students in science and engineering professions will have a strong base for using computers in industry or to develop even stronger competency in graduate studies.

In assessing performance in these areas, IT^2 reviewers will consider indicators such as:

- The number of faculty participating in training opportunities, the quality of their experiences, and the extent to which their teaching is affected by them

- The quality of education experienced by students going into information technology careers or making the transition to higher levels of education and the numbers of such students

* The quality and variety of Information Technology curricular materials for majors and non-majors and the extent of their use

Graduate and postdoctoral training opportunities

All agencies participating in IT^2 provide opportunities for graduate and postdoctoral participation on research awards. In response to needs for broader training that incorporates concepts from a strong multidisciplinary base, several agencies participating in IT^2 will provide focused opportunities for training.

NIH and DOE will expand programs in disciplinary areas to include training in advanced computational techniques in the biomedical sciences, in biology, and in simulation and modeling of physical phenomenon in combustion and climate. In addition, NIH will develop focus center grants requiring cross-disciplinary training using the tools of, and training in, bioinformatics. DOE will provide advanced education for software research, particularly in software engineering and interoperability of complex systems.

NASA will expand opportunities for graduate and postdoctoral training in collaborative technologies. New approaches for learning and teaming will be tailored to take advantage of a highly communicative, immersive environment. The emphasis will be on advanced engineering environments in which team decisions based on rapid integrated analyses and simulations replace current document-based sequential system design processes. A distributed active learning environment (facilities) for training in a collaborative environment will call on the full potential of NASA's Intelligent Synthesis Environment. Universities will work with industry, Government laboratories, and professional societies in developing effective instructional and training facilities for the new design approach afforded by the ISE.

NSF will focus its efforts on graduate traineeships in areas such as computer science, bioinformatics, computer engineering, and information technology. Graduate traineeships are competitive awards to departments or other equivalent administrative units to support between five and 20 students in innovative research and education settings.

NSF and NIH awards will be coordinated with existing programs. To attract and retain students, stipends may be significantly higher than is currently the practice in many disciplines. This would provide financial incentives to retain students for research and teaching at the Ph.D. level.

In assessing performance of this aspect of the IT^2 portfolio, reviewers would consider indicators such as:

- The number of students obtaining advanced degrees in information technology fields and the quality of their educational experiences; and

- The number of students obtaining advanced degrees in other fields with strong credentials in information technology applications and the quality of their educational experiences.

Workforce retraining and access to distributed learning

The need for continuous education and access to distributed modes of learning is exceptionally strong for the Information Technology workforce and those scientists and engineers who make frequent use of Information Technology in the course of their work. In FY 2000, IT^2 will emphasize:

- *Expanded Outreach through Centers:* Many agencies have outreach programs operated through existing programs and centers. For example, NASA has a Learning Technologies program that introduces students to the people, instruments, and data that are the core of its Earth science programs. All agencies will enhance and expand such activities, with emphasis on newly funded centers and new Information Technology-related projects funded at existing centers.

- *Developing the technology and research base for advanced distributed learning:* Activities will include: software development and distribution (for software that implements distributed learning); research on using technology in education (delivery, assessment, collaboration) and assistance to new sites in its adoption; and developing a clearinghouse for information on distance learning in Information Technology. Additionally, new courses might be developed with Federal support when great need and innovation are evident.

- *Building a digital library for information technology education:* A digital library for course work in information technology areas and in training for information technology professionals will build upon existing digital library research for K-12 and undergraduate levels.

The goals of these several activities are similar — to keep the knowledge base of those active in Information Technology-related activities current so that we have the most capable workforce possible. In assessing the performance of IT^2 activities, reviewers will include the following indicators for consideration:

- Number of individuals involved in outreach activities and the influence of those activities on their actions

- Level of incorporation of research results into new technology for distributed learning

- Quality and variety of new distributed learning media and materials

- Size and level of use of the Information Technology digital library

Increasing representation of women, minorities, and persons with disabilities in the information technology workforce

The low participation of women, minorities, and persons with disabilities in the Information Technology workforce is cause for concern and even alarm. The goal of increasing their level of participation is implicit across all IT^2 activities. Two activities address this issue explicitly.

- *Developing a Research Base on Academic and Career Choices:* IT^2 will increase the base of research to understand why students choose Information Technology, science or mathematics majors, why they drop out, and how economic, sociological, and psychological as well as racial and gender factors affect those decisions. Projects will also explore the effects of current best practices on recruiting and retaining students, particularly students from under-represented groups. Additionally, a base of research on career choices and factors is needed.

- *Best Practices:* Research into curricula and teaching methods will incorporate exploration of factors affecting the lack of diversity and encourage the use of best practices to recruit and retain information technology majors.

In assessing IT^2 performance, reviewers will consider indicators such as the following:

- Quality of research on academic and career choices and its usefulness in practical situations;
- The use of best practices in recruiting and retaining underrepresented groups by the participating agencies; and
- The level of increase in participation of underrepresented groups in the Information Technology workforce.

Agency Specifics

DOE

In response to the PITAC recommendations on social, economic, and workforce factors, the SSI will increase research funding in relevant areas of information technology and will develop strategies to retrain working scientists and information technology workers as well as new undergraduate and graduate students in these disciplines.

Execution Plan

The management plan for this element is the same as the DOE execution plan described in the Fundamental Information Technology Research and Development section of this document.

Milestones

The overall milestone is to provide advanced education for software research, particularly software engineering and interoperability of complex systems.

FY 2000 ($2 million)

- Support undergraduate and graduate fellowships in computational science. ($1 million)
- Initiate a program for retraining scientists working in areas where there are few jobs for work in computational science and computer science areas of importance to DOE ($1 million)

NASA

Rapid change in technology for engineering design requires constant updating of engineering curricula in universities and continuous re-education of the technical workforce. Training and education in the use of new computational engineering and science tools will be an important part of NASA's contribution to this initiative.

NASA's goal is to revolutionize engineering and science culture to enhance the creative process by:

- Enhancing and augmenting the practical experience of new engineering graduates
- Eliminating technical obsolescence in the workforce through education and training
- Removing cultural management barriers

Execution Plan

NASA will develop and apply new approaches for learning and teaming to take advantage of a highly communicative immersive environment. Emphasis will be on changing the engineering culture to take full advantage of advanced engineering environments in which team decisions based on rapid integrated analyses and simulations replace current document-based, sequential system design processes. A distributed active learning environment for training in a collaborative environment will be developed and processes will be established to train the current and future workforce in state-of-the-art techniques.

The full potential of the Intelligent Synthesis Environment will be realized in educating and training science and engineering teams in the component technologies and in new approaches for collaborative distributed synthesis and virtual product development. Funding will allow universities to work with industry, government laboratories, and professional societies in developing effective instructional and training facilities for the new design approach afforded by the ISE.

Milestones

FY 2000 ($1 million)

- Support training for advanced science and engineering programs

In FY 2001, FY 2002, FY 2003, and FY 2004, NASA will continue to support training for advanced science and engineering programs.

NIH

NIH continues to be a leader in the training of biomedical and behavioral researchers. The need to attract cross-disciplinary researchers from information technology into biology is imperative, as biology becomes a more information-based science. The paucity of researchers with abilities in both biology and information technology will soon become an impediment in the advancement of scientific research.

Execution Plan

NIH proposes to:

- Stimulate the integration of individuals with quantitative expertise, such as physicists, engineers, mathematicians, and computer scientists, into the biological and imaging sciences. Career development awards would be expanded to include programs to provide the requisite skills to investigators through cross-disciplinary training. (The Individual Mentored Research Scientist Development Award in Genomic Research and Analysis is an example.) The development and successful introduction of new technologies require both an understanding of biology and the fundamentals of the technologies to be applied, so that the solutions derived are appropriate and can be implemented.

- Provide grant funding to refocus existing resource centers where cross-disciplinary information technology training takes place, employing the tools of bioinformatics and teaching bioinformatics to be used in the scientists' workplaces.

Milestones

FY 2000 ($2 million)

- Integrate quantitative experts, such as physicists, engineers, mathematicians, and computer scientists, into the biological and imaging sciences

- Refocus existing resource centers where cross-disciplinary information technology training takes place, employing the tools of bioinformatics and teaching bioinformatics to be used in the scientists' workplaces

FY 2001 - FY 2004

- Future NIH activities will be guided by the report of the Advisory Committee to the NIH Director (ACD) Working Group on Biomedical Computing, released in June 1999.

NSF

Information technology advances have changed the way all people conduct their work and live their lives. The potential impact of information technology advances in the social, economic, ethical, and legal arenas are immense. The workforce must be trained in the effective use of information technology to remain competitive in the 21st century.

Past NSF activities have focused on:

- Increasing the pool of people with the knowledge, skills, and insights to lead research contributing to high performance computing and information science
- New course and curriculum development
- Collaborative research opportunities for undergraduates
- PACI outreach to enable all citizens to use emerging computing technologies to advance their ability to understand and solve problems in education, science, business, government, and society
- Learning and intelligent systems, including research on learning technologies

In responding to IT^2, NSF will encourage proposals that will extend these ideas to:

- Expanding understanding of the impact of social, ethical, economic, political, and legal factors on the development of information technology and vice versa in order to mitigate negative socio-economic impacts of information technology technologies
- Developing a more skilled American workforce for the future global marketplace, using the many powerful tools available because of information technologies

NSF's unified announcement of opportunities (discussed in Section 2 of this document) will also serve as a vehicle to solicit proposals in these areas. Activities in the first area will be pursued in concert with the Directorate for Social, Behavioral, and Economic Sciences. Some of those in the second area will be done jointly with the Directorate for Education and Human Resources. These are areas where joint activities with other agencies involved in IT^2 are also appropriate. Other aspects of the execution plan and NSF milestones described in Section 2 are also relevant. Special attention will be focused on digital government, developing diversity in the information technology workforce, and providing teaching experiences for CISE students.

Execution Plan

NSF will address information technology workforce issues in connection with existing programs and new activities. Program announcements will emphasize the goals of ensuring an adequate and diverse workforce, and workforce retraining for information technology activities. Support is foreseen for individual investigators, for teams of three to five investigators, and for larger centers. This initiative will strengthen ongoing programs including Minority Institution Research Infrastructure Improvement and Graduate Fellowships for Women in Computer Science. In addition, NSF will announce new programs to ensure adequacy and diversity of the workforce, ensuring representation of women, minorities, and the disabled.

To coordinate research and implementation, NSF will establish a distributed center for information technology workforce studies and information technology related research on learning technologies. The center will provide a repository of information on research studies of the information technology workforce, be an active disseminator of results, and coordinate workshops and seminars, as well as perform select research studies.

NSF will address SEW issues in three ways:

- Within the context of general information technology centers and research projects
- In interdisciplinary research grants involving SEW as one focus area
- In new centers for a national effort to address SEW research areas

NSF will ensure that research on SEW issues is part of every information technology center activity and most large information technology research projects by considering social, economic and workforce implications as part of the review process for every proposal. Interdisciplinary research and centers will be supported under the management of multiple NSF directorates in areas such as the link between computation and social systems, digital government, decision science, electronic commerce, simulation and modeling, and resources for education.

Milestones

FY 2000
- Hold a comprehensive series of workshops on SEW issues
- Release program announcements following from the workshop
- Hold competitions for centers and other projects

FY 2001
- Establish centers and research projects
- Initiate competitions for individual and interdisciplinary projects

FY 2002
- Hold workshops to evaluate ongoing SEW efforts, and recommend new kinds of centers

FY 2003
- Release new program announcements following from FY 2002 workshops

FY 2004
- Interim center reviews
- Establish next round of centers
- Hold competitions for individual and interdisciplinary projects

5. Management Plan

IT^2 Interagency Management Objectives

IT^2 will enable agencies to work together collaboratively to meet the primary objective of the initiative: to achieve the maximum National benefit of long term fundamental information technology research while meeting the mission-oriented goals of participating Government agencies. Two overriding management principles will be:

- **Open competition and/or peer review** will ensure that:
  - The best ideas for research in information technology and computing for science and engineering are identified and pursued
  - The best individuals and teams are selected to carry out the research
  - The best providers of computational infrastructure resources are selected and linked to the current resource base
  - Academic, national laboratory, and private sectors — including institutions from all geographic areas, historically minority institutions, and small colleges and universities — have equal opportunities to compete for access to available computational resources
- **Interagency coordination** will ensure that:
  - Participating agencies work together to address issues that transcend mission objectives
  - Resources are leveraged to maximize benefit to IT^2, related ongoing research, and agency missions
  - The computing infrastructure is optimized to meet IT^2 and agency objectives

IT^2 Interagency Management Structure

The IT^2 initiative is coordinated within the National Science and Technology Council (NSTC) as part of an integrated program that incorporates related ongoing Federal information technology R&D programs. Policy guidance and leadership is provided by a senior NSTC management team. The operational working group developing the initial IT^2 research agenda is developing a transition plan to provide integrated coordination of IT^2 with the ongoing interagency HPCC and NGI programs. The National Coordination Office for Computing, Information, and Communications R&D, which assures coordination of the base HPCC and NGI programs, is working with the operational working group to assure a smooth transition to integrated coordination of IT^2 with the ongoing programs. Integration of the management structures should be completed by September 1999.

The Senior Principals Group

A senior management team, chaired by and reporting to the Assistant to the President for Science and Technology, has been established within the NSTC to provide policy leadership, guidance, and oversight of the IT^2. Referred to as the Senior Principals Group, this team will initially consist of senior officials from the IT^2 funding agencies: the Director of the National Science Foundation; the Administrator of the National Aeronautics and Space Administration; the Under Secretary of Energy, the Under Secretary of Commerce for Oceans and Atmosphere and Administrator, National Oceanic and Atmospheric Administration; the Director of the National Institutes of Health; and the Under Secretary of Defense (Acquisition and Technology); as well as senior officials from OMB and NEC. The Senior Principals Group helps the Assistant to the President for Science and Technology establish program goals and monitor the progress of the initiative, balancing IT^2 goals with agency missions and capabilities. It also ensures that Federal efforts are coordinated and managed in a way that allocates funds in an open, competitive process that supports the best ideas. Once the transition to a fully integrated management and coordination structure for Federal Information Technology research programs has been completed, the Senior Principals Group will establish priorities and assure balance in the entire national information technology research portfolio. Membership on the Senior Principals Group may be broadened to include other agencies who contribute to the funding of the multi-agency information technology research program.

The IT^2 Working Group

An operational working group chaired by the NSF Assistant Director of Computer and Information Science and Engineering (CISE) reports directly to the Senior Principals Group. NSF has been chosen to lead this group in response to the PITAC's recommendation that NSF assume a leadership role in basic information technology research. Each participating IT^2 agency has appointed a representative to the working group who has operational authority over information technology research or information technology infrastructure within his or her agency. The National Coordination Office provides coordination support as it has for the HPCC programs.

The working group is charged with:

- Coordinating research in all major IT^2 areas
- Facilitating the planning of advanced infrastructure made available under IT^2 funding by DOE and NSF, and promoting competitive processes in the purchase, siting, and availability of this infrastructure
- Promoting policies to ensure availability of systems to appropriate research teams, as consistent with agency missions and priorities

The IT^2 Working Group is developing a transition plan to provide for integrated coordination of IT^2 with ongoing research programs, including HPCC and NGI, by September 1999. Some of the integrated management structure will use existing groups and teams, which will be reorganized to provide new focus and strengthened management to the existing base programs related to IT^2. New groups or teams may be established to provide management for some IT^2 tasks, such as the development and operation of the new advanced infrastructure funded by NSF and DOE under this initiative. This group will develop a program to ensure that the new infrastructure is purchased, sited, and made available on an open, competitive basis. It will also ensure that the systems are made available to research teams with the most compelling research concepts and the best ideas for building partnerships between experts in information and computational science and groups familiar with research challenges in biochemistry and climate modeling that can benefit from access to ultra-fast machines.

Overall, the integrated management structure for coordinating IT^2 and existing programs will facilitate coordination among agencies in projects that require close and continued partnerships. Agencies will retain control over the budgets appropriated to them and support the coordinated effort only in areas where they have authority. Each agency will use its own method for inviting proposals (for example, solicitations, broad area announcements, and calls for proposals). The agencies will coordinate their activities by using program managers and external experts to review resulting proposals, as well as through interagency program manager coordination.

Relationship to base program

The IT^2 initiative proposed as a part of the President's budget builds on a strong history of Federally-funded, multi-agency research programs by funding extensions of ongoing activities and expansions into entirely new areas. The transition to an integrated management structure for coordinating IT^2 along with related base research programs will permit better leverage of Federal investments in complementary information technology research programs and enhance the strength and balance of the national research portfolio.

Current base program

The 1991 High performance Computing Act (15 U.S.C. 5501) authorized an interagency research and development program built on ambitious technical goals and encouraging coordination of agency computer and communications research budgets. By 1998 the highly successful HPCC programs had achieved their technical goals. For example, by 1998 Federally sponsored R&D produced a supercomputer capable of conducting one trillion calculations each second — a remarkable achievement that was only a remote possibility ten years ago. The HPCC programs also revealed new potential areas of research.

Last year's Next Generation Internet Act (P.L. 105-305) expanded on the 1991 HPC Act by authorizing a broad range of new networking research designed to achieve 100- to 1,000-fold increases in network speed and advanced network services (including reliability, security, mobility, multicasting). The Act also documented Congressional interest in a broad array of additional research topics in high confidence systems, human centered systems, education training and human resources, and the next generation of advanced computing.

In 1996, DOE established ASCI to provide the supercomputing, modeling, and simulation capabilities to maintain the safety, reliability, and performance of the Nation's nuclear weapons stockpile. ASCI is purchasing the world's most powerful supercomputers and conducting related R&D to benefit not only nuclear weapons stewardship, but also a variety of other scientific disciplines. The table on the next page provides an integration of the budgets for the base and initiative described in this section.

Relationship of IT^2 programmatic themes to current base program

The IT^2 initiative builds on the following Federal information technology R&D programs:

1. The HPCC programs ($919 million in FY 2000), a multi-agency activity. The FY 2000 HPCC crosscut is comprised of ten agencies (six of which are proposed to participate in the IT² initiative), reporting three Program Component Areas (PCAs) as defined in OMB Circular A-11:

   - *Large Scale Networking (including the Next Generation Internet Initiative):* R&D in high performance network components; technologies to enable wireless, optical, mobile, and wireline communications; large-scale network engineering, management, and services; and systems software and program development environments for network-centric computing.

   - *High end Computing and Computation:* R&D in high end computing hardware and software innovations, and in algorithms and software for modeling and simulation needed to address Grand Challenge-class applications, including associated infrastructure.

   - *High Confidence Systems:* R&D in information security, information survivability, and assurance technologies to achieve high levels of availability, reliability, and restorability of information systems.

2. PCAs of HPCC coordinated by the interagency management team, but not reported in the President's budget (approximately $170 million in FY 2000):

   - *Human Centered Systems:* R&D to increase the accessibility and usability of computing systems and communications networks, including research on data visualization tools, speech recognition and synthesis, and information integration.

   - *Education, Training, and Human Resources:* Computing and communications R&D to advance education and training technologies at all levels from K-12 to graduate-level university programs and life-long learning.

3. DOE/ASCI ($543 million in FY 2000):

- The President's fiscal year 2000 budget reports ASCI along with the HPCC programs. ASCI has built the world's fastest supercomputers and conducts related research and development to enable computer simulation of nuclear weapons detonations. The Administration has worked to ensure that innovations and new technologies emerging from the ASCI program will benefit a broad range of scientific disciplines.

Table 5.1 provides the HPCC programs crosscut of the President's FY 2000 budget. Table 5.2 (page 74) lists the six agencies participating in the IT² initiative the HPCC program budgets; the PCAs of HPCC coordinated by the interagency management team but not reported in the President's budget for FY 1999 and FY 2000 (shown in the "Other" columns); and the proposed IT² budgets mapped into the IT² initiative categories.

The proposed $366 million in FY 2000 will fund extensions of ongoing activities and expansions into new areas:

1. Fundamental Information Technology Research ($228 million):

   - *New Work:* Areas of emphasis will include software engineering, systems protected against accidental failures or intentional intrusion, autonomous

Table 5.1
FY 2000 HPCC Budget Request

(Dollars in Millions)

Agency	HECC	LSN [a]	HCS	Other PCAs [b]	TOTAL
NSF	216.8	76.4	21.1		**314**
NASA	108.2	20.4	7.8		**136**
DARPA	38.4	75.5	16.0		**130**
DOE Office of Science	82.6	33.8			**116**
NIH	27.7	69.0	5.3	2.0	**104**
NSA	27.5	1.7	47.8		**77**
NIST	3.5	5.2	5.5		**14**
NOAA	10.3	2.7			**13**
AHCPR		5.2		5.5	**11**
EPA	4.2				**4**
SUBTOTAL	**519.2**	**289.9**	**103.5**	**7.5**	**919**[c]
DOE ASCI					543
TOTAL					**1,462**

[a] The NGI budget is part of the LSN budget.

[b] Several agencies included some HuCS and ETHR budgets in the President's FY 2000 HPCC Budget. This does not represent the full funding for HuCS and ETHR in FY 2000.

[c] Subtotals do not add to total due to roundoff.

controls, and basic mathematical methods. Research topics include new tools for visualizing complex data, language translation, devices for ensuring information access to people with physical disabilities, and intelligent data searching. New activities will also include very high-risk, long-term research on quantum computing and DNA computing.

- *Expansion of Ongoing Work:* While most of these activities are new, they build on the HPCC programs' High Confidence Systems and Human Centered Systems PCAs, as well as portions of HPCC's High End Computing and Computation and Large Scale Networking PCAs. For example, the Next Generation Internet initiative supports experiments on relatively small, high-speed networks, while new research will focus on enormous and complex networks that may someday link billions of separate elements — some mobile, some stationery, some that enter and leave the network continuously. Designing such enormous deeply networked systems adds a range of problems not included in the original NGI initiative.

2. Advanced Computing for Science and Engineering Applications ($123 million):

   - *New Work:* These activities will provide the civilian research community a network of supercomputers approaching the capabilities of those built for

TABLE 5.2
IT^2 PARTICIPATING AGENCIES HPCC AND IT^2 BUDGETS
(Dollars in Millions)

	FY 1999			FY 2000			
	HPCC (a)	**Other (b)**	**Total**	**HPCC (a)**	**Other (b)**	**IT^2**	**Total**
Fundamental IT Research							
Basic IT (software)(c)	0	0	0	0	0	84	84
(High Confidence Systems)	0	27	27	50	0	0	50
Human Computer Interaction (d)	0	136	136	0	132	26	158
Large Scale Networking and Scalable infrastructure(e)	211	0	211	209	0	67	276
High End Computing (f)	160	0	160	145	0	48	193
Advanced Computing for S&E							
Advanced Infrastructure (g)	182	0	182	207	0	67	274
Advanced Science and Engineering Computation	114	0	114	113	0	47	160
Computer Science and Enabling Technology	28	0	28	29	0	8	37
National Information Infrastructure Applications	57	0	57	60	0	5	64 (i)
Economic & Social Implications (h)	0	30	30	0	31	15	46
TOTAL	751	192	943	813	163	366	1,342

(a) HPCC as reported in the President's FY 1999 Budget for the six agencies proposed to participate in IT^2; the President's HPCC Budget includes 11 agencies. Similarly for the President's FY 2000 HPCC Budget that includes 10 agencies.

(b) HPCC R&D coordinated as part of the HPCC programs but not reported in the President's HPCC Budget per Circular A-11 guidance. In FY 1999 HECC and LSN were reported in the President's HPCC Budget, but not HCS, HuCS, or ETHR; in FY 2000 HECC, LSN, and HCS were included but not HuCS or ETHR.

(c) Software engineering, component based software, active software (not in direct support of other program elements)

(d) Includes the Human Centered Systems in previous HPCC accounts

(e) Including NGI and "scalable infrastructure" (deeply networked systems, network modeling). Does NOT include networking infrastructure now included in LSN

(f) Includes only System software technology (HECC thrust 1) and Leading-edge research for future generations of computing (HECC thrust 2)

(g) Includes new DoE and NSF programs and the "incorporating technology into applications" (HECC thrust 3) and "infrastructure for research" (HECC thrust 4) in existing HPCC. Also includes LSN "large scale networking engineering, management and services"

(h) Includes existing work on Education Training and Human Resources

(i) Total does not add due to roundoff.

nuclear weapons research through DOE's ASCI Program. This computing infrastructure will involve creating new computer architectures, operating systems, and data management capabilities. It will be orders of magnitude more powerful than that currently available to the civilian science community and will give U.S. research teams throughout substantial new abilities to attack problems in biochemistry, materials, climate change, and many of the fundamental riddles of science.

- *Expansion of Ongoing Work:* Portions of HPCC's High End Computing and Computation and Large Scale Networking PCAs contribute to the new, advanced infrastructure by developing civilian supercomputing and networking facilities at NSF, DOE, and other Federal agencies.

3. Social, Economic, and Workforce Implications of Information Technology ($15 million):

   - *New Work:* This element will greatly expand research into social, economic, and workforce impacts of information technology, including transformation of social institutions, impact of legislation and regulation, electronic commerce, barriers to information technology diffusion, and effective use of technology in education, ensuring that all Americans have the education to take advantage of high-wage jobs created in the new economy.

   - *Expansion of Ongoing Work:* New research will leverage the HPCC's Education, Training, and Human Resources PCA, which develops computing and communications technologies for education, training, and lifelong learning.

These proposed expansions make fiscal year 2000 an important transition year for the Federal government's information technology R&D portfolio. The President's budget highlights the major new areas of emphasis that comprise the IT^2 initiative. In future years, the Administration will coordinate and report base programs and the IT^2 initiative as an integrated whole.

General Implementation Guidelines

IT^2 participants are committed to using peer-reviewed competitions to select teams to develop the generic computer science and enabling technologies will be open to permit the broadest participation possible. In joint solicitations, selected proposals will be distributed among the funding agencies according to the best match to agency missions and the related software development budgets of each of the participating agencies. The balance of activities between jointly- and independently-solicited proposals will also be considered.

- Peer-reviewed competitions to select science teams will require two different approaches:

  - NSF will use competitions to determine which areas of science and engineering will form the initial set of IT^2 activities.

  - DOE, NOAA, NASA, and NIH have selected an initial set of scientific applications areas, as noted in the sections above, and will hold competitions consistent with agency needs and mission requirements to determine the performers.

- Joint solicitations and review processes may be issued where agency interests overlap.

Computational Infrastructure

The goal is to have the initial hardware available in FY 2000 - FY 2001. The hardware acquisition schedule enables the development of scalable systems and application software, as well as initial scientific and engineering applications, to proceed on these machines in preparation for the larger machines to be available in FY 2003 - FY 2004. NSF and DOE plan to develop the initial high end infrastructure, while NASA and NOAA plan smaller more specialized acquisitions. To meet these deadlines, the acquisition processes must commence in the spring of 1999.

The development and use of the computational infrastructure must support:

- **Generic, broad science and engineering applications** that encompass a range of hardware resources. For NSF, the needs of the broader scientific community will drive computational infrastructure requirements, emphasizing generic simulation infrastructure.

- **Large-scale scientific and engineering applications** that require extensive use of the hardware resources toward a focused goal. For DOE, NOAA, and NASA, the needs of the scientific community in the initial applications areas will drive the computational infrastructure requirements and its subsequent development.

- **Research and development of large-scale simulation software** that requires flexibility in use as well as rapid response

- **Experimentation** in high end computing and computational scalability

The IT^2 Working Group will ensure there are no gaps in the development of the infrastructure and will avoid unnecessary duplication of effort. Each agency will retain responsibility for its own procurements. The working group will coordinate the activities related to management of hardware resources and management of the process of bringing cutting edge research to bear on development and deployment. The generic infrastructure will be developed through partnerships between sites that operate the hardware and the national research community.

Site selection

After general specifications for the high end hardware centers are determined, the agencies involved will issue parallel solicitations for proposals to establish these centers. The criteria for decision-making will be stated at the time of the solicitations.

Resource allocation

Resource allocation, such as machine time on the new very-large-scale computing facilities, will be handled in an openly competitive manner appropriate to the participating agencies.

- The NSF PACI program has used a program committee of external peer reviewers to select among requests for access to PACI resources.

- The Environmental Molecular Science Laboratory at DOE's Pacific Northwest National Laboratory uses peer review for assessing the merits of proposals from large, nationwide teams focused on complex problems in environmental molecular science to allocate its supercomputer resources.

- DOE's National Energy Research Supercomputer Center has begun using a mix of program committee advice and mission program direction.

Each computing facility will ask its review committees to determine whether a particular request makes appropriate use of its resources or whether the application is better suited for another facility. The committees will include outside experts in terascale computing for the principal scientific and engineering applications as well as personnel from the computing facility. An interagency task group comprised of the directors of the computer facilities, along with experts in the relevant computational and computer science areas, will work with the IT^2 Working Group (discussed in section 5.3.2) to provide overall coordination and guidance on optimizing resource usage.

Reporting

Agencies will report annually to the Senior Principals Group and the IT^2 Working Group on progress made toward meeting their IT^2 milestones. In addition, OMB has indicated that the IT^2 plans, budgets, and activities coordinated along with the base HPCC and NGI programs will be reported in the annual Supplement to the President's Budget. Individual agencies also have their own reporting procedures, which will include agency IT^2 planning and assessment, as appropriate.

PITAC and Other External Advisory Groups

As mandated by the Next Generation Internet Act of 1998 (PL 105-305), the PITAC will review and "assess the extent to which Federal support of fundamental research in computing [through IT^2 and ongoing base research] is sufficient to maintain the Nation's critical leadership in this field." Advice and comment on developing the best strategies and policies to implement the IT^2 initiative and ensure balance in the entire Federal information technology research portfolio will also be sought from other experts in the scientific community.

The senior management team, the IT^2 Working Group, the agencies that propose to participate in the IT^2 initiative, and the NCO meet frequently with members of Congress and Congressional staff, academia, industry, and professional societies, to exchange technical and programmatic information about areas of common interest. These meetings include workshops in which suggestions about Federal information technology research are solicited.

National science policy is the purview of OSTP and the NSTC. In areas of particular national concern, review and allocation procedures will be instituted as appropriate to the area of concern. Specific organizations may be asked to review some or all aspects of the IT^2 initiative. For example, in support of climate research, the U.S. Global Change Research Program (USGCRP) will be asked to participate in or review resource allocation decisions, as appropriate.

Finally, each agency involved in the initiative has its own advisory group that may be invited to provide comments concerning the individual agency's IT^2 process.