INFORMATION TECHNOLOGY
FOR THE TWENTY-FIRST CENTURY:

A BOLD INVESTMENT IN AMERICA'S FUTURE

January 24, 1999

Working Draft

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

Executive Summary

- As part of their fiscal year 2000 budget, President Clinton and Vice President Gore are proposing a $366 million, 28 percent increase in the government's investment in information technology research.

- This initiative, known as IT^2 (Information Technology for the Twenty-First Century), will support three activities:

  - **Long-term information technology research** that will lead to fundamental breakthroughs in computing and communications, in the same way that government investment beginning in the 1960s led to today's Internet;

  - **Advanced computing for science, engineering, and the Nation** including software, networks, supercomputers, and research teams needed to support it. This will support applications such as reducing the time required to develop life-saving drugs; designing cleaner, more efficient engines; more accurately predicting hurricanes and tornadoes as well as long-term climate change; and accelerating scientific discovery; and

  - **Research on the economic and social implications of the Information Revolution** and efforts to help train additional information technology workers at our universities.

- The potential benefits of IT^2 are compelling:

  - The results of past government research (e.g., the Internet, the first graphical Web browser, advanced microprocessors) have helped strengthen American leadership in the information technology industry, which now accounts for 1/3 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.

  - Information technology is changing the way we live, work, learn, and communicate with each other. 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 health care for rural Americans through telemedicine. U.S. leadership in information technology is also essential for our national security.

  - Information technology has a mutually reinforcing relationship with our "golden age" of science and engineering. Advances in supercomputers, simulations, and networks are creating a new window into the natural world — making computing as valuable as theory and experimentation as a tool for scientific discovery. At the same time, challenging scientific problems such as predicting the impact of climate change, designing more efficient and cleaner energy systems, and gaining new insights into the fundamental nature of matter push the frontiers of information technology capability.

- The initiative builds on previous and current programs in computing and communications, including the High Performance Computing and Communications Program (authorized by legislation introduced by then-Senator Gore), the Next Generation Internet, which was authorized by the Congress in 1998, and the Department of Energy's Accelerated Strategic Computing Initiative (ASCI). It responds to recommendations made by an external advisory committee requested by the Congress (The President's Information Technology Advisory Committee), which concluded that the government was underinvesting in long-term information technology research relative to its importance to the Nation. This committee, which is comprised of leaders from industry and academia, concluded that the private sector was unlikely to invest in the long-term, fundamental information technology research needed to sustain the Information Revolution. The initiative also reflects a strong belief held by the research community regarding the potential of information technology to accelerate the pace of discovery in all science and engineering disciplines.

- The agencies that will be involved in IT^2 include the National Science Foundation, the Department of Defense (including the Defense Advanced Research Projects Agency), the Department of Energy, the National Aeronautics and Space Administration, the National Institutes of Health, and the National Oceanic and Atmospheric Administration. Roughly 60 percent of the funding will go to support university-based research, which will also help meet the growing demand for workers with advanced information technology skills.

- Some of the potential breakthroughs that may be possible as a result of IT^2 include:

  - Computers that can speak, listen, and understand human language, are much easier to use, and accurately translate between languages in real-time;

  - "Intelligent agents" that can roam the Internet on our behalf, retrieving and summarizing the information we are looking for in a vast ocean of data;

  - A wide range of scientific and technological discoveries made possible by simulations running on supercomputers, accessible to researchers all over the country;

  - Networks that can grow to connect not only tens of millions of computers, but hundreds of billions of devices;

  - New ways of developing complex software that is more reliable and easy to maintain, and can be depended on to run the phone system, the electric power grid, financial markets, and other core elements of our infrastructure; and

  - Computers that are thousands of times faster than today's supercomputers, and are based on fundamentally different technology, such as biological or quantum computing.

This document is available at http://www.ccic.gov/.

Background

With this ambitious Information Technology for the Twenty-First Century (IT^2) initiative, the Federal Government is making a dramatic new commitment to research in information technology, building on existing and previous government investments and accomplishments in this area, including the High Performance Computing and Communications Program and the Next Generation Internet as coordinated by the National Science and Technology Council. This initiative also responds directly to the concerns and recommendations of The President's Information Technology Advisory Committee (PITAC).

The support for this initiative is substantial.

- President Clinton and Vice President Gore believe that strengthening America's leadership in information technology is critical to our economic prosperity, our quality of life, and our national security. President Clinton called for a significant increase in information technology research in his June 1998 MIT Commencement address. Vice President Gore has been a long-time supporter of information technology research, popularizing the term "information superhighway" over 20 years ago and authoring the High Performance Computing Act of 1991 as a United States Senator.

- The President's Information Technology Advisory Committee has called for a significant increase in information technology research, and has provided the Congress and the Administration with recommendations on research priorities.

Government agencies have convened public workshops to obtain more detailed input on the opportunities and challenges in information technology research. The agencies and the broader scientific and engineering community have stated that advances in high-end computing, modeling and simulation, and information management could lead to scientific and technological breakthroughs in widely differing areas such as the design of more efficient engines, more rapid development of life-saving drugs, the analysis and prediction of climate change, and a better understanding of the fundamental nature of matter.

Industry leaders credit Federally-funded research with building the foundation for nearly all the fundamental innovations in computers and computation, including networking, the windows and Web browsers that make computers accessible to everyone, software for displaying 3-D images (now used in a wide range of applications, from accelerating aircraft design and testing to simulating dinosaurs in movie thrillers), and the basic architecture of today's powerful computers.

Government-supported developments in scientific computing have led to major advances in areas as disparate as weather and climate prediction, genomics, cosmology, and materials research, and have laid the groundwork for applications such as safer buildings and telemedicine.

Federal investment has been essential for the flow of basic inventions that provide nourishment for today's exploding information technology industry. Perhaps more importantly, however, Federal investments have provided training for talented students who emerged from colleges and graduate programs to provide the intellectual leadership for today's information technology revolution.

In addition to supporting industrial innovation and the economic benefits that it brings, Federal support for computing research has enabled government agencies to accomplish their missions better. For example, the Department of Defense's (DOD) concept for warfighting in the future rests on the foundations of information dominance and technological innovation. The Department of Energy (DOE) today relies on the simulation capabilities of large supercomputers to certify the nation's nuclear stockpile. The National Aeronautics and Space Administration (NASA) has expanded the frontiers in air and space through the innovative use of information technology to enable the design of complex vehicles, achieve high reliability, and manage massive amounts of data. The National Institutes of Health (NIH) are taking advantage of new Web technologies and improved networking capabilities to expand access to medical information. The National Oceanic and Atmospheric Administration (NOAA) has exploited information technology to build a better, faster severe weather and climate forecast system resulting in fewer weather-related deaths, lower property damage, and better understanding of global warming issues.

Under the leadership of the President's Advisor on Science and Technology and the umbrella of the National Science and Technology Council (NSTC), the Federal research and development (R&D) agencies have worked cooperatively to craft the IT^2 to meet the following objectives:

- The IT^2 will support a balanced research portfolio in computer and information science and engineering that will strengthen the national investment in fundamental research;

- By supporting development of dramatically increased and broadly accessible computing capabilities, the IT^2 will provide strong new links between research in information science and other critical fields like biology, chemistry, physics, and materials, initiating a new era of simulation-based research; and,

- The IT^2 will fund research on the legal, economic, and social impacts of information technology and ways to assess these impacts, and will grow the intellectual base needed to ensure continued innovation in information technology.

Rationale and Benefits

The Administration believes that there are a number of compelling reasons to increase the Nation's investment in long-term information technology research and promote applications of information technology that advance science, engineering, and a broad range of national goals.

1. **Past government support for information technology research has resulted in a huge economic return on investment, and continues to play a pivotal role in promoting innovation.**

Many of the most important breakthroughs in computing and communications (e.g., the Internet, the first graphical Web browser, high-speed networks, artificial intelligence, supercomputers, databases, and the graphical user interface) have come from government-sponsored research. The Internet alone has created hundreds of billions of dollars in new wealth — vastly exceeding the government's investment in networking research. Simulation-based science is also having an impact in industry. For example simulation of the properties of molecules and processes was critical to the timely introduction of environmentally-benign replacements for freon, a nearly $2B per year global market.

The private sector spends much more than the government to commercialize new ideas. But, as Vice President Gore has noted, government-supported research often provides the initial, critical “spark” for innovation. Data from the Council on Competitiveness suggest that only 6-9 percent of the information technology industry’s R&D expenditures go toward long-term, fundamental research. It is difficult for companies to justify investment in fundamental research to their shareholders, given the difficulty that any one firm has in rapidly capturing and commercializing the benefits from this kind of research.

2. Information technology is becoming increasingly important to the U.S. economy, making technological leadership in this sector even more critical in the 21st century.

- Information technology has accounted for one-third of U.S. economic growth from 1995 to 1997.

- The global information and communications industry is now roughly $2 trillion in size. Business-to-business electronic commerce in the U.S. is projected to grow to $1.3 trillion by the year 2003.

- In 1996, 7.4 million Americans worked in information technology industries and information technology-related occupations, earning, on the average, almost $46,000 per year, 60 percent more than the average private sector wage of $28,000.

- Information technology now accounts for 45 percent of private equipment investment, up from 3 percent in the 1960s. All firms are using information technology to reduce the time required to develop new products, increase productivity, deliver “just-in-time” training to their employees, forge closer relationships with suppliers, and tailor products and services to the needs of individual customers.

3. The broader national payoff from improvements in information technology will be enormous.

Information technology is not only important for its impact on America’s economic growth, job creation, productivity, and global competitiveness. Information technology is a powerful tool to achieve virtually every important national goal:

- *A world-class educational system:* Information technology can enhance parent-teacher interaction, allow students to take “virtual” field trips, practice science as opposed to merely reading about it (e.g., learn astronomy by using a remote telescope), or perhaps someday interact with, and learn from, an infinitely patient, virtual version of Einstein, Lincoln, or Socrates.

- *Life-long learning for all Americans:* Information technology can enable adults to learn at a time, place, and pace that is convenient for them, allowing them to acquire the skills they need to succeed in the new economy while balancing the needs of work and family.

- *A strong defense:* U.S. military strategy, as described in the Joint Chiefs of Staff's Joint Vision 2010, now relies on information superiority, which is defined as "the capability to collect, process, and disseminate an uninterrupted flow of information while exploiting or denying an adversary's ability to do the same." In addition, the ASCI program is a key element in assuring the safety, reliability, and performance of our nuclear arsenal.

- *Expanded access to high-quality, affordable health care:* Simulation enables the design of more affordable and effective drugs; electronic claims-filing can reduce administrative costs; secure patient records can improve privacy and ensure that patient records can only be viewed by authorized personnel; telemedicine can enhance the quality of care in rural areas; public health information systems can spur an increase in childhood immunization rates; and consumer health information can allow individuals to make more informed decisions about their own health care needs.

- *Improving the quality of life for the 55 million Americans with disabilities:* New technologies such as speech recognition and intelligent robotic devices can make it easier for people with disabilities to interact with computers, lead more independent lives, and gain new employment opportunities.

- *Protection of life and property:* Through better forecasts of hazardous weather and more timely and effective communications, Americans are better able to avoid life-threatening storms and to protect property.

- *A more efficient and open government:* Better use of information technology is at the heart of the Administration's efforts to "reinvent government" by making it more responsive to its citizens.

- *Fighting crime:* Information technology can be a powerful tool in law enforcement by, for example, allowing law enforcement officials to create "virtual" swat teams with digital wireless networks, and access databases of mug shots and fingerprints of criminals twenty-four hours a day.

- *Decreasing cost and time to market:* Information technology can reduce the cost and time to market by replacing costly engineering practices, such as pilot plants and test crashes, with scientific and engineering simulations.

- *Improving the environment:* Advanced simulation will lead to better understanding of environmental processes on scales from global to microscopic, testing cause and effect, and more accurately forecasting near-term climate changes such as El-Niño. It will also build the scientific basis for understanding long-term climate changes including ocean and ecological effects. Advanced simulation will lead to improvements in combustion technologies that could result in cost savings and reduction in carbon emissions.

4. Experts believe that the Federal government is underinvesting in R&D in fundamental information technology research.

The government's investment in computing, information, and communications research is roughly $1.5 billion — or less than 2 percent of the Federal government's R&D budget. The committee of experts in industry and academia created by the Congress and appointed by the President (The

President's Information Technology Advisory Committee) has concluded that we are significantly under-investing in information and communications R&D. They have recommended doubling the government's investment in information technology research.

5. As our economy and society become increasingly dependent on information technology, we must be able to design information systems that are more secure, reliable, and dependable.

The software systems that lie at the core of worldwide financial systems, air traffic management, defense command and control — indeed, virtually all parts of our economy — are the most complex human inventions ever created. As a result, however, our society now faces unknown hazards both from hostile attacks on these systems and from the even greater threat that simple mistakes or system failures will bring wholesale collapse of critical systems. The small software failures that have shut down large parts of the Nation's phone systems and air traffic control systems and the "millennium bug" are examples of what can go wrong in our current environment. We do not know how to design and test complex software systems with millions of lines of code in the same way that we can verify whether a bridge or an airplane is safe. More R&D in this area is desperately needed.

6. Information technology will revolutionize our national science and engineering R&D strategy.

There has been near-miraculous increase in the power of computers with speeds and capacities approximately doubling every 18 months for over two decades. We are on the verge of an era where scientific computation will be on a par with laboratory experiment and mathematical theory as a tool for research in science and engineering. The computer is literally providing a new, exquisitely detailed view of the natural world.

Scientific research in the 20th century has been characterized by a search for the fundamental laws and principles that govern our natural world. But as profound as the scientific discoveries of the 20th century have been, the benefits of this knowledge are only now beginning to be tapped. Complex problems of the 21st century will be solved by scientific simulation that will synthesize the knowledge we have gained. For example, supercomputing and simulation technology allows researchers to develop life-saving drugs more rapidly, better understand the functions of our genes once they have been sequenced, more accurately predict tornadoes, and design engines that are cleaner and more fuel-efficient.

7. Funding research also helps alleviate the shortage of information technology workers.

Many information technology companies are concerned that they cannot hire enough skilled workers — in particular, computer science graduates. Increased information technology research activities in colleges and universities will create more opportunities for student participation and enhance the skills levels of graduating students.

8. The Information Revolution is just beginning.

As impressive as technology is today, it is important to realize that the Information Revolution is just beginning. Researchers in the field have identified many research challenges that can profoundly change our lives, just as the computer and the Internet have in recent times.

R&D Agenda and Goals

As part of the President Clinton's fiscal year 2000 budget, the Administration is proposing a $366 million, multi-agency information technology research initiative that would increase Federal investment in the following three areas:

- Fundamental information technology research in software, human-computer interaction and information management, scalable information infrastructure, and high-end computing;

- Advanced computing for science, engineering, and the Nation, including supercomputers, software, networks, and research teams needed to support it; and

- Research in the ethical, social, and economic implications of the Information Revolution, and support for the education and training of America's information technology workforce.

The major components of the initiative and its research agenda are described below. Please note:

- The description of the research agenda is intended to highlight a few of the most significant information technology research challenges and opportunities to be addressed. In many of these technical areas, agencies have convened workshops of the leading researchers in the field in government, industry, and academia to define more fully the research agenda. The results of these efforts are reflected in the planning documents being prepared by the agencies.

- Although it is important to set goals and a research agenda, the government must be open to new ideas and new opportunities. The IT^2 will elicit many bold and creative ideas from the research community, and government agencies must be sufficiently agile and flexible to act upon them.

- Although the components of the initiative are described separately, they are interdependent. For example, purchasing high-end machines only makes sense in the context of software research that will make it much easier to develop applications, or that will allow users to see, understand, and manipulate the huge quantities of data generated by simulations and experiments. New insights into the way that people actually use computers at home or at work will help set the research agenda for collaboration tools.

Fundamental Information Technology Research

Fundamental information technology research addresses long-term, high risk investigations of the underlying issues confronting computer science and engineering. There are dramatic new opportunities and challenges to be addressed by fundamental research in information technology – opportunities on which the Nation must capitalize.

The fundamental information technology research component of IT^2 consists of four research focal points: software, human computer interfaces and information management, scalable information infrastructure, and high-end computing. Collectively, these research areas provide a diversified program for long-term exploration of important research problems on how to make computing and information systems easier to use, more reliable and secure, more effective, and more productive.

1. Software

Software research was judged by The President's Information Technology Advisory Committee to be the highest priority area for fundamental research. From the desktop computer to the phone system to the stock market, our economy and society have become increasingly reliant on software. This Committee concluded that not only is the demand for software exceeding our ability to produce it; the software that is 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" problem are but two significant examples of what can go wrong. There are several research areas where Federally-funded R&D could assist in creating productive, reliable, and useful software.

- *Software engineering:* Currently, we lack the understanding of how to design and test complex software systems with millions of lines of code in the same way that we can verify whether a bridge is safe. Research in this area could increase software productivity, make software more reliable and easy to maintain, and automatically discover errors. Research is also needed to reduce the cost and time penalties currently required to make systems safe and reliable. Progress on safety and reliability is particularly important for 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 the shortage of programmers is to make programming so easy that users with little or no programming expertise can do it. An example of this is the spreadsheet, which allows business and financial analysts to manipulate numbers or conduct "what if" scenarios in a way that previously required customized programs. Making end-user programming more widespread will require advances in areas such as 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 interchangeable parts that are used in the manufacturing sector. Research is needed to make it easier to find the right software component, to predict accurately 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, but this is just the beginning. Active software will eventually be able to update itself, monitor its progress toward a particular goal, discover a new capability that is needed for the task at hand, and safely and securely download the piece of software needed to perform that task.

- *Autonomous software:* Increased research in this area could result in more intelligent software and robots. Unmanned vehicles could keep our troops from harm's way. "Intelligent agents," or "knowbots," could search the Internet on our behalf. Robots and knowbots could plan, react appropriately to unpredicted changes, and cooperate with humans and other robots. Cars could drive themselves, or automatically avoid collisions. Robots could also explore planets (e.g. Voyager, Deep Space 1) or places on earth where it is not safe for people to travel.

2. Human-Computer Interaction and Information Management

- *Computers that Speak, Listen, and Understand Human Language:* Today, more than 40 percent of American households own computers. However, computers are still too hard to use for most Americans, and surveys show that computer users waste over 12 percent of their time because they can't understand what their computers are doing. Better human-computer interfaces could make computers easier and more enjoyable to use for more Americans, resulting in an increase in productivity. Given the number of people who now use computers routinely, the payoff would be enormous.

  Ideally, people would be able to have a conversation with a computer, as opposed to being limited to today's WIMP (windows, icons, mouse, pointer) interface. Research is required to make computer speech more intelligible, to increase the accuracy of speech recognition, and to give computers the capability to ask questions to confirm or clarify something a user has said. This capability will be particularly useful for people who don't have access to a keyboard (e.g. mobile professionals, doctors) or for people with disabilities who are sight-impaired or cannot use a keyboard. It will also enable simultaneous translation from one language to another. For example, it could make possible a real-time translating telephone, or allow people to search accurately foreign language databases in their native language. The potential benefits for global electronic commerce and international collaboration are enormous. Coupling this technology with robotic extenders will increase people's productivity beyond our imagination.

- *Information visualization:* Computer users are now trying 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. Without improvements in our ability to see, understand, and manipulate huge quantities of data, our ability to tackle some of the most important science and engineering challenges will be limited.

3. 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 and more pervasive, the networks of tomorrow will have to support a billion or more users. These users will be sending and receiving voice, video, and high-speed data; accessing the network while traveling; sending tens of millions of simultaneous requests for information from a popular Web site; and depending on the Internet to run a business, deliver government services, or respond to a medical emergency. The Internet of the future will also connect billions or even trillions of devices. Computers will be combined with sensors, wireless modems, 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. Enabling the Internet to grow (or "scale") to support these new demands requires basic information research in a number of areas.

- *Deeply Networked Systems:* Such systems, allowing a dramatic increase in the number of devices that can be attached to the network, will result in compelling new applications. For example: low-cost wireless sensors could give us real-time information on air and water pollution, improving our ability to monitor the environment and respond to man-made disasters. "Guardian angels" could monitor the health and safety of individuals (e.g., firemen, law enforcement officials, soldiers, home health care patients). Crisis management centers could use sensors carried by response teams and airplanes to improve responses to forest fires, floods, and hurricanes. Enabling this to occur requires advances such as new mechanisms for naming, addressing, and network configuration, and much cheaper network interfaces.

- *Anytime, anywhere connectivity:* Improvements in wireless technology could bring high-speed "anytime, anywhere" connectivity to all U.S. citizens. One of the major benefits of higher-speed wireless networks is their ability to expand access and reduce the "cost penalty" associated with deploying advanced telecommunications in rural areas. This is critical to rural economic development, since companies are increasingly basing site selection on the quality of the telecommunications infrastructure. Wireless networks could also extend services such as distance learning and telemedicine to remote rural areas in the U.S. and to the markets of developing countries.

- *Network modeling and simulation:* Since we have no experience in building a network that is as complex as the future Internet, we need better tools – proven through experience – to model the behavior of networks. The capability for "faster than real-time" simulation that allows network operators to prevent congestion or collapse is also needed.

4. High-End Computing

Advances in high-end computing benefit the country, the economy, and the lives of all U.S. citizens. High-end computers can be used to gain new medical insights, to forecast the weather with greater accuracy, to design advanced weapons systems, and to predict climate change more accurately.

As a result of investments by the High Performance Computing and Communications Program and the Accelerated Strategic Computing Initiative, the Department of Energy's national labs are currently running computers capable of sustaining a trillion calculations per second on 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. Long-term research is also required to develop computers that are capable of a thousand trillion (10^{15}) calculations per second. The Government has a particularly important role to play in high-end computing. While 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 in the private sector. Research topics will include:

- *Improving the performance and efficiency of supercomputers:* U.S. companies can now make 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. R&D is needed to make these high-end machines easier to program, and to improve the efficiency of typical programs. This research will focus on developing programs that can be easily moved from one high-end computer to another, and creating new algorithms, problem-solving environments, and compilers that are designed with massively parallel machines in mind.

- *Creating a computational grid:* Clearly, many researchers and computer users will not have local access to all the computing power and storage they need. In the same way that the electric power grid provides universal access to electrical power, a computational grid could provide more 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 half-way around the world — in three-dimensional, virtual environments. Building these computational grids will require new programming tools: software that can translate the requirements of an application 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.

- *Revolutionary computing:* Eventually the technology that we have used to make computers smaller, faster and cheaper since the 1960s will run into physical limitations. We need to begin now to explore fundamentally new computer technologies (e.g. biological computing, single electron transistors, computers that take advantage of quantum logic, and devices based on carbon nanotubes). This could lead to computers that are well beyond a thousand times faster than today's fastest supercomputers.

Advanced Computing for Science, Engineering, and the Nation

During the past decade, fueled by the exponential increases in computing power provided by information technology, we have made dramatic progress in the 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. We still face the challenge of using this capability to predict the behavior of larger molecules such as proteins, with their thousands of atoms in constant motion, all tugging at each other from different directions. Since subtle differences in the shape of a protein determine its biological function and its potential role in medical treatments, we have fundamental interest in rapid, accurate, complete models of the protein's behavior.

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. This will require: (i) obtaining computers that are 100-1000 times more powerful than those now available to the research community and making them available on a competitive basis, (ii) developing the scientific and engineering simulation software and other tools needed for these new machines to be useful for research (such as new mathematical algorithms and parallel programming environments, as well as tools for collaboration, visualization and data management systems), and (iii) building multidisciplinary teams in which scientists working on the most challenging research areas benefit from advances in fundamental information technology research funded through this initiative and in which computer scientists explore interesting and difficult new problems.

1. Advanced Infrastructure

The fastest computers now available to the civilian research community on a competitive basis are capable of doing 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 5 trillion (a thousand billion) computations per second by the end of fiscal year 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 aspects of the advanced infrastructure will benefit from the fundamental information technology research funded in other parts of IT^2.

2. Advanced Science and Engineering Computation

New supercomputers require fundamentally new approaches to scientific computation. Most computers in wide use by the research community today operate with a single computer processor. This processor does tasks one after the other. The ultra-fast machines purchased under this initiative will have thousands of separate computer processors all operating at the same time. This can be of great value since many processes in the real world operate simultaneously (weather changing across the globe, atoms moving in a complex material). Effective use of such machines requires separating a problem into pieces to be worked on by different small computers and then reassembling the pieces – something very difficult to do in practice.

Thus, researchers involved in IT^2 will need to be prepared to take vastly different approaches from those currently in wide use. Either individually or as a group, they will need to demonstrate deep understanding of the underlying science and engineering fields, 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. What they learn 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.

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 some extraordinary advances in solid state physics. Information science will be no exception. IT^2 is designed to build a strong working relationship between teams working on scientific and engineering simulations — whose primary interest is attacking a problem of climate change or research physics — and teams interested in the broad field of information research, whether it is in visualization or basic aspects of software design. Both groups should grow and benefit from this relationship.

The following provides examples of potential science and engineering problems that will become tractable with the range of computational power currently under discussion for IT^2. Note that this is in no way a comprehensive list.

- *Predicting Climate Change:* Significant improvements are needed in both the accuracy of the forecasts and our 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.

- *Severe Weather Prediction:* 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 the function of a protein is intimately connected to its three dimensional structure, the connection is only poorly understood. A detailed understanding of this structure-function relationship, reached through computational simulation, 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. The IT^2 will result in better understanding of the ground motion of large sedimentary basins during earthquakes by allowing scientists to study this phenomenon with much greater accuracy than is possible now.

- *Simulation of Combustion*: Predictive simulations of the performance of automobile engines hold the promise of simultaneously decreasing the nation's dependence on foreign oil and increasing the quality of the environment (by reducing emissions, including carbon dioxide). Predictive simulation of the performance of diesel engines will help the transportation industry meet the proposed emission standards in the year 2004.

- *Materials Simulation*: The nation has a continuing need for new materials — such as 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 computing-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. Such tests are essential to capitalize on the major investments that have been made in the new observational instruments.

Projects funded under IT^2 will be selected because they are of importance to science and engineering, they are in a position to take advantage of enormously more powerful computational tools, and they involve collaboration between information scientists and experts in the research area. Access to high-end capabilities will be important for both the development work needed to make use of these extraordinary new tools and for the large-scale modeling efforts envisioned. Projects from all areas of science and technology will be considered, with each agency providing support for projects within the scope of its mission. And — because it cannot be emphasized enough — ties with the information technology research community will be the key to success.

3. Computer science and enabling technology

The effort described above puts the use of massively parallel computing on an aggressive trajectory similar to that of the ASCI program. The agencies deploying the infrastructure for computation will partner with ASCI to focus on the development of 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 technology 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 point of synergism between the fundamental information technology R&D and advanced science and engineering computation elements of IT^2 as it is currently structured.

4. National Information Infrastructure Applications

There are many large-scale, distributed applications of high social and economic impact that are not explicitly part of the IT^2 but can take advantage of advanced information infrastructure. These applications may require one or more features of advanced computing and networking infrastructure, such as high-speed communications or distributed access to a large number of computers, devices, or people, or high performance processors. The R&D agencies will look for opportunities to partner with and leverage other programs to identify applications and testbeds to stress the advanced information infrastructure along one or more dimensions.

Economic and Social Implications of Information Technology, and Training of the Information Technology Workforce

1. Economic and social impacts of information technology

Information and communications technologies are having a pervasive impact on our economic, political, social, and cultural life. As we enter the 21st century, we should expect this trend to accelerate, as information technology becomes more powerful and widely available. Despite the significant economic, legal, social, ethical, political, and cultural impacts of this technology, the Federal government has sponsored little social science research in this area. For example, there is no equivalent in the information technology area of the National Institutes of Health ELSI (Ethical, Legal and Social Implications) Research Program. Increased research in this area with greater interaction between computer and social scientists, would be helpful for a number of reasons. It could contribute to the design of information systems by providing insights into how information systems are actually used; identify barriers to the adoption of information technology and its applications; help policymakers by providing more empirical data on the impact of information technology; and encourage the development of solutions to problems caused by information technology (e.g. the erosion of personal privacy). Some of the research topics in this area might include:

- the effectiveness of industry self-regulation and other forms of non-traditional governance for the Internet;
- the impact of information technology on the nature of work, including telecommuting, quality of worklife, and participation in "virtual" teams;
- the effect of disparities in access to information technology by race, income, ethnicity, and geography; and
- the impact of electronic commerce on market structure, supplier-customer relations, productivity, price stability, and employment.

2. Information technology workforce

The science agencies also need to do more to address the rapidly growing demand for workers with information technology skills at the undergraduate, graduate and post-graduate levels. Colleges and universities are reporting increased undergraduate enrollments at the same time as applications for graduate study are declining. The National Science Foundation is proposing a new initiative to ensure that all faculty (2 year colleges, 4 year colleges, research universities) have access to modern curricula and instructional materials. Additional funds are also required for graduate traineeships to support graduate and post-graduate students. Furthermore, a large portion of the research grants provided to universities will go to support research assistant positions for graduate students.

Of course, these initiatives by the science agencies will only address one component of the much broader information technology workforce issue. The demand for information technology workers is increasing across-the-board. Many workers now require basic computer literacy to compete for the jobs of the 21st century. There are also many information technology-related jobs that do not require an undergraduate or graduate education in computer science, or even any undergraduate degree, such as network administrators, entry-level computer programmers, help desk operators, systems analysts, and Web page designers. Agencies such as the Department of Labor, the Department of Education and the Department of Commerce are working with industry, professional societies, and labor organizations to ensure that the United States has the best information technology workforce in the world.

Management of the IT^2

While information research is a critical national research priority, no cabinet agency has information research as its central mission. As a result, a new senior management team, reporting directly to the President's Advisor for Science and Technology, has been formed to set policy and coordinate the work of this new initiative. This group will consist initially of 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). It will fall under the umbrella of the NSTC and will also include senior officials from the Office of Management and Budget and the National Economic Council. This group will assist the President in establishing and monitoring goals for the program, allocating research tasks to agencies on the basis of agency missions and capabilities, ensuring tight coordination of Federal efforts, and ensuring that the entire research program is managed in a way that allocates funds in an open, competitive process aimed at funding the best possible ideas.

This senior group will be supported by a working group consisting of an individual appointed by each of the principal agencies. It will be chaired by the Assistant Director of the National Science Foundation for Computer and Information Science and Engineering. The National Science Foundation has been chosen to lead this group since it is alone in having all three thrusts of this initiative as part of its primary mission. This working group will be charged with preparing research plans and budgets for the entire Federal effort, preparing annual reports which can describe progress and plans, and facilitating coordination between 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. The working group will ensure that these separate research projects, taken together, provide a sound and balanced national research portfolio. The working group is comprised of individuals with operational authority over information technology research programs within their home agencies. The National Coordination Office will assist in coordinating this working group, as it has in the past for high performance computing and communications activities.

The working group will oversee research in all the major areas of the initiative. The working group is developing plans for allocating this diverse set of tasks to subgroups that will focus on specific components or areas of the initiative. The working group will utilize existing groups and teams to the maximum extent possible.

One requirement of the new organization will be to develop and operate the new advanced infrastructure that will be made available to the research community by the National Science Foundation and the Department of Energy under funds provided by this initiative. A subgroup, chaired by these two agencies, will be charged with devising a program that will ensure that these new computers are purchased, sited, and made available on an open, competitive basis. This subgroup will also develop a process to 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 state-of-the-art information and computational science and groups familiar with research challenges in areas like biochemistry or climate modeling that can benefit from access to ultra-fast machines.

The proposed fiscal year 2000 budget is provided in the table below:

Agency	Fundamental Information Technology Research	Advanced Computing for Science, Engineering, and the Nation	Ethical, Legal, and Social Implications and Workforce Programs	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

Next Steps

In the coming months, the Administration, working closely with Congress, will pursue several next steps for the initiative, including: (1) refining the management structure for the initiative; (2) continuing to refocus and strengthen ongoing related programs to complement the initiative; (3) continuing to seek external advice from industry and academia; and (4) developing a detailed technological and programmatic roadmap for the initiative.

APPENDIX

AGENCY ROLES IN THE INFORMATION TECHNOLOGY FOR THE TWENTY-FIRST CENTURY (IT^2) INITITATIVE

Table of Contents

Department of Defense (DoD)

The Department of Defense's (DoD) participation in this initiative will be supervised by the Under Secretary of Defense for Acquisition and Technology in his role as principal staff assistant to the Secretary of Defense for research and development programs within the DoD. DoD's activities will involve support for basic long-range research and related equipment necessary to facilitate advances in information technology managed through the University Research Initiative program, focused research activities to be managed by the Defense Advanced Research Projects Agency (DARPA), and a new Advanced Research and Development Activity (ARDA) in Information Technologies in support of the Intelligence and Information Security Communities.

The Initiative is Critical to DoD

This initiative will attack problems that will enable revolutionary capabilities of importance to the DoD and the nation as a whole. 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. In short, the DoD is participating because the military that stays ahead in information and its timely dissemination is likely to dominate in the future.

University Initiatives

The University Research Initiative is an established, well-regarded DoD program designed to enhance universities' capabilities to perform basic science and engineering research and related education in science and engineering areas critical to national defense. The Information Technology Initiative will include support through the University Research Initiative focused on stimulating fundamental advances in information science and engineering areas of interest to DoD. DoD plans to support multi-disciplinary research and the purchase of equipment necessary to perform fundamental research work that relates to information science and technology.

DARPA Focused Research Programs

I. Software Research

A particular area of concern to the DoD is the degree to which today's interactive software, much of which was pioneered by DARPA-sponsored researchers, requires the constant attention of human users, who both supply the inputs and consume the results. Although this interactive approach to computation seemed appropriate during the transition from mainframes to personal computers, we must now lay the software foundations for a new era in which our computers will vastly outnumber our population. To accommodate this next step forward, we must develop new approaches to software that will allow individual users to leverage hundreds, and eventually thousands, of networked processors. Such new software approaches include active software that is capable of deploying itself to nodes within the network and autonomous software that allows robots (and knowbots) to react to changes in their environment, thereby freeing humans from the need to painstakingly monitor the actions of every device and agent under their control. In addition, tolerant software technologies will allow systems built from parts that are imperfect to deliver answers that are good enough and given soon enough, as opposed to perfect answers provided too late.

II. Human-Computer Interaction and Information Management

DoD’s efforts to free our nation’s soldier and citizenry from the tyranny of the keyboard will involve the development of critical technologies to enable more effective interaction between human users and the vast data of future information systems. Technologies to be developed include multi-modal interfaces and the ability to access information regardless of the language the data was originally entered in.

III. Scalable Networks

Current Internet technology targets only two percent of all computers, i.e., the personal computers, servers and supercomputers that are directly responsive to human operators. The remaining 98 percent of computers are stranded within devices whose sensors and actuators are in direct contact with the physical world. Imagine the productivity gains that could be achieved if we were to extend the “depth” of the network to reach these embedded devices and close the chasm between the virtual world, of networked bits and bytes, and the physical world, which we strive to monitor and manage. DARPA will develop methods for accessing, tasking, and managing dynamic, multi-node networks of sensors and fusing and analyzing the resulting data. DARPA will also focus on the network technology required to enable these dynamic networks of sensors, actuators, processors and data storage devices.

IV. High End Computing

DoD’s focus at the high end will be on the development of 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 that is directly acquired from the physical world. These include new approaches to computer component technologies and system architectures.

DARPA’s Role

DARPA has a tradition of deliberate entrepreneurship and risk-taking to achieve high payoff goals — especially in areas where the level of technical risk would inhibit others. The focused research areas of DoD’s Information Technology Initiative participation are aligned with DARPA’s investment strategy, which: (1) concentrates on high-payoff technologies and concepts; (2) sometimes invest in those critical technology developments, such as the ARPANET and internetworking technology, that require an extended focus and long-term development; (3) is idea-driven and outcome oriented, i.e., creates a vision of what could be done in 10 to 20 years to solve a specific problem and defines the technology development that must be pursued to make the vision a reality. DARPA will channel DoD resources into selected technologies and pursue them with vigor and determination. Just as critical, however, are the basic research programs of DoD and elsewhere in the Federal Government that sustain the research base on which DARPA’s efforts must be built.

ARDA

The Defense Department and the Intelligence Community are also launching a program known as ARDA, the Advanced Research and Development Activity in Information Technology for the Intelligence and Information Security Communities. ARDA was created because both of these communities believe that they are currently underfunding high-risk/high-payoff research in information technology. In addition to supporting fundamental research and development in key areas of information technology, ARDA will also work to recruit “research fellows” to supplement the Intelligence Community’s R&D workforce, transfer technology that is developed or discovered as a result of the initiative, and form strategic partnerships with world-class researchers in industry and academia.

ARDA will focus its long-term research on a small number of problems that are relevant to the Information Security Community, the Intelligence Community, or support enabling technologies that are relevant to both communities. Examples of potential research areas and objectives include, but are not limited to:

A security management infrastructure that protects data from both active attack or from unintended, passive access;

Securing Intelligence Community and Information Security Community systems and communications as they become increasingly integrated into the Global Information Infrastructure;

Screening, filtering, identifying, categorizing, mining, clustering, selecting, and retrieving multimedia data and documents based on the information needs of intelligence analysts and customers of intelligence;

Automating (partially or fully) selected human analytic processes and cognitive activities;

Enabling information technologies such as advanced computing, mass storage, networking, microelectronics, reverse engineering, mathematics, cryptology, and computational linguistics; and

Methodologies for testing and evaluating information technology research results for efficiency, accuracy, flexibility, portability, usability and scalability across a wide problem set.

Department of Energy's (DOE)

Background. DOE is a science-based agency that contributes to the future of the nation by ensuring our energy security, maintaining the safety and reliability of our nuclear stockpile, cleaning up the environment from the legacy of the Cold War, and enhancing our understanding of the physical world through research in fundamental science.

To accomplish its mission objectives and continue to effectively solve challenging problems of national importance in the 21st century, the DOE has proposed an investment to take it into a new era of computational scientific simulation—to rapidly develop and deploy a new generation of computational simulation capabilities along with the supporting national terascale computing infrastructure and to apply this new capability to large-scale civilian science and engineering problems. This capability will not only revolutionize the Department's approach to solving the most demanding, mission-critical problems, but will stimulate our national system of innovation.

Computer-based scientific simulation is one of the most significant developments in the practice of scientific enquiry in the 20th Century. Computational simulation has dramatically advanced our understanding of the fundamental properties of matter (e.g., the structure and interactions of molecules) and is beginning to have a major impact on critical engineering problems (e.g., automobile design). In many cases, simulation is the only economically and technically viable means to accomplish the DOE missions due to the complexity of the problems and the expense or duration of the needed experiments and tests. In some instances, simulation is the only solution because it is not possible to carry out the needed experiments or tests.

As an example, in 1996, the Clinton Administration launched DOE's Accelerated Strategic Computing Initiative. This program fundamentally changes the way our nation ensures the safety, reliability, and performance of our nuclear weapons stockpile by shifting from an emphasis on nuclear tests to one based on computational simulation validated by experiment.

DOE Role. DOE's participation in the President's IT^2 Initiative is built on a long and successful history as a lead agency in computational science, computer science, applied mathematics, and high-performance computing; in the design, development and management of large-scale user-facilities, including a number of national computing facilities; and in the development of cross-cutting computing and engineering technologies such as data analysis and visualization, collaborative environments, and other related technologies.

DOE's unique set of research capabilities, mission goals, and high-performance computing and advanced applications expertise are critical to three primary goals of the President's IT^2 Initiative:

GOAL #1 - SCIENTIFIC APPLICATIONS. To design, develop, and deploy computational simulation capabilities to solve scientific and engineering problems of extraordinary complexity.

GOAL #2 - COMPUTER SCIENCE AND ENABLING TECHNOLOGY. To discover, develop, and deploy crosscutting computer science, applied mathematics, and other enabling technologies.

GOAL #3 - SCIENTIFIC SIMULATION INFRASTRUCTURE. To establish a national terascale distributed scientific simulation infrastructure.

Goal #1. SCIENTIFIC APPLICATIONS. DOE will initially focus on two major simulation projects that are critical to the agency's mission, have urgent deadlines, are of high scientific impact, and are well positioned to take advantage of terascale computing—global systems and combustion systems. In addition, a number of applications in basic science (e.g., genomics) that fulfill these criteria will be initiated.

Applications	Project Goal	Benefits
Atomic, molecular, gas/fluid, and device-level simulation of chemical reactions, fluid dynamics, and energy transfer in real-world **combustion devices**	To enable **predictive** simulation of complex, practical combustion systems that underlie 80% of global energy usage.	Design of improved combustion devices (e.g., automobile engines) leading to: • Reduced pollution–close to zero for nitrogen oxides and soot. • Improved efficiency–reduces fuel consumption, minimizes carbon dioxide emissions
Advanced modeling of the impact of human activities on **global systems** such as weather patterns and long-term climate change	To predict, by the year 2005, regional climate under scenarios of global energy usage, in a multi-agency, international effort to support the IPCC 4th Assessment.	• Improved management of water resources • Estimates of regional energy demand and use • Anticipating damage caused by natural disasters
Basic sciences including, but not limited to, material science, structural genomics, plasma physics, high energy physics, and subsurface transport	Project goals are specific to the research areas (to be selected by peer review), but will in all cases dramatically alter research progress though revolutionary applications of simulation.	• New concepts for and understanding of scientific problems of national importance that cannot be obtained from theory or experiment alone • Accelerated development of new energy technologies from knowledge gained in scientific simulations

Goal #2. COMPUTER SCIENCE AND ENABLING TECHNOLOGY. DOE will focus on the development of 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 technology 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.

Goal #3. SCIENTIFIC SIMULATION INFRASTRUCTURE. The hardware strategy will be driven by the applications requirements and will be based on the acquisition of a balanced system of advanced computers for computational methods and software development as well as the demanding applications listed in GOAL #1. Major computing platforms are planned for acquisition by DOE with an aggregate of 5 teraflops in FY 2000 and 40 teraflops by the year 2003. Local random-access memory, on-line disk, and archival storage needs will increase from 1-2 terabytes, 30-40 terabytes, and approximately 5 petabytes, respectively, in FY 2000 to 10-15 terabytes, 300-500 terabytes, and 15-20 petabytes, respectively, in FY 2003. The greatly expanding computational capacity of these resources coupled with the distributed nature of the scientific teams collaborating on the applications requires advanced communications capabilities to support distributed collaborations, distributed simulations, remote visualizations, remote steering of computations, and the sharing of huge datasets.

While the DOE's major contribution to the President's IT^2 Initiative relates to the advanced computation and engineering component, its efforts in crosscutting and enabling technologies are expected to both drive and profit from basic information technology research.

National Aeronautics and Space Administration
(NASA)

Background: NASA is an investment in America's future. As explorers, pioneers, and innovators, we expand frontiers in air and space to inspire and serve America and to benefit the quality of life on Earth. NASA's bold missions in space and aeronautics require significant advances in information technology. Potential benefits of NASA's participation in the Information Technology for the Twenty-first Century Initiative are:

- To develop **autonomous spacecraft and rovers**, that will enable us to explore the universe with less risk and at lower cost**.**

- To increase our ability to transform large and distributed streams of data into **scientific data understanding and knowledge**.

- To promote new concepts in **aviation operations** that will enable aircraft to safely and effectively negotiate clearances, routings and sequencing.

- By enabling **an intelligent synthesis environment** that will revolutionize the nation's science and engineering infrastructure – the tools, technologies, capabilities, and work practices we use to deliver cutting-edge capabilities

NASA's participation in the President's initiative is built on a long and successful history in computational science, computer science, and high-performance computing. NASA is an innovator and early user of information technologies. NASA's unique mission and research capabilities will be applied to support the three goals of the President's IT^2 initiative:

Goal #1 - Long-term information technology. To discover, develop and deploy fundamental computer science, computational science and applied mathematics to serve the national economy, national security, education, the global environment and broad societal goals.

Goal #2 - Advanced computing for science, engineering, and the Nation. Empower computational discovery in all areas and ensure tera-scale computational resources are available to a broad research community to enable scientists and engineers to solve problems of extraordinary complexity.

Goal #3 - Training of the information technology workforce. Understand 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.

NASA Role Goal # 1. In FY 2000, the first year of a planned presidential initiative, NASA will support fundamental intelligent systems research focused on several areas, including:

Software research – NASA will conduct research in automated reasoning which requires advances in model-based reasoning, high assurance software, biologically motivated adaptive systems, and planning and scheduling. The goal is to demonstrate coordinated planning and execution of self-tasking and self-repairable robotic networks early in the next decade to support future decisions on establishing permanent, sustainable robotic outposts at key points of scientific interest in the solar system. Developed in close coordination with the research objectives of ongoing and future programs, these revolutionary capabilities will dramatically increase the return of NASA's space and Earth science missions.

Human-centered computing and information management – NASA will develop intelligent systems for data understanding including geographically distributed computing, reconfigurable computer architectures, biologically-motivated hardware and software, and knowledge discovery and data mining. NASA will advance human-centered computing including knowledge management and institutional knowledge capture, immersive/haptic environments, internet-based knowledge representation and cognitive architectures. The goal is to provide the technology to understand and manage large and distributed streams of data to support applications such as Earth system science.

High-end computing – NASA will explore revolutionary computing concepts including quantum mechanical computing, neurally-inspired computing, holographic memory devices, optoelectronic systems, biological computing, biomimetics and ubiquitous intelligence.

NASA Role Goal #2. NASA will support relevant research in applications and use of tera-scale infrastructure in implementing this initiative.

Applications – NASA will advance the Intelligent Synthesis Environment, which includes rapid synthesis and simulation on tera-scale systems and distributed collaborative engineering capability. The goal is to demonstrate design and synthesis of vehicles and missions with greatly reduced redesign and rework and improved operational reliability. The first Intelligent Synthesis Environment testbed will apply these benefits to the design of future reusable launch vehicles that promise to greatly reduce the cost of space transportation. By 2003, substantial improvement will be made towards the following objectives as measured against a standard vehicle and mission baseline: reduce design and development time to 12 to 18 months, reduce testing requirements by 75%, predict life-cycle cost to within 10%, and predict life-cycle risks with a confidence level of 95%.

Access to tera-scale infrastructure - NASA applications which require tera-scale capability include understanding seasonal-to-interannual climate change and demonstration of intelligent synthesis environment.

NASA Role Goal #3. NASA will support training and education efforts to revolutionize cultural change in the scientific and engineering community to exploit the power of an intelligent synthesis environment.

National Institutes of Health
(NIH)

The NIH mission is to uncover new knowledge that will lead to better health for everyone. NIH works toward that mission by conducting research in its own laboratories; supporting the research of non-Federal scientists in universities, medical schools, hospitals, and research institutions throughout the country and abroad; helping in the training of research investigators; and fostering communication of biomedical information. One of the world's foremost leaders in medical research, NIH funds almost 30 percent of the Nation's health-related research and development.

To achieve its mission, the NIH relies on bioinformatics, high-end computer 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, will be dependent upon information technology.

The reliance of the medical sciences on the application of information technology can be exemplified by several recent advances and current projects:

The full blueprint for a multi-celled organism *C. elegans*, or roundworm, was announced in December, 1998. This marked a historic step for the Human Genome Project as it is the first time scientists have deciphered the genetic instruction of a complete organism that shares many of the same basic life functions as humans. Because *C. elegans* shares 40% of its genes with humans, it can serve as a model for human genetic research and provide entrée into the discovery of cures to human diseases.

Many of today's most effective interventions are the direct result of knowledge gained through clinical trials–studies that evaluate the safety and effectiveness of new drugs and other interventions. NIH is creating an internet accessible database that will include all federally and privately funded clinical trials for drugs for serious or life threatening diseases and conditions submitted under Investigational New Drug applications.

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 robust and improved software. There are great requirements in bioinformatics, imaging and statistical analysis as well as development of methods in ontologies (knowledge bases) and three-dimensional image labeling and searching. NIH proposes to:

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

- Develop simulation methods for implementing realistic models of molecular, cellular, organ and epidemiological systems on parallel architectures and create component libraries that will provide a framework so that researchers can 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. This initiative will provide the capability to quantify and track the expression of tens of thousands of genes in space and time, and will generate enormous amounts of data.

High-End Computing. Supercomputing technology is used across the government for applications such as imaging, and processing the volumes of data such as that generated by the Human Genome Project. To permit labs to establish high-throughput/low-cost processor clusters, NIH proposes to:

- Develop an initiative to permit small labs 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, off-the-shelf, processors instead of expensive "supercomputer" processors. This funding will provide the kernel for enabling development, software conversion and the creation of a cookbook approach for other labs implementing these systems.

IT Workforce Training. 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 increases to become an 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. 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 excellent example. The development and successful introduction of new technologies require both an understanding of biology and the fundamentals of the technology to be applied, so that the solutions derived are appropriate and can be implemented.

National Oceanic and Atmospheric Administration (NOAA)

The **National Oceanic and Atmospheric Administration's mission** is to describe and predict changes in the Earth's environment, and to conserve and manage wisely the Nation's coastal and marine resources to ensure sustainable economic opportunities. To support this mission, NOAA develops and uses complex computer models of the atmosphere and oceans, maintains databases approaching a petabyte in size, collects environmental data in real-time both routinely and in response to disasters, and disseminates information products to emergency managers, business and the public for the protection of life and property.

A fundamental approach that NOAA has used over its entire history is computer simulation of the atmosphere and oceans, treated either separately or as interacting systems. Computer modeling is central to its mission, both for weather forecasting and climate research. Weather forecasting requires rapid and robust input of observational data via data collection and communications from sources around the globe, from satellites, and, in the near future, from high density Doppler radar data. These data are assimilated, using advanced variational techniques, into weather models with increasingly sophisticated physics and higher spatial and temporal resolution. A fundamental requirement of weather forecast models is that they produce forecast results within extremely strict wall-clock constraints. This can only be accomplished with the fastest computers available using the most efficient software techniques and algorithms. Climate modeling by its nature requires long running model experiments with increasingly complex model physics. At these time scales, the oceans have a profound influence on the atmosphere, requiring fully interacting representation of the atmosphere, oceans, and land-surface hydrology. As climate research advances and begins to resolve regional climate effects, climate model sophistication and complexity will increase dramatically, not only to address the demands of higher model resolutions, but also to support the demands of more elaborate representations of model physics that these higher resolutions require. Finally, NOAA must be more effective in providing forecast and research information to its many customers. Advanced communications technologies are providing NOAA with new opportunities to enhance this information dissemination, not only within its organization, but also to the public at large in normal times and to emergency managers in times of crisis. This work leverages existing resources to most effectively advance the use of state-of-the-art information technologies to accomplish NOAA's mission.

As NOAA faces new demands for increasingly accurate weather forecasts and more robust predictions of near- and long-term climate, NOAA scientists must address important computer science challenges in the development and implementation of climate/weather applications for scalable architectures. First, as models become increasingly complex, the need for a modular or component-based design becomes more critical, both to improve usability by a broad range of users and to assure maintainability. While NOAA scientists have been incorporating this approach into the design of some of their major climate model codes, further work is needed to refine and test this strategy for its value for collaborative geophysical research. In addition, as NOAA organizations acquire and install highly scalable parallel systems for use in both weather forecasting and climate research, end users are in the process redesigning and optimizing all of their major weather and climate for highly to massively parallel architectures. The acquisition by the National Weather Service of an IBM SP system for its operational weather forecasting mission is an example of this transition and its importance to NOAA's mission. For these transitions to be effective, NOAA will need to develop efficient scalable versions of current NOAA applications, which will require the development of new algorithms and the redesign of model codes to optimize cache use rather than vector performance. The most effective way to achieve this will be for model developers to work side-by-side with computer scientists, both to create effective model designs and to facilitate the development of software tools. Finally, as the resolution and complexity of their models increase, NOAA scientists will require much more effective ways to visualize and therefore better understand the multi-dimensional geophysical data sets that these models produce.

At the beginning of this initiative, NOAA will push the state of the art in the use of advanced high-speed computing, efficient parallel algorithms, and high bandwidth data communications to provide the Nation with the best-available weather forecasts and climate predictions. As the initiative progresses, NOAA will be among the first to adopt technologies to benefit the Nation in large-scale computing, data management for petabyte data systems, data visualization as an aid in the diagnosis of the output of advanced geophysical numerical models, and nomadic real-time streaming of data, sound, and video over robust adaptive networks.

Benefits
Using the results of the FY 2000 IT^2 initiative, over the next 5 to 10 years NOAA expects to realize many critical benefits including the following:

HURRICANES
Reduce 72-hour forecasts of hurricane track error by 20% from 220 to 150 nautical miles.

Recognize hurricane forecast situations that have inherently low predictability through the use of compute-intensive ensemble techniques.

Improve forecasts of hurricane intensity by 20-30% over current forecasts through use of higher resolution models and more sophisticated physics made possible by advanced computing and atmospheric science.

SEASONAL TO INTERANNUAL CLIMATE
NOAA will improve forecasts of seasonal to interannual climate by including the full global air-sea interaction and hydrology, and enhanced probabilistic guidance using ensembles.

TORNADOES
NOAA forecasters expect to be able to forecast tornadoes as much as 2 hours in advance (the current standard is 15 minutes) through the use of ultra-high resolution, limited-area models made possible by advanced information technology as well as advanced atmospheric understanding.

ROUTINE FORECASTS
NOAA will provide 5-day forecasts with the accuracy of current 4-day forecasts, a 20% improvement.

LONG-TERM CLIMATE
NOAA will produce a 50% reduction in current modeling uncertainty through a more complete treatment of clouds. These science-based improvements will produce a more fact-centered basis for optimal policy decisions in the future.

Milestones:

FY 2000

Release two additional features of the Scalable Modeling System, a suite of advanced modules which take advantage of emerging component software technologies to support rapid development and easy porting of atmospheric, climate and ocean models across scalable architectures.

FY2001

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. The system will also be optimized for performance for the Weather Service's scalable architecture using advanced programming and cache management techniques.

FY2002

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. A key aspect of this model development activity will be the use of component-based design to facilitate the exchange of physics packages and maximize the usability of the modeling system.

FY2003

Isolate some sources of climate "drift" and define a strategy for reducing their effect on long-running, higher resolution coupled climate models. A key aspect of the diagnostic process will be robust, flexible data visualization and advanced diagnostic strategies including data mining techniques to locate the sources of drift.

FY2004

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. It is expected that these enhancements will be considerably simpler to implement as a result of the component-based design implemented in the early years of the initiative. These complex and subtle research results will be communicated using collaborative visualizations techniques not readily available today.

National Science Foundation
(NSF)

The National Science Foundation funds a very broad spectrum of fundamental research in science, mathematics, and engineering at the Nation's academic institutions. Equally important is the NSF's responsibility for educational activity in the sciences across the full spectrum from K-12 to graduate school to informal education and workforce training. The Foundation's activities in all these areas contribute to, and are aided by, its strong support of research activity in basic information technology, including computer science, computation, and information transfer.

In FY 2000, NSF will support basic efforts focused on several areas of software research, including:

- *Building "no-surprise" performance-engineered software-systems* in a scalable way *t*o deliver 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 lie at 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.

Efforts in human-computer interactions and information management will focus on research in areas such as:

- *Multiplying individual physical and mental capabilities* via computer sensors and actuators to control devices which 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-dimensional, virtual environment.

- *Building a ubiquitous content infrastructur*e that enables seamless retrieval of text, data, visual and other available information in all subject areas by all citizens.

In the area of scalable information infrastructure, NSF-supported research will focus on:

- *Broadband tetherless communications* to help realize exciting new technologies such as telemedicine, crisis-management applications, and expanded distance-learning.

- *Understanding, modeling, and predicting the behavior of networks.* Research to simulate and model million-node to billion-node networks, taking into account such tradeoffs as bandwidth, reliability, latency, and different network modalities.

- *Integrating end-to-end performance component*s (networking, data manipulation, algorithms, simulation and modeling, human-computer interfaces, visualization) to achieve overall system performance.

High-end computing research has direct impact on both computer and information science, and areas such as biology, engineering, geosciences, mathematical and physical sciences, and social, behavioral, and economic sciences. Research activities will include:

- *Algorithms related to computational complexity*, particularly algorithms for solving the partial differential equations arising from models of physical phenomena using high-end systems architectures.

- *Access to terascale systems* for computer science researchers, recognizing that many important problems emerge only at large scale. Examples include research on file systems coordinating terabyte data transfers on thousands of disks, and programming systems for managing multilevel memory hierarchies on thousands of processors.

- *Empowering computational discovery* in all of these areas by coupling advanced computation, communications, and data technologies. Research will include work on computational manipulation, modeling, and representation of physical phenomena, as well as on how to achieve the scalability of thousands of processors, million-way parallelism, and distributed and parallel data archives.

Relevant infrastructure is key to implementing this initiative.

- NSF will provide open, competitive access to cutting-edge computing resources for the science and engineering community. Plans are: putting in place a 5-teraflop capability at one site in FY 2000; potentially developing a second site with 5-10 teraflop capability in FY 2001; and providing upgrades. Careful coordination will be undertaken with other agencies to leverage the software, tools and technology investments for mutual benefit. The terascale computing resources will be connected to the PACI network, the core of NSF's computational infrastructure, enabling an acceleration of capabilities planned for the future.

- To make these computing resources available to the broad research community ubiquitous high-bandwidth network connectivity must be available. This will make the location of data and computing resources irrelevant, allowing placement where facilities are available and operating costs are minimized.

There are important ethical, social and information technology workforce issues to confront. These include:

- *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 IT technologies

- *Developing a more skilled American workforce* for the future global marketplace, utilizing the many powerful tools available because of information technologies.