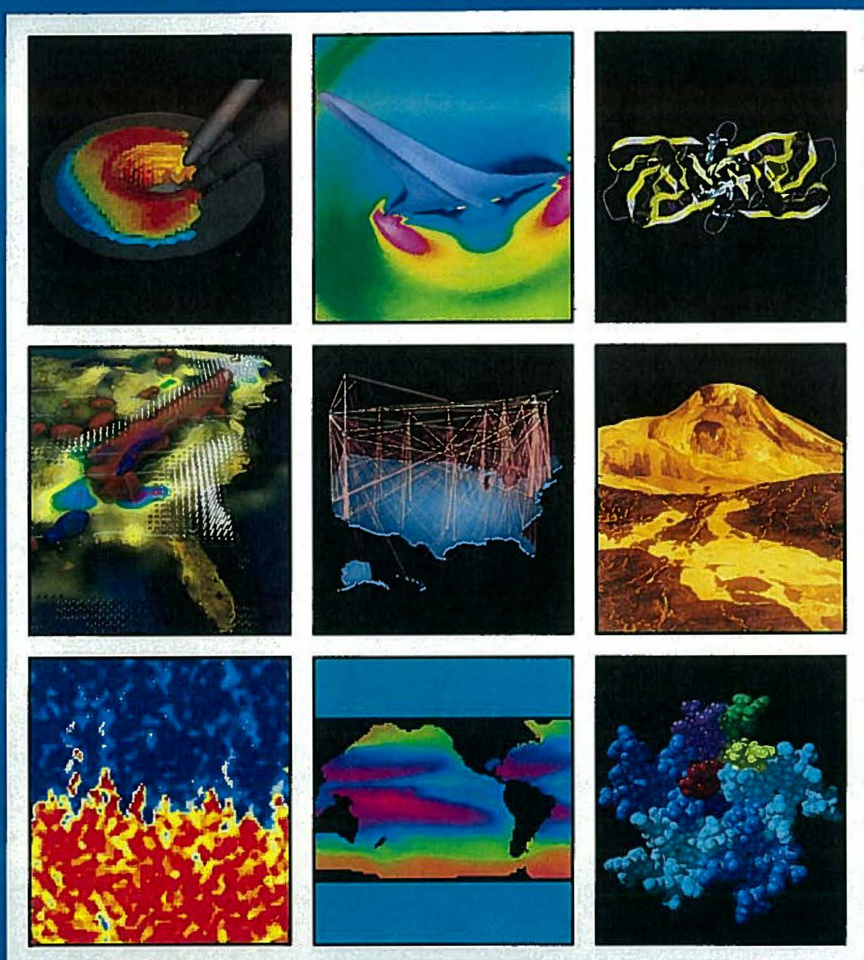
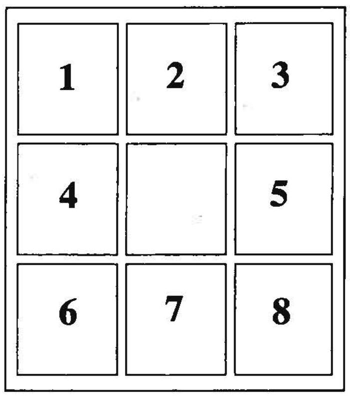
Grand Challenges 1993:
High Performance Computing and Communications

The FY 1993 U.S. Research and Development Program



A Report by the Committee on
Physical, Mathematical, and Engineering Sciences
Federal Coordinating Council for Science, Engineering,
and Technology

To Supplement the President's Fiscal Year 1993 Budget



On the Cover:

1. Fuel Combustion
2. High Speed Civil Transports
3. Rational Drug Design
4. Air Pollution
5. Venus Imaging
6. Magnetic Recording Technology
7. Ocean Modeling
8. Design of Protein Structures

The images used in this report were produced by ongoing scientific projects in areas of the HPCC Program. They were selected to illustrate the breadth of subject matter of the HPCC Program, and are elaborated upon in Chapter 4. The center figure is a visualization of the emerging National Research and Education Network, showing the interconnection of the network backbone with selected regional and campus-area networks.

Grand Challenges 1993:
High Performance Computing and Communications

The FY 1993 U.S. Research and Development Program

A Report by the Committee on
Physical, Mathematical, and Engineering Sciences

Federal Coordinating Council for Science,
Engineering, and Technology

Office of Science and Technology Policy

To Supplement the President's Fiscal Year 1993 Budget

Office of Science and Technology Policy
Federal Coordinating Council for Science, Engineering, and Technology
Committee on Physical, Mathematical, and Engineering Sciences

Chairman

Walter Massey, National Science Foundation

Members

William Happer, Department of Energy, Vice Chairman
Victor H. Reis, Department of Defense, Vice Chairman
Scott Sewell, Department of the Interior
Harry C. Mussman, Department of Agriculture
Robert White, Department of Commerce
Donald A. B. Lindberg, National Institutes of Health
John B. Childers, Department of Education
Norine Noonan, Office of Management and Budget
Karl Erb, Office of Science and Technology Policy
Richard H. Petersen, National Aeronautics and Space Administration
Erich Bretthauer, Environmental Protection Agency

Vernon Pankonin, National Science Foundation, Executive Secretary

FCCSET Directorate

Vickie V. Sutton, Assistant Director
Charles H. Dickens, Executive Secretary

High Performance Computing and Communications and Information Technology Subcommittee

David Nelson, Department of Energy, Chairman
Stephen L. Squires, Defense Advanced Research Projects Agency, Co-Chairman
Lee Holcomb, National Aeronautics and Space Administration, Co-Chairman
A. Nico Habermann, National Science Foundation, Co-Chairman
Norman H. Kreisman, Department of Energy, Executive Secretary
James Burrows, National Institute of Standards and Technology
Steven Isakowitz, Office of Management and Budget
Daniel R. Masys, National Institutes of Health
Fred Scoresby Long, National Oceanic and Atmospheric Administration
Joan H. Novak, Environmental Protection Agency
Kirk Winters, Department of Education
Eugene Wong, Office of Science and Technology Policy

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20506

MEMBERS OF CONGRESS:

I am pleased to forward with this letter "Grand Challenges 1993: High Performance Computing and Communications, The FY 1993 U.S. Research and Development Program" prepared by the Committee on Physical, Mathematical, and Engineering Sciences of the Federal Coordinating Council for Science, Engineering, and Technology (FCCSET), to supplement the President's Fiscal Year 1993 Budget. This report, the second in a series begun in 1991, describes the Presidential Initiative in High Performance Computing and Communications.

This Initiative is a well-coordinated, interagency research and development effort designed to sustain and extend U.S. leadership in all advanced areas of computing and networking. The Initiative, which represents over four years of planning, not only provides a far-sighted vision for investment in technology but also recognizes the importance of human resources and applications that serve major national needs. This sound national investment will bring both economic and social dividends, including advances in education, productivity, basic science, and technological innovation.

High performance computing and communications are pervasive tools for science and technology. Therefore, planning for this Initiative has included close cooperation with other Presidential Initiatives developed through the interagency FCCSET process, including Global Change Research, Advanced Materials and Processing, Biotechnology Research, and Mathematics and Science Education.

During the past year, substantial technical advances have been made in several areas of advanced computing, all of which strengthen the rationale for moving forward with this Initiative as rapidly as possible. The strong support for the President's Initiative by the Congress, academia, and leaders of the U.S. computer industry is greatly appreciated.

The coordination and integration of the interagency research and development strategy for this Initiative was led very ably by Dr. Walter E. Massey, Chairman of the Committee on Physical, Mathematical, and Engineering Sciences, and his interagency committee members, associates, and staff. They are all to be commended on the excellent work that is manifest in both the Initiative and the report.

D. Allan Bromley
Director

Table of Contents

List of Figures

Grand Challenges 1993:
High Performance Computing and Communications

The FY 1993 U.S. Research and Development Program

EXECUTIVE SUMMARY

High performance computing and computer communications networks are increasingly important to scientific advancement, economic strength, and national security. The goal of the Federal High Performance Computing and Communications (HPCC) Program is to accelerate significantly the availability and utilization of the next generation of high performance computers and networks in a manner consistent with the Strategic and Integrating Priorities shown in Figure 1.

The HPCC Program is the result of several years of effort on the part of senior government, industry, and academic scientists and managers to design a research agenda to extend U.S. leadership in high performance computing and networking technologies. The program is planned, funded, and managed with close cooperation among Federal agencies and laboratories, private industry, and academe to ensure that the fruits of this research program are brought into the educational and commercial marketplaces as rapidly as possible.

There have been steady accomplishments over the past year. Several participating agencies have begun to fund HPCC research groups, centers, and consortia on various grand challenge application problems. Major new scalable high performance systems have been announced and delivered. New software applications have been developed or ported to the emerging high performance systems. Traffic on the operational network has doubled in the past year, as has the number of interconnected local and regional networks. In the past year, a large number of researchers, scholars, students, educators, scientists, and engineers have been trained to use these emerging technologies.

Numerous organizations have recently undertaken studies that have provided valuable feedback on the structure and content of the program. Some of the organizations that supported studies include the National Academy of Sciences; EDUCOM and the Computing Research Association, each representing numerous universities; the Computer Systems Policy Project, representing leading U.S. computer companies; the Office of Technology Assessment; professional societies including Association for Computing Machinery, Institute of Electrical and Electronic Engineers, American Mathematical Society, and others; and the Federal Networking Advisory Committee.

For FY 1993 the HPCC Program proposes to invest $803 million in the four complementary and coordinated components shown in Figure 1. This investment represents a $148 million, or 23%, increase over the FY 1992 enacted level.

The HPCC Program is driven by the recognition that unprecedented computational power and its creative use are needed to investigate and understand a wide range of scientific and engineering "grand challenge" problems. These are fundamental problems whose solution requires significant increases in computational capability and is critical to national needs. Progress toward solution of these problems is essential to fulfilling many of the missions of the participating agencies. Examples of grand challenges addressed include: prediction of weather, climate, and global change; determination of molecular, atomic, and nuclear structure; understanding turbulence, pollution dispersion, and combustion systems; understanding the structure of biological macromolecules; improving research and education communications; understanding the nature of new materials; and solving problems applicable to national security needs.

Other Presidential Initiatives including Global Change Research, Advanced Materials and Processing, Biotechnology, and Mathematics and Science Education depend on the capabilities that this initiative will produce. In addition, the generic technologies developed will make possible advances in many other areas of direct benefit to millions of Americans.

The HPCC Program nurtures the educational process at all levels by improving academic research and teaching capabilities. Advanced computing and computer communications technologies will accelerate the research process in all disciplines and enable educators to integrate new knowledge and methodologies directly into course curricula. Students at all levels will be drawn into learning and participating in a wide variety of research experiences in all components of this program.

The FY 1993 Program and this document were developed by the High Performance Computing and Communications and Information Technology Subcommittee under the direction of the Committee on Physical, Mathematical, and Engineering Sciences (PMES) of the Federal Coordinating Council for Science, Engineering, and Technology (FCCSET).

FIGURE 1

The High Performance Computing and Communications Program

Goals: Strategic Priorities

- Extend U.S. technological leadership in high performance computing and computer communications.
- Provide wide dissemination and application of the technologies both to speed the pace of innovation and to serve the national economy, national security, education, and the global environment.
- Spur gains in U.S. productivity and industrial competitiveness by making high performance computing and networking technologies an integral part of the design and production process.

Strategy: Integrating Priorities

- Support solutions to important scientific and technical challenges through a vigorous R&D effort.
- Reduce the uncertainties to industry for R&D and use of this technology through increased cooperation among government, industry, and universities and by the continued use of government and government-funded facilities as a prototype user for early commercial HPCC products.
- Support the underlying research, network, and computational infrastructures on which U.S. high performance computing technology is based.
- Support the U.S. human resource base to meet the needs of industry, universities, and government.

Program Components

High Performance Computing Systems (HPCS)
- Research for Future Generations of Computing Systems
- System Design Tools
- Advanced Prototype Systems
- Evaluation of Early Systems

Advanced Software Technology and Algorithms (ASTA)
- Software Support for Grand Challenges
- Software Components and Tools
- Computational Techniques
- High Performance Computing Research Centers

National Research and Education Network (NREN)
- Interagency Interim NREN
- Gigabits Research and Development

Basic Research and Human Resources (BRHR)
- Basic Research
- Research Participation and Training
- Infrastructure
- Education, Training, and Curriculum

1. Program Goals and Overview

Introduction

High performance computing continues to gain recognition as a powerful technology for industrial design and manufacturing, scientific research, communications, and information management. A robust U.S. high performance computing and computer communications capability contributes to leadership in critical technology and national security areas. Improved computational and communications technologies contribute to more effective approaches to problem solving, new products and services, and enhanced national competitiveness across broad sectors of the economy. The High Performance Computing and Communications (HPCC) Program was initiated last year in the President's Fiscal Year 1992 Budget. It will accelerate the development of a thousand-fold improvement in useful computing capability and a hundred-fold improvement in available computer communications capability by 1996, and it will enhance the range of scientific and engineering disciplines that can effectively exploit this computational capability.

Based on several years of planning under the auspices of the Federal Coordinating Council for Science, Engineering, and Technology (FCCSET), Federal agencies in collaboration with scientists and managers from U.S. industry, universities, and laboratories have developed the HPCC Program to meet the challenges of advancing computing and associated communications technology and practices. Agencies have realigned and enhanced their high performance computing research and development programs, coordinated their activities with other agencies, and shared common resources to develop the program described in this document.

Planning for the HPCC Program has proceeded under the leadership of the President's Office of Science and Technology Policy. This program received added impetus by the passage of a multi-year authorization in the High Performance Computing Act of 1991 (P.L. 102-194), signed by President Bush on December 9, 1991. On that occasion, the President stated, "The development of high performance computing and communications technology offers the potential to transform radically the way in which all Americans will work, learn and communicate in the future. It holds the promise of changing society as much as the other great inventions of the 20th century, including the telephone, air travel and radio and TV."

There have been steady accomplishments over the past year. Several participating agencies have begun to fund application groups, centers, and consortia on various grand challenge problems. Major new scalable high performance systems have been announced and delivered. New software applications have been developed or ported to the emerging high performance systems. Traffic on the operational network has doubled in the past year, as has the number of interconnected local and regional networks. In addition, large numbers of researchers, scholars, students, educators, scientists, and engineers have been trained to use these emerging technologies.

Needs and Benefits

High performance computing has become a vital enabling force in the conduct of science and engineering research over the past three decades. Computational science and engineering has joined, and in some areas displaced, the traditional methods of theory and experiment. This trend has been powered by computing hardware and software, computational methodologies and algorithms, availability and access to high performance computing systems and infrastructure, and the growth of a trained pool of scientists and engineers. This process has been nurtured by Federal investment in advanced research, agency supercomputer centers, and national networks through DARPA, DOD, DOE, ED, EPA, NASA, HHS/NIH, DOC/NIST, DOC/NOAA, NSA, and NSF. These facilities have contributed to national mission areas such as energy, space, health, defense, environment, weather, and basic science and technology that could not be effectively addressed without the use of such advanced facilities.

High performance computing technology is knowledge and innovation intensive. Its development and use engages the entire scientific and engineering community. Building upon fundamental research of the early 1980's, a new computing technology of advanced parallel processing computers has emerged. This innovative approach to high performance computing systems promises to achieve sustained performance improvements of a thousand-fold compared to the systems at the start of the HPCC Program.

In a growing number of science and technology fields, progress and productivity in modern research are increasingly dependent on the close interaction of people located in distant places, sharing and accessing computational resources across networks. Although the U.S. is the world leader in most of the critical aspects of computing and computer communications technology, this lead is being challenged.

The HPCC Program is a strategic Federal investment in the frontiers of computing and computer communications technologies and is formulated to satisfy national needs from a variety of perspectives including: technology, science applications, human resources, and technology transition. Needs are derived from the agency missions and based on the underlying science, engineering, and technology base required to carry out these missions. Many of these mission needs are related to solving very intensive large scale computing problems. These fundamental problems often cut across agencies and missions and are called "grand challenge" problems.

The industrial and academic sectors provide major sources of innovation, cost-effective development, and support of information technologies and their application to grand challenge problems. As these technologies are developed, the results support the Federal agency missions and become available nationally. The program provides for development of these revolutionary technologies within a framework of a partnership among government, industry, and academe, which promotes the rapid transition of laboratory results into new products that can then be applied within the program.

During the past year substantial advice and recommendations have come from industrial, professional, and scientific organizations including the National Academy of Sciences; EDUCOM and the Computing Research Association, each representing numerous universities; the Computer Systems Policy Project, representing leading U.S. computer companies; the Office of Technology Assessment; professional societies including Association for Computing Machinery, Institute of Electrical and Electronic Engineers, American Mathematical Society, and others; and the Federal Networking Advisory Committee. This has helped focus planning and management activities and strengthen coordination with these other sectors.

Program Description

The Program consists of four integrated components representing the key areas of high performance computing and communications:

> *High Performance Computing Systems* (HPCS) - the development of the underlying technology required for scalable high performance computing systems capable of sustaining trillions of operations per second on large problems. Research in very high performance systems is focusing both on increasing the absolute level of performance attainable and on reducing the cost and size of these very high performance systems in order to make them accessible to a broader range of applications.
>
> *Advanced Software Technology and Algorithms* (ASTA) - the development of generic software technology and algorithms and the deployment of the most innovative systems for grand challenge research applications in a networked environment.
>
> *National Research and Education Network* (NREN) - the development of a national high speed network to provide distributed computing capability to research and educational institutions and to further advanced research on very high speed networks and applications.
>
> *Basic Research and Human Resources* (BRHR) - support for individual investigator and multidisciplinary long term research drawn from diverse disciplines, including computer science, computer engineering, and computational science and engineering; initiation of activities to significantly increase the pool of trained personnel; and support for efforts leading to accelerated technology transition.

Advances in high performance computing enable advances in almost every other science and engineering discipline. There is a complex web of research interdependencies among the four components, and each area contributes to progress in other areas. Because of these dependencies, achieving and maintaining balance between the research components is a primary goal and the most important priority in the current context and environment. The HPCC Program is designed to provide balanced support both for technology areas (including components, systems, software, and algorithms) and for applications, infrastructure, and human resources to achieve rapid overall research progress and productivity.

The component activities are planned to produce a succession of intermediate benefits on the way to meeting the long range programmatic goals. The HPCC Program builds on Federal programs already in place, providing additional resources in selected areas. Computational science and engineering grand challenges that emerge from all science and engineering disciplines are the focal points for these efforts.

Goals

The goals of the High Performance Computing and Communications Program are to:

- Extend U.S. technological leadership in high performance computing and computer communications.
- Provide wide dissemination and application of the technologies both to speed the pace of innovation and to serve the national economy, national security, education, and the global environment.
- Spur gains in U.S. productivity and industrial competitiveness by making high performance computing and networking technologies an integral part of the design and production process.

These goals will be realized by achieving: computational performance of one trillion operations per second (10^{12} ops, or teraops) on a wide range of important applications; development of associated system software, tools, and improved algorithms for a wide range of problems; a national research network capable of one billion bits per second (10^9 bits, or gigabits); sufficient production of Ph.D.s and other trained professionals per year in computational science and engineering to permit effective use and application of these new technologies.

Strategy

The goals will be met through coordinated government, industry, and university collaboration to:

- Support solutions to important scientific and technical challenges through a vigorous R&D effort.
- Reduce the uncertainties to industry for R&D and use of this technology through increased cooperation among government, industry, and universities and by the continued use of government and government-funded facilities as a prototype user for early commercial HPCC products.
- Support the underlying research, network, and computational infrastructures on which U.S. high performance computing technology is based.
- Support the U.S. human resource base to meet the needs of industry, universities, and government.

At the program component level the strategy will: exploit and extend scalable parallelism and engage in intensive system software development in HPCS; use common requirements of the grand challenges to foster software progress in ASTA; evolve from the current Internet and other Federally managed networks supporting research and education to the NREN using an evolutionary series of systems; and strengthen academic and human resources development activities in computer science and computational science and engineering as part of BRHR.

Program Management Approach

The High Performance Computing and Communications Program is implemented as a partnership among Federal agencies and other organizations. Major portions of the program are cost-shared and leveraged by the participation of industry and universities.

Leadership for the Program is provided by the Office of Science and Technology Policy, through the Federal Coordinating Council on Science, Engineering, and Technology (FCCSET) Committee on Physical, Mathematical, and Engineering Sciences (PMES). The membership of PMES includes senior executives of many Federal agencies. Planning for the HPCC Program is coordinated by the PMES High Performance Computing, Communications, and Information Technology (HPCCIT) Subcommittee. The HPCCIT, currently led by DOE, meets regularly to coordinate agency HPCC programs through information exchanges, the common development of interagency programs, and the review of individual agency plans and budgets.

Under the HPCCIT there are currently several task groups that coordinate activities in specific areas, including Networking, Research, Education, and Applications. From time to time, individual agencies are assigned responsibility to lead the coordination of the HPCCIT and the task groups. The Applications group, currently led by NASA, coordinates activities related to grand challenge applications, software tools needed for applications development, and software development at high performance computing centers. The Networking group, currently led by NSF, coordinates network integration activities and works closely with the Federal Networking Council (FNC), which consists of representatives from interested federal agencies. The FNC is responsible for coordinating the efforts of government HPCC participants and other NREN constituents, in addition to providing a liaison to others interested in the Federal program. The Research group, currently led by DARPA, addresses a broad range of issues including basic research progress, technology trends, and alternative approaches to deal with technology limits in information technology. The Education group, currently led by NIH, coordinates education and training aspects of HPCC and provides liaison with other education-related activities under the FCCSET.

In order to improve collaboration with U.S. industry, universities, state and local governments, and other non-Federal communities, a Federal Networking Advisory Committee has been created to support the FNC, and pursuant to P.L. 102-194, a

High Performance Computing Advisory Committee will be established to support the overall HPCC Program.

The participating agencies have altered their organizational or management structure in order to facilitate participation in the HPCC Program. Each agency has established a focal point for matters related to the HPCC Program. (See inside back cover for agency contact addresses.) DARPA has established a Computing Systems Technology Office which serves as the DOD focal point for HPCC. In addition, DARPA has created a High Performance Computing Joint Program Office to coordinate advanced technology development within the DOD and cooperatively with other agencies. NSF has established an HPCC Coordination Committee with budget, planning, and oversight responsibilities. The DOE has expanded its Scientific Computing Staff to respond to this initiative and has assigned to it overall responsibility for the DOE HPCC program component. NASA has responded by creating an HPCC Program Office within the Office of Aeronautics and Space Technology to coordinate, plan, and implement HPCC technology development.

Most of the participating agencies have published solicitations for research in HPCC areas. U.S. industry, academia, and other developers and users of HPCC technology are involved in program planning and execution through advisory committees, commissioned reviews, and self-generated commentary; as well as through direct participation as performers of R&D and as suppliers of technology.

The HPCC Program has benefitted from the interest, advice, and specific recommendations that have come from various industrial, professional, and scientific organizations. Some of these organizations include the National Academy of Sciences; EDUCOM and the Computer Research Association, each representing numerous universities; the Computer Systems Policy Project (CSPP), representing leading U.S. computer companies; the Office of Technology Assessment; professional societies including Association for Computing Machinery, Institute of Electrical and Electronics Engineers, American Mathematical Society, and others. As an example of the dynamism and flexibility of the HPCC program, many of these recommendations are already being incorporated and responded to in the current HPCC Program plans. For example, the CSPP has conducted an intensive study of the HPCC Program structure and published their views in December 1991. The CSPP has noted that the HPCC program is a significant, critical, and necessary undertaking by the Federal government, and has provided recommendations that have helped to focus discussion on four important issues below:

Expanding the vision of the HPCC Program to include research on generic, enabling technologies to support a wider range of applications such as better health care, lifelong learning, improved services for senior citizens, enhanced industrial design, and broad access to public and private databases. The HPCC program emphasizes investment in many of the supporting advanced technologies such as open scalable systems, portable software, mass data storage, and advanced network protocols. In addition, the HPCC program is making additional efforts to inform industry and academia to propose and enter collaborations on specific grand challenge applications.

Establishing a technology and policy foundation for an information and communications infrastructure for the future. To address this issue, the Federal Networking Council, an interagency coordination group that supports the development of the National Research and Education Network, has formed a policy committee to address critical network issues such as security and privacy, intellectual property rights, and network access, interoperability, and technology transfer.

Improving management and governance of the HPCC program and increasing opportunities for industry participation. The HPCC program will continue to strengthen its current coordinated management structure as the program matures and more activities shift from planning to execution phase. This should provide for validation of policies and plans, enable detailed monitoring of progress towards goals and objectives, better satisfy Congressional reporting requirements, and serve as a single point-of-contact for industry, academia, and other government agencies. Over the coming months, the Administration will be exploring alternative management approaches that might bring additional coordination and accountability to the program. To improve technology transfer, the Department of Energy is supporting the development of a model Cooperative Research and Development Agreement (CRADA) that will ease significantly the exchange of computing technology between private industry and government laboratories.

Formulating the HPCC budget priorities to achieve a balanced program. The budget proposes the largest dollar increase for ASTA, the program component which addresses grand challenge applications and software research. The program will continue to support basic research, generic technologies, and broad applications, but will work in the coming years to sustain an appropriate funding balance between software research and computer equipment and facilities.

The HPCC Program is executed according to management principles that have proven effective in other focused research programs in government and industry. Some of these principles include:

> *Sustain balance and critical mass.* The program must achieve sufficient scope and balance among the components. The HPCC Program must operate at a sufficient scale and coverage of technology areas that the new technologies can be effectively applied to grand challenge problems with acceptable levels of risk.
>
> *Build on agency strengths.* The strategy builds on agency strengths by giving appropriate agencies the responsibility to coordinate activities in areas of demonstrated capability. Joint programs are developed where agencies have complementary strengths in particular technology areas.
>
> *Accelerate technology transfer.* The transfer of technology from research to development and to application is a principal challenge addressed by the HPCC Program. Barriers to acceptance of new technologies can include high initial cost, inadequate and user-unfriendly software, and lack of standards. The HPCC Program addresses barriers such as these by relying on substantial industry participation, by sharing development among agencies, and by involving users early in the process. Risk is minimized by adopting a

mixed strategy, in which multiple approaches are pursued to technology issues such as high performance architecture, microelectronics, and system software.

Seek broadly applicable technology. The program emphasizes general-purpose solutions with wide applicability, rather than promoting the development of special-purpose systems tailored to specific problem domains. Wide diffusion of new technologies into the national technology base will result in cost-effective support for the broadest range of applications on a diverse base of computing resources, facilitating the application of high performance technologies to major national problems such as health and education.

Overcome barriers. The program emphasizes scalable, interoperable technologies in order to overcome costs in creating successive generations of high performance computers and software, thereby facilitating evolution of applications to increasingly powerful and flexible computing, networking, and software technologies.

In the coming year, the HPCC Program will continue to strengthen its management approach for improving planning, assessment, budgeting, and reporting.

2. High Performance Computing and Communications Program Components

The HPCC Program is composed of four integrated and coordinated components that are designed to enhance scientific and engineering productivity and support long term agency needs. The emerging computing environment is that of advanced workstations with high resolution displays connected to a high speed computer communications network and high performance computing resources. The regional and national networks provide access to additional high performance computing systems and research resources. Realizing the full potential of teraops computing and gigabits networking systems will require advanced software technology and algorithms and people educated and trained to use these tools and resources in this dynamic field.

High Performance Computing Systems (HPCS)

The HPCS component produces scalable parallel computing systems through the development of prototype systems. The program is designed to attack computational science problems by developing innovative systems that will provide a one-hundred- to one-thousand-fold increase of sustained computing capability over machines that were in conventional use at the start of the program.

The program is structured to focus on technological challenges in the earliest stages of the product development cycle. Critical underlying technologies are developed in prototype form along with associated design tools. This allows empirical evaluation of alternative solutions as the prototype systems mature. Evaluation is performed throughout the development cycle, with experimental results being fed back into the design process to refine successive generations of systems. There is risk inherent in creating new technologies, and each project will be managed according to its proximity to commercial introduction. For instance, larger projects which are close to yielding commercial products are performed on a cost-shared basis with industry.

HPCS is composed of the four subcomponents described below:

Research for Future Generations of Computing Systems. This activity focuses on the underlying architecture, components, packaging, systems software, and scaling concepts to achieve affordable high performance systems. These projects ensure that the required advanced technologies will be available for the new systems while providing a foundation for the more powerful systems that need to follow. This activity also produces system software to support heterogeneous configurations of workstations and high performance servers.

System Design Tools. This activity develops computer aided design tools and frameworks for enabling multiple tools to work together to enable design, analysis, simulation, and testing of systems components. The tools will enable rapid prototyping of new system concepts.

Advanced Prototype Systems. Systems capable of scaling to 100 gigaops performance have already been produced. The teraops level will be reached by the mid-1990's. Research in very high performance systems is focusing both on increasing the absolute level of performance attainable and on reducing the cost and size of these very high performance systems in order to make them accessible to a broader range of applications.

Evaluation of Early Systems. Experimental systems will be placed at sites with high levels of expertise to provide feedback to systems architects and software designers. Performance evaluation criteria for systems and results of evaluation will be made widely available. Because of scalability, early systems can be acquired in smaller configurations to evaluate their potential performance. As noted below, the ASTA component will support, on a cost shared basis, the acquisition of large scale systems for experimental use in grand challenge applications.

Accomplishments:

- High performance computing systems capable of scaling to 100 gigaops have been announced. Systems in this performance class will eventually be available in multiple architectures, many of which represent hybridizations of computational models used in earlier systems.
- New models of computation that may yield higher performance for many classes of problems, including more advanced hybrid models, are being explored through joint university-industry projects.
- Scalable mass storage systems have been designed, prototyped, and demonstrated in collaboration with industry. The scalable mass storage approach is now being used in most new high performance computing systems. Very large capacity (10^{15} bits) archival storage systems have also been developed.
- Operating system technology has been developed that is capable of efficiently supporting parallel and heterogeneous distributed applications involving configurations of high performance systems and workstations.

Advanced Software Technology and Algorithms (ASTA)

The ASTA component of the HPCC Program produces generic software technology, new algorithms, and early prototype applications software to support a broad range of high performance applications of heterogeneous and networked high performance computing systems. Significant improvements in algorithm design and software technology are essential to achieving sustained high levels of computing system performance. In addition, improvements in algorithms and software will open up major new areas of application of high performance computing technology in science, engineering, health care, education, national security, and other areas. Parallel and distributed systems provide new opportunities and challenges for software and algorithms. Advances in generic software are needed to minimize the time researchers in other scientific disciplines spend to master advanced, complex computing science skills.

A typical user environment can include various kinds of advanced workstations with high resolution color displays connected to a high performance network with diverse local high performance computing resources. Access to additional high performance computing systems is available through regional and national networks. Systems software will support these heterogeneous hardware and network configurations, with software support for library, user interaction, and other functions common across applications.

Support for Grand Challenges. Grand challenge problems currently addressed by the HPCC Program are of such a magnitude and complexity that they will require continuous advances in computational power and improvements in computational models for the next decade or more. Grand challenge problems will be selected for their national importance, for their potential to challenge computational technologists to advance their solution, for the extent of cost sharing with sources directly concerned with the specific applications, and for the leverage they can exert on other related technology developments. Grand challenge problems that will take advantage of the capabilities developed by the HPCC Program are contained in many other Federal initiatives, including Global Change Research, Advanced Materials and Processing, Biotechnology, and Mathematics and Science Education. Support will be provided to improve algorithms for hybrid computational models, multidisciplinary applications, and heterogeneous and distributed problems.

Software Components and Tools. Users of high performance systems have common needs in software technology and programming environments, including software tools, systems software, and support for reuse and for software libraries. Software tools include, for example, higher level languages, advanced compiler technology, optimization and parallelization tools, interoperability support, data management, visualization, debugging and analysis, instrumentation and performance measurement, and other tools. The availability of supported applications software is generally less for emerging new architectures than it is for established ones. Increasing support for software components and tools, with an emphasis on generic multi-architecture capability, will help accelerate software product development. (See Figure 2.)

Because of this very wide range of software tools required to develop and manage high performance applications, tool development efforts in HPCC will emphasize the integration of multiple tools developed by different organizations for use on multiple hardware platforms. Integration technology increases opportunity for commercial participation in tool development. Systems software efforts will focus on high performance support for heterogeneous configurations of systems in order to maintain flexibility and growth potential in user environments.

Finally, software library efforts will focus on development of common architectures and interfaces to increase the potential for reusability despite multiple underlying models of computation, the diversity of programming languages in use, and the varying degree of assurance provided with software components. Library efforts will lead to a software interchange capability that can support exchange of software components among researchers, enabling collaborative development of systems through appropriate sharing of software assets.

Computational Techniques. Scalable and portable software libraries are being developed to enable software to move from one generation to the next and across different computational platforms. A variety of experimental applications are being developed to evaluate the performance and generality of the new computing technologies. Research will include support for new parallel algorithms, numerical and mathematical analyses, parallel languages, and conceptual models of scientific applications. To support visualization of data and processes, standard systems level tools will be developed. Data interpretation and communications will be vigorously addressed to enable real-time visualization in heterogeneous workstation/server configurations.

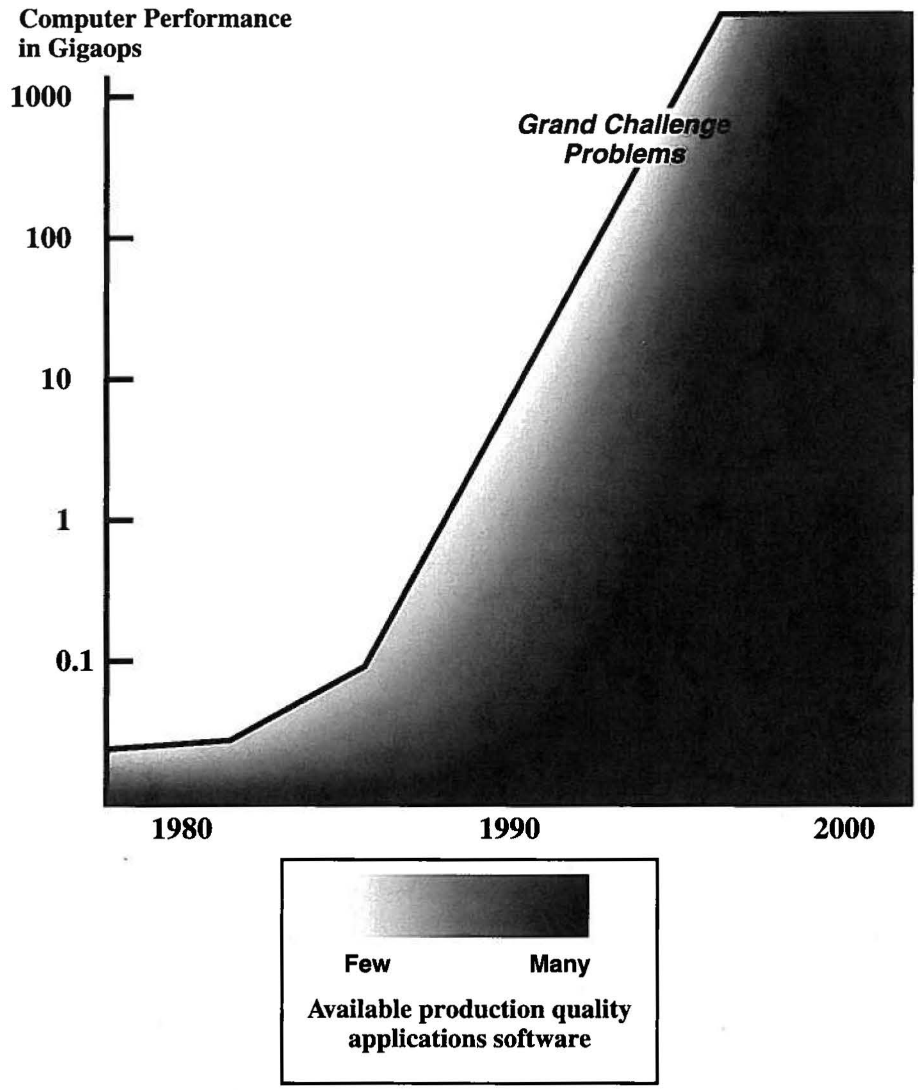
High Performance Computing Research Centers. The HPCC Program supports the deployment of innovative high performance computing architectures to high performance computing research centers for use by computational scientists working on grand challenge and disciplinary applications. High performance computing research centers generally acquire early high performance computing systems which are integrated into prototype heterogeneous computing configurations and connected to the developing NREN. This allows for a wide spectrum of experiments for scalability studies, as well as an opportunity to provide grand challenge researchers access to the largest possible advanced systems. As part of their efforts, the high performance computing research centers also support researchers in areas such as algorithms development, software environment and tool development, mass storage, and scientific visualization.

Accomplishments:

- Collaborations between agencies and industry have been initiated on grand challenge problems in quantum chemistry, environmental remediation, groundwater simulation, materials science problems, air pollution models, design of catalysts for chemical processes, and other areas.
- Consortia, centers, and groups are being established involving industrial, academic, and government participants to address applications software development, generic systems software, and effective sharing of leading-edge computational resources.
- Progress in algorithms and applications with direct applications to improved health care has been significant, ranging from the design of pharmaceuticals to enhancing understanding of fundamental biological processes. Already, order of magnitude speedups have been achieved for codes such as three-dimensional reconstruction of two-dimensional electron micrographs of such viruses as the herpes simplex virus—the cause of cold sores or fever blisters. Other healthcare related applications include prosthesis design.
- A parallel high performance computing system has been used successfully to develop a method for enhancing the data from the Hubble Space Telescope. This method, called the Maximum Entropy Method, increases resolution by a factor of four over the degraded data.

FIGURE 2

Availability of Applications Software for High Performance Computing Systems



National Research and Education Network (NREN)

The NREN is both a goal of the HPCC Program and a key enabling technology for success in the other components. As used in this report, the NREN is the future realization of an interconnected gigabit computer network system supporting HPCC. The NREN is intended to revolutionize the ability of U.S. researchers and educators to carry out collaborative research and education activities, regardless of the physical location of the participants or the computational resources to be used. (See Figure 3.) As its name implies, NREN is a network for research and education, not general purpose communication. Nonetheless, its use as a testbed for new communications technologies is vital. A fundamental goal of the HPCC is to develop and transfer advanced computing and communications technologies to the private sector of the U.S. as rapidly as possible, and to enhance the nation's research and education enterprise. The development and deployment of advanced applications, such as image visualization and distributed computing, will be applied to problems such as medical diagnosis, aerodynamics, advanced materials, and global change, and will provide the impetus necessary for transfering the supporting technologies and capabilities throughout the U.S. science, technology, and education infrastructure. These capabilities and technologies will be developed through the cooperative effort of U.S. industry, the Federal government, and the educational community.

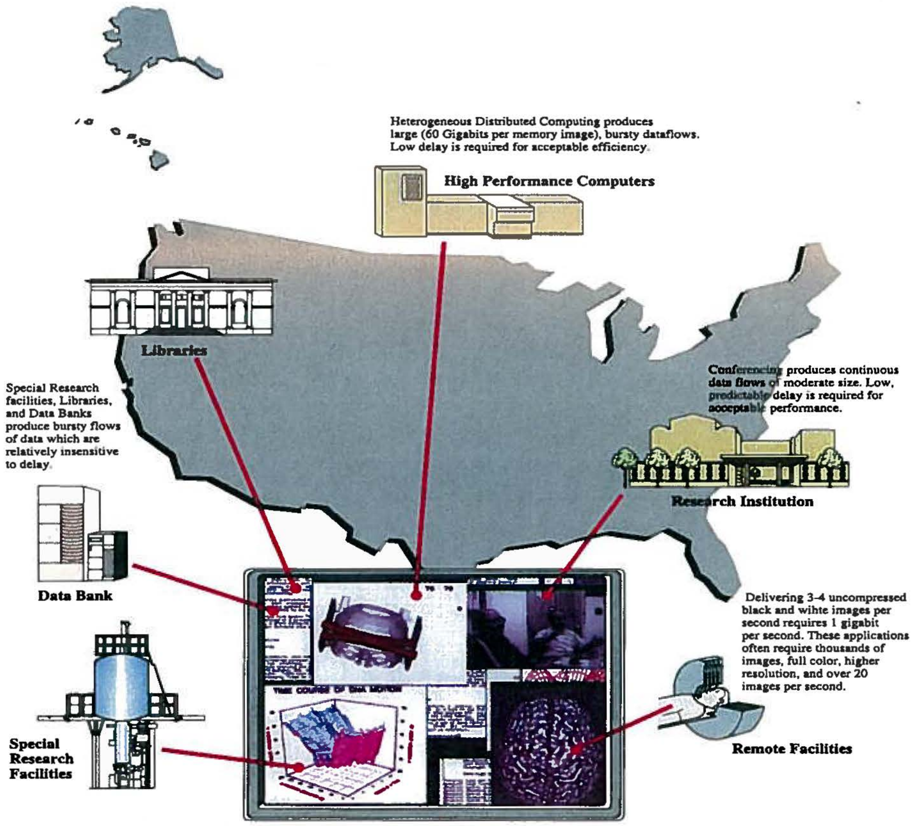
In order to ensure a coordinated Federal program in this area, a Federal Networking Council (FNC) has been established consisting of agency representatives. The FNC has the primary responsibility for coordinating the efforts of government HPCC participants and other NREN constituents, in addition to providing a liaison with others involved or interested in the program. To further collaboration with non-Federal communities such as state and local government, U.S. industry, and private education, the FNC has established the Federal Network Advisory Committee, consisting of distinguished members from those communities.

The NREN component of the HPCC Program is comprised of two interrelated and complementary subprograms, the Interagency Interim NREN subcomponent, and the Gigabit Research and Development subcomponent.

Interagency Interim NREN. The Interagency Interim NREN is an evolving operating network system. Near term (1992-1996) research and development activities will provide for the smooth evolution of this networking infrastructure into the future gigabit NREN. Interagency Interim NREN activities will achieve this goal by expanding the connectivity and enhancing the capabilities of the Federally funded portion of today's research and education networks, and by deploying advanced technologies and services as they mature. The Interagency Interim NREN, which is based on DARPA's internet technology, builds on NSF's NSFNET, DOE's Energy Sciences Network, NASA's Science Internet and other networks supporting research and education. (See Figure 4.) During 1992, the Interagency Interim NREN will accelerate the introduction of commercial 45 megabit transmission technologies and services into operational use, including the Switched Multi-Megabit Data Services (SMDS). These advances will support

FIGURE 3

Network Applications



NREN supported activities will promote enhancements in networking infrastructure both for the connectivity to many geographically distributed science facilities, information resources, and research teams and also for the technologies and tools to perform the distributed computing, data access, and multi-media processes as illustrated above.

both HPCC and non-HPCC related research and education, extend our understanding of the uses of moderately high capability networks and promote the evolution of Interagency Interim NREN performance toward gigabit speeds.

Additional research and development activities include the following:

- Assist upgrades of regional and community-of-interest networks, where appropriate, especially where these upgrades enhance end-to-end reliability.
- Produce improved user-level tools to enable scientists and educators to take advantage of the network's capabilities.
- Refine the understanding of the requirements for high capability networks: bandwidth, latency, predictability, and stability.
- Improve the technologies necessary for policy control, resource allocation, accounting, security, and routing coordination. Integrate and interconnect commercial services and other relevant networks where appropriate.
- Provide for a network information service that acts as a primary source of information on access to and use of the network.
- Enhance the current interconnected multi-agency architecture to provide for the interoperability of Federal and non-Federal networks, to the extent appropriate, in a way that allows for the autonomy of each network component.

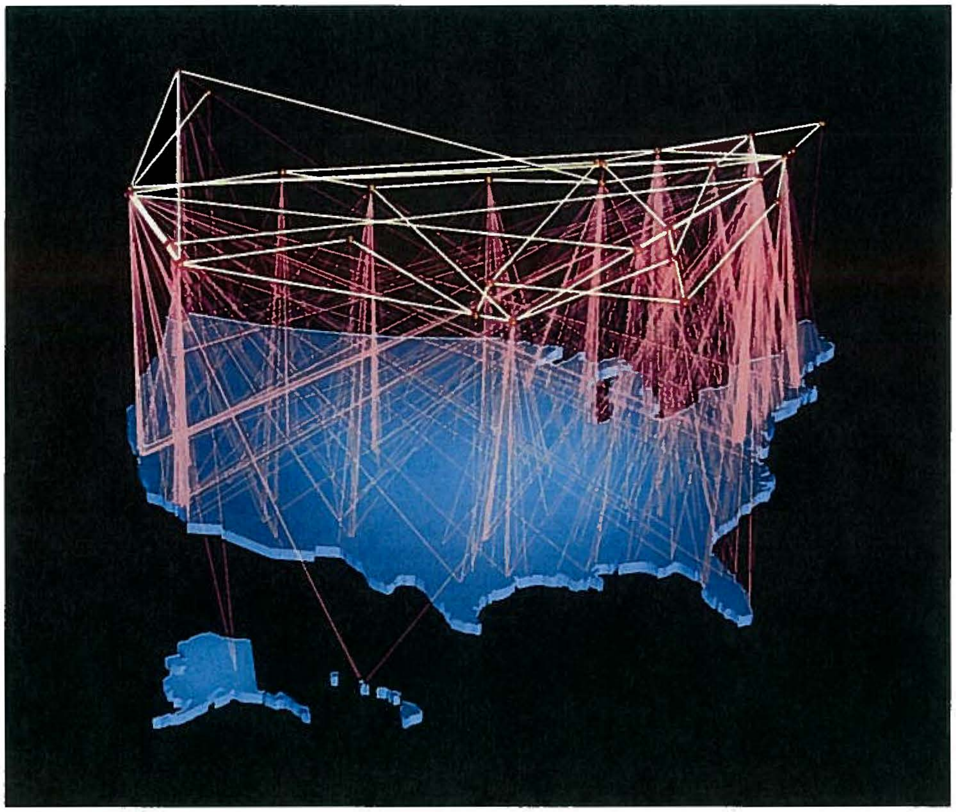
Gigabit Research and Development. The Gigabit Research and Development subcomponent is aimed at providing the technology base needed to achieve, at a minimum, gigabit speeds and advanced capabilities in the NREN. It will include efforts to develop technology to implement, operate, and effectively use a very high capability communications infrastructure.

Specific research and development activities include the following:

- Develop technology for advanced high throughput, low latency switches.
- Develop advanced computer-network interface technology to enable end users to access the capabilities of the network.
- Develop methods for control of congestion and network flow for high capability networks where the transit time across the network has a significant impact on the problem.
- Support gigabit testbed networks to provide early evaluation of research concepts and advanced data communications equipment in the context of applications. These testbeds include collaborations of government laboratories, university researchers, and industrial partners.
- Incorporate the advanced communications technology now under development in coordination with industry, in testbeds and gigabit research efforts. Transfer these advanced technologies to the Interagency Interim NREN as they mature and become stable.

FIGURE 4

Network Connectivity



The image represents the interconnected "backbone" networks of NSF, NASA, and DOE, together with selected client regional and campus-area networks. Nodes of the backbones are represented as connected spheres on a plane above the outline of the United States; the client networks are represented as dendritic lines from the backbone nodes to the geographic locations where the client networks attach.

This program subcomponent will include industrial partners, thereby assisting them to develop prototypes for a future high capability commercial communications infrastructure and reduce the risks for U.S. communications vendors in several important ways: defining and developing the market for the high capability network services; implementing, operating, and managing leading edge networks; promoting the prompt and timely adoption of standards for advanced communication protocols; and providing commercial communications vendors with successful examples of high capability networks which can form a basis for economic analyses of their future offerings.

Accomplishments:

- Packet traffic on the major Interagency Interim NREN backbone has doubled in the past year, as has the number of attached networks.
- Interagency collaboration has been initiated on gigabit networking research, and several testbeds have been established, each focusing on specific applications for high capacity networks.
- Multi-protocol routers are now in preliminary use to support interoperation and integration of diverse networks.

Basic Research and Human Resources (BRHR)

This component addresses long term national needs related both to the development of high performance computing and communications technology and to its effective application. The BRHR component includes activities designed to increase the base of skilled personnel, to enhance the basic research foundation for high performance computing and communications, and to more effectively apply high performance computing and communications technologies to education and training.

Basic Research. These activities support increased participation of investigators in innovative basic and multidisciplinary research in computer science, computer engineering, mathematics, and computational science and engineering. The strategy is to increase the number of basic research and multidisciplinary awards across all disciplines for which computational methods are critical to achieving advances or scientific breakthroughs. Example program activities in basic research include: research on scientific algorithm development; research on scalable numeric and symbolic parallel algorithms for common computational problems; performance prediction techniques for concurrent systems; fault tolerant strategies for parallel and distributed systems; and development of optimization techniques that support heterogeneous software configurations involving multiple programming languages. Basic research may relate to any of the other three HPCC components (HPCS, ASTA, NREN).

Research Participation and Training. These activities address the human resource needs in the computer and computational sciences at the postdoctoral, graduate, and undergraduate levels. In addition, senior investigators and practitioners are trained in the use of advanced computing and network communications to aid in their work. Program activities include: workshops and seminars; postdoctoral fellowships in computational science and engineering; institutional training and postdoctoral programs; internships; exchange programs for knowledge transfer through national laboratories, centers, universities, and industry; and software dissemination through centers, national databases and libraries.

Infrastructure. These activities will improve university facilities for computer science, computer engineering, and computational science and engineering research related to high performance computing and its underlying basic research disciplines. Program activities include improvement of equipment in computer science, computer engineering, and computational science and engineering academic departments, centers, groups, and institutions; development and expansion of scientific databases; and distribution of integrated system-building kits and toolsets.

Education, Training and Curriculum. These activities will expand and initiate activities to improve undergraduate and precollege education and training opportunities in high performance computing and computational science and engineering, for both students and educators. The introduction of associated curriculum and training materials at all levels is an integral part of this effort. Program activities include training scientists at national centers; supercomputer workshops for established investigators to gain "hands on" expertise and tools for their research problems; summer internship programs for promising undergraduate and high school students; and integrating the experience of professional scientists and engineers in the development of curriculum and instruction materials for high school students and teachers.

Accomplishments:

- Initiated Computational Science Graduate Fellowships with 22 recipients in the first year to schools in 14 states.
- Significantly increased the number of awards for computational science and engineering postdoctoral research associateships.
- Provided multiagency support for the supercomputing applications contest programs for teams of high school students, such as SuperQuest.
- Augmented University Centers of Excellence and Research Institutes to pursue HPCC basic research. A Historically Black Colleges and Universities program has been established.
- Increased the Graduate Student Researchers Program to support fellowships in high performance computing and computation at major centers.

3. Program Development and Agency Budgets

Program Planning

Leadership for the HPCC Program is provided by the Office of Science and Technology Policy (OSTP), through the Federal Coordinating Council for Science, Engineering, and Technology (FCCSET) Committee on Physical, Mathematical, and Engineering Sciences (PMES). The membership of PMES includes the senior executives of many Federal agencies. The High Performance Computing, Communications, and Information Technology Subcommittee (HPCCIT) of PMES, the successor to the Executive Committee of the High Performance Computing and Communications Working Group, was chartered in March 1991 by PMES and FCCSET. The basic duties of the Subcommittee are to establish task groups, prepare budgets, and coordinate program planning and execution. The membership of the HPCCIT Subcommittee includes a representative from each member agency. One new member agency, Department of Education, was added for a total of nine. Currently there are four task groups that address Networking, Research, Education, and Applications. In addition, because of the close coordination required in the implementation of the NREN, a Federal Networking Council consisting of agency representatives has been established.

Evaluation Criteria

Each participating agency's HPCC contribution was reviewed against formal evaluation criteria during the planning and budget process. A review of participating agencies was performed using these evaluation criteria to develop the FY 1993 agency requests for the Program. The evaluation criteria are shown in Figure 5.

FIGURE 5

Evaluation Criteria for the Federal High Performance Computing and Communications Program

Relevance/Contribution. The research must significantly contribute to the overall goals and strategy of the Federal High Performance Computing and Communications (HPCC) Program, including computing, software, networking, and basic research, to enable solution of the grand challenges.

Technical/Scientific Merit. The proposed agency program must be technically/scientifically sound and of high quality, and must be the product of a documented technical/scientific planning and review process.

Readiness. A clear agency planning process must be evident, and the organization must have demonstrated capability to carry out the program.

Timeliness. The proposed work must be technically/scientifically timely for one or more of the HPCC program components.

Linkages. The responsible organization must have established policies, programs, and activities promoting effective technical and scientific connections among government, industry, and academic sectors.

Costs. The identified resources must be adequate, represent an appropriate share of the total available HPCC resources (e.g., a balance among program components), promote prospects for joint funding, and address long-term resource implications.

Enhancements to Existing Program Research. Existing agency HPCC programs will receive adequate support before new initiatives are funded.

Agency Approval. The proposed program or activity must have policy-level approval by the submitting agency.

Agency Budgets

For FY 1993, $803 million is being proposed, a $148 million or 23% increase over the FY 1992 enacted level. The budget is shown in Figure 6.

The funds proposed from Federal sources are not intended to carry out the entire HPCC Program. Portions of this Program will be cost shared by organizations from the participating sectors. Cost sharing will occur with various U.S. industrial and university partners to a large extent in the HPCS and the NREN components, and to a lesser but increasing extent, in ASTA. Cost sharing will occur in the ASTA component with agency programs and other computational applications programs, for example, in specific grand challenge areas, via multidisciplinary collaborations. This component also includes deployment of high performance systems to HPCC research centers. The close coupling of support in the Program will result in significant leverage, accelerate technology transfer, stimulate U.S. industry and markets, and enable the solution of computationally intensive applications. In addition, although the HPCC Program is not intended to include national security applications, the technology produced will have an important impact in these national security areas.

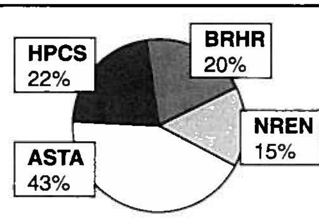
FIGURE 6

High Performance Computing and Communications Budgets by Agency and Program Component

(Dollars in millions)

Agency	FY 1992	FY 1993	FY 1993 HPCC Component			
			HPCS	ASTA	NREN	BRHR
DARPA	**232.2**	**275.0**	119.5	49.7	43.6	62.2
NSF	**200.9**	**261.9**	28.6	125.6	45.1	62.6
DOE	**92.3**	**109.1**	10.9	69.2	14.0	15.0
NASA	**71.2**	**89.1**	14.1	61.4	9.8	3.8
HHS/NIH	**41.3**	**44.9**	4.2	22.6	7.2	10.9
DOC/NOAA	**9.8**	**10.8**	0.0	10.4	0.4	0.0
EPA	**5.0**	**8.0**	0.0	6.1	0.4	1.5
DOC/NIST	**2.1**	**4.1**	1.1	1.0	2.0	0.0
Total	**654.8**	**802.9**	**178.4**	**346.0**	**122.5**	**156.0**

FY 1993 Component Funding Comparison



HPCC — High Performance Computing and Communications

Components of HPCC:
- **HPCS** — High Performance Computing Systems
- **ASTA** — Advanced Software Technology and Algorithms
- **NREN** — National Research and Education Network
- **BRHR** — Basic Research and Human Resources

Agency Program Descriptions

The agency responsibilities under the HPCC Program are outlined in Figure 7. Several agencies have been assigned a coordinating responsibility in specific technical areas:

- DOE and NASA will coordinate activities in high performance computing system evaluation, testbed development, and applications software capabilities.
- NASA will coordinate the accumulation and access to the high performance computing software base. This will be facilitated by a wide area file system technology currently being deployed for early experimental use by DARPA, which will be extended to include the NREN as it matures and is deployed by NSF.
- DARPA will coordinate activities in the development of high performance computing systems and associated operating systems software. NASA and DARPA will coordinate the development of associated software tools. DARPA will also coordinate activities in gigabit network technology research.
- NSF will coordinate the broad deployment of the Interagency Interim NREN working with universities, industry, and agencies having mission specific requirements. NSF will also coordinate basic research and infrastructure activities to foster balance in research participation and training.
- HHS/NIH will coordinate educational activities.
- DOC/NIST will coordinate standards, network security, and systems instrumentation and methodology for performance measurement.

Each agency participates in all of the identified activities to ensure that the resulting capabilities are a good match to user needs. DOC/NOAA, ED, HHS/NIH, and EPA bring distinctive applications areas of broad interest and network user bases.

FIGURE 7

HPCC Program: Agency Responsibilities: HPCS, ASTA

Agency \ Activity	High Performance Computing Systems	Advanced Software Technology & Algorithms
DARPA	• Technology development and coordination for affordable teraops systems	• Technology development for parallel algorithms and software tools
NSF	• Basic architecture research • Prototyping experimental systems	• Research in: - Software tools, databases - Grand challenges • Computer access
DOE	• Technology development • Systems evaluation	• Energy applications research centers • Energy grand challenge and computation research • Software tools
NASA	• Aeronautics and space application testbeds	• Software coordination • Computational research in: - Aerosciences - Earth and space sciences
HHS/NIH	• System evaluation and performance analysis	• Medical application testbeds for medical computation research
DOC/NOAA		• Ocean and atmospheric computation research • Software tools • Computational techniques
EPA		• Research in environmental computations, databases, and application testbeds
DOC/NIST	• Research in systems instrumentation and performance measurement • Research in interfaces and standards	• Research in: - Software indexing and exchange - Scalable parallel algorithms

FIGURE 7 (continued)

HPCC Program: Agency Responsibilities: NREN, BRHR

National Research and Education Network	Basic Research and Human Resources	Activity / Agency
• Technology development and coordination for gigabits networks	• University programs	**DARPA**
• Facilities coordination and deployment • Gigabits research	• Programs in: - Basic research - Education/training/curricula - Infrastructure	**NSF**
• Gigabits applications research • Access to energy research facilities and databases	• Basic research and education programs	**DOE**
• Access to aeronautic and spaceflight research centers	• Research institutes and university block grants	**NASA**
• Development of intelligent gateways • Access for academic medical centers	• Basic Research • Internships for parallel algorithm development • Training and career development	**HHS/NIH**
• Ocean and atmospheric mission facilities • Access to environmental databases		**DOC/NOAA**
• Environmental mission assimilation by the States	• Technology transfer to States • University programs	**EPA**
• Coordinate performance assesment and standards • Programs in protocols and security		**DOC/NIST**

Defense Advanced Research Projects Agency (DARPA)

The DARPA HPCC Program is balanced across the four HPCC Program components and is designed to stimulate the aggressive advance of the high performance computing and communications technology base. These technologies include scalable technologies for high performance computing and networking along with system software to support the effective application of the high performance technologies to large scale problems.

DARPA is the lead DOD agency for advanced technology research and has the leadership responsibility for HPCC within DOD. Complementing the DARPA HPCC program is the DARPA Strategic Computing Program, which addresses defense specific needs such as embedded systems, accelerators for specific problems domains, and grand challenge problems related to defense. DARPA executes its portion of the HPCC program in close cooperation with the Office of Naval Research, the Air Force Office of Scientific Research, the Army Research Office, Service Laboratories, the National Security Agency, and other DOD organizations and Federal agencies.

High Performance Computing Systems efforts are organized into four areas: research for future generations of computing systems, system design tools, advanced prototype systems, and evaluation of early systems. Technologies are being developed to support multiple models of computation, high performance interconnection, scalable mass storage, and supporting system software. Systems capable of 100 gigaops for large problems have already been announced and several of these systems are expected to be operational during FY 1993. The research challenge is not just to develop architecture and systems capable of scaling to teraops performance and beyond, but to do this in such a way that very large scale systems will become more affordable over time, thus increasing the range of applications that can effectively exploit the high performance technologies. DARPA has also developed and demonstrated system software approaches that support open heterogeneous workstation/server approaches to high performance computing.

In *Advanced Software Technology and Algorithms,* DARPA is producing scalable libraries for Defense-related problem domains and programming and analysis tools for scalable parallel and distributed heterogeneous systems.

DARPA efforts in the *National Research and Education Network* focus on the development of high performance networking technologies to satisfy gigabit technology needs for NREN and for Defense. Several gigabit networking testbeds have already been established to explore issues related to switch design, protocols, and interoperability.

The *Basic Research and Human Resources* component focuses on the basic scientific issues underlying the other three areas, in cooperation with other basic research programs. The DARPA program, along with NSF, provides the major Federal support to universities in computer science. The program addresses the fundamental scientific challenges underlying HPCC technology, including, for example, very high performance systems, flexible and reliable software, and high performance algorithms.

National Science Foundation (NSF)

The NSF HPCC program supports activities in all four components through individual investigator, group, center, and infrastructure awards. The activities are supported and coordinated through existing program structures to provide close cooperation with ongoing computational science and engineering programs throughout the NSF disciplinary directorates. The NSF HPCC program supports technology transfer by developing HPCC technologies through cooperative projects involving industries and academic research institutions.

The NSF *High Performance Computing Systems* component activities focus on support of basic research on heterogeneous and distributed high performance computing systems, coupled with prototyping and evaluation of promising new systems. NSF will emphasize research on new architectures, novel computing systems optimized for specific research applications, and new systems level tools for automated design and evaluation.

In *Advanced Software Technology and Algorithms,* NSF emphasizes developing the elements of a scientifically productive computational research environment by supporting research on generic software and programming environments, programming and applications tools for disciplinary scientists, computational techniques, and new conceptual models for scientific computing. Support will be provided for multidisciplinary grand challenge applications groups centered at academic institutions, in cooperation with other agencies and industry. NSF high performance computing research centers will: provide researchers access to a variety of high performance computing systems; be the focus for technology transfer and training activities; and provide a rich environment for coordinated approaches to grand challenge problems.

The NSF *National Research and Education Network* component activities include: coordinating the upgrading of the Interagency Interim NREN backbone network; assisting regional networks to upgrade facilities, capacity, and bandwidth; interconnecting the backbone networks of other agencies; and developing safeguards to enhance security of access to the connected computer systems. NSF supports the coordinated research and development effort that will culminate in initial deployment of a gigabit networking capability, including work on very high speed switches. NSF educational activities will expand grants for high school educational activities, upgrade operational services for educational users, and extend information access facilities and tools.

The NSF *Basic Research and Human Resources* activities support innovative basic research in scientific algorithm development for highly parallel computers; prediction techniques for concurrent systems; and heterogeneous computing. Research participation and training activities will address HPCC human resources at the postdoctoral, graduate, and undergraduate levels. Infrastructure activities will provide equipment and develop general use scientific databases, system-building kits, and toolsets. Education, training, and curriculum activities will be expanded to improve undergraduate and precollege education and opportunities will be provided for students and educators to participate in ongoing computational science and engineering research activities, such as the NSF-led SuperQuest.

Department of Energy (DOE)

In order to sustain a balanced program the DOE has taken steps to ensure that it makes contributions to all four components of the HPCC Program.

The DOE *High Performance Computing Systems* component emphasizes the evaluation and characterization of the applicability of advanced computers to large scientific and engineering problems. To perform this evaluation, DOE through its laboratories and other contractors and grantees is integrating the latest computers, mass storage devices, and other tools to further the knowledge base in computational science.

The DOE *Advanced Software Technology and Algorithms* component supports, in conjunction with energy related industrial partners, the application of high performance computers in computational chemistry, porous flow, groundwater transport, and high temperature superconductors, catalysis, quantum chemistry-based pharmaceutical design, and environmental remediation. The DOE recently conducted an evaluation of proposals for research centers which will support collaborations of computational scientists and engineers. These centers will provide early evaluation of advanced computer architectures and a coordinated program focused on supporting grand challenge computational research which is cost shared with other scientific research programs in DOE or with industrial partners to ensure the diffusion of technology. In addition, to ensure that technology transfer occurs effectively between DOE and its industrial collaborators, the DOE laboratories have initiated over 10 HPCC Cooperative Research and Development Agreements in the past year.

The DOE *National Research and Education Network* activity will ensure the availability of advanced network capabilities for access to its researchers and facilities. DOE plans to improve connectivity and network capabilities available to DOE researchers and research facilities nationwide through an upgrade to the Energy Sciences Network and through collaboration with other Federal agencies and regional networks. DOE will also help to define, jointly fund with other agencies, and in some cases carry out the research which will be required to evolve from today's infrastructure to the NREN. This will include participation in gigabit network testbeds to evaluate how they impact the solution of large science and engineering programs, development of tools for building and testing large distributed applications, and research on network congestion and flow control.

The DOE *Basic Research and Human Resources* component is expanding its program in basic research in the broad range of mathematical and computer sciences necessary to underpin all its other scientific research. DOE will expand the graduate fellowship programs and postdoctoral support in computational sciences. The first year of the computational science fellowship activity was highly competitive, with awards at 15 universities in 14 states. DOE will continue to expand its programs for access to high performance computers for students and teachers at high schools as well as its summer internship programs for promising undergraduate students at DOE laboratories.

National Aeronautics and Space Administration (NASA)

NASA's HPCC Program participates in all four components of the Federal HPCC Program through an integrated program focused on NASA's grand challenges in Computational Aerosciences and in Earth and Space Sciences. The goal of the NASA program is to accelerate the development and application of high performance computing technologies to meet NASA science and engineering requirements.

NASA's program brings together teams of computer and computational scientists to develop the necessary technologies within two vertically integrated NASA grand challenge projects that are unique to the NASA mission. These technologies include applications algorithms and programs, systems software, peripherals, networking, and high performance computing hardware.

In cooperation with the other Federal agencies, NASA's *High Performance Computing Systems* component will deploy experimental scalable computer capabilities essential for computational design of integrated aerospace vehicle systems and for predicting long term global changes.

In *Advanced Software Technology and Algorithms*, NASA will acquire hardware for computational science testbeds and develop software tools to enhance productivity, including load balancing tools, run time optimizers, monitors, parallelization tools, as well as data management and visualization tools. NASA has supported planning for an HPCC Software Sharing Experiment to identify issues related to software sharing and reuse. Cooperation and coordination with academia has been enhanced through a series of workshops recently conducted by NASA. These workshops addressed the Federal HPCC program and NASA's needs in software technologies and algorithms, computational aerosciences, and supercomputer testbed.

The NASA *National Research and Education Network* activity will provide high-speed network connections among NASA, industry and academic researchers.

NASA's *Basic Research and Human Resources* component will encourage research into the underlying theory and concepts of high performance computing and will foster research in high performance computing at NASA centers, research institutes, and universities.

Several computational advances have already been achieved using high performance computing testbeds at NASA centers. A unique computational method for enhancing imagery from the Hubble Space Telescope (HST) has been developed. This technique, called the Maximum Entropy Method, increases resolution by a factor of four over that of an unprocessed HST image. A new turbomachinery simulation code has been implemented on an advanced parallel system. This academically-developed code simulates three-dimensional inviscid, time-accurate fluid flow through multistage turbomachinery (i.e., jet engines). A multidisciplinary approach to simulating the flow field surrounding a powered-lift aircraft has been accomplished. The simulation used over 2.8 million grid points and incorporates an engine simulator to model the effects of experimental variables.

National Institutes of Health (HHS/NIH)

The HHS/NIH HPCC Program integrates high performance computing with computationally intensive biomedical research applications, links academic health centers via computer networks, creates advanced methods to retrieve information from life sciences databases, and provides training in biomedical computer sciences. NIH operates the first supercomputer dedicated to biomedical research in the public domain. NIH is evaluating advanced supercomputer architectures and system software for applicability to biomedical problems.

The *High Performance Computing Systems* component focuses on instrumentation, performance analysis, and evaluation of high performance systems, with emphasis on algorithms important to biomedical research.

The *Advanced Software Technology and Algorithms* component focuses on two areas of importance in biomedical research and education: biotechnology computing and digital imaging. In biotechnology computing, the program will support development of molecular sequence comparison algorithms, new database methods, structure determination from x-ray and NMR data and algorithms to predict biological structure and function from genetic code. The HPCC program will complement the Human Genome Project by providing new methods for computer based analysis of normal and disease associated genes. Accomplishments already realized in this area include the development of a similarity searching algorithm capable of rapidly analyzing hundreds of millions of molecular units, and the development of a promising new retrieval system to access DNA and protein sequence databases, along with literature citations in which the sequences were published. The biomedical images area will support new methods for representing, linking, and rendering images of biological structure.

The *National Research and Education Network* component has two subcomponents: connections among academic medical centers and their associated computerized information sources, and development of intelligent gateways that link conceptually related databanks in the life sciences. The academic medical centers of the country are confronted with a growing array of disconnected computer-based information sources, ranging from patient records, x-rays, and laboratory systems, to basic research tools such as protein and DNA sequence databanks. The focus of the Gateways program is the building of systems that are capable of translating a user's request into queries to appropriate databases from a broad range of widely available resources.

The *Basic Research and Human Resources* component addresses the need to train biomedical researchers and health care providers in the use of advanced computing and network communications to aid in their work. The successful predoctoral and postdoctoral grants program for career training in medical informatics will be expanded and a new postdoctoral training program will be started for the cross-training of biomedical investigators in advanced computer techniques and communications systems.

National Oceanic and Atmospheric Administration (DOC/NOAA)

DOC/NOAA formulates and carries out operational and research programs in weather prediction, ocean sciences, the Climate and Global Change Research Program, and the Coastal Oceans Program, together with data management activities for all agency programs. The HPCC Program will allow extensive development of new forecast models, studies in computational fluid dynamics, and the incorporation of evolving computer architectures and networks into the systems that carry out agency missions. These activities are carried on in collaboration with both the HPCC agencies and others.

The NOAA HPCC Program focuses on two program components:

Advanced Software Technology and Algorithms supports grand challenges in atmospheric and oceanic sciences; development of advanced numerical models simulating the general circulation of the oceans and atmosphere in support of NOAA missions and the activities of collaborating research groups; development of new computational methods for solving atmospheric, oceanic, and related problems on new computer architectures; data management R&D; support for basic research in strategies, techniques, and tools required for the management and analysis of large-scale distributed scientific databases and distributed data handling, including quality control; and development of algorithms for massively parallel processors, together with their standardized software component libraries and tools for the solution of oceanic and atmospheric analysis and forecasting problems.

NOAA uses advanced computational facilities for the evaluation of near operational prototype computers having novel architectures, and is developing algorithms appropriate for these new architectures. As the program continues to grow, advanced studies of software systems will be undertaken to determine the most advantageous production deployment of these systems.

The *National Research and Education Network* component provides networking for the climate and global change research community and a wide range of agency missions in oceanography, weather prediction, and environmental sciences research including distribution of high quality data sets and information. NOAA will expand use of advanced network strategies and hardware technologies related to its missions.

Environmental Protection Agency (EPA)

The EPA research program has three main goals: incorporation of advances in computing and communications technology into EPA's environmental assessment applications, demonstration of the advanced techniques on several grand challenge problems, and development of techniques to transfer these high performance tools to key State, Federal, and industrial users with decision making responsibility. These advanced environmental assessment tools will be capable of handling multiple pollutant interactions including the air/water interchange and optimization of control strategies.

The *Advanced Software Technology and Algorithms* component focuses on development and optimization of key environmental algorithms for computers with several to thousands of parallel processors. An advanced modeling framework is being designed to allow rapid integration of these new algorithms and emerging high performance technology to address grand challenge problems of air and water quality. Advances in scientific visualization center on provision of critical means for scientists and policy analysts to better interpret and understand results of environmental assessment tools. EPA has demonstrated the use of an object oriented applications building package to control air pollution model computation and visualization in a distributed, heterogeneous computing environment.

The *National Research and Educational Network* component was added to EPA's FY 1993 program to ensure that a critical mass of environmental problem solving groups are connected to the network to support use of advanced environmental assessment tools and collaboration research studies.

The *Basic Research and Human Resources* component includes the establishment of technology training programs targeted to educate State users, industrial environmental groups, and Federal policy analysts on the use of HPCC technology for solving environmental problems. The approach provides early prototype applications to key user communities to obtain feedback on ease of use and functionality for environmental assessment analysis. The program also supports cross-discipline career training through universities and encourages high school students to explore HPCC technology through programs like SuperQuest.

National Institute of Standards and Technology (DOC/NIST)

DOC/NIST supports three components of the HPCC Program.

In *High Performance Computing Systems*, NIST develops instrumentation and methodology for performance measurement of high performance networks and massively parallel computer systems. During the past year, two VLSI based instrumentation and monitoring methods were developed for shared memory and distributed memory computer architectures. Both chip sets provide a passive, programmable, non-perturbing means to capture events and trace system level activities. Data acquired is time stamped and sent by hardware in real-time to a separate computer via an integrated network interface for subsequent processing and transformation into performance metrics. Both were designed to be incorporated in scalable, massively parallel systems. Data analysis has led to identification of a means for succinct representation of state information of dominant modes of operation within massively parallel computer systems, and has established the importance of the succinct representation of these performance metrics. This information is important in system component load balancing and support of user built analytic tools.

In *Advanced Software Technology and Algorithms*, NIST will focus on: scalable, parallel algorithms in the field of computational geometry, and software classification and exchange via the Guide to Available Mathematical Software (GAMS) system. The GAMS system has been incorporated in the NASA Software Sharing Experiment as a means to classify, select and exchange software for scientific and engineering applications. Work has been initiated to improve the utility and portability of the system.

For the *National Research and Education Network*, NIST will support, coordinate and promote the development of standards within the Federal Government to provide interoperability, common user interfaces to systems, and enhanced security. As advanced requirements flow from HPCC research and development, NIST will promote "Open Systems" standards to assist industry commercialization of the R&D, with the aid of other agencies. NIST also conducts research and development on new high performance communications protocols and will deploy a gigabit testbed to assess protocol performance. NIST will coordinate and contribute to the development of security technology, guidelines and standards which are related to components of the Program. A gigabit network testbed is being deployed to support this activity and includes a high performance parallel interface switch, a massively-parallel, video supercomputer and a vector class supercomputer. Related communciations hardware interfaces and high performance communications protocols are in the design stage.

4. Grand Challenge and Supporting Technology Case Studies

This section incorporates examples of high performance computing and computer communications technologies and several illustrative grand challenge applications in computational science and engineering that are presented on the cover of this report. The list of grand challenges is too long to allow an example from each possible area, and lack of representation of certain areas does not imply lack of importance. In addition to these examples, there are many important applications of high performance computing in national security areas that are supported by agency mission funding.

The examples below were chosen to illustrate the diversity and significance of application areas that have been addressed to date.

- Magnetic Recording Technology
- Rational Drug Design
- High Speed Civil Transports
- Catalysis
- Fuel Combustion
- Ocean Modeling
- Ozone Depletion
- Digital Anatomy
- Air Pollution
- Design of Protein Structures
- Venus Imaging
- Technology Links Research to Education

Magnetic Recording Technology

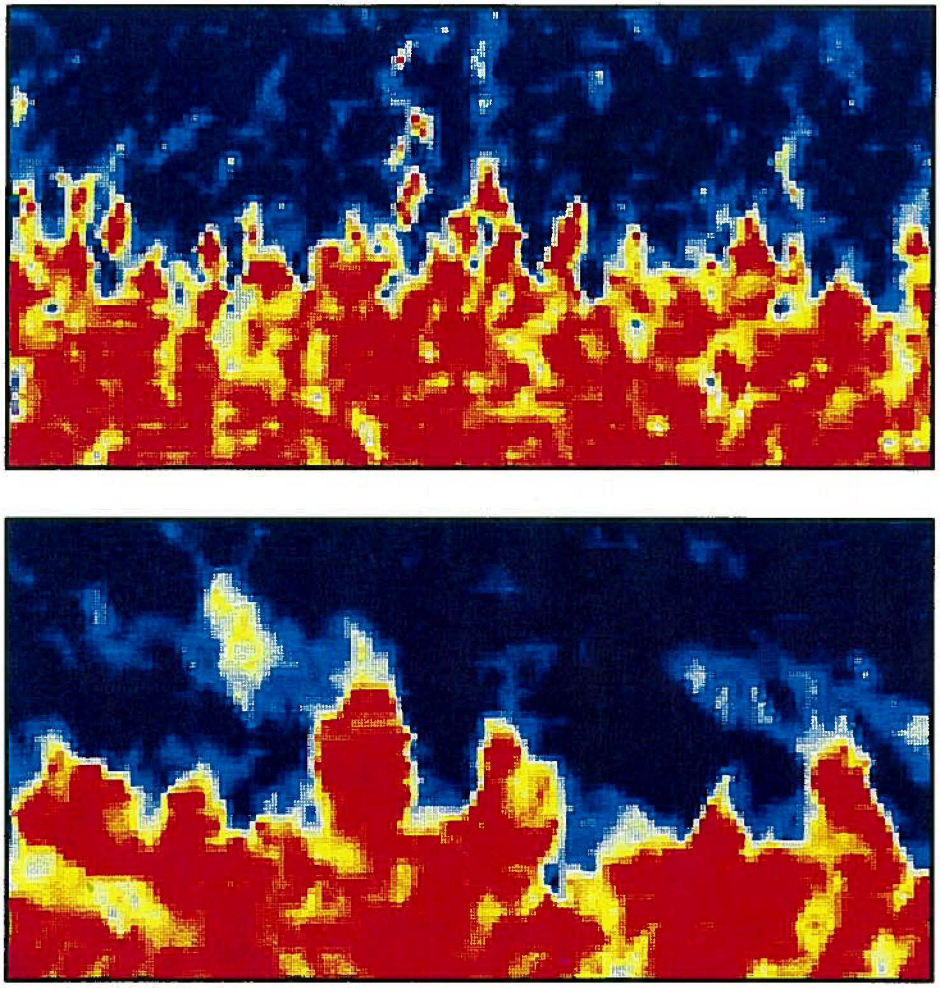
The knowledge and information explosion in every aspect of modern life is creating rapidly increasing requirements for efficient storage of information. Currently, the dominant methods for information storage employ magnetic recording media such as the thin magnetic films that are used in computer disk drives and tape systems.

Future advances in magnetic recording technology will underlie new technological advances and scientific breakthroughs, from laptop supercomputers to data collection from earth orbiting satellites downloading terabits of information per day.

One of the goals of magnetic media research is to pack more and more information onto available media. The fundamental problem is to avoid or compensate for magnetic noise in the metallic thin-films used to coat high density disks. The films are crystalline, rather than continuous, and are composed of tiny magnetic grains of cobalt alloys deposited on a substrate of chromium. The grainy, particulate nature of the medium is the fundamental cause of noise.

Researchers are using high performance computers to study the effects of the two primary magnetic interactions: magnetostatic and exchange. First, all magnetic dipoles (grains) in the films interact with each other via long range magnetostatic coupling. It is this long-range interaction that makes the analysis computationally intensive. Second, exchange coupling occurs between neighboring grains when, for example, cobalt diffuses into the intergranular boundaries, causing direct magnetic interaction. Even minuscule exchange coupling, the simulations show, dramatically increases noise.

The smaller that the distinctly separable grains can be made, the higher will be the density at which digital bits can be recorded and retrieved. The key is to devise alloys and fabrication processes to assure nonmagnetic grain boundaries, minimizing exchange-coupling noise. Based on theoretical and computational studies, researchers have recently shown that one billion bits can be packed onto a square inch of disk surface, nearly 30 times the storage density of current media.



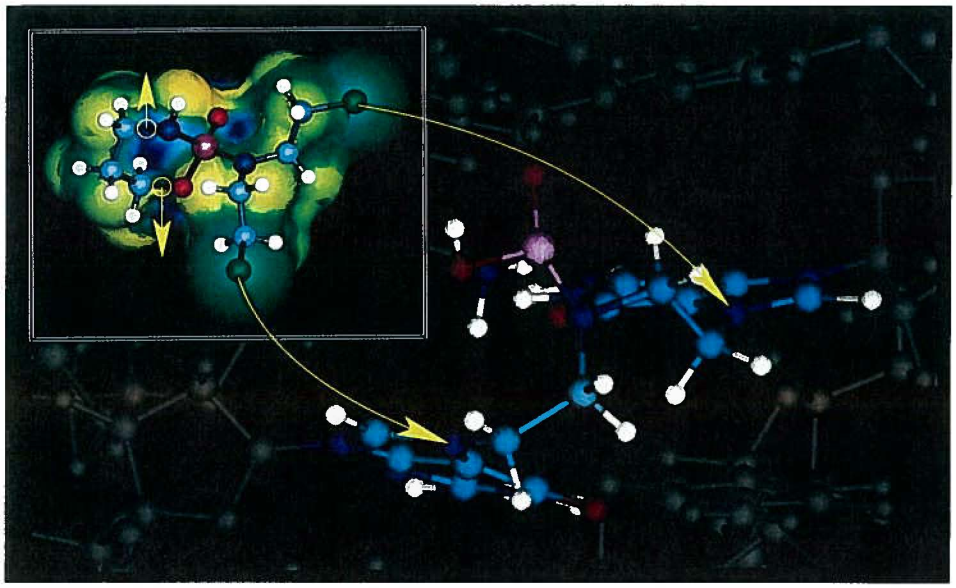
The pictures represent magnetization patterns for recorded bit boundaries, or transitions, in thin film recording media. The color, from blue to red, represents the magnetization component in the recording direction. The bottom picture is the magnetization pattern for a typical film medium with intergranular exchange coupling, which is a noisy medium. The top picture is the magnetization pattern for a typical film medium with zero intergranular exchange coupling and the resulting transition noise is much less.

Rational Drug Design

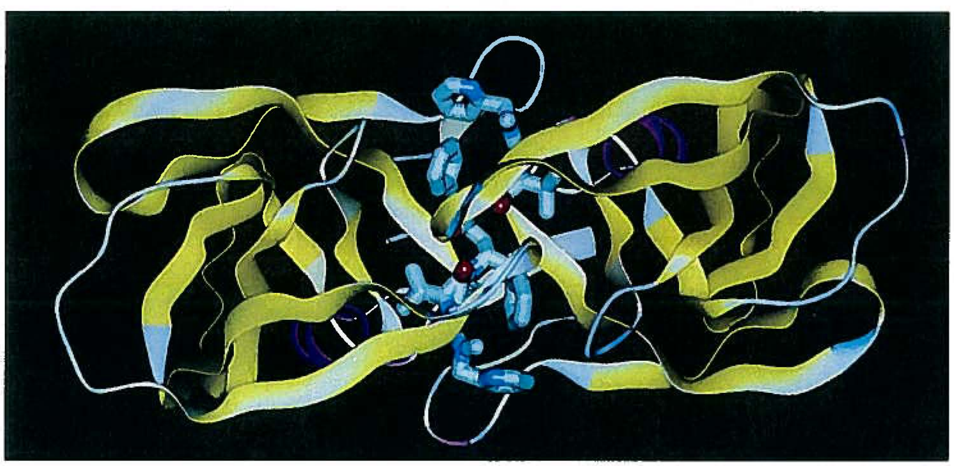
Recent progress in the development of therapeutic agents for two medical problems — cancer and acquired immunodeficiency syndrome (AIDS) — illustrates the value and potential of modern computer-assisted, receptor-based drug design as a means to markedly reduce the time and cost of the drug development cycle. Parallel developments in x-ray crystallography and nuclear magnetic resonance spectroscopy along with high performance computing have provided extraordinary and rapid progress in our understanding of the three-dimensional structure determinations of biological macromolecules. This provides the basis for rational drug design, the logical search for a specific therapeutic agent developed from an improved understanding of molecular structure and function.

The most commonly used alkylating agent to treat some cancers is cyclophosphamide. The active metabolite of that drug, phosphoramide mustard (PM), induces the anticancer effects. It crosslinks with DNA molecules in cancerous cells to inhibit their growth. The mechanism through which PM determines which cells are cancerous is unknown, but it is widely believed to be due to a unique biochemical property of the cancer cells. More detailed chemical calculations have challenged existing concepts of how PM induces its effects on cancer cells. The complete elucidation of the mechanisms for that anticancer drug and others requires a computing capability many times greater than that which currently exists. This may lead to the development of more effective drugs to treat several different types of cancer.

The development of therapeutics for AIDS is another example of active-site-based drug design. The structure of the human immunodeficiency virus (HIV) includes a protease, an enzyme which catalyzes a mandatory step in the maturation of the virus so that it can infect cells. Using a high performance computer, new potential agents have been identified to block the action of the HIV protease. These compounds have anti-HIV activity, but further modifications of candidate agents need to be made to induce a greater therapeutic potential and lower unwanted side effects. Higher performance systems are needed to fully understand these new candidate agents and to develop and design new drugs.



The most widely used anti-cancer drug, cyclophosphamide (inset) acts by binding (cross-linking) two (guanine) bases of the DNA of a cancerous cell. Two chlorine atoms from the drug (green in the inset) are stripped off in the body and replaced by two nitrogen atoms (shown by arrows) of the DNA. This forms the bonding that prevents DNA replication, thus killing the cancerous cell. When broken down by the body, the side chain shown on the lower left of the inset is released, causing toxic side effects. Though this research required massive amounts of computation, calculations thousands of times more complex are needed to fully understand the activity of this drug.

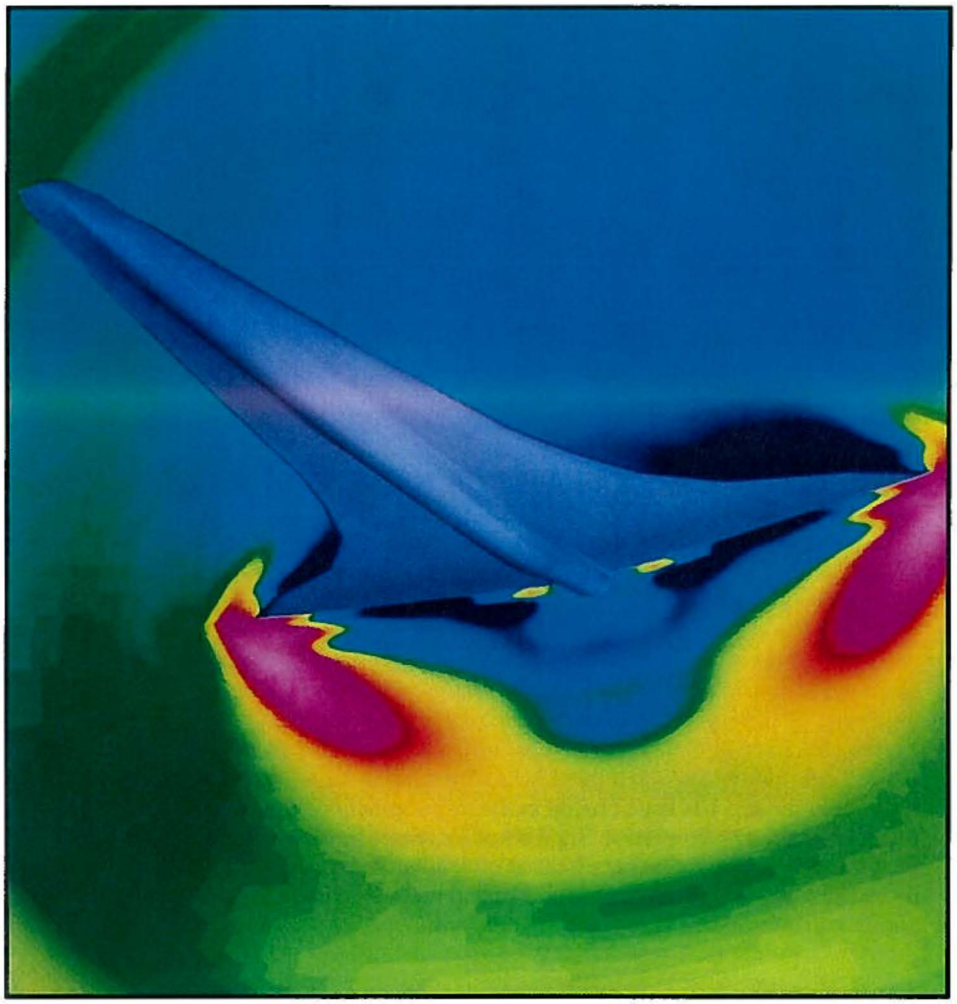


Atomic structure of a rationally-designed inhibitor bound to the active site of HIV protease ("ribbon" structure). The structure of the complex was determined by x-ray crystallographic analysis using a supercomputer networked to a high speed graphics workstation.

High Speed Civil Transports

The U.S. has initiated a research program to examine the feasibility of a new High Speed Civil Transport (HSCT). This type of aircraft will significantly impact the world's transportation system. Significant technological challenges must be addressed to assure environmental compatibility and to realize the commercial potential. Among the most important of these are: aerodynamic efficiency, advanced materials, and engines with acceptable environmental impact. High performance computing systems are essential to bringing HSCT to reality at the earliest possible date.

Computational fluid dynamics codes running on high performance computing systems will enable the rapid evaluation of alternative HSCT designs for aerodynamic efficiency and stability, as well as predicting the associated sonic boom. In addition, other programs will account for the atmospheric heating of the vehicle and the physical stresses of supersonic flight to provide the necessary information for the development of advanced materials and fabrication of new structures to commercialize HSCTs. High performance computing technologies are crucial to the design and evaluation of advanced propulsion systems that will efficiently and safely power HSCTs while significantly reducing community noise and exhaust emissions harmful to the upper atmosphere.



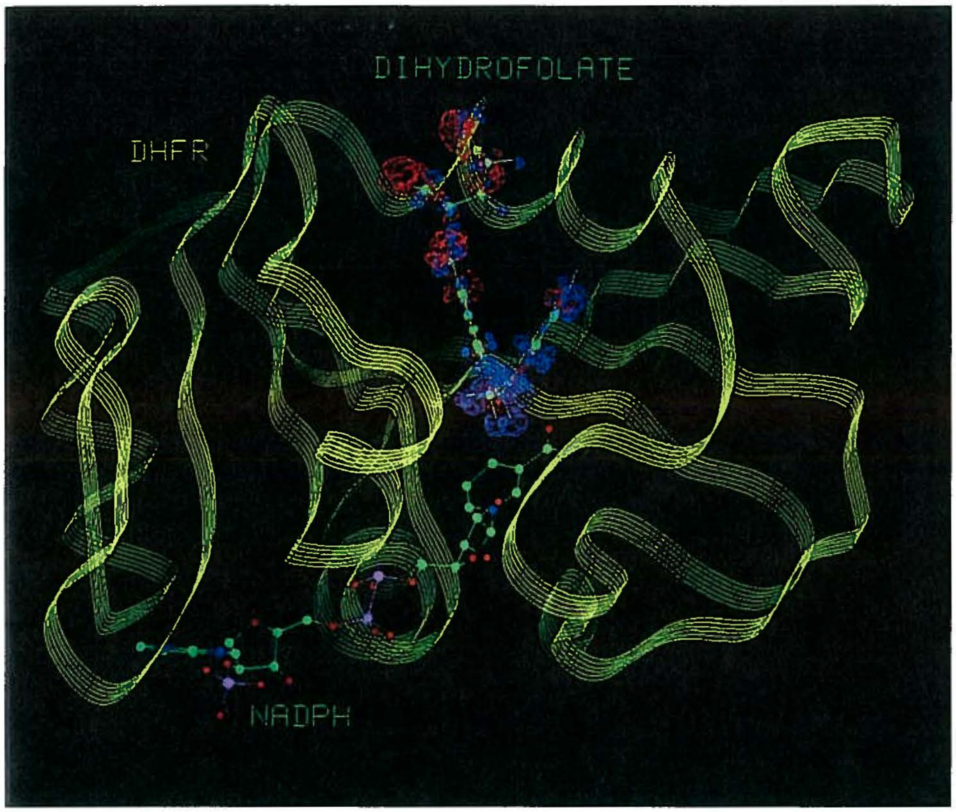
The figure above represents a numerical simulation model of a proposed low sonic boom commercial High Speed Civil Transport. The program is simulating almost level flight at twice the speed of sound (Mach 2). The intensity and directional characteristics of the sonic boom are depicted by coloring the atmospheric pressure surrounding the aircraft. Highest intensity is shown in red; lowest is in blue.

Catalysis

Catalysts for chemical reactions are currently used in one-third of all manufacturing processes, producing hundreds of billions of dollars worth of products each year. While most catalysts are still designed empirically, high performance computing, along with state-of-the-art algorithms and physical models, is being used to develop a computer design and analysis capability for commercially important catalytic processes and catalysts. The goal is to reduce the time required to design catalysts and to optimize their properties. Important catalysis applications include enzymes, biomimetic catalysts, microporous solids, and metals.

Many biological processes are catalytically controlled by enzymes, which give a high level of selectivity to a desired reaction over competing reactions. While enzymes are fundamental in medical science, environmental restoration, and food processing, the way in which enzymes function in their catalytic role is not completely understood for any system. High performance computing technology will enable larger simulations that more realistically model the substrate, cofactor, and local portions of the enzyme.

Enzymatic reactions are attractive because they exhibit an extremely high degree of selectivity to specific reactants and products. However, they also can be limited to chemical and physical environments different from those in many industrial processes. Alternatives to enzymes that may be more useful industrially are biomimetic catalysts, so called because they mimic biological catalysts (enzymes). These catalysts are synthetic materials that generally are much simpler chemically and can function at higher temperatures than their biological counterparts. Computer modeling of these biomimetic catalysts using relatively simple force field models has been crucial in the development of molecules with the correct size and shape properties. Significant increases in computational power will be necessary to determine the effect of molecular structure on electronic states in these catalysts, and to evaluate the chemical reactions that will occur. Massively parallel quantum models are being developed to analyze these complex systems.



The enzyme dihydrofolate reductase catalyses the reaction of dihydrofolate to tetrahydrofolate, an important one-carbon donor in a variety of biosynthesis reactions including the synthesis of amino acids and nucleic acids. The figure shows the effect of the enzyme in shifting the electron density in dihydrofolate towards that required to make the reaction take place. Red indicates increases of electron density, blue indicates decreases of electron density.

Fuel Combustion

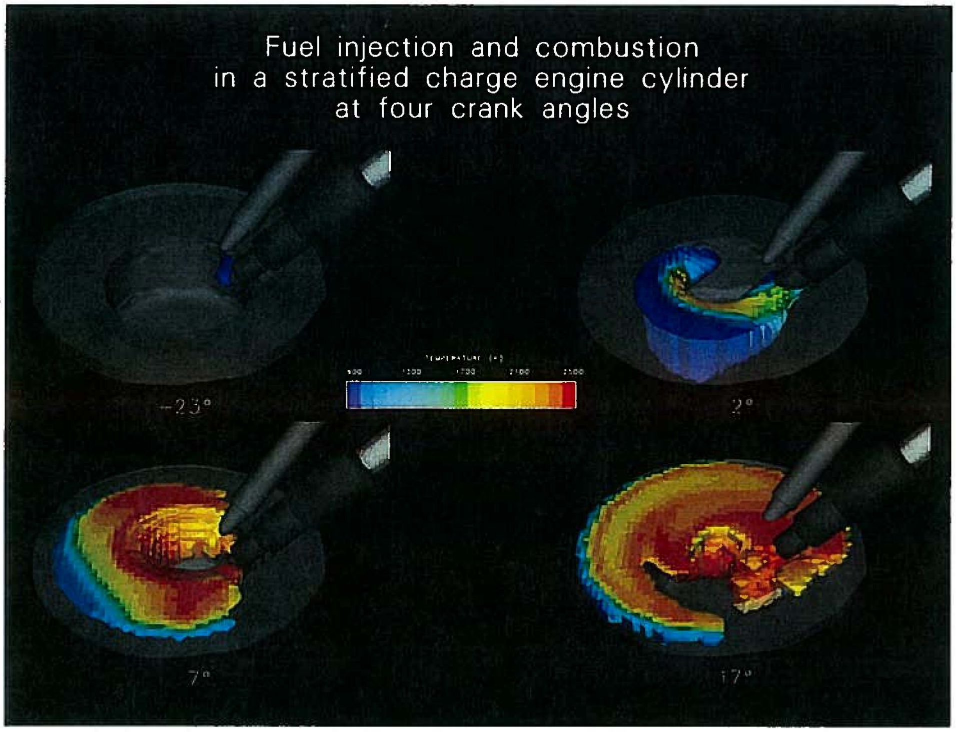
A detailed understanding of fuel burning inside an internal combustion engine requires the comprehension of a number of complex processes. One example is an experimental engine design whose goal is to run leaner and cooler than conventional engines, thereby reducing emissions while increasing the compression ratio and control of ignition for greater efficiency and fuel economy. The burning of the fuel-air mixture in such an engine brings into play a few hundred different chemical reactions between numerous short and long-lived chemical species. The rates at which these reactions occur depend on the temperature and concentrations of the species. For instance, as combustion proceeds, various pollutants may be formed, such as nitrogen oxides and unburned hydrocarbons. These hydrocarbons may form solid particles known as soot. Also, the fuel may autoignite in some region before the flame arrives, giving rise to the small explosion known as engine knock.

Because engine operation is so complex, designing an optimum engine by experimentation (varying all design parameters independently) is prohibitively expensive. A computer model of the engine allows us to effect design changes numerically on a scale of hours at a fraction of the cost.

The engine components depicted in the illustration are a cupped piston, fuel injector, and spark plug. The piston draws in fresh air and expels exhaust gases as the combustion cycle progresses. At an appropriate point in the cycle (upper right in the figure), liquid fuel is injected into the engine cylinder. The fuel jet breaks up into small droplets, which interact with the air. The droplets alternately elongate and flatten; some of them break up into multiple new droplets, others collide and coalesce to form larger drops. The spray is caught up by the air flow and swirls around in the cylinder. It simultaneously evaporates to form a gaseous fuel-air mixture.

Present engine models are deficient in one way or another because of limitations in computing resources. Computer codes can provide information on large scale fluid flows, but are limited in their ability to describe the small scale flow structures in turbulent flow. Furthermore, fuel chemistry models are necessarily crude. Current chemical kinetics calculations approach the detail necessary for an adequate description of the chemistry, but even these models require serious improvements. Such calculations are capable of only the most rudimentary fluid mechanical effects.

The envisioned increase in computing power will make improved models of turbulence and chemistry computationally tractable. These improvements are critical to the effective development of clean, efficient engines.



Shown above is a computer simulation depicting four different times during the combustion cycle of a Direct-Injection Stratified Charge engine. The multicolored surface indicates the fuel's location. When the spray has evaporated sufficiently a spark ignites the mixture, and combustion begins as a flame propagating through the chamber. As the engine proceeds through its cycle (indicated by the increase in crank angle from -23 to +17 degrees), the fuel burns, resulting in high temperatures (see color code).

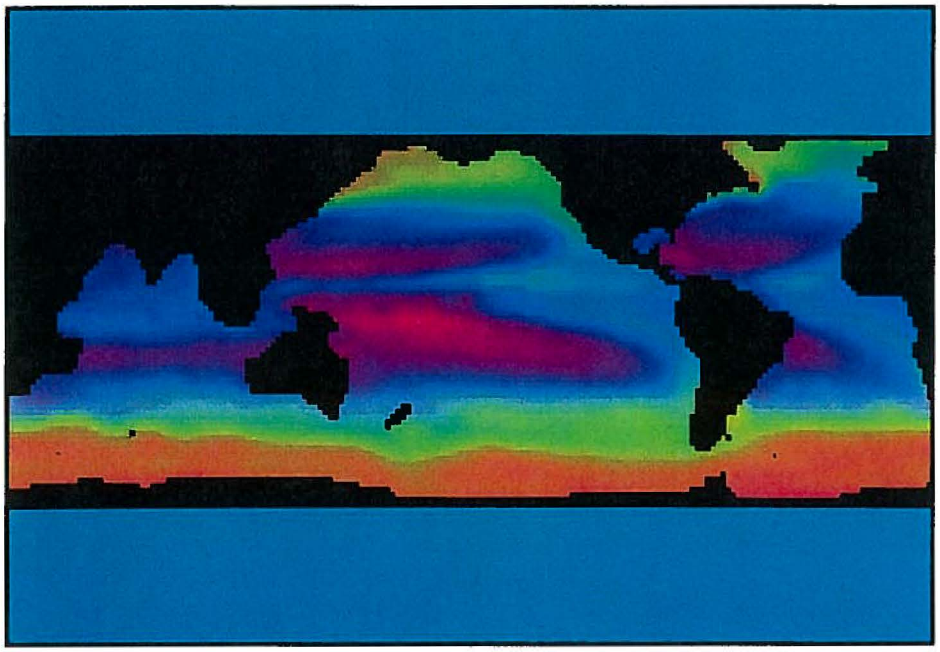
Ocean Modeling

The ocean is the other turbulent fluid in the climate system. By contrast with the atmosphere, the ocean has much smaller spatial structures and much longer time responses. The ocean currents, which transport almost as much heat as atmospheric flows, are only about 50 kilometers across, and ocean eddies that are the counterparts of synoptic weather systems are one-hundredth the size of those in the atmosphere. These oceanic phenomena evolve over periods from years to centuries to profoundly influence the climate. The small space scales and long time scales of the ocean present a massive challenge to high performance computing.

The ocean is particularly important in maintaining the climate of the Northern Hemisphere through an organized system of currents which has been likened to a conveyor belt. Currents near the surface transport heat from the Tropical Pacific all the way to the North Atlantic, where the heat enters the atmosphere through evaporation. The evaporation makes the surface waters heavy enough to sink and return to the Tropical Pacific, where they rise and are freshened by rains and warmed by the sun for another trip to the North Atlantic. However, the entire system is very sensitive to changes in atmospheric conditions, such as might occur from increasing greenhouse gases.

The first global ocean simulation which is capable of barely resolving the eddies and strong currents, including those that act as a conveyor belt, has only recently been accomplished. An integration of just a decade required 3000 processor hours on a gigaflop supercomputer. If multi-century ocean simulations with adequate spatial resolution are to be performed in support of climate prediction, then teraflop machines will be required.

Because of the turbulent nature of oceanic flows, there should be demands for archiving, visualization, and communication that parallel those of the most ambitious atmospheric and hydrodynamic applications. For example, the first global simulation with barely resolved eddies produced more than 250 gigabytes of processed output. A two hour animation was made to help understand the resulting phenomena, and the animation required the transfer of gigabytes of model output across high-speed networks.

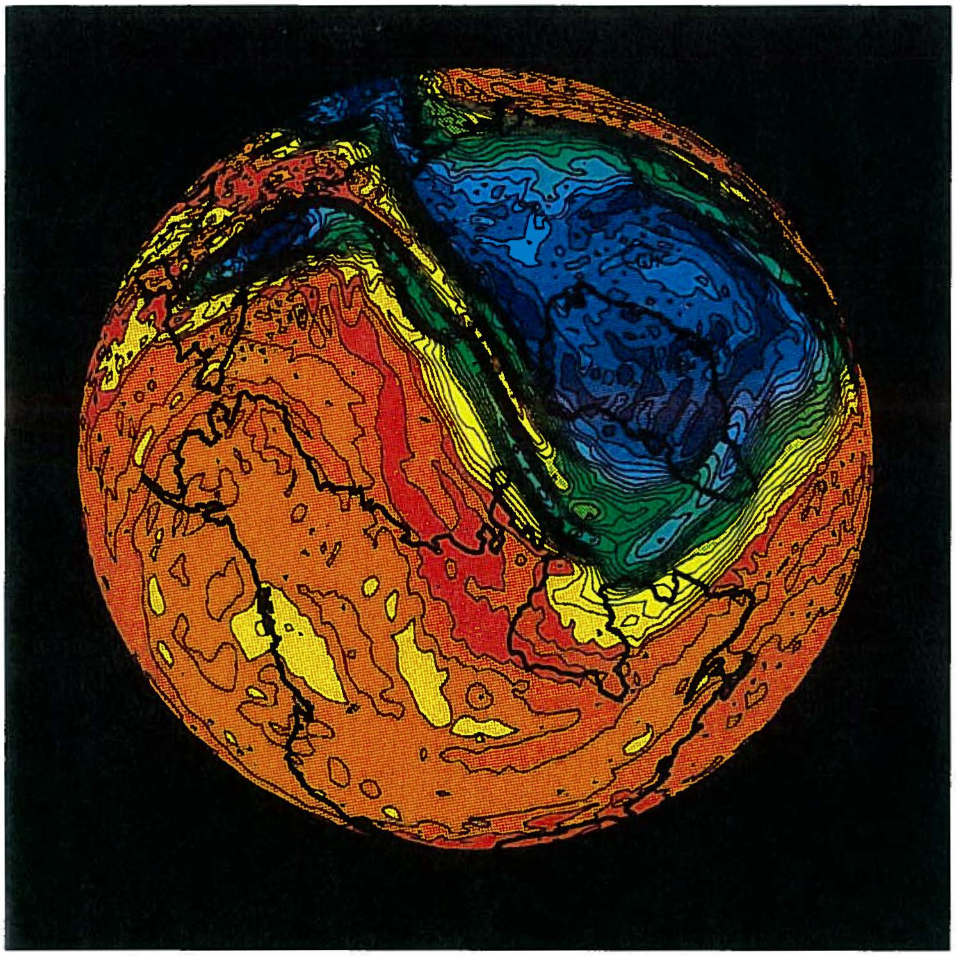


Salinity of the upper Pacific Ocean.
The atmospheric hydrogical cycle strongly influences the North Pacific, where an excess of precipitation freshens (and lightens) the water that will transport heat all the way to the North Atlantic. This effect is seen in upper-ocean salinity from a global ocean model with barely resolved eddies. The blue zones in the figure are relatively fresh, while the red ones are more saline. Steady ocean currents set up a sharp frontal barrier along the equator. However, transient flows, such as a vigorous current reversal near New Guinea and waves on either side of the equator, transport salty water toward the North Pacific to restore a balance. This figure is part of an animation that graphically portrays these features, as well as other vigorous flows along the path to the North Atlantic.

Ozone Depletion

The depletion of the Earth's ozone has an important impact on animal and plant life, including higher incidence of skin cancer in humans, because of the increased ultraviolet radiation reaching the surface through the reduced ozone shield. Recent observations of unpredicted ozone losses in the Earth's stratosphere have led to intense research efforts to understand the cause. The "ozone hole" that has been observed in the stratosphere above Antarctica during the southern hemisphere winter is a dramatic example of this ozone depletion process. Recently, reduced ozone levels have been observed in the northern hemisphere stratosphere at high latitudes around the perimeter of the polar vortex. Scientists are now studying the dynamical and chemical mechanisms associated with this newly discovered ozone depletion to determine whether an ozone hole similar to that over Antarctica can also form in the northern hemisphere.

The chemical and dynamical mechanisms that control the ozone depletion process are extremely complex. For example, the chemistry of the problem involves hundreds of different chemical processes. Moreover, the efficiency of this chemical activity depends upon the detailed transport and mixing of air containing very different chemical compositions. Because of the complex interaction between dynamics and chemistry, an essential element in this investigation is the use of comprehensive global, high-resolution computer models to simulate the ozone chemistry and associated dynamics that occurs in the Earth's atmosphere. For example, current models require as much as ten supercomputer processor hours to model one day of interaction. One of the major requirements for continuing advancements in this field will be the incorporation of new technologies to meet the tremendous computational demands of future global climate models.



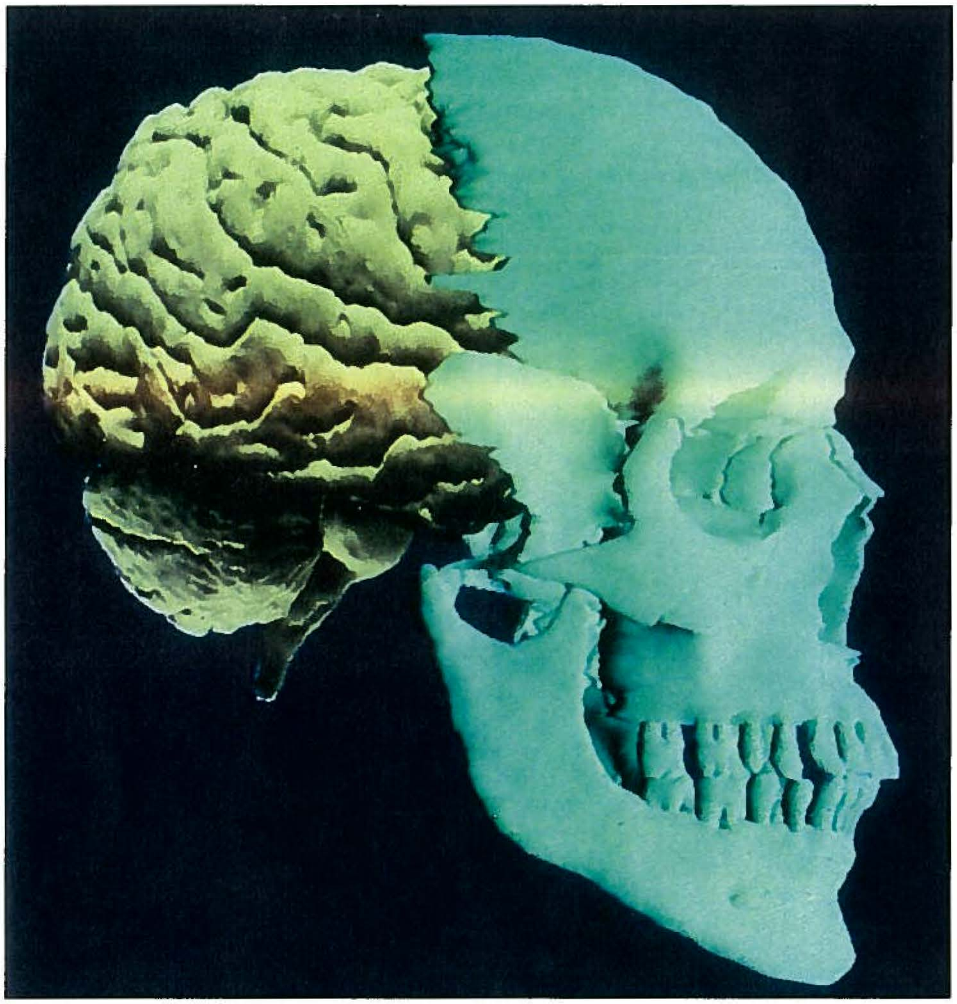
The figure shows an instantaneous field of nitrous oxide (N_2O) for a model date in February on a constant entropy surface corresponding to an altitude of roughly 20 kilometers. Nitrous oxide is a biologically produced trace gas that is a fundamental determinant of the natural ozone balance. Its long lived character also makes it an excellent tracer of atmospheric motions. The figure shows very low values of N_2O in the polar regions; the N_2O structure essentially defines the winter polar vortex that determines the chemistry typical of the polar ozone hole phenomenon. Note the virtual wall of N_2O gradient that occurs between the polar vortex and the outside region. At times, the wall structure becomes complicated and folds upon itself, forming what appear to be double and triple wall regions.

Digital Anatomy

New computer-based imaging techniques are making it possible to explore the structure and function of the human body with unprecedented accuracy. Clinical imaging methods such as computed tomography, magnetic resonance imaging, and positron emission tomography yield two-dimensional images which may be transformed by computer algorithms into three-dimensional pictures. Reference standard normals, such as the Visible Human project, provide three-dimensional numerical coordinates from which internal as well as external structure can be depicted, rotated, viewed from any angle, and reversibly "dissected."

Automated correlation of images derived from different imaging techniques is a formidable technical problem whose solutions are computation-intensive. These image datasets are also large (the Visible Human project, for example, will generate a four trillion byte image library), and high speed networks are the only feasible method for allowing researchers and health professionals to browse such electronic image collections interactively, in a manner analogous to today's interactive browsing of textual databases.

The computer and network tools applied to the challenge of multimodal imaging promise an era of new insight into the anatomy and physiology of human health and disease in the three dimensions of biological structure and the fourth dimension of time.



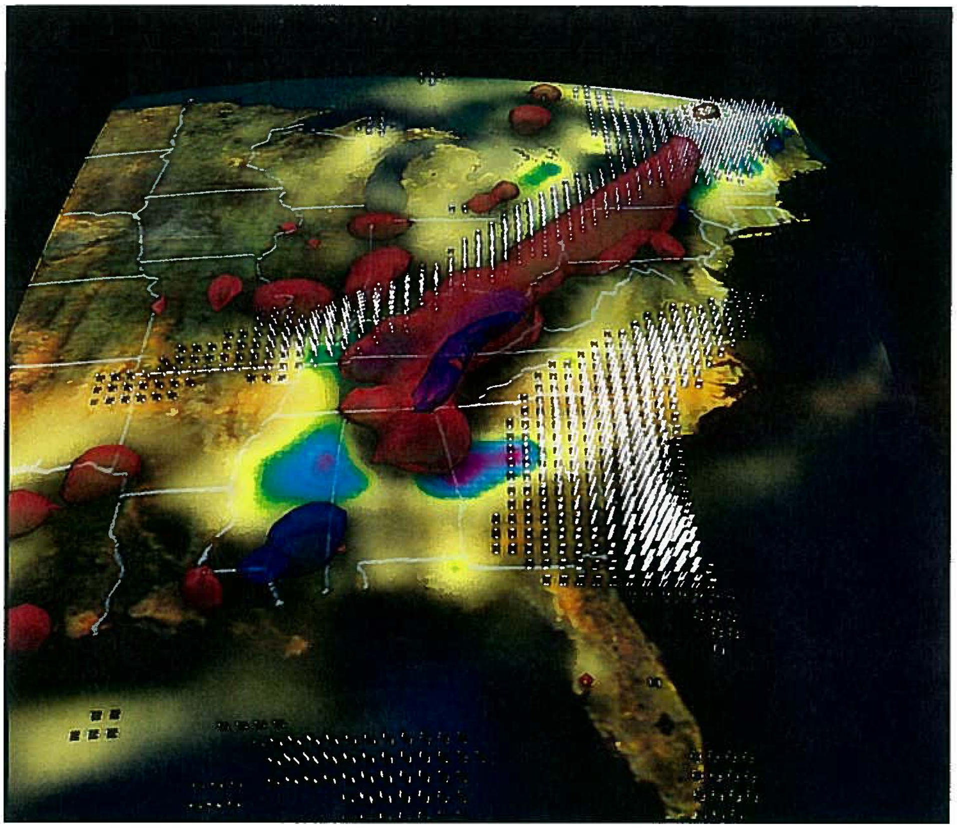
The volumetric rendering of the human head is capable of depicting deep as well as surface structures. This three-dimensional view was reconstructed from two-dimensional Magnetic Resonance Imaging data acquired from a living person. Conventional computers require several seconds to several hours to compute a single three-dimensional view; advanced architecture machines developed under HPCC will allow such images to be rotated, "dissected," and viewed from any angle under real-time control.

Air Pollution

Reduction of pollutant and pollutant precursor emissions is costing billions of dollars. The early optimism of the past two decades about the ease of controlling and reducing air pollution, especially ozone, has given way to a realization of the complexity and extreme difficulty of achieving the reductions necessary to attain the health standards in many urban areas. In addition, connections between pollutant media are increasingly being recognized. For example, nitrogen deposition from air pollutants is now thought to be a major contributor to eutrophication of coastal estuaries, adding to the agricultural and urban sources that have been studied and characterized.

Computational models are powerful and necessary tools to study the transport and transformation of pollutants and to provide guidance on the expected effectiveness of emission control strategies. These air quality models must incorporate descriptions of physical, chemical and meteorological processes that encompass scales from regional (several thousands of kilometers) to local (several kilometers) to adequately represent the production of pollutants. Interactions between many physical processes, such as clouds and precipitation, must be incorporated in the computational descriptions, leading to models that are, of necessity, highly sophisticated and which must be run on supercomputers. Yet, the complexity and completeness of the air quality models must be advanced to more accurately simulate the effectiveness of emissions controls and find effective and efficient solutions to air pollution.

High performance computing will enable additional and more complete descriptions of the physical and chemical processes to be simulated in the air quality models. Advanced visualization techniques will allow the scientists to better study and understand the complex, nonlinear interactions of the atmospheric pollutant system, leading to more informed assessments. High performance computing will enable scientists to explore the complex interactions in the models more quickly and provide the type of guidance needed to reduce air pollution to meet the air quality standards.



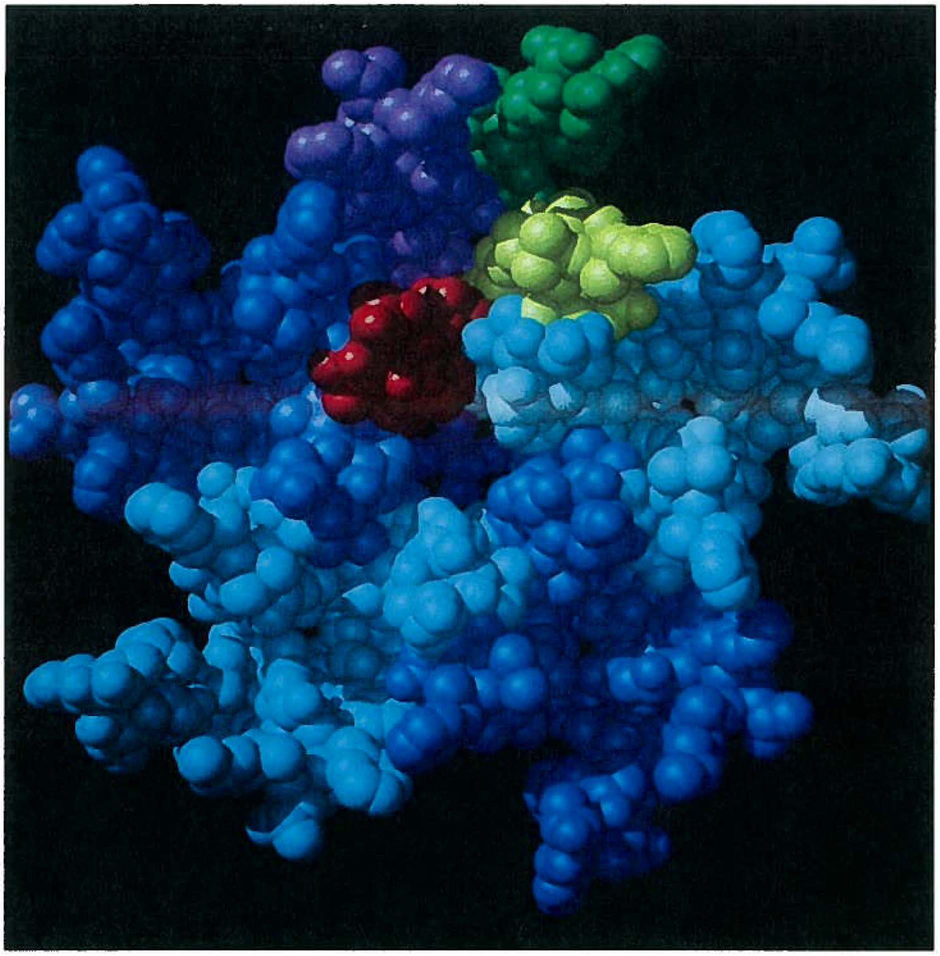
Depicting the interactions between several processes simulated by air quality models leads researchers to more understanding of the complex atmospheric system. High concentrations of sulfate (shown as a grey cloud) are followed as they are simultaneously transported and scavenged by precipitation (the silver pill boxes). The resultant accumulated wet deposition of sulfur due to several rain storms passing over the eastern United States is shown as a color-coded, two-dimensional concentration map "painted" on the landscape.

Design of Protein Structures

Although the information contained in the amino acid sequence of a protein is sufficient to determine the three-dimensional or "folded" structure of a protein, the rules by which the protein assumes its final configuration — its three-dimensional structure — are not yet understood. Success, measured even by a partial understanding of these rules, is often of immediate importance to biomedical research, and scientists are using high performance computing to learn more about the process.

An understanding of protein folding could help to design new proteins with a specific structure, for a specific need, perhaps to catalyze a biochemical reaction. If the design of a protein *de novo* seems far away, a more reasonable goal of intelligently modifying a protein to yield a desired change in stability, reactivity, or specificity seems more achievable. Simulation of protein activity, tracking the movements as the protein goes from unfolded to a folded state, has proved particularly fruitful. But, since most biochemical reactions proceed in fractions of microseconds, current computationally intensive modelling methods are thousands to millions of times short of simulating interesting kinetic reactions.

Computational biologists are looking to massively parallel computing architectures to overcome this barrier, anticipating order of magnitude improvements in the calculations, and the ability to engineer the complete structural properties of a protein to desired specifications.



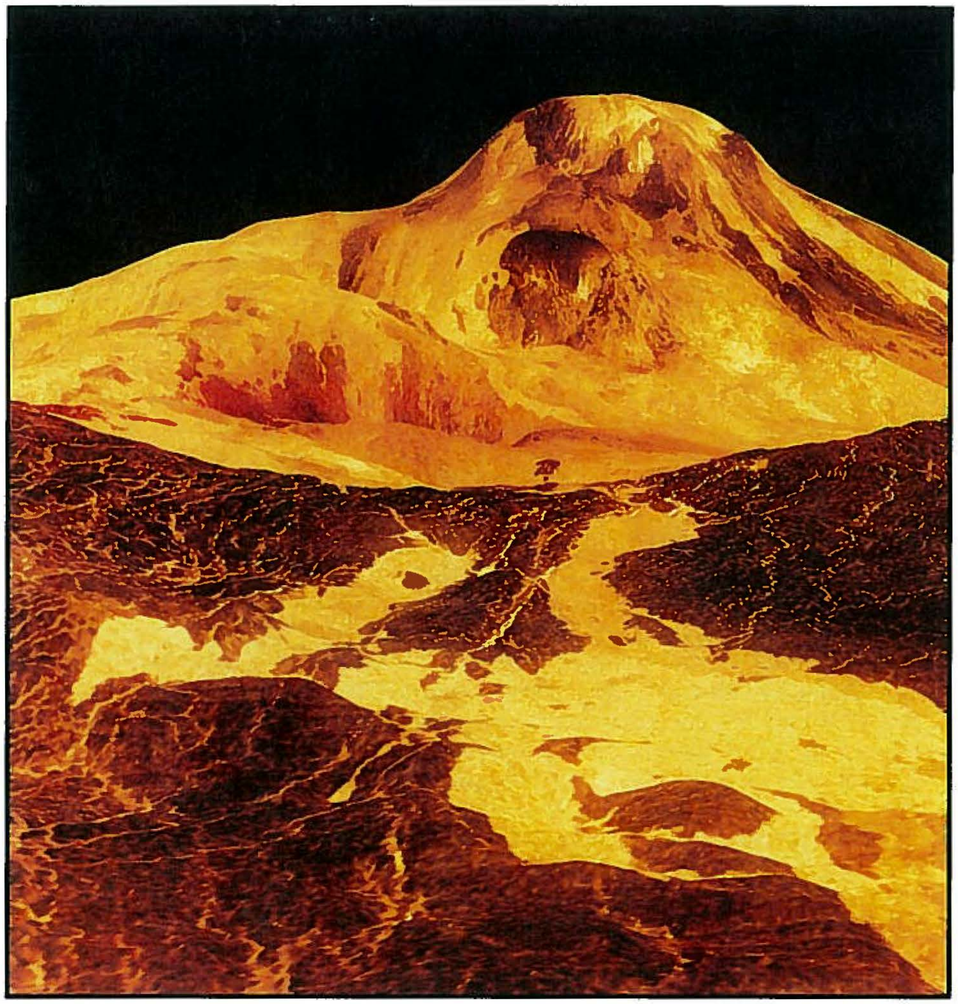
Scientists have been forming models of channels, like the one found in the acetylcholine receptor, to see how alpha-helices (a pattern for folding amino-acids residues in proteins in a circular pattern) might pack together to control the flow of ions through the protein. The figure shows a channel with a space filling technique. In the center, an ion surrounded by several water molecules is shown in blue. Construction of the model was based on a four-helix bundle.

Venus Imaging

The Magellan spacecraft has been mapping the surface of Venus since September 15, 1990. Magellan uses a radar instrument to penetrate the thick clouds surrounding the planet to produce 120-meter resolution images of the Venus surface. Venus is the most Earth-like of the planets of the inner solar system, with about the same size, mass and density. However, the "greenhouse effect" due to the clouds surrounding Venus results in a surface temperature of about 900° F and pressure at the surface about 90 times that of Earth. Despite these differences, Magellan data for Venus have revealed a wide range of geologic features from volcanoes to mountain ranges that are providing insight into how geologic processes on Earth and Venus operate. Magellan has provided images and topographic information for over 90 percent of the Venus surface. The total volume of data returned by Magellan exceeds 3×10^{12} bits or 3 terabits (all previous U.S. planetary missions combined have returned about 9×10^{11} bits).

The Magellan radar data go through several processing steps, starting at a Synthetic Aperture Radar (SAR) processor that turns the echoes returned from the planet's surface into images. The data then go through various image processing steps, where Magellan image strips are combined to form mosaics which can then be color enhanced or combined with topographic data to provide scientists with tools for understanding the geology of Venus. The large data volume associated with Magellan has necessitated new and increasingly more complex methods of data processing, handling and storage that will be significant to future data-intensive projects such as NASA's Mission to Planet Earth. Magellan images, such as that shown in the figure, provide insights into the physics of volcanism and its significance to planetary evolution.

Using the large parallel supercomputers for producing the rendered images dramatically increases our capabilities. For example, a single image produced on a conventional high performance workstation takes several minutes to complete. Animations, even short ones, typically take weeks to produce. Massively parallel supercomputers hold the promise of being able to store the huge datasets on-line and render animations in real-time.



Maat Mons, the largest volcano on Venus, is shown in the three-dimensional perspective view of the Venus surface rendered on a 528 node parallel supercomputer. The viewpoint is located 560 kilometers north of the volcano at an elevation of 1.7 kilometers. Vertical exaggeration is about 22 times. Lava flows extend for hundreds of kilometers across the fractured plains shown in the foreground, to the base of Maat Mons. The volcano is about 8 kilometers high and over 300 kilometers across, similar in size to Mauna Loa on Earth. To produce the image, Magellan SAR data were combined with radar altimetry to develop a three-dimensional map of the surface. Rays cast in the computer intersect the surface to create a three-dimensional perspective view. Simulated color and a digital elevation map are used to enhance small-scale structure. The simulated hues are based on color images recorded by Soviet landers.

Technology Links Research to Education

It can take ten years or more for new ideas in science to be incorporated into the curriculum of schools. Given the rapid advances in science, new ways must be found to link researchers, educators and learners.

The technologies that make possible the accelerated pace of scientific advances are also a key to creating the needed linkages with education. High performance computer applications such as simulations and visualizations used to advance scientific knowledge are now being tailored to help teachers and students understand both the methodologies of modern scientific inquiry and specific concepts in these scientific domains. The models and simulations used by researchers are mathematically complex. Through advances in visualization technology, it becomes possible for students who are not yet sophisticated in quantitative reasoning to get a qualitative understanding of a model of complex phenomena. The phenomena under investigation are often too small (e.g., molecules, ions) or too abstract (e.g., fractals, chaos) for students to experience directly in the "real world."

The tools of high performance computers and workstations, computer networks, and software make it possible for researchers to share their current research activities with high school teachers and students. While high performance computers are not physically located at the schools, communications networks enable students and teachers working at their own workstations to access the very large and computationally intensive simulation software and data bases operating at a remote university or research center. The network also makes it feasible for teachers and students in schools to interact with the scientific research teams located at the remote sites.

For example, researchers in computational physics develop and use software models of microscopic molecular dynamics. Students in a variety of different junior high and high schools are now using those same software models to develop their understanding of molecular behavior as well as the underlying probabilistic nature of such phenomena. Connected to the simulations is special front end software designed to guide students' learning, investigations, and experiments. One such simulation concerns the complex web of interactions between individual water molecules. Parameters such as temperature and pressure directly affect the hydrogen bonds in the network.

Students using an advanced workstation linked to a massively parallel supercomputer. Shown is a real-time molecular dynamics simulation of 216 water molecules. The blue lines represent a view of the spanning network of hydrogen bonds in liquid water. The simulation provides a "hands on" experience enabling the student to discover for herself the principles involved in hydrogen bond network formation in liquid water.

GLOSSARY

Algorithm
: A procedure designed to solve a problem. Scientific computing programs contain algorithms.

ASTA
: Advanced Software Technology and Algorithms

Bit
: Acronym for binary digit

BRHR
: Basic Research and Human Resources

Byte
: A group of adjacent binary digits operated upon as a unit (usually connotes a group of eight bits).

CAD
: Acronym for "computer aided design."

Computer Engineering
: The creative application of engineering principles and methods to the design and development of hardware and software systems.

Computer Science
: The systematic study of computing systems and computation. The body of knowledge resulting from this discipline contains theories for understanding computing systems and methods; design methodology, algorithms, and tools; methods for the testing of concepts; methods of analysis and verification; and knowledge representation and implementation.

Computational Science and Engineering
: The systematic application of computing systems and computational solution techniques to mathematical models formulated to describe and simulate phenomena of scientific and engineering interest.

DARPA
: Defense Advanced Research Projects Agency

Distributed systems
: A distributed system consists of multiple interconnected computers, each potentially containing data and programs, that can together carry out a single computation.

DOC
: Department of Commerce

DOD
: Department of Defense

DOE
: Department of Energy

ED
: Department of Education

EPA
: Environmental Protection Agency

FCCSET
: Federal Coordinating Council for Science, Engineering, and Technology

Flops
: Acronym for floating point operations per second. The term "floating point" refers to that format of numbers which is most commonly used for scientific calculation. Flops is used as a measure of a computing system's speed of performing basic arithmetic operations such as adding, subtracting, multiplying, or dividing two numbers.

Giga-
: 10^9 or billions of ... (e.g.: gigabits)

Grand challenge
: A grand challenge is a fundamental problem in science and engineering, with broad economic and scientific impact, whose solution could be advanced by applying high performance computing techniques and resources.

Heterogeneous system
: A heterogeneous system is a distributed system that contains more than one kind of computer.

HHS
: Health and Human Services

HPCC
: High Performance Computing and Communications

HPCCIT
: High Performance Computing, Communications, and Information Technology

HPCS
: High Performance Computing Systems

High Performance Computing
: High performance computing encompasses advanced computing, communications, and information technologies, including scientific workstations, supercomputer systems, high speed networks, special purpose and experimental systems, the new generation of large scale parallel systems, and applications and systems software with all components well integrated and linked over a high speed network.

Interagency Interim NREN
: Interagency Interim National Research and Education Network. The Interagency Interim NREN is an evolving operating network system. Near term (1992-1996) research and development activities will provide for the smooth evolution of this networking infrastructure into the future gigabit NREN.

Interoperability
: Interoperability is the effective interconnection of two or more different computer systems, databases, or networks in order to support distributed computing and/or data exchange.

Mega-
: 10^6 or millions of ... (e.g.: megaflops)

NASA
: National Aeronautics and Space Administration

Network
: Computer communications technologies that link multiple computers to share information and resources across geographically dispersed locations.

NIST
: National Institute of Standards and Technology

NOAA
: National Oceanic and Atmospheric Administration

NREN
: National Research and Education Network. The NREN is the realization of an interconnected gigabit computer network system devoted to HPCC.

NSA
: National Security Agency

NSF
: National Science Foundation

Ops
: Acronym for operations per second. Ops is used as a rating of the speed of computer systems and components. In this report ops is generally taken to mean the usual integer or floating point operations depending on what functional units are included in a particular system configuration.

Parallel processing
: Simultaneous processing by more than one processing unit on a single application.

Peta-
: 10^{15} or thousands of trillions of ... (e.g.: petabytes)

PMES
: Committee on Physical, Mathematical, and Engineering Sciences

Portable
: Portable computer programs can be run with little or no change on many kinds of computer systems.

Prototype
: The original demonstration model of what is expected to be a series of systems. Prototypes are used to prove feasibility, but often are not as efficient or well-designed as later production models.

Scalable
: Scalable computing systems have the property that the size of the system can be increased with an expected increase in performance proportional to the extent of the increase in size.

Supercomputer
: At any given time, that class of general-purpose computers that are both faster than their commercial competitors and have sufficient central memory to store the problem sets for which they are designed. Computer memory, throughput, computational rates, and other related computer capabilities contribute to performance. In addition, performance is often dependent on details of algorithms and on data throughput requirements. Consequently, a quantitative measure of computer power in large-scale scientific processing is difficult to formulate and may change as the technology progresses.

High Performance Computing and Communications Agency Points of Contact

Director
Computing Systems Technology Office
Defense Advanced Research
Projects Agency
3701 North Fairfax Drive
Arlington, VA 22203

HPCC Coordinator
Office of the Assistant Director
Computer & Information Science
& Engineering
National Science Foundation
1800 G Street, N.W.
Washington, DC 20550

HPCC Program Coordinator
Scientific Computing Staff
MS-G217, GTN
Office of Energy Research
Department of Energy
Washington, DC 20585

Director for HPCC
Office of Aeronautics and
Space Technology
National Aeronautics and
Space Administration
600 Independence Avenue, S.W.
Code RC
Washington, DC 20546

Associate Director
Division of Extramural Programs
National Library of Medicine
Bethesda, MD 20892

Office of the Chief Scientist
National Oceanic and
Atmospheric Administration
Herbert C. Hoover Building
Room 5128
14th and Constitution Avenue
Washington, DC 20230

HPCC Program Coordinator
Atmospheric Research and Exposure
Assessment Laboratory
ACMD, MD-80
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Computer Systems Laboratory
Office of the Director
HPCC Coordinator
B166 Technology Building
National Institute of Standards
and Technology
Gaithersburg, MD 20899

To obtain a copy of this document — send request to:

Federal Coordinating Council for Science, Engineering, and Technology
Committee on Physical, Mathematical, and Engineering Sciences
c/o National Science Foundation
Computer and Information Science and Engineering Directorate
1800 G Street, N.W.
Washington, D.C. 20550