DOCUMENT RESUME

ED 368 330 IR 016 567

TITLE	High Performance Computing and Communications: Toward a National Information Infrastructure.
INSTITUTION	Federal Coordinating Council for Science, Engineering and Technology, Washington, DC.
PUB DATE	94
NOTE	190p.; Illustrations and photographs may not copy adequately.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	Computer Centers; Computer Networks; *Educational Technology; Government Role; Information Networks; Program Evaluation; Research and Development; *Technological Advancement; *Telecommunications; Training
IDENTIFIERS	*High Performance Computing; High Performance Computing Act 1991; *National Information Infrastructure; National Research and Education Network

ABSTRACT

This report describes the High Performance Computing and Communications (HPCC) initiative of the Federal Coordinating Council for Science, Engineering, and Technology. This program is supportive of and coordinated with the National Information Infrastructure Initiative. Now halfway through its 5-year effort, the HPCC program counts among its achievements more than a dozen high-performance computing centers in operation nationwide. Traffic on federally funded networks and the number of local and regional networks connected to these centers continues to double each year. Teams of researchers have made substantial progress in adapting software for use on high-performance computer systems; and the base of researchers, educators, and students trained in HPCC technologies has grown substantially. The five HPCC program components in operation at present are: (1) scalable computing systems ranging from affordable workstations to large-scale high-performance systems; (2) the National Research and Education Network; (3) the Advanced Software Technology and Algorithms program; (4) the Information Infrastructure Technology and Applications program; and (5) the Basic Research and Human Resources program. Components of each are reviewed. Twenty-five tables present program information, with 72 illustrations. (SLD)

High Perf
Computing & Commu

Toward a
Information Infr

A. Nico Habermann (1932-1993)

This "Blue Book" is dedicated to the memory of our friend and colleague, A. Nico Habermann, who passed away on August 8, 1993, as this publication was going to press. Nico served as Assistant Director of NSF for Computer and Information Science and Engineering (CISE) for the past two years and led NSF's commitment to developing the HPCC Program and enabling a National Information Infrastructure. He came to NSF from Carnegie Mellon University, where he was the Alan J. Perlis Professor of Computer Science and founding dean of the School of Computer Science. He was a visionary leader who helped promote collaboration between computer science and other disciplines and was devoted to the development and maturation of the interagency collaboration in HPCC. We know Nico was proud to be part of these efforts.

HIGH PERFORMANCE COMPUTING AND COMMUNICATIONS

TOWARD A NATIONAL INFORMATION INFRASTRUCTURE

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

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 "High Performance Computing and Communications: Toward a National Information Infrastructure" prepared by the Committee on Physical, Mathematical, and Engineering Sciences (CPMES) of the Federal Coordinating Council for Science, Engineering, and Technology (FCCSET), to supplement the President's Fiscal Year 1994 Budget. This report describes the FCCSET Initiative in High Performance Computing and Communications.

The interagency HPCC Initiative is developing computing, communications, and software technologies for the 21st century. It is making progress toward providing the high performance computing and communications capabilities and advanced software needed in critical research and development programs. The HPCC Program is fully supportive of and coordinated with the emerging National Information Infrastructure (NII) Initiative, which is part of the President's and Vice President's Technology Initiative released February 22, 1993.

To enable the NII Initiative to build on the HPCC Program, the FCCSET CPMES High Performance Computing, Communications and Information Technology (HPCCIT) Subcommittee has included a new program component, Information Infrastructure Technology and Applications (IITA). It will provide for research and development needed to a address National Challenges, and it will also address problems where the application of HPCC technology can provide huge benefits to America. Working with industry, the participating agencies will develop and apply high performance computing and communications technologies to improve information systems for National Challenges in areas such as health care, design and manufacturing, the environment, public access to government information, education, digital libraries, energy management, public safety, and national security. IITA will support the development of the NII and the development of the computer, network, and software technologies needed to provide appropriate privacy and security protection for users.

The coordination and integration of the interagency research and development strategy for this Initiative and its coordination with other interagency FCCSET Initiatives has been led very ably by the CPMES and its HPCCIT Subcommittee. Donald A. B. Lindberg, Chair of the HPCCIT Subcommittee, his interagency colleagues, their associates, and staff are to be commended for their efforts in the Initiative itself and in this report.

John H. Gibbons
John H. Gibbons
Director

Office of Science and Technology Policy

Federal Coordinating Council for Science, Engineering, and Technology

Committee on Physical, Mathematical, and Engineering Sciences

Department of Agriculture
Department of Commerce
Department of Defense
Department of Education
Department of Energy
Department of Health and Human Services
Department of Interior
Environmental Protection Agency
National Aeronautics and Space Administration
National Science Foundation
Office of Management and Budget
Office of Science and Technology Policy

FCCSET Directorate

Charles H. Dickens, Executive Secretary
Elizabeth Rodriguez, Senior Policy Analyst

High Performance Computing and Communications and Information Technology Subcommittee

Name	Agency or Department	HPCCIT Subcommittee	HPCCIT Executive Committee *(see p. iii)*
Donald A.B. Lindberg	National Coordination Office	Chair	Chair
James Burrows	National Institute of Standards and Technology	Representative	Representative
John S. Cavallini	Department of Energy	Alternate	Alternate
Melvyn Ciment	National Science Foundation	Alternate	Alternate
George Cotter	National Security Agency	Representative	
Robin L. Dennis	Environmental Protection Agency	Alternate	

Name	Agency or Department	HPCCIT Subcommittee	HPCCIT Executive Committee *(see p. iii)*
Norman Glick	National Security Agency	Alternate	
A. Nico Habermann *(Deceased)*	National Science Foundation	Representative	Representative
Lee Holcomb	National Aeronautics and Space Administration	Representative	Representative
Paul E. Hunter	National Aeronautics and Space Administration	Alternate	Alternate
Steven Isakowitz	Office of Management and Budget	Representative	
Charles R. Kalina	National Coordination Office	Executive Secretary	
Norman H. Kreisman	Department of Energy	Alternate	Alternate
R.J. (Jerry) Linn	National Institute of Standards and Technology	Alternate	Alternate
Daniel R. Masys	National Institutes of Health	Representative	
Bruce McConnell	Office of Management and Budget	Alternate	
William L. McCoy	Federal Aviation Administration	Observer	
James A. Mitchell	Department of Education	Representative	
David Nelson	Department of Energy	Representative	Representative
Michael R. Nelson	Office of Science and Technology Policy	Representative	
Joan H. Novak	Environmental Protection Agency	Representative	

Name	Agency or Department	HPCCIT Subcommittee	HPCCIT Executive Committee (See Note below)
Merrell Patrick	National Science Foundation	Alternate	Alternate
Alex Poliakoff	Department of Education	Alternate	
Thomas N. Pyke	National Oceanic and Atmospheric Administration	Representative	Representative
John Silva	Advanced Research Projects Agency	Alternate	Alternate
Paul H. Smith	National Aeronautics and Space Administration	Alternate	Alternate
Stephen L. Squires	Advanced Research Projects Agency	Alternate	Alternate
John C. Toole	Advanced Research Projects Agency	Representative	Representative
Judith L. Vaitukaitis	National Institutes of Health	Alternate	

Note: The Advanced Research Projects Agency, the National Science Foundation, the Department and Energy, and the National Aeronautics and Space Administration hold permanent positions on the HPCCIT Executive Committee; the other two positions rotate among the other agencies.

Editorial Group for High Performance Computing and Communications 1994

Editor
Sally E. Howe
Assistant Director for Technical Programs, Nationa' Coordination Office

Program Component and Section Editors

Melvyn Ciment, *NSF, Past Co-Editor*	HPCC Program and Program Components
Stephen Griffin, *NSF*	Case Studies
Daniel R. Masys, *NIH*	Video
Merrell Patrick, *NSF*	BRHR

Calvin T. Ramos, *NASA*	ASTA
Stephen L. Squires, *ARPA*	HPCC Program, HPCS, IITA
Roger Taylor, *NSF*	NREN

Agency Editors

Norman Glick	National Security Agency
Stephen Griffin	National Science Foundation
Frederick C. Johnson	National Institute of Standards and Technology
Thomas Kitchens	Department of Energy
Daniel R. Masys	National Institutes of Health
James A. Mitchell	Department of Education
Joan H. Novak	Environmental Protection Agency
Thomas N. Pyke	National Oceanic and Atmospheric Administration
Calvin T. Ramos	National Aeronautics and Space Administration
Stephen L. Squires	Advanced Research Projects Agency

Copy Editor

Patricia N. Williams	National Coordination Office

Acknowledgments

In addition to the FCCSET, CPMES, HPCCIT Subcommittee, and Editorial Group, many other people contributed to this book, and we thank them for their efforts. We explicitly thank the following contributors to the Component and Agency Sections:

Robert Aiken, Lawrence Livermore National Laboratory, DOE
Robin Dennis, EPA
Darleen Fisher, NSF
Michael St. Johns, USAF, ARPA
Anthony Villasenor, NASA

Contributors to the Case Studies are acknowledged at the end of the Case Studies section.

We thank Joe Fitzgerald and Troy Hill of the Audiovisual Program Development Branch at the National Library of Medicine for their artistic contributions and for the preparation of the book in its final form, and we thank Patricia Carson and Shannon Uncangco of the National Coordination Office for their contributions throughout the preparation of this book.

Table of Contents

Table of Contents (cont.)

List of Tables

List of Illustrations

List of Illustrations (cont.)

High Performance Computing and Communications: Toward a National Information Infrastructure

The FY 1994 U.S. Research and Development Program

Executive Summary

The goal of the Federal High Performance Computing and Communications (HPCC) Program is to accelerate the development of future generations of high performance computers and networks and the use of these resources in the Federal government and throughout the American economy. Scalable high performance computers, advanced high speed computer communications networks, and advanced software are critical components of a new National Information Infrastructure (NII). This infrastructure is essential to our national competitiveness and will enable us to strengthen and improve the civil infrastructure, digital libraries, education and lifelong learning, energy management, the environment, health care, manufacturing processes and products, national security, and public access to government information.

The HPCC Program evolved out of the recognition in the early 1980s by American scientists and engineers and leaders in government and industry that advanced computer and telecommunications technologies could provide huge benefits throughout the research community and the entire U.S. economy. The Program is the result of several years of effort by senior government, industry, and academic scientists and managers to initiate and implement a program to extend U.S. leadership in high performance computing and networking technologies and to apply those technologies to areas of profound impact on and interest to the American people.

The Program is planned, funded, and executed through the close cooperation of Federal agencies and laboratories, private industry, and academia. These efforts are directed toward ensuring that to the greatest extent possible the Program meets the needs of all communities involved and that the results of the Program are brought into the research and educational communities and into the commercial marketplace as rapidly as possible.

Now halfway through its five-year effort, the Program's considerable achievements include:

- More than a dozen high performance computing centers are in operation nationwide. New scalable high performance systems are in operation at these centers, more advanced systems are in the pipeline, and new systems software is making these systems increasingly easy to use. Benchmark results improve markedly with each new generation of hardware and software and bring the Program closer to its goal of achieving sustained teraflop (trillions of floating point operations per second) performance.

- Traffic on federally-funded networks and the number of new local and regional networks connected to these networks continue to double every year.

  More than 6,000 regional, state, and local IP (Internet Protocol) networks in the U.S., and more than 12,000 worldwide, are connected; more than 800 of the approximately 3,200 two-year and four-year colleges and universities in the Nation are interconnected; and an estimated 1,000 high schools also are connected to the Internet. Traffic on the NSFNET backbone has doubled over the past year and has increased a hundred-fold since 1988.

  Already, HPCC research in the next generation of networking technologies indicates that the Program goal of sustained gigabit (billions of bits) per second transmission speeds will be achieved by no later than 1996.

- Teams of researchers have made substantial progress in adapting software applications for use on scalable high performance systems and are taking advantage of the increased computational throughput to solve problems of increasing resolution and complexity.

  Many of these problems are "Grand Challenges," fundamental problems in science and engineering with broad economic and scientific impact whose solution can be advanced by applying high performance computing techniques and resources. These science and engineering Grand Challenge problems have motivated both the creation and the evolution of the HPCC Program. Solution of these problems is critical to the missions of several agencies participating in the Program.

- T. e base of researchers, educators, and students trained in using HPCC technologies has grown substantially as agencies have provided training in these technologies and in application areas that rely on them.

The HPCC Program fully supports and is closely coordinated with the Administration's efforts to accelerate the development and deployment of the NII. The Program and its participating agencies will help provide the basic research and technological development to support NII implementation. To this end, several strategic and programmatic modifications have been made to the HPCC Program. The most significant of these is the addition of a new program component, Information Infrastructure Technology and Applications (IITA).

IITA is a research and development effort that will enable the integration of critical information systems and their application to "National Challenge" problems. National Challenges are major societal needs that computing and communications technology can help address in key areas such as the civil infrastructure, digital libraries, education and lifelong learning, energy management, the environment, health care, manufacturing processes and products, national security, and public access to government information. The IITA component will develop and demonstrate prototype solutions to National Challenge problems.

IITA technologies will support advanced applications such as:

- Routine transmission of an individual's medical record (including X-ray and CAT scan images) to a consulting physician located a thousand miles away.
- The study of books, films, music, photographs, and works of art in the Library of Congress and in the Nation's great libraries, galleries, and museums on a regular basis by teachers and students anywhere in the country.
- The flexible incorporation of improved design and manufacturing to produce safer and more energy-efficient cars, airplanes, and homes.
- Universal access by industry and the public to government data and information products.

The five HPCC Program components and their key aspects are:

High Performance Computing Systems (HPCS)

- Scalable computing systems, with associated software, including networks of heterogeneous systems ranging from affordable workstations to large scale high performance systems
- Portable wireless interfaces

National Research and Education Network (NREN)

- Widened access by the research and education communities to high performance computing and research resources
- Accelerated development and deployment of networking technologies

Advanced Software Technology and Algorithms (ASTA)

- Prototype solutions to Grand Challenge problems through the development of advanced algorithms and software and the use of HPCC resources

Information Infrastructure Technology and Applications (IITA)

- Prototype solutions to National Challenge problems using HPCC enabling technologies

Basic Research and Human Resources (BRHR)

- Support for research, training, and education in computer science, computer engineering, and computational science, and infrastructure enhancement through the addition of HPCC resources

HPCC Program agencies work closely with industry and academia in developing, supporting, and using HPCC technology. In addition, industrial, academic, and professional societies provide critical analyses of the HPCC Program through conferences, workshops, and reports. Through these efforts, Program goals and accomplishments are better understood and Program planning and management are strengthened.

The National Coordination Office (NCO) for High Performance Computing and Communications was established in September 1992 to provide a central focus for Program implementation. The Office coordinates the activities of participating agencies and organizations, and acts as a liaison to Congress, industry, academia, and the public. National Library of Medicine Director Donald A. B. Lindberg concurrently serves as Director of the NCO, in which capacity he reports directly to John H. Gibbons, the Assistant to the President for Science and Technology and the Director of the Office of Science and Technology Policy.

In the past year, the National Security Agency, in the Department of Defense, and the Department of Education have joined the HPCC Program, bringing to 10 the number of participating agencies. The total FY 1993 HPCC budget for these 10 agencies is $805 million. For FY 1994, the proposed HPCC Program budget for the 10 agencies is $1.096 billion, representing a 36 percent increase over the appropriated FY 1993 level

The HPCC Program is one of six multiagency programs under the Federal Coordinating Council for Science, Engineering, and Technology (FCCSET). The other five programs are Advanced Manufacturing; Advanced Materials and Processing; Biotechnology Research; Global Change Research; and Science, Mathematics, Engineering, and Technology Education. Each of these depends on the capabilities provided by HPCC.

The FY 1994 Program and this document are the products of the High Performance Computing, Communications, and Information Technology Subcommittee (HPCCIT) under the direction of the FCCSET Committee on Physical, Mathematical, and Engineering Sciences (CPMES).

Some Grand Challenge Problems

Build more energy-efficient cars and airplanes

Design better drugs

Forecast weather and predict global climate change

Improve environmental modeling

Improve military systems

Understand how galaxies are formed

Understand the nature of new materials

Understand the structure of biological molecules

Some National Challenge Application Areas

The civil infrastructure

Digital libraries

Education and lifelong learning

Energy management

The environment

Health care

Manufacturing processes and products

National security

Public access to government information

HPCC Program Overview

High performance computing has become a critical tool for scientific and engineering research. In many fields, computational science and engineering have become as important as 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 the growth of a trained pool of scientists and engineers.

The High Performance Computing and Communications (HPCC) Program has accelerated this progress through its investment in advanced research in computer and network communications hardware and software, national networks, and agency high performance computing centers. The 10 Federal agencies that participate in the HPCC Program (listed on page 12), along with their partners in industry and academia, have made significant contributions to addressing critical areas of national interest to both the Federal government and the general public.

High Performance computing is knowledge and technology intensive. Its development and application span all scientific and engineering disciplines. Over the last 10 years, a new approach to computing has emerged that can support a broad range of needs ranging from workstations for individuals to the largest scale highest performance systems that are used as shared resources. The workstations may also be small scale parallel systems and connect by high performance networks into clusters. Through the combination of advanced computing and computer communication networks with associated software, these systems may be scaled over a wide performance range, may be heterogeneous, and may be shared over large geographic distances by interdisciplinary research communities. The largest scale parallel systems are referred to as massively parallel when hundreds, thousands, or more processors are involved. Networks of workstations provide access to shared computing resources consisting of other workstations and larger scale higher performance systems.

High performance computing refers to the full range of supercomputing activities including existing supercomputer systems, special purpose and experimental systems, and the new generation of large scale parallel architectures.

A Research and Development Strategy
for High Performance Computing

Executive Office of the President
Office of Science and Technology Policy
November 20, 1987

The uses of and demand for advanced computer networking funded in part by the HPCC Program continue to expand. Progress and productivity in many fields of modern scientific and technical research rely on the close interaction of people located at distant sites, sharing and accessing computational resources across high performance networks. Their use of networks has provided researchers with unexpected and unique capabilities and collaborations. As a result, the scientific community is demanding even higher performance from networks. This increased demand includes increasing numbers of users; increasing usage by individual users; the need to transmit more information at faster rates; more sophisticated applications; and the need for increased security, privacy, and the protection of intellectual property.

The solution of "Grand Challenge" problems is a key part of the missions of many agencies in the HPCC Program. Grand Challenges are fundamental problems in science and engineering with broad economic and scientific impact whose solution can be advanced by applying high performance computing techniques and resources. These problems have and will continue to tax any available computational and networking capabilities because of their demands for increased spatial and temporal resolution and increased model complexity. The fundamental physical sciences, engineering, and mathematical underpinnings are similar for many of these problems. To this end, a number of multiagency collaborations are underway. (Examples of these problems are identified in Some Grand Challenge Problems on page 5 and in HPCC Research Areas on page 14.)

Although the U.S. remains the world leader in most of the critical areas of computing and computer communications technology, this lead is being threatened by countries that recognize the strategic nature of these technology developments. The HPCC Program leads the Federal investment in the frontiers of computing and computer communications technologies, formulated to satisfy national needs in science and technology, the economy, human resources, and technology transfer.

The HPCC Program will help provide the technological foundation for the National Information Infrastructure (NII). The NII will consist of computers and information appliances (including telephones and video displays), all containing computerized information, linked by high speed telecommunication lines capable of transmitting billions of bits of information in a second (an entire encyclopedia in a few seconds). A Nation of users will be trained to use this technology.

The computing and networking technology that will make the NII possible is improving at an unprecedented rate, expanding its effectiveness and even further stimulating our imaginations about how it can be used. Using these technologies, a doctor who seeks a second opinion could transmit a patient's entire medical record – X-rays and ultrasound scans included – to a colleague thousands of miles away, in less time that it takes to send a fax today. A school child in a small town could come home and through a personal computer reach into an electronic Library of Congress or a great art gallery or museum to view thousands of books, photographs, records, videos, and works of art, all stored electronically. At home, viewers could use equivalent commercial services to choose at any time to view one of thousands of films or segments of television programming.

The Administration is committed to accelerating the development and deployment of the NII, which the U.S. will need to compete in the 21st century. This infrastructure of "information superhighways" will revolutionize the way we work, learn, shop, and live, and will provide Americans the information they need, when they need it, and where they need it - whether in the form of text, images, sound, or video. It promises to have an even greater impact than the interstate highways or the telephone system. The NII will be as ubiquitous as the telephone system, but will be able to carry information at least 1,000 times faster. It will be able to transmit not only voice and fax, but will also provide hundreds of channels of interactive high-definition TV programming, teleconferencing, and access to huge volumes of information.

Thanks in part to the HPCC Program, this technology is already in use in many of our research laboratories where it is transforming the way research is done. Scientists and engineers can access information from computer databases scattered throughout the country and use high performance computers and research equipment thousands of miles away. Perhaps most importantly, researchers can collaborate and share information and tools with colleagues across the country and around the world as easily as if they were in the same room.

This same telecommunications and computing technology could soon be available to all Americans, provided there is adequate public and private investment and forward-looking government policies that promote its deployment and use.

The Administration believes that the Federal government has several important roles to play in the development of this infrastructure, which will be built and operated primarily by the private sector. The HPCC Program is a key part of the Administration's strategy for the NII. On February 22, 1993, the President and the Vice President unveiled a Technology Initiative that outlined the five parts of the Administration's strategy for building the National Information Infrastructure:

1. Implement the HPCC Program.

This Program is helping develop the basic technology needed for the NII.

2. Develop NII technologies.

Through the new Information Infrastructure Technology and Applications (IITA) component of the HPCC Program, industry, universities, and Federal laboratories will collaborate to develop technologies needed to improve effective use of the NII.

3. Fund networking pilot projects.

The Federal government will provide funding for networking pilot projects through the National Telecommunications and Information Administration (NTIA) of the Department of Commerce, which currently plays a key role in developing Federal communications policy. NTIA will provide matching grants to states, school districts, libraries, and other non-profit entities to purchase the computers and network connections needed for distance learning and for linking into computer networks such as the Internet. These pilot projects will demonstrate the benefits of networking in the educational and library communities In addition, to the extent that other agencies undertake networking pilot projects, NTIA will coordinate such projects, as appropriate.

4. Promote dissemination of Federal information.

Every year, the Federal government spends billions of dollars collecting and processing information (e.g., economic data, environmental data, and technical information). Unfortunately, while much of this information is very valuable, many potential users either do not know that it exists or do not know how to access it. The Administration is committed to using new computer and networking technology to make this information more available to the taxpayers who paid for it. This will require consistent Federal information policies designed to ensure that Federal information is made available at a fair price to as many users as possible while encouraging the growth of the information industry.

5. Reform telecommunications policies.

Government telecommunications policy has not kept pace with new developments in telecommunications and computer technology. As a result, government regulations have tended to inhibit competition and delay deployment of new technology and services. Without a consistent, stable regulatory environment, the private sector will hesitate to make the investments necessary to build the high speed national telecommunications network that this country needs to compete successfully in the 21st century. To address this and other problems, the Administration has created a White House-level interagency Information Infrastructure Task Force that will work with Congress, the private sector, and state and local governments to reach consensus on and implement policy changes needed to accelerate deployment of the NII.

Although the HPCC Program began as a research and development program, its impact is already being felt far beyond the research and education communities. The high performance computing technology developed under this Program has allowed users to improve understanding of global warming, discover more effective and safer drugs, design safer and more fuel-efficient cars and aircraft, and access huge "digital libraries" of information. The high speed networking technology developed and demonstrated by the HPCC Program has accelerated the growth of the Internet computer network and enabled millions of users not just to exchange electronic mail, but to access computers, digital libraries, and research equipment around the world. This technology, which allows Internet users to hold a video conference from their desks, is enabling researchers across the country to collaborate as effectively as if they were in the same room. The new IITA component of the HPCC Program will accelerate the deployment of HPCC technology into the marketplace and ensure that all Americans can enjoy its benefits.

Federal investment in new technologies is one of the best investments the government can make, one that will provide huge, long-term benefits in terms of new jobs, better health care, better education, and a higher standard of living. This is particularly true in the case of the National Information Infrastructure, which will provide benefits to all sectors of our economy. Few initiatives offer as many potential benefits to all Americans.

Several strategic and programmatic modifications have been made to the HPCC Program in order to enable the NII Initiative to build on the Program's original four components. The most significant of these is the addition of the new IITA program component. IITA consists of research and development to enable the integration of critical information systems and the application of these systems to "National Challenges," problems where the application of HPCC technology can provide huge benefits to all Americans.

These efforts will develop and apply high performance computing and communications technologies to improve information systems for National Challenges such as the civil infrastructure, digital libraries, education and lifelong learning, energy management, the environment, health care, manufacturing processes and products, national security, and public access to government information. Working with industry, IITA will support the development of the NII and the development of the computer, network, and database technology needed to provide appropriate privacy and security protection for users.

Overview of the Five HPCC Components

Five equally important, integrated components represent the key areas of high performance computing and communications:

HPCS – High Performance Computing Systems

Extend U.S. technological leadership in high performance computing through the development of scalable computing systems, with associated software, capable of sustaining at least one trillion operations per second (teraops) performance. Scalable parallel and distributed computing systems will be able to support the full range of usage from workstations through the largest scale highest performance systems. Workstations will extend into portable wireless interfaces as technology advances.

NREN – National Research and Education Network

Extend U.S. technological leadership in computer communications by a program of research and development that advances the leading edge of networking technology and services. NREN will widen the research and education community's access to high performance computing and research centers and to electronic information resources and libraries. This will accelerate the development and deployment of networking technologies by the telecommunications industry. It includes nationwide prototypes for terrestrial, satellite, wireless, and wireline communications systems, including fiber optics, with common protocol support and applications interfaces.

ASTA – Advanced Software Technology and Algorithms

Demonstrate prototype solutions to Grand Challenge problems through the development of advanced algorithms and software and the use of HPCC resources. Grand Challenge problems are computationally intensive problems such as forecasting weather, predicting climate, improving environmental monitoring, building more energy-efficient cars and airplanes, designing better drugs, and conducting basic scientific research.

IITA – Information Infrastructure Technology and Applications

Demonstrate prototype solutions to National Challenge problems using HPCC enabling technologies. IITA will support integrated systems technology demonstration projects for critical National Challenge applications through development of intelligent systems interfaces. These will include systems development environments with support for virtual reality, image understanding, language and speech understanding, and data and object bases for electronic libraries and commerce.

BRHR – Basic Research and Human Resources

Support research, training, and education in computer science, computer engineering, and computational science, and enhance the infrastructure through the addition of HPCC resources. Initiation of pilot projects for K-12 and lifelong learning will support expansion of the NII.

HPCC Program Goals

Extend U.S. technological leadership in high performance computing and computer communications.

Provide wide dissemination and application of the technologies to speed the pace of innovation and to improve the national economic competitiveness, national security, education, health care, and the global environment.

Provide key parts of the foundation for the National Information Infrastructure (NII) and demonstrate selected NII applications.

HPCC Agencies

ARPA – Advanced Research Projects Agency, Department of Defense

DOE – Department of Energy

ED – Department of Education

EPA – Environmental Protection Agency

NASA – National Aeronautics and Space Administration

NIH – National Institutes of Health, Department of Health and Human Services

NIST – National Institute of Standards and Technology, Department of Commerce

NOAA – National Oceanic and Atmospheric Administration, Department of Commerce

NSA – National Security Agency, Department of Defense

NSF – National Science Foundation

HPCC Program Strategies

Develop, through industrial collaboration, high performance computing systems using scalable parallel designs and technologies capable of sustaining at least one trillion operations per second (teraops) performance on large scientific and engineering problems such as Grand Challenges.

Support all HPCC components by helping to expand and upgrade the Internet.

Develop the networking technology required for deployment of nationwide gigabit speed networks through collaboration with industry.

Demonstrate the productiveness of wide area gigabit networking to support and enhance Grand Challenge applications collaborations.

Demonstrate prototype solutions of Grand Challenge problems that achieve and exploit teraops performance.

Provide and encourage innovation in the use of high performance computing systems and network access technologies for solving Grand Challenge and other applications by establishing collaborations to provide and improve emerging software and algorithms.

Create an infrastructure, including high performance computing research centers, networks, and collaborations that encourage the diffusion and use of high performance computing and communications technologies in U.S. research and industrial applications.

Work with industry to develop information infrastructure technology to support the National Information Infrastructure.

Leverage the HPCC investment by working with industry to implement National Challenge applications.

Enhance computational science as a widely recognized discipline for basic research by establishing nationally recognized and accepted educational programs in computational science at the pre-college, undergraduate, and postgraduate levels.

Increase the number of graduate and postdoctoral fellowships in computer science, computer engineering, computational science and engineering, and informatics, and initiate undergraduate computational sciences scholarships and fellowships.

HPCC Research Areas

- *Aerospace*
  - *Aircraft*
  - *Spacecraft*

- *Basic Science and Technology*
  - *Astronomy*
  - *Computers and network communications*
  - *Earth sciences*
  - *Molecular, atomic, and nuclear structure*
  - *The nature of new materials*

- *Education*

- *Energy*
  - *Combustion systems (in automobile engines, for example)*
  - *Energy-efficient buildings*

- *Environment*
  - *Pollution*
  - *Weather, climate, and global change prediction and modeling*

- *Health*
  - *Biological molecules*
  - *Improved drugs*

- *Library and Information Science*

- *Manufacturing*

- *Military Systems and National Security Systems*

HPCC Program Management

The HPCC Program is planned, funded, and executed with the close cooperation of Federal agencies and laboratories, private industry, and academia. These efforts are directed toward ensuring that to the greatest extent possible the Program meets the needs of all communities involved and that the results of the Program are brought into the research and educational communities and into the commercial marketplace as rapidly as possible.

The National Coordination Office (NCO) for High Performance Computing and Communications was established in September 1992. Donald A. B. Lindberg was selected to direct the NCO, while continuing to serve as Director of the National Library of Medicine. The NCO coordinates the activities of participating agencies and organizations, and serves as a liaison to Congress, industry, academia, and the public.

As Director of the NCO, Lindberg reports to John H. Gibbons, the Assistant to the President for Science and Technology and the Director of the Office of Science and Technology Policy. The Director of the NCO also chairs the CPMES High Performance Computing, Communications, and Information Technology (HPCCIT) Subcommittee. The Subcommittee 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. It is also informed by presentations by other Federal working groups and by public bodies.

Several HPCCIT working groups coordinate activities in specific areas. Individual agencies are responsible for coordinating these efforts:

- The Communications group, led by NSF, coordinates network integration activities and works closely with the Federal Networking Council (FNC). The FNC consists of representatives from interested Federal agencies, coordinates the efforts of governmen' HPCC participants and other NREN constituents, and provides liaison to others interested in the Federal Program.

- The Applications group, 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 Research group, led by ARPA, focuses on basic research, technology trends, and alternative approaches to address the technological limits of information technology. Its activities are integrated into the overall research program through meetings with the various technical communities.

- The Education group, led by NIH, coordinates HPCC education and training activities and provides liaison with other education-related efforts under FCCSET.

A Federal Networking Advisory Committee (FNCAC) supports the FNC by providing input and recommendations from broad communities and constituencies of the NREN effort. Pursuant to P.L. 102-194, the High Performance Computing Act of 1991, a High Performance Computing Advisory Committee will be established to support the overall HPCC Program. The Committee will improve communications and collaborations with U.S. industry, universities, state and local governments, and the public.

Each participating agency has focal points for addressing matters related to the HPCC Program. Organizational and management structures facilitating participation in the HPCC Program are described in the sections presenting individual agency programs. Many participating agencies have published documents about their HPCC programs and solicitations for research in HPCC areas; requests should be directed to the respective HPCC contacts listed at the end of this document.

U.S. industry, academia, and other developers and users of HPCC technology are involved in agency program planning and execution through advisory committees, commissioned reviews, self-generated commentary, and through direct participation in HPCC research and development efforts and as suppliers of technology.

The HPCC Program has benefited from the interest, advice, and specific recommendations of a variety of governmental, industrial, academic, professional, trade, and other organizations. These include:

- Federal organizations
  - Commerce, Energy, NASA, NLM, Defense Information (CENDI) Group
  - Congressional Research Service (CRS)
  - Department of Commerce HPCC Coordinating Group
  - Department of Health and Human Services Agency Heads
  - Federal Aviation Administration (FAA)
  - Federal Information Resources Management Policy Council (FIRMPOC)
  - Federal Library and Information Center Committee (FLICC) Forum on Information Policies
  - Food and Drug Administration (FDA) Center for Drug Evaluation and Research
  - House Armed Services Committee Staff
  - House Subcommittee on Science; Committee on Science, Space, and Technology
  - House Subcommittee on Technology, Environment, and Aviation; Committee on Science, Space, and Technology
  - House Subcommittee on Telecommunications and Finance; Committee on Energy and Commerce
  - NASA Advisory Council
  - National Telecommunications and Information Administration (NTIA)
  - NIH Information Resources Management Council
  - Securities and Exchange Commission (SEC)
  - Senate Science Subcommittee; Committee on Commerce, Science, and Transportation

- Federally-chartered institutions
  - National Academy of Sciences (NAS)/Computer Science and Technology Board (CSTB)
  - NAS/CSTB/NRENaissance Study Group
  - NAS/Institute of Medicine
  - NAS/National Research Council (NRC) Executive Board

- State organizations
  - Texas Education Network
  - Wisconsin Governor's Council for Science and Technology

- University organizations
  - Computing Research Association (CRA)
  - EDUCOM
- Professional societies
  - American Association of Engineering Societies (AAES)
  - American College of Cardiologists
  - American Institute of Medical and Biological Engineering (AIMBE)
  - Association of American Medical Colleges
  - Coalition of Academic Supercomputer Centers (CASC)
  - Computer Professionals for Social Responsibility (CPSR)
  - International Medical Informatics Association (IMIA)
- Industrial organizations
  - American Electronics Association (AEA)
  - Computer Systems Policy Project (CSPP)
  - Information Industry Association (IIA)
- Local organizations
  - Montgomery County High Technology Council
  - The Suburban Maryland Technology Council
- Other organizations with interest in HPCC
  - Coalition for Networked Information (CNI)
  - Foundation for Educational Innovation
  - Microelectronics and Computer Technology Corporation (MCC)
  - Science, Technology, and Public Policy Program
    John F. Kennedy School of Government, Harvard University
  - Supercomputing Center Directors
- Representatives from individual corporations, publishers, research laboratories, and supercomputing centers
- Foreign governments and organizations, including the British and Canadian governments

The HPCC Program has sponsored several major conferences and workshops. HPCC representatives have given presentations at a number of related conferences as well. These are itemized below.

Major HPCC Conferences and Workshops During FY 1993

Event	*Sponsor*
High Performance Computing Industry Presentations	*NASA*
Improving Medical Care: Vision of HPCCIT	*National Library of Medicine*
HPCC Applications in Chemistry	*NIH*
HPCC Comprehensive Review	*NASA*
Blue Ribbon Panel on HPCC	*NSF*
Workshop and Conference on Grand Challenge Applications and Software Technology	*Nine HPCC Agencies*

HPCC-Related Conferences

American Institute of Medical and Biological Engineering 1993 Annual Meeting

Coordinating Federal Health Care: Progress and Promise Workshop

Council on Competitiveness – Forum on Information Infrastructure

Electronic Industries Association Fifth Annual Federal Information Systems Conference

Georgetown University Medical School Computer Health Care Conference

HPCC-Related Conferences (cont.)

Institute of Electrical and Electronics Engineers (IEEE) Information Exchange Conference

International Council for Scientific and Technical Information (ICSTI) Annual Meeting

NII Round Table sponsored by 3Com Corporation

Research Consortium Inc. (RCI) North American Annual Member Executive Conference

Scientific Computing and Automation Conference

Supercomputing '92

3Com Interop Conference

Wide Area Information Services (WAIS) Conference

The dynamism and flexibility of the HPCC Program is illustrated by the incorporation of many of these recommendations into current Program plans. For example, the CSPP conducted its second intensive study of the HPCC Program structure and in January 1993 published "Perspectives on the National Information Infrastructure: CSPP's Vision and Recommendations for Action." Many of their recommendations formed the basis for plans in the HPCC Program's new IITA component.

Interdependencies Among Components

A complex web of interdependencies exists among the five components of the HPCC Program; success in each component is critical to the success of the others. Because of these interdependencies, maintaining balance among the components is crucial for Program success. The current balance is designed to foster that success as rapidly as possible. Some examples of the large number of interdependencies are given below.

Examples of two-component interdependencies:

HPCS and NREN – The development of routers for the NREN component and other advanced component technologies depend on HPCS research in scalable computing and component technologies.

HPCS and ASTA – The advanced computing systems with associated systems software are used by ASTA for Grand Challenge research.

HPCS and IITA – HPCS systems are used by IITA for National Challenge research.

NREN and ASTA – ASTA's Grand Challenge research helps to determine requirements for high performance networks for which NREN must provide new capabilities and technologies. Examples are distributed heterogeneous computing, scientific visualization, and the NSF metacenter research efforts.

NREN and IITA – NREN provides the networking technology base for IITA. An example is interactive video.

One example of a three-component interdependency:

HPCS, NREN, ASTA – Networks developed under the NREN component are used to access testbeds developed under the HPCS component in ASTA Grand Challenge research.

An example of a four-component interdependency:

HPCS, NREN, ASTA, IITA – Some computationally intensive IITA applications will use systems developed under the HPCS component, connected by NREN-developed networks, and Grand Challenge software developed under ASTA. An example is managing emergencies such as hurricanes.

Two five-component interdependencies:

BRHR dependencies – BRHR depends on each of the other components, both individually and in combination, for research subjects.

BRHR provides training and education in the four other components and their interrelationships.

Coordination Among Agencies

The participating agencies cooperate extensively in their efforts toward accomplishing HPCC Program goals, in part through the management vehicles described earlier and through the use of HPCC products from those other agencies wherever feasible. There are many other collaborations, including:

- Evaluation of early systems – these systems are procured primarily by NSF, DOE, and NASA; together with NIH, NSA, NOAA, and EPA, they are evaluated for mission-specific computational and information processing applications.
- DOE and NASA are coordinating testbed development to ensure that a diverse set of computing systems are evaluated.
- NIST is developing guidelines for measuring system performance, performance measurement tools, and software needed to monitor and improve the performance of advanced computing systems at HPCC-sponsored high performance computing centers.
- The gigabit testbeds (described on pages 37-39 in the NREN section).
- High speed networking experiments – ARPA, NASA, and NSA collaborate.
- Network security – ARPA, NSA, NIST, and other agencies collaborate.
- The NSF Supercomputer Centers. Other agencies jointly support and use these environments for their own missions and constituencies. One example is the NIH Biomedical Research Technology program in biomedical computing applications.
- The Concurrent SuperComputing Consortium (described on page 44 in the ASTA section).
- The National Consortium for High Performance Computing established by ARPA in cooperation with NSF (described on pages 44-45 in the ASTA section).
- The High Performance Software Sharing Exchange uses ARPA's wide area file system, NASA's distributed access to electronic data, and software repositories from DOE and NIST. These repositories are accessed by the other agencies.
- Joint agency workshops (for example, the recent "Workshop and Conference on Grand Challenge Applications and Software Technology").
- Representation on research proposal review panels (for example, DOE uses other agency experts in its Grand Challenge Review Committee).

Membership on the HPCCIT Subcommittee

The National Coordination Office and the HPCCIT Subcommittee actively encourage other Federal agencies (Departments, agencies within Departments, or independent agencies) to consider joining HPCCIT either as Official Members or as Observers.

Official Membership

If an agency proposes a program and the HPCCIT Subcommittee determines that the program meets the Evaluation Criteria (see below) and approves it, then the Subcommittee will recommend to the Committee on Physical, Mathematical, and Engineering Sciences (CPMES) that the agency be added to the HPCC Program and participate in the budget crosscut.

Observer Status

Upon request from the agency and approval by the HPCCIT Subcommittee, an agency may participate in the technical program and attend HPCCIT Subcommittee meetings.

Requests for membership or observer status should be directed to the National Coordination Office (listed on page 166 in the Contacts section).

Evaluation Criteria for the HPCC 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, information infrastructure, and basic research, to enable solution of the Grand Challenges and the National 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.*

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

Technology Collaboration with Industry and Academia

HPCC agencies work in partnership with each other and industry and academia to develop cost-effective high performance computing and communications technologies. The HPCC Program fosters the development and use of advanced off-the-shelf technology so that research results and products are simultaneously available to support both Federal agency missions and the computational needs of the academic and private sectors.

HPCC research is carried out in close collaboration with industry, particularly manufacturers of computer and communications hardware, software developers, and representatives from key applications areas. A number of mechanisms are used to facilitate these interactions: consortia, contracts, cooperative agreements such as Cooperative Research and Development Agreements (CRADAs), grants, and other transactions. Some examples of these collaborations are:

- The gigabit testbeds (described on pages 37-39 in the NREN section).
- The Concurrent SuperComputing Consortium (described on page 44 in the ASTA section).
- The National Consortium for High Performance Computing established by ARPA in cooperation with NSF (described on pages 44-45 in the ASTA section).
- NSF's National Supercomputer Centers and Science and Technology Centers and their metacenter efforts (described on page 66 in the NSF section).
- The Computational Aerosciences Consortium (described on page 85 in the NASA section).
- The Consortium on Advanced Modeling of Regional Air Quality (CAMRAQ) (described on pages 45-46 in the ASTA section).

The National Coordination Office for High Performance Computing and Communications meets with representatives from industry and academia. Major HPCC conferences and workshops held by HPCC agencies or addressing HPCC issues have included a number of representatives from industry and academia as well. Further details about these activities are presented in the Program Management section.

Agency Budgets by HPCC Program Components

FY 1993 Budget (Dollars in Millions)

Agency	HPCS	NREN	ASTA	BRHR	TOTAL
ARPA	119.5	43.6	49.7	62.2	**275.0**
NSF	25.9	40.5	108.0	50.8	**225.2**
DOE	10.9	10.0	65.3	14.8	**101.0**
NASA	11.1	9.0	59.1	2.9	**82.1**
NIH	3.0	4.1	31.4	8.0	**46.5**
NSA	34.8	3.2	5.4	0.2	**43.6**
NOAA	-	0.4	9.4	-	**9.8**
EPA	-	0.4	6.0	1.5	**7.9**
NIST	0.3	1.2	0.6	-	**2.1**
ED	-	2.0	-	-	**2.0**
TOTAL	**205.5**	**114.4**	**334.9**	**140.4**	795.2

FY 1994 Budget (Dollars in Millions)

Agency	HPCS	NREN	ASTA	IITA	BRHR	TOTAL
ARPA	151.8	60.8	58.7		71.7	**343.0**
NSF	34.2	57.6	140.0	36.0	73.2	**341.0**
DOE	10.9	16.8	75.1		21.0	**123.8**
NASA	20.1	13.2	74.2	12.0	3.5	**123.0**
NIH	6.5	6.1	26.2	24.0	8.3	**71.1**
NSA	22.7	11.2	7.6		0.2	**41.7**
NIST	0.3	1.2	0.6	24.0	-	**26.1**
NOAA	-	1.6	10.5		0.3	**12.4**
EPA	-	0.7	9.6		1.6	**11.9**
ED	-	2.0	-		-	**2.0**
TOTAL	**246.5**	**171.2**	**402.5**	**96.0**	**179.8**	**1,096.0**

Agency Responsibilities by HPCC Program Components

	HPCS	NREN	ASTA	IITA	BRHR
DOD/ARPA	• Scalable computing systems including networks of workstations, large scale parallel systems, & heterogeneous configurations • Portable scalable systems software • Advanced components, packaging, design tools, & rapid prototyping for VLSI systems	• Scalable Internet technologies with integration of different bitways • Gigabit technologies & testbeds	• Scalable algorithms & libraries • Software & system development environments for parallel & distributed heterogeneous systems	• *Scalable Internet technologies for universal & ubiquitous access* • *Mobile & wireless systems* • *Hypermedia systems with intelligent user interfaces*	• Support university research in computer & computational science
NSF	• Research on systems architecture • Early systems prototyping & evaluation • Design & prototyping tools • Distributed design & intelligent manufacturing • Optical communications systems & devices	• Coordinate NREN component • Manage & upgrade NSFNET backbone • Provide enhanced network security • Increase Internet connectivity • Provide Internet information services • Gigabit network testbeds	• R&D in algorithms & software technologies, distributed & heterogeneous environments • Grand Challenges research, including interdisciplinary collaboration • Scientific databases • Deploy parallel systems in NSF Supercomputer Centers, testbeds, etc. • Implement metacenter	• Support development of tools for human communication & collaboration • Develop multimedia software & educational materials • Support digital libraries research • Pilot projects for information-intensive applications • Support virtual reality applications	• Support university research in computer & computational science & engineering • Improve HPCC infrastructure at colleges & universities • Support education & training
DOE	• Procure & evaluate early systems • Research in parallel systems • Hierarchical data storage development	• Manage & upgrade ESnet: provide access to energy research facilities • Procure fast packet services • Research in gigabit networks. packetized video & voice	• Energy Grand Challenges research • Research in database technology, computational techniques • High Performance Computing Research Centers	• *Develop prototype energy-related applications*	• Support educational programs • Develop computational science texts
NASA	• Procure & evaluate early systems	• Manage & upgrade AERONet & NSInet • Procure fast packet services • Research in gigabit networks. satellite-based applications	• Grand Challenges research in aeronautics & the Earth's environment • HPCC Software Exchange	• Develop prototype applications for accessing NASA data. & for digital libraries. education. & manufacturing	• Support research in aeronautics & Earth sciences & high performance computing • Support educational programs

40

39

BEST COPY AVAILABLE

	HPCS	NREN	ASTA	IITA	BRHR
HHS/NIH	• Procure & evaluate early systems for biomedical applications	• Extend gigabit speed communications • Medical Connections – increase Internet connectivity • Access to X-ray images • Research in Unified Medical Language Systems & digital anatomy	• Grand Challenges research in molecular biology & biomedical imaging • Scalable software libraries • Scientific databases • "Visible Human" project • HIV research	• Develop testbeds linking medical facilities & sharing medical information • Develop anatomy visualization software • Develop virtual reality technology for simulating operations, etc. • Develop medical database technology	• Medical informatics fellowships • Biomedical science education • Collaboration among biomedical organizations • New and upgraded high performance computing resource centers
DOD/NSA	• Deploy experimental scalable systems, emphasizing interoperability & open systems • Develop mass archival storage • Promote interoperability	• Provide Internet connectivity • Develop network security technology • Gigabit network testbed	• Develop systems software for distributed parallel systems • Research in algorithms, programming languages, visualization, systems modeling	• *Investigate developmental technologies supportive of information infrastructure applications* • *Research in security, mass storage, database management, transaction processing*	• Support basic research in high performance computing & networking
DOC/NIST	• Develop instrumentation & methodology for performance measurement of systems & networks	• Coordinate development of standards for interoperability, user interfaces, security	• Algorithms & generic software for manufacturing & other applications • Guide to Available Mathematical Software	• Develop manufacturing applications	• Train users of scalable systems • Support university collaborations
DOC/NOAA		• Increase Internet access to systems & data • Provide information & assistance services	• Environmental Grand Challenges research • Acquire scalable system	• *Develop environmental monitoring, prediction, & assessment applications* • *Expand access to environmental data*	• Develop & evaluate materials for training users of environmental assessment tools • Support university research & educational programs
EPA		• Increase & upgrade Internet connectivity	• Environmental assessment Grand Challenges research • Acquire high performance system	• *Expand access to environmental data* • *Develop intelligent user interfaces to environmental applications*	
ED		• Develop PATHWAYS for Internet access by teachers & parents • Increase Internet connectivity			

Note: Italicized ARPA IITA entries are being funded as part of their FY 1994 HPCC budget; italicized DOE, DOD/NSA, DOC/NOAA, and EPA IITA entries are candidate activities and are not funded in the FY 1994 HPCC budget.

High Performance Computing Systems (HPCS)

The HPCS component produces scalable parallel computing systems in collaboration with industry and academia. Unlike dedicated, single processor architectures of the past, scalable parallel systems have the property that increases in size result in proportional improvement in performance. This is achieved by connecting multiple processors and memory units through a scalable interconnection structure. Scalable systems can be configured over a wide range that can deliver high performance computing to users at both small and very large scales.

Because the computing system designs are scalable, they can be used in smaller scale workstations. Such workstations may also have high performance graphics capabilities to enable visualization of a computational result and provide interactive interfaces to the user. These workstations may be linked to local networks connected to the Internet, a network of networks that includes high performance subnets linking higher performance and larger scale computing systems throughout the country.

HPCS focuses on the fundamental scientific and technological challenges of accelerating the advance of affordable scalable parallel high performance computing systems. Critical underlying technologies are developed in prototype form along with associated design tools. This allows evaluation of alternatives as the prototype systems mature. Evaluation continues throughout the research and development process, with experimental results used to refine successive generations of systems.

Scalable computing technologies used in combination with scalable networking technologies provide the technology base needed to address the Grand Challenges and the National Challenges. The necessary software technologies are developed by the ASTA and IITA components.

HPCS is composed of four elements to produce progressively more advanced and mature systems:

I. Research for Future Generations of Computing Systems

This element develops the underlying architecture, components, packaging (integration of electronics, photonics, power, cooling, and other components), systems software, and scaling concepts to achieve affordable high performance computing systems. These efforts ensure that the required advanced technologies will be available for the new systems and provide a foundation for the more powerful systems to follow. This element also produces the basic approaches for systems software, programming languages, and environments for heterogeneous configurations of workstations and high performance servers.

II. System Design Tools

This element develops computer aided design tools and the technology to allow multiple design tools to work together in order to enable the design, analysis, simulation, and testing of system components and modules. These tools make rapid prototyping of new system concepts possible. New design tools will be produced to enable the design of more advanced prototype systems using new technologies as they emerge.

III. Advanced Prototype Systems

Systems capable of scaling to 100 gigaops (billions of operations per second) performance have begun to emerge. Teraops (trillions of operations per second) performance designs will be demonstrated by the mid 1990s. Research in high performance systems focuses on reducing the cost and size of these systems so they can be used for a broader range of applications.

IV. Evaluation of Early Systems

Experimental systems will be placed at sites where researchers can provide feedback to systems and software designers. Performance evaluation criteria for systems and results of evaluations will be made widely available. Scalability enables small to medium size systems to be used for early performance evaluation and software development in preparation for larger scale applications. Larger scale systems are included in the ASTA component for applications such as the Grand Challenges.

HPCS Accomplishments

- Small, medium, and large scale systems developed under the HPCS component have been deployed and are being used in the ASTA component. This includes large systems deployed in various high performance computing centers and some systems installed in heterogeneous configurations.

  The small and medium scale systems are being used to develop algorithms and software, including fundamental building blocks for Grand Challenge problems and a wide variety of new scientific computation models. These prototypes are characterized by very fast routing and component technology, capable of scaling up to 100 gigaops system configurations.

- Scalable systems continue to be evaluated and refined, providing early feedback on hardware, operating systems, compilers, software development tools, input/output systems, and mass storage systems. This process has resulted in rapid upgrades in a commercial system to a scalable operating system based on very small and efficient software called microkernel technology. Extensions such as real time services and distributed and replicated file systems are under development.

- New technologies are providing a scalable, modular approach to mass storage performance and archiving needed in the new large scale parallel computing systems:

  - Prototype scalable mass storage systems that use parallel arrays of inexpensive disk drives to achieve both high aggregate data transmission rates and large storage capacity have been demonstrated. These systems demonstrate an approach that is the basis for a new generation of high performance file servers and mass storage systems that are internal to scalable parallel computing systems.

  - Petabyte mass storage systems, which can hold images from about 50 university libraries, are now available using commodity storage modules with automated robotic transfer to multiple read/write units. These systems help meet the dramatically increasing requirements for mass storage – from storing library information to storing remotely sensed satellite imagery.

- Evolving advanced component technology is being employed in early experimental computer systems. This technology will form the basis for a new generation of higher performance, physically smaller, and more affordable computing systems. Examples include the following:

  - Single chip nodes that integrate processing, storage and communications, new systems software, and new development environments, have been demonstrated. These have the potential of providing very cost-effective scalable computing using these single chip or fine grained nodes.

  - Multichip modules are being studied in experiments to determine the optimal design for future scalable units.

- Supporting technologies that enable the rapid design, prototyping, and manufacturing of HPCS systems have made an important contribution to HPCS progress. Examples of rapid prototyping facilities used by researchers include:

  - A laser direct write multichip module tool and associated design capability has been developed to reduce the prototyping time of new modules from months to two weeks. This enables designs to be developed more rapidly, and allows for the exploration of more effective and cost-effective alternatives.

  - New algorithms have been incorporated into design systems that extend synthesis to be applied to new technologies such as field programmable gate arrays and various integrated circuit technologies.

  - A model "factory of the future," linking advanced design technologies from workstations to large scalable computing, was completed and coupled to a prototype factory (described on pages 152-153 in the Case Studies section). These technologies form the basis of a new generation of computational prototyping, exploiting networked and distributed design processes for rapidly prototyping future generations.

- New competitive contractual mechanisms have been developed to enable the timely purchase of experimental systems. These joint government-industry research projects allow experimental use and early evaluation by a variety of user communities, which in turn provides early feedback to the hardware vendors and to developers of associated software technology. Such projects accelerate the maturation of these complex technologies in preparation for their larger scale use by the ASTA component.

National Research and Education Network Program (NREN)

The NREN* component will establish a gigabit communications infrastructure to enhance the ability of U.S. researchers and educators to perform collaborative research and education activities, regardless of their physical location or local computational and information resources. This infrastructure will be an extension of the Internet, and will serve as a catalyst for the development of the high speed communications and information systems needed for the National Information Infrastructure (NII).

The emerging NII will require: advances in the underlying foundations of networking technology and in generic networking services; the development and deployment of major new networking technologies; broader access to state-of-the-art high performance computing facilities; and early testing of new commercial products and services so that these can be effectively integrated into NREN associated networks.

The principal objectives of the NREN component are to:

- Establish and encourage wide use of gigabit networks by the research and education communities to access high performance computing systems, research facilities, electronic information resources, and libraries.

- Develop advanced high performance networking technologies and accelerate their deployment and evaluation in research and education environments.

- Stimulate the wide availability at reasonable cost of advanced network products and services from the private sector for the general research and education communities.

- Catalyze the rapid deployment of a high speed general purpose digital communications infrastructure for the Nation.

The NREN component's Interagency Internet and Gigabit Research and Development elements contribute to reaching these goals:

I. Interagency Internet

Near-term enhanced network services will be developed on the Nation's evolving networking and telecommunications infrastructure for use by mission agencies and the research and education communities. Interagency Internet activities include expansion of the connectivity and enhancement of the capabilities of the federally funded portion of today's research and education networks, and deployment of advanced high performance technologies and services as they mature. Coordinated among Federal agencies in cooperation with the private sector, this effort succeeds the Interim Interagency NREN element identified in previous reports about the HPCC Program.

* NSF has applied for registration of the service mark "NREN" with the U.S. Patent and Trademark Office.

The Interagency Internet is a network of networks, ranging from high speed cross-country networks, to regional and mid-level networks, to state and campus network systems. Its major Federal components are the national research agency networks listed below. When these agencies' "backbone networks" are upgraded, together they will form a national gigabit network to support research and education. This network may in turn serve as a prototype for broader national gigabit networks.

Major Federal Components of the Interagency Internet

NSFNET

NSF-funded national backbone network service

ESnet

DOE's Energy Sciences Network

NSI

NASA's Science Internet

ARPA's exploratory networks

The Interagency Internet and the other, non-federally-supported, portions of the Internet connect the Nation's communities of researchers and educators to each other; to facilities and resources such as computation centers, databases, libraries, laboratories, and scientific instruments; and to supporting organizations such as publishers and hardware and software vendors. The Interagency Internet also provides international connections that serve the national interest. These services will be continually enhanced as the Interagency Internet evolves.

The Interagency Internet also provides a testbed to stimulate the market for advanced network technologies such as Synchronous Optical Network (SONET) transmission infrastructure, Asynchronous Transfer Mode (ATM) cell switches, high speed routers, computer interfaces, and other communications hardware and software. These technologies are being developed by the telecommunications industry, routing vendors, and computer manufacturers, in collaboration with government and academia, as part of the NREN component of the HPCC Program. Through these efforts, the HPCC agencies will provide expertise in the systems integration of key technologies to form an integrated and interoperable high performance network system that will continue to meet the needs of the Nation's research and education communities. Once the initial development risks are reduced through this collaboration among government, industry, and academia, the U.S. communications community can build on these experiences and develop new products and services to serve the broader marketplace of NII applications.

II. Gigabit Research and Development

The Gigabit Research and Development element is a comprehensive program to develop the fundamental technologies needed for a national network with advanced capabilities and with a minimum gigabit per second (Gb/s) transmission speed. Gigabit research and development takes place in two ways: through a basic research program that provides the building blocks to move data at increasingly faster rates with novel techniques such as all optical networking; and through the deployment of testbed networks that use and prove the viability of these techniques. The testbeds provide an environment for the development of advanced applications targeted toward the solution of HPCC Grand Challenges.

As these technologies for networking hardware and software are developed and shown to be viable and cost-effective, they will be incorporated into the Interagency Internet. They will also provide a foundation for supporting Grand and National Challenges and their further extension to the National Information Infrastructure. Building on this foundation, the government and industrial partners will develop prototypes for a future high capability commercial communications infrastructure for the Nation.

NREN Component Management

Each agency implements its own NREN activities through normal agency structures and coordination with OMB and OSTP. All 10 agencies participate in the NREN component as users. Multiagency coordination is achieved through FCCSET, CPMES, HPCCIT, and the HPCCIT High Performance Communications working group.

Operation of the Interagency Internet is coordinated by the Federal Networking Council (FNC), which consists of agency representatives. The FNC and its Executive Committee establish direction, provide further coordination, and address technical, operational, and management issues through working groups and ad hoc task forces. The FNC has established the Federal Networking Council Advisory Committee, which consists of representatives from several sectors including library sciences, education, computers, telecommunications, information services, and routing vendors, to assure that program goals and objectives reflect the interests of these broad sectors.

Accomplishments

Increased Connectivity and Use of the Interagency Internet

The Interagency Internet has experienced tremendous growth in the number of connections (and hence the number of researchers) it supports, and in the amount of traffic that it carries. Significant leveraging of the Interagency Internet activities have resulted in the following:

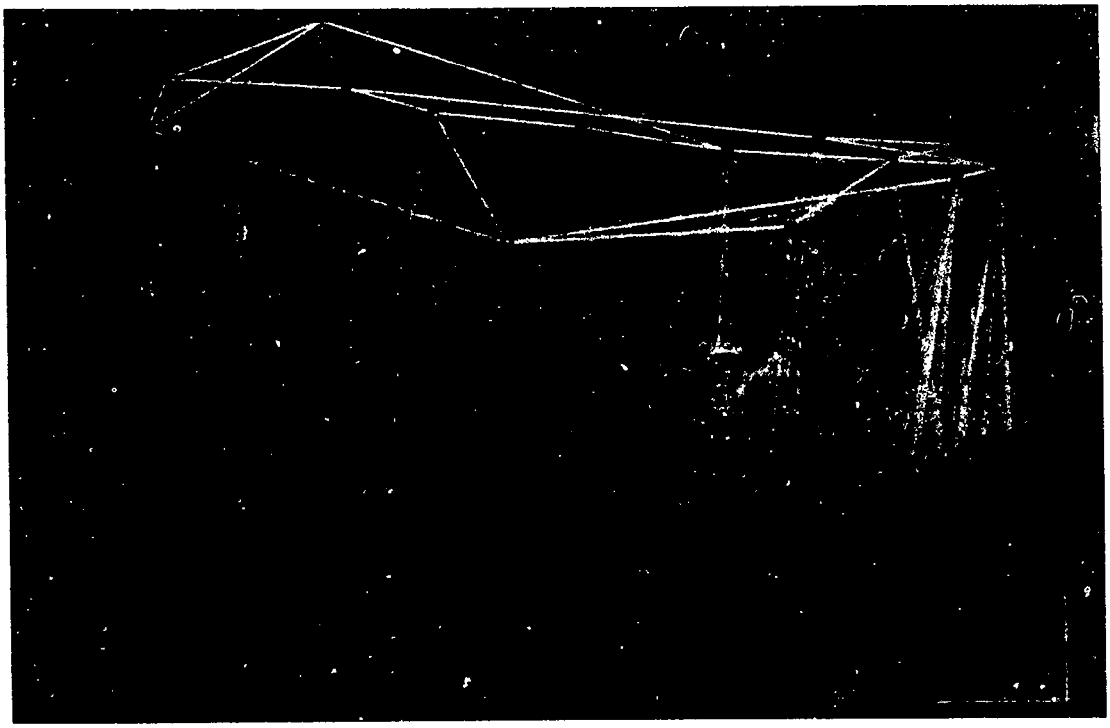
- More than 6,000 regional, state, and local IP (Internet Protocol) networks in the U.S. are connected as of March 1993. More than 12,000 such networks are connected worldwide.

- More than 800 of the approximately 3,200 two-year and four-year colleges and universities in the Nation are interconnected, including all of the schools in the top two categories ("Research" and "Doctorate") of the Carnegie classification.

- An estimated 1,000 U.S. high schools also are connected. The exact number is difficult to determine since regional networks have widely leveraged NSF and other agency funds to connect such institutions without direct agency involvement. For example, state initiatives in Texas and North Carolina proceed with little or no Federal funding or involvement.

Traffic on the NSFNET backbone has doubled over the past year, and has increased a hundred-fold since 1988. Improvements and upgrades to the network made by NSF have kept pace with the increased traffic and have advanced the state of network technology and operations.

ARPA, DOE, NASA, and NSF provide international connectivity to the Pacific Rim, Europe, Japan, United Kingdom, South America, China, and the former Soviet Union, for mission-specific scientific collaborations and general research and education infrastructure requirements. These links are of varying speeds, with many of the larger "fat pipes" cost-shared and co-managed by agencies requiring high speed connectivity.



Schematic of the interconnected "backbone" networks of NSF, NASA, and DOE, together with selected client regional and other networks. The backbone topology is shown on a plane above the outline of the U.S. Line segments connect backbone nodes with geographic locations where client networks attach.

Procurement of Fast Packet Services

The need for advanced networking services has been driven by the requirements of distributed scientific visualization and remote experiment control, and more recently by the phenomenal growth of multimedia applications. In order to satisfy these needs, DOE and NASA are in the process of jointly acquiring fast packet services based on new telecommunications industry-provided services (for example, ATM/SONET). The initial deployment will provide a 45 megabits per second (Mb/s) backbone service; upgrades to higher speeds are planned as soon as technology and budgets permit.

A key feature of this procurement is the use of as yet untariffed telecommunications provider services in an alpha test mode. In order to meet this procurement's deployment schedule, the telecommunications industry has accelerated its prototyping and deployment plans. The IITA component of the HPCC Program will later use these technologies via commercial services from the telecommunications industry and router vendor market.

NASA and ARPA will use NASA's geostationary Advanced Communications Technology Satellite (ACTS), launched in 1993, to provide even higher speed (for example, 622 Mb/s) ATM/SONET transmission to remote sites such as Alaska and Hawaii. The deployment will allow the HPCC agencies to gain experience in interfacing both terrestrial and satellite high speed communications systems. ARPA manages the development and deployment of the ACTS High Data Rate Terminals.

Internet Network Information Center (InterNIC)

In 1992, NSF issued a competitive solicitation for an Internet Network Information Center (InterNIC) to provide a variety of services to the worldwide Internet community. Awards were made to three organizations listed below to collaborate in providing these services. Information about how to connect to the Internet, pointers to network tools and resources, and seminars on various topics held across the country is available from the InterNIC Information Services (listed on page 167 in the Contacts section).

InterNIC Service Awards

Service	*Description*	*Award Made To*
Information	*General information about the Internet and how to use it*	*General Atomics/CERFnet*
Directory and Database	*Coordinated directory of the growing number of resources available on the Internet*	*AT&T*
Registration	*Registry of the growing number of networks connected to the Internet*	*Network Solutions Inc.*

Prompted by the recent development of network-based tools to seek out information by querying remote databases, NSF has established a Clearinghouse for Networked Information Discovery and Retrieval tools for assembling, disseminating, and enhancing such publicly available network tools. The clearinghouse complements the InterNIC.

Solicitation of the Next Generation NSFNET

Now that basic network services are readily and economically available commercially, NSFNET will, beginning in 1994, evolve into a very high speed national backbone for research applications requiring high bandwidth. In a new solicitation, NSF is requesting proposals to:

- Establish an unspecified number of Network Access Points (NAPs) where regional and other service providers will be able to exchange traffic and routing information.
- Establish a Routing Arbiter to ensure coherent and consistent routing of traffic among NAP participants.
- Establish a very high speed Backbone Network Service (vBNS) linking the NSF-supported Supercomputing Centers.
- Allow existing or realigned regional networks to connect to NAPs or Network Service Providers, who would connect to NAPs, for interregional connectivity.

The NAPs will provide connectivity to mid-level or regional networks serving both commercial and research and education customers and will also provide access to the vBNS.

With respect to regional networks, this solicitation addresses only interregional connectivity. Ongoing complementary intraregional support will continue and will be funded at constant or rising levels. These efforts include the Connections Program, which provides grants either to individual institutions or to more effective or more economical aggregates. A separate announcement to address intraregional connection of high-bandwidth users to the vBNS is planned for FY 1994.

Interconnecting the NSF Supercomputer Centers, the vBNS will be part of the Interagency Internet. It is expected that the vBNS will run at a minimum speed of 155 Mb/s and that low speed connections to NAPs will be routed elsewhere.

Gigabit Research Projects

By 1996, gigabit research will lead to an experimental nationwide network able to deliver speeds up to 2.4 billion bits per second to individual end user applications.

Ongoing research and development addresses communications protocols, resource allocation algorithms, network security systems, exploration of alternative network architectures, hardware and software, and the validation of that research by the deployment of several wide-area testbed networks. Several high data rate local area network testbeds will allow Federal agencies, industry, and academic researchers to explore innovative approaches to advanced applications such as global change research, computer imagery, and chip design.

In 1990, ARPA and NSF jointly began sponsoring five gigabit network research testbeds; all are expected to be operational by the end of 1993. The research at the five testbeds and at testbeds initiated subsequently (e.g., MAGIC sponsored by ARPA), focuses on network technology and network applications, with alternative network architectures, implementations, and applications of special interest.

Each testbed explores at least one aspect of high performance distributed computing and networking; together they seek to create and investigate a balanced high performance computing and communications environment.

Testbed teams consist of several government agencies (ARPA, DOE, Department of the Interior, NASA, NSF, state centers, and supercomputer centers), a number of universities, computer companies, and various local and long distance telephone companies that participate both as service providers and experimenters.

Other Projects

ARPA-sponsored consortia and individual projects are implementing novel networks that minimize or eliminate electronic content and replace it with optical technology. These efforts use alternative optical schemes for data rates in excess of 10 Gb/s. Industry partnerships guarantee a rapid transition of the most promising technologies into the commercial sector.

ARPA's Washington Area Bitway is a multiple-technology testbed in the Washington-Baltimore area that enables early experience with advanced network technologies. The first phase, called the Advanced Technology Demonstration Network (ATDnet) uses the best commercial prototypes of SONET/ATM technology to provide 100Mb/s-1Gb/s services to several DOD agencies and NASA. Applications include ACTS ground connections, imaging, and gigabit encryption. Later phases will demonstrate advanced optical technologies over the same optical fiber paths.

Another ARPA demonstration project will show the utility of asymmetric rate/asymmetric path (Cable TV and dialup) network access. Planned for the San Francisco Bay Area and for the Washington, D.C. area, the project is designed to explore a relatively inexpensive alternative to satisfying the "last mile" – that is, connections to homes and businesses – in high speed networking implementations.

NSF also supports Project ACORN, a collaborative research effort with an NSF Engineering Research Center and its industrial consortia, that is investigating lightwave networks of the 21st century. The project's TeraNet, a laboratory implementation and feasibility demonstration, will lead to a campus-wide field experiment involving leading-edge users. NSF has also begun to support research in all-optical networks.

Gigabit Testbeds

Testbed	*Description*	*Sites*	*Principal Research Participants and Collaborating Telecommunications Carriers*
Aurora*	Explore alternative network technologies, new network management and distributed system paradigms, and quality of service techniques for gigabit network multimedia applications.	Bellcore -- Morristown, NJ IBM -- Hawthorne, NY MIT -- Cambridge, MA U. Pennsylvania -- Philadelphia	Bell Atlantic Bellcore IBM MCI MIT NYNEX U. Arizona U. Pennsylvania
Blanca*	Investigate network control, real-time protocols, and distributed interactive applications including remote thunderstorm visualization, radio astronomy imaging, and multimedia digital libraries.	Lawrence Berkeley Laboratory (LBL) -- Berkeley, CA NCSA -- Champaign-Urbana, IL U. California at Berkeley U. Illinois -- Champaign-Urbana U. Wisconsin at Madison	Ameritech Astronautics AT&T Bell Atlantic LBL NCSA Pacific Bell U. California at Berkeley U. Illinois at Champaign-Urbana U. Wisconsin at Madison
Casa*	Investigate distributed large-scale supercomputing over wide-area gigabit networks using chemical reaction dynamics, geology, and climate modeling applications.	Caltech -- Pasadena, CA Jet Propulsion Laboratory (JPL) -- Pasadena, CA Los Alamos National Laboratory (LANL) San Diego Supercomputer Center (SDSC) UCLA	Caltech JPL LANL MCI Pacific Bell SDSC UCLA USWest
Nectar*	Investigate software and interfacing environments for gigabit-based heterogeneous computing and explore chemical processing and combinatorial optimization applications.	Carnegie-Mellon U. (CMU) -- Pittsburgh, PA Pittsburgh Supercomputer Center (PSC)	Bell Atlantic/Bell of Pennsylvania Bellcore CMU PSC
VISTANet*	Evaluate the application of gigabit networks and distributed computing techniques to interactive radiation therapy medical treatment planning.	BellSouth -- Chapel Hill, NC GTE -- Durham, NC MCNC (formerly Microelectronics Center of North Carolina) -- Research Triangle Park, NC U. North Carolina at Chapel Hill (UNC-CH)	BellSouth GTE MCNC North Carolina State U UNC-CH
Magic	Early demonstration of high speed terrain visualization with longer range plans to incorporate real-time sensor data into a real-time virtual world model and display.	Minnesota Supercomputer Center -- Minneapolis U. of Kansas in Lawrence U.S. Army's Future Battle Laboratory -- Fort Leavenworth, KS U.S. Geological Survey (USGS) -- Sioux Falls, SD	Digital Equipment Corp. DOE Earth Resources Observation Systems Data Center LBL MITRE Minnesota Supercomputer Center Northern Telecom Split Rock Telecom Sprint Southwestern Bell SRI International U. Kansas U.S. Army High Performance Computing Research Center U.S. Army's Future Battle Laboratory USGS

* *Information summarized from "A Brief Description of the CNRI Gigabit Testbed Initiative," Corporation for National Research Initiatives, 1895 Preston White Drive, Suite 100, Reston, Virginia, 22091-5434, (703) 620-8990.*

NREN FY 1994 Milestones

Bring MAGIC testbed into operation. Conduct initial terrain visualization demonstrations.

Demonstrate prototypes of gigabit ATM/SONET technology operating over fiber and satellite (using ACTS) media. Install initial gigabit network interconnection.

Bring all-optical testbed networks into operation.

Put medical, terrain visualization, and modeling applications on 100 megabit and gigabit class networks.

Complete ESnet and NSI fast packet upgrades.

Beta test high speed LAN interconnects with ESnet's fast packet WAN services.

Make awards to establish a series of NAPs, a Routing Arbiter, and a vBNS that links NSF supercomputer sites and is accessible from the NAPs.

Formulate programs and solicit proposals to support high bandwidth applications on the vBNS.

Continue improvements in U.S.-to-international connectivity.

Advanced Software Technology and Algorithms (ASTA)

The purpose of the ASTA component is to develop the scalable parallel algorithms and software needed to realize the potential of the new high performance computing systems in solving Grand Challenge problems in science and engineering. The early experimental use of this software on the new systems accelerates their maturation and identifies and resolves scaling issues on the most challenging problems.

The principal objectives of the ASTA component are to:

- Demonstrate large-scale, multidisciplinary computational results on heterogeneous, distributed systems, using the Internet to access distributed file systems and national software libraries.
- Establish portable, scalable libraries that enable transition to the new scalable computing base across different systems and their continued advance through successive generations.
- Develop a suite of software tools that enhance productivity (e.g., debuggers, monitoring and parallelization tools, run-time optimizers, load balancing tools, and data management and visualization tools).
- Promote broad industry involvement.

The ASTA component is composed of four elements:

I. Support for Grand Challenges

Prototype applications software will be developed to address computationally intensive problems such as the Grand Challenges. Solution of these problems is not only critical to the missions of agencies in the HPCC Program, but has broad applicability to the national technology base. Continuing increases in computational power enable researchers in government, industry, and academia to address problems of greater magnitude and complexity. Increased computational power enables:

- More realistic models as a result of higher resolution computational models. An example is weather models that show features on a local or regional scale, not just on a continental or global scale.
- Reduced execution times. Models that took days of execution time now take hours, enabling the user to modify the input more frequently, perhaps interactively and graphically, thus gaining insight faster. Reduced execution times also enable modeling over longer time scales (for example, 100 year climate models can now be executed in the same time it used to take for 10 year models).
- More sophisticated models, including models formerly too time consuming. The radiative properties of clouds can be included in climate models, for example.
- Lower cost solutions to specific problems, resulting in availability to a larger user community.

Multidisciplinary teams of computational scientists and disciplinary specialists from one or more Federal agencies and from industry and academia are working together to address these problems. Many of these Grand Challenges projects are cofunded and cosponsored by industry.

Significant progress has already been made toward solving many of these problems, and is expected to continue as the Program progresses. These new computational approaches are already being incorporated by industry into new products, services, and industrial processes such as testing and manufacturing.

II. Software Components and Tools

A complete collection of software components and tools is essential for a mature scalable parallel computing environment that supports portable software. These software components and tools must include standard higher level languages; advanced compilers; tools for performance measurement, optimization and parallelization, debugging and analysis; visualization capabilities; and interoperability and data management protocols.

As scalable parallel computing extends to a distributed computing environment, greater demands will be made upon the advanced network technologies developed and deployed through the NREN component. Operating systems and database management software for heterogeneous configurations of workstations and high performance servers will be developed along with remote procedure calls and interprocess communication protocols to support gigabit per second interconnections.

Broad industry involvement will be promoted through the identification and development of system structures using open interfaces. Industry involvement will also increase as computational approaches become available for more problem domains and more individuals are trained in high performance computing.

III. Computational Techniques

Portable scalable software libraries are being developed to enable software to move across different computational platforms and from one generation to the next. The performance and generality of the new computing technologies will be evaluated using a variety of experimental applications. Standard systems-level tools will be developed to support visualization of data and processes.

IV. High Performance Computing Research Centers (HPCRC)

The HPCRCs will deploy prototype large scale parallel computing facilities accessible over the Internet through the integration of advanced and innovative computing architectures (both hardware and software). Computational scientists working on Grand Challenge applications, software components and tools, and computational techniques will be able to access the largest advanced systems available in order to conduct a wide spectrum of experiments and scalability studies. Through the HPCRCs, the HPCC Program will evaluate prototype system and subsystem interfaces, protocols, advanced prototypes of hierarchical storage subsystems, and high performance visualization facilities. This work is done in cooperation with the Evaluation of Early Systems element of the HPCS component.

Major FY 1993 activities and accomplishments and FY 1994 plans

Grand Challenges Research

The majority of ASTA Grand Challenge research has focused on key computational models using first generation HPCC computers. This research will be extended to include applications software for multidisciplinary applications, hybrid computational models, and heterogeneous and distributed problems on second generation computer testbeds.

HPCC Agency Grand Challenge Research Teams

NSF – Computational science and engineering in academic disciplines

DOE – Energy and materials

NASA – Aerosciences and Earth and space sciences

NIH – Biomedical applications

NOAA – Atmospheric and oceanic computational modeling

EPA – Environmental modeling

Collaboration with industry and academia, primarily through CRADAs and consortia, is a high priority. Collaborative efforts are currently underway in the areas of environmental and Earth sciences; computational physics; computational biology; computational chemistry; material sciences; and computational fluid and plasma dynamics.

A four day "Workshop and Conference on Grand Challenge Applications and Software Technology" was held during May 1993 in Pittsburgh, PA. Funded by nine HPCC agencies, it brought together some 30 Grand Challenge teams. Discussions centered on multidisciplinary computational science research issues and approaches, the identification of major technology challenges facing users and providers, and refinements in software technology needed for Grand Challenges research applications. A special session was devoted to industrial applications, including the aeronautics, automotive, chemical, energy, financial, health care, and textile industries.

Software Sharing

The large collection of software needed to address the Grand Challenges and other computationally intensive problems is certain to grow at a rapid rate. Effective and efficient mechanisms to manage and reuse this software are essential.

Toward this end, NASA coordinates the collection of and access to a high performance computing software repository. The High Performance Computing Software Exchange uses ARPA's wide area file systems and NASA's distributed access to electronic data to connect software repositories in several Federal agencies. These repositories include netlib from NSF and DOE, and NIST's Guide to Available Mathematical Software (GAMS) (described on pages 148-149 in the Case Studies section). They will be expanded to include databases and bibliographic archives.

Mosaic, a hypermedia interface to repositories throughout the Internet, including gopher and WAIS, has been developed by the National Center for Supercomputer Applications (NCSA). In addition to providing a means to browse the Internet, it enables research results to be electronically published with text, images, image sequences, software libraries, and other resources. (Mosaic screens are shown on page 67 in the NSF section.)

An HPCC Software Exchange Prototype System is being built to establish foundations and guidelines for software submission standards, directories, indices, and Unix client/server interfaces. The critical processes and procedures to be used are derived from a 1992 Software Sharing Experiment, which identified and reviewed current software and set priorities and mechanisms for creating needed software.

Supercomputing Centers and Consortia

Two examples illustrate the collaborative efforts initiated through the HPCC Program:

- The Concurrent SuperComputing Consortium (CSCC) is an alliance of universities, research laboratories, government agencies, and industry that pool their resources to gain access to unique computational facilities and to exchange technical information, share expertise, and collaborate on high performance computing, communications, and data storage issues. CSCC members are:

  ARPA
  Argonne National Laboratory
  California Institute of Technology
  The Center for Research on Parallel Computation (an NSF Science and Technology Center)
  Intel's Supercomputer Systems Division
  Jet Propulsion Laboratory
  Los Alamos National Laboratory
  NASA
  Pacific Northwest Laboratory
  Purdue University
  Sandia National Laboratory

- The National Consortium for High Performance Computing was initiated by ARPA to accelerate advances in high performance computing technology. It focuses on 1) software and applications development, and 2) the fostering of interdisciplinary collaboration among DOD and other Federal agencies and laboratories, industry, academic partners, and other research and development organizations to solve important problems in defense and national security. Initiated in coordination and consultation with NSF using non-HPCC funds, this Consortium is an example of how HPCC technologies are being deployed on a national scale.

Scalable parallel systems are also located at the NSF sponsored Supercomputer Centers – the Cornell Theory Center (CTC), the National Center for Supercomputer Applications (NCSA) at Champaign-Urbana, the Pittsburgh Supercomputer Center, and the San Diego Supercomputer Center (SDSC). Plans are underway to establish a "metacenter" in which these Centers will be interconnected via high performance networks, allowing their supercomputing resources to be used as if they were an integrated high performance computing system.

NSF's Supercomputer Centers offer an interdisciplinary and collaborative environment for industrial and academic researchers. More than 100 corporations have formal affiliations with the Centers resulting in transition of enabling technologies and expertise. The Centers also work directly with vendors to identify and predict the needs of the computational science research community and to develop and test hardware and software systems. Special programs at the Centers are funded by other agencies, including the NIH Biomedical Research Technology program in biomedical computing applications at CTC and NCSA. Other federally-supported high performance computing activities not coordinated or budgeted through the HPCC Program, such as the National Center for Atmospheric Research (NCAR) in Boulder, CO, maintain important connections with HPCC.

DOE and NASA have also established HPCRCs. These include the DOE facilities at Los Alamos National Laboratory and Oak Ridge National Laboratory and NASA facilities at Ames Research Center and the Goddard Space Flight Center.

The Consortium on Advanced Modeling of Regional Air Quality (CAMRAQ) is working to develop pollution modeling systems, such as the regional environmental impact of pollutants including ozone, sulfate, nitrates, and articulates. Participants include:

Federal organizations
- Defense Nuclear Agency
- EPA
- NASA
- NOAA
  - Aeronomy Laboratory
  - Atmospheric Research Laboratory
  - National Meteorological Center
- National Park Service
- U.S. Army Atmospheric Sciences Laboratory

Federally-chartered institutions
- National Academy of Sciences/National Research Council

State and local organizations
- California Air Resources Board
- Northeast States for Coordinated Air Use Management
- South Coast Air Quality Management District

Industrial organizations

- American Petroleum Institute
- Chevron Research Corporation
- Electric Power Research Institute
- PG&E
- Southern California Edison Company

International organizations

- Environment Canada, Atmospheric Environment Service
- EUROTRAC/EUMAC
- Ontario Ministry of the Environment

DOE AND NASA are responsible for coordinating Grand Challenges applications software development, and are coordinating testbed development to ensure that a diverse set of computing systems are evaluated. Applications software and high performance computing benchmarks will be used by participating agencies to evaluate high performance computing system options. A key component of this effort will be to provide feedback to developers of teraops systems.

ASTA Milestones

FY 1993 - 1994

- Demonstrate initial multidisciplinary Grand Challenge applications.
- Deploy 100-gigaops systems to major high performance computing centers.

FY 1994

- Complete initial software components and tools for large scale systems.
- Deploy HPCC prototype libraries to the NII.

Beginning in FY 1994

- Deploy 300-gigaops systems to major high performance computing centers and enable their use.

Information Infrastructure Technology and Applications (IITA)

The IITA component of the HPCC Program is a research and development program intended to:

- Develop the technology base underlying a universally accessible National Information Infrastructure (NII).

- Work with industry in using this technology to develop and demonstrate prototype "National Challenge" applications.

National Challenges are major societal needs that high performance computing and communications technology can address in areas such as the civil infrastructure, digital libraries, education and lifelong learning, energy management, the environment, health care, manufacturing processes and products, national security, and public access to government information. The list of National Challenges is dynamic and will expand as the technologies and other applications mature.

Solution of these National Challenges requires a three-part technology base consisting of the following services, tools, and interfaces:

- Services that are common to and necessary for the efficient operation of the NII. For example, conventions and standards are needed to handle different formats of data such as text, image, audio, and video.

- Tools for developing and supporting common services, user interfaces, and NII applications.

- Intelligent user interfaces to NII applications.

The IITA component depends critically on the technologies already developed by the HPCC Program, and places its own set of demands on the Program. IITA efforts will strengthen the underlying HPCC technology base, broaden the market for these technologies, and accelerate industry development of the NII.

The Federal agencies that participate in the HPCC Program will work with industry and academia to develop these technologies. The NII, however, will be built and operated primarily by the private sector, which can form new markets for products and services enabled by the emerging NII. This joint effort by government, industry, academia, and the public to develop the NII and address the National Challenges will require:

- Deployment of substantially more high performance computing systems of increasingly higher performance.

- A nationwide communications network with vastly greater capacity for connections and throughput than today's Internet.

- Further development and wider deployment of applications software for computationally intensive problems such as the Grand Challenges.

- Education of large numbers of developers of NII technologies and training of a Nation of users.

Users of the early NII will be able to take advantage of small to moderate capacity computers and slow to medium speed communications, provided they have high quality user interfaces and access to the applications. As user interfaces improve, more computing and communications performance may be required. This can be achieved through the continual advances in the underlying technology developed under the HPCC Program.

The HPCC Program's original focus on research and development will continue to play a pivotal role in enhancing the Nation's computing and communications capabilities. For example, the Grand Challenges will continue to provide the scientific focus for critical computing technologies because of their profound and direct impact on fundamental knowledge. The IITA component will enable the extension of these technologies and the development of National Challenge applications that have immediate and direct impact on critical information systems affecting every individual in the Nation. Distinctions between National Challenges and Grand Challenges are shown in the table on the next page.

The following two examples illustrate potential applications of the NII.

A Medical Emergency

Having taken ill, a traveler is hospitalized and undergoes tests, including X-rays, CAT scans, and MRI. At the same time, the attending medical professionals quickly retrieve test results from the traveler's last physical examination. The images are compared, diagnoses made, and treatments prescribed. This scenario is difficult if not impossible to implement today, in part because diagnostic images are commonly not in computer-readable form and network speeds are generally too slow to transmit large three-dimensional image data sets.

Truly remote medical care will depend on services, standards, tools, and user interfaces to store, find, transmit, manipulate, display (and superimpose), compare, and analyze three-dimensional image data from several sources. Diagnostic test results and large image data sets from the physical examination must be available on computers that can be accessed from the hospital's computers over a communications network; they must be retrieved quickly; the scientific data used in guiding the diagnosis and treatment must also be available from electronic libraries and must be quickly retrieved; and the privacy of these patient records must be protected. All of this supposes completion of rather extensive and complex inter-professional medical arrangements. In addition, it must be done using a user interface customized for the practice of "distance medicine," including collaborations among different sources of expertise.

Contrasts Between Grand Challenges and National Challenges

	Grand Challenges	National Challenges
Focus	*Computation intensive*	*Information intensive*
Users	*Computer scientists and computational scientists, extending to scientists and engineers (numbering in the millions)*	*Information and application users in major national sectors, extending to all sectors (numbering in the hundreds of millions)*
Distribution of HPCS Resources	*Focus on largest systems at "centers" and smaller systems for software development*	*Focus on distributed systems of moderate scale with many users, scaling with increasing number of users*
Main HPCS Use	*Workstation client/server systems for computing and systems development, scaling with scientific needs*	*Workstation client/server systems for access to and processing of information with extensions to mobile and wireless systems*
Main NREN Use	*Share computing resources, data, and software; collaboration*	*Support distributed systems*
Main ASTA Use	*Scientific software for computationally-intensive applications*	*Computationally-based services*
Privacy and Security	*Desirable but not critical to deployment; essential in long term*	*Essential for deployment*
Copyright	*Desirable but not critical to deployment*	*Essential for deployment*
Network Performance	*High among largest scale computer systems and users needing visualization*	*Nominal to all involved, moderate to most when needed, high to those with greatest need*

A Weather Emergency

Using a real-time weather forecast for the area 20 miles directly ahead, a trucker diverts to an alternate route and reduces by hours the potential delay in delivering critically needed parts to a company that uses a just-in-time inventory system. Relayed from a weather forecast center to the truck's on board navigation system, this highly accurate forecast that pinpoints developing adverse weather conditions is made possible by the use of new computer weather prediction models that exploit the capabilities of high performance, massively parallel computing systems.

Already, advances in computing and communications technologies have led to significant improvements in weather forecasting. As illustrated in the recent case of Hurricane Andrew in Florida, this improved forecasting can save lives as well as millions of dollars in evacuation costs through better targeting of evacuation efforts.

Services, standards, tools, and user interfaces are required to build and support systems for acquiring large amounts of three-dimensional environmental data from different sources (e.g., in situ and remote sensing observations). These systems must also support high resolution modeling using these data, incorporating improved representations of the physical environment, and a real-time information dissemination capability to provide detailed forecast information for hundreds of different locations to thousands of users.

Unlike the first example, the user community for environmental information is larger and more diverse. Weather forecasts are needed by the general public, and for aviation, ship navigation, and agriculture, for example, while HPCC-funded researchers use much of the same observational data to model global change. Starting with user interfaces tailored to different kinds of users, individual users customize them for their own needs. The delivery and use of environmental information for this broad range of applications is performed by a partnership of government and value added private sector information companies, all part of the NII.

The IITA component will enable the development of an integrated infrastructure so that these two apparently unrelated applications can work together efficiently. This infrastructure includes:

- A networked computing base that provides appropriate performance.
- Methods to provide security, privacy, and copyright protection, and other services such as "digital signatures" to authenticate transactions.
- Technical conventions and standards (especially for databases).
- Tools to build and support the user interfaces.
- Tools to build and support the applications themselves.

The IITA component will demonstrate these technologies through testbed and pilot projects. These projects will evaluate new technologies, provide training in their use, and demonstrate specific National Challenge applications. Successful projects will serve as models to be further refined and engineered for larger scale deployment.

In order to facilitate the deployment of the NII by the private sector, the government will work closely with industry, academia, and users on all aspects of the IITA component. The private sector is expected to deploy many of its own applications, in areas such as commerce and entertainment. The government's goal is that from the user's point of view these applications will be integrated as seamlessly as possible into a single NII, with appropriate use restrictions and protections incorporated as needed.

The IITA component is organized into four interrelated elements. Each builds on the foundation of the HPCC Program and, in large measure, builds on its predecessor.

I. Information Infrastructure Services

These are the basic services and interfaces, and the underlying technical conventions and standards, that provide the common foundation for a broad range of information technology-based applications. Building upon the Interagency Internet, these modular units will in turn be the building blocks of the NII applications. These include:

- Data formats and object structures (including single- and multimedia formats such as image, audio, and video).
- Methods for managing distributed databases.
- Services that provide access to electronic libraries and computer-based databases.
- Methods for exchanging data (e.g., data compression) and integrating data (e.g., merging and overlaying images) from one or more electronic libraries and computer-based databases.
- Services to search for and retrieve data and objects.
- Protocols and processes such as "digital signatures" needed to obtain appropriately secure and legal access to information (including protection of copyrighted material).
- Usage metering mechanisms to enable implementation of payment policies.
- High integrity, fault-tolerant, trusted, scalable computing systems.
- Protocols and processes needed to obtain the appropriate communications speed and bandwidth.

II. Systems Development and Support Environment

This element includes a comprehensive suite of software tools and applications methods such as software toolkits and software generators for use by computer programmers, tools and methods for integrating elements of virtual reality systems, collaboration software systems, and applications-specific templates and frameworks. They will be used to:

- Interface to existing services (for example, existing search services).
- Develop new distributed services for the Information Infrastructure Services element and for the NII.

- Develop generic user interfaces, including templates and frameworks, to facilitate the use of the services provided by the Information Infrastructure Services element in the development of advanced and customized user interfaces by the Intelligent Interfaces element described below.

- Develop generic applications, including architectures and frameworks, for use in the Intelligent Interfaces element and for use by applications developers in implementing applications such as the National Challenges and other applications in the NII.

This element also includes the systems simulators and modeling methods to be used in designing the technology underlying the NII.

III. Intelligent Interfaces

In the future, high level user interfaces will bridge the gap between users and the NII. A large collection of advanced human/machine interfaces must be developed in order to satisfy the vast range of preferences, abilities, and disabilities that affect how users interact with the NII.

Intelligent interfaces will include elements of computer vision and image understanding; understanding of language, speech, handwriting, and printed text; knowledge-based processing; and multimedia computing and visualization. In order to enhance their functionality and ease of use, interfaces will access models of both the underlying infrastructure and the users. Just as people now do their own "desktop publishing," they will have their own "desktop work environments," environments that will extend to mobile and wireless networking modes. Users will be able to customize these environments, thereby reducing reliance on intermediate interface developers.

IV. National Challenges

National Challenges are fundamental applications that have broad and direct impact on the Nation's competitiveness and well-being. They will enable people to handle the increasing amounts of information and the increasing dynamics of the 21st century.

Using selected HPCC enabling technologies and the technologies developed by the other IITA elements, this element will use pilot projects to develop "customized applications" in areas such as the civil infrastructure, digital libraries, education and lifelong learning, energy management, the environment, health care, manufacturing processes and products, national security, and public access to government information. Detailed goals of four of these applications areas are as follows:

Digital Libraries

Develop systems and technology to:

- Enable electronic publishing and multimedia authoring.

- Provide technology for storing petabytes of data for nearly instantaneous access by users numbering in the millions or more.

- Quickly search, filter, and summarize large volumes of information.

- Quickly and accurately convert printed text and "pictures" of all forms into electronic form.
- Categorize and organize electronic information in a variety of formats.
- Use visualization to quickly browse large volumes of imagery.
- Provide electronic data standards.
- Simplify the use of networked databases in the U.S. and worldwide.

Prototypic scientific data bases, including remote-sensing images, will be developed. Librarians and other users will be trained in the development and use of this technology.

Education and Lifelong Learning

Conduct pilot projects that connect elementary and secondary schools to networks through which students and teachers can:

- Communicate with their peers and with students and faculty at colleges and universities across the country.
- Access information databases and other computing resources.
- Use authoring tools to embody the experiences of the best teachers in systems that others can use.
- Have greater access to the NII technologies, enabling them to develop and use them more effectively.
- Enable future generations to be literate in information technology so that they will be prepared for the 21st century and beyond.

Health Care

Develop and provide:

- Access to networks that link medical facilities and enable health care providers and researchers to share medical information.
- Technology to visualize and analyze human anatomy and to simulate medical procedures such as operations.
- The ability to treat patients in remote locations in real-time by having "distance collaborations" with experts at other medical facilities.
- Technology by which health care providers can readily access databases of medical information and literature.

- Technology to store, access, and transmit patients' medical records and protect the accuracy and privacy of those records when doing so.

Manufacturing

Research and development to:

- Prototype advanced computer-integrated manufacturing systems and computer networks linking these systems.
- Work with industry in implementing standards for these advanced manufacturing operations, and process and inventory control.
- Transition the manufacturing process to the new scalable computing and networking technology base.
- Train management and employees in advanced manufacturing.

Basic Research and Human Resources (BRHR)

The BRHR component is designed to increase the flow of innovative ideas by encouraging investigator-initiated, long-term research in scalable high performance computing; to increase the pool of skilled and trained personnel by enhancing education and training in high performance computing and communications; and to provide the infrastructure needed to support these research and education activities.

The BRHR component is organized into four elements:

I. Basic Research

This element supports increased participation by individual investigators in conducting disciplinary and multidisciplinary research in computer science, computer engineering, and computational science and engineering related to high performance computing and communications. Research topics include:

- Foundations of future high performance computing systems.
- High performance hardware components and systems, high density packaging technologies, and system design tools.
- Mathematical models, numeric and symbolic algorithms, and library development for scalable and massively parallel computers.
- High level languages, performance prediction models and tools, and fault tolerant strategies for parallel and distributed systems.
- Large scale database processing; knowledge based processing; image processing; digital libraries; visualization; and multimedia computing.
- Resource management strategies and software collaboratory environments for scalable parallel and heterogeneous distributed systems.

II. Research Participation and Training

This element addresses the human resources pipeline in the computer and computational sciences, at undergraduate, graduate, and postdoctoral (training and re-training) levels. Activities include:

- Workshops, short courses, and seminars.
- Fellowships in computational science and engineering and experimental computer science.
- Career training in medical informatics through grants to young investigators.

- Institutional training and postdoctoral programs; knowledge transfer exchange programs at national laboratories, centers, universities, and industry.

- Software dissemination through national databases and libraries.

III. Infrastructure

This element seeks to improve university and government facilities for computer science, computer engineering, and computational science and engineering research related to high performance computing. Activities include:

- Improvement of equipment in computer science, computer engineering and computational science and engineering academic departments, centers, and institutions; development of scientific databases and repositories.

- Distribution of integrated system building kits and software tools.

IV. Education, Training, and Curriculum

This element seeks to expand existing activities and initiate new efforts to improve K-12, undergraduate, and graduate level education and training opportunities in high performance computing and communications technologies, 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. Activities include:

- Bringing people, especially teachers, to national centers and laboratories, for summer institutes and other training, technology transfer, and educational experiences.

- Utilizing professional scientists and engineers to provide curriculum development materials and instruction for high school students in the context of high school supercomputer programs, supercomputer user workshops, summer institutes, and career development informatics for health sciences.

BRHR Component Implementation

Each agency that participates in the BRHR component sponsors research participation and education/training programs designed to meet specific mission needs. Some of these activities are as follows.

- ARPA supports basic research in such areas as high performance components, high density packaging, scalable concepts, system design tools and foundations of petaops systems.

- NSF's basic research programs promote innovative research on the foundation sciences and technologies of HPCC as well as specific disciplinary activities in HPCC. NSF coordinates its basic research and infrastructure activities to foster balance in the multiagency HPCC Program.

Through its "research experiences for undergraduates," SuperQuest, postdoctoral, graduate fellowship, educational and minority infrastructure programs, NSF addresses long term national needs in HPCC.

- DOE supports basic research to advance the knowledge of mathematical, computational, and computer sciences needed to model complex physical, chemical, and biological phenomena involved in energy production and storage systems. DOE also is actively involved in education and training activities at all levels.

- NASA conducts basic research through NASA research institutes and university block grants, including support at the graduate and postdoctoral level.

- NIH supports basic research and training in the use of advanced computing and network communications. Predoctoral and postdoctoral grants for career training in medical informatics are being expanded.

- NOAA conducts basic research in computational fluid dynamics applications in atmospheric and oceanic processes.

- EPA sponsors targeted fellowships and basic research activities, and develops and evaluates training methods and materials to support transfer of advanced environmental assessment tools to Federal, state, and industrial sectors.

Partnerships with industry, universities, and government help accomplish BRHR objectives.

FY 1993 Accomplishments

In FY 1993, more than 1,000 research awards fund the following activities:

- Basic research in high performance computational approaches to materials processing, molecular structures, fluid dynamics, and structural mechanics.

- Basic research on scalable parallel systems in fundamental areas of mathematical models, algorithms, performance evaluation techniques, databases, visualization and multimedia computing, digital libraries, and collaboratory technologies.

- An increased number of computer and computational science and engineering postdoctoral awards and graduate fellowships.

- High school honors programs, teacher training programs and "research experience for undergraduates" in HPCC.

- The introduction of the computational science textbook, which involved 24 authors in 10 different disciplines, into classrooms as part of a pilot project for Computer Science for High School Teachers.

- Institutional infrastructure awards to support experimental and novel high performance computing research at universities and national laboratories.

- Educational and minority infrastructure awards in undergraduate institutions.

FY 1994 Plans

- Develop a program to apply the principles of artificial intelligence to advanced intelligent manufacturing.

- Develop a program to integrate virtual reality technology into high performance computing and communications systems.

- Develop an initiative in digital libraries.

- Increase research in real time, three-dimensional imaging and multimedia computing.

- Increase support of information intensive applications of HPCC technologies in health care, information libraries development, education, and manufacturing.

- Increase support for scalable parallel computers.

- Increase education and training activities in HPCC through establishment of network-based educational testbeds.

- Increase number of postdoctoral and graduate fellowship awards.

Advanced Research Projects Agency (ARPA)

As the HPCC Program reaches the middle of its initial five-year phase, the ARPA program is shifting focus from stimulating the development of the new scalable computing technology base and early experimental use toward developing the technologies needed to enable a broad base of applications and users, including their extension to a National Information Infrastructure. In addition, the foundations for future generations of computing systems involving even more advanced technologies are being developed.

The current scalable computing technology base is characterized by the first 100 gigaop class computing systems that are being experimentally used on a wide variety of problems in the scientific and engineering research communities. Scalable operating systems are enabling software development on high performance workstations connected to higher performance servers through networks.

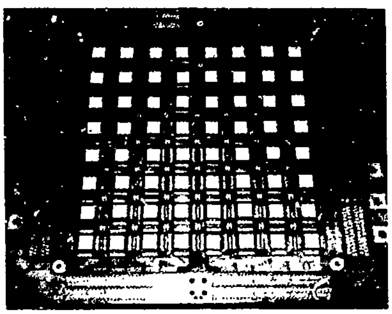
The experience gained in the early experimental use of these new computing technologies is used to refine the next generation and guide the development of more advanced software and system development technologies. The combination of scalable computing and scalable networking technologies provides the foundation for solving both Grand Challenges with large scale parallel systems and National Challenges with large scale distributed systems. This enables HPCC to progress toward an NII.

ARPA is the lead DOD agency for advanced technology research and has the leadership responsibility for High Performance Computing (HPC) within DOD. The ARPA HPC Program develops dual use technologies with broad applicability to enable the defense and intelligence communities to build on commercial technologies with rapid development of more specific characteristics when needed. ARPA has no laboratories or centers of its own and executes its programs 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. ARPA participates in joint projects with other agencies in the Federal HPCC Program, a variety of Defense agencies, the Intelligence community, and other Federal institutions.

Joint projects with other agencies are established to accelerate technology development and transition. ARPA joint projects with NSF include foundations for scalable systems, visualization, Grand Challenges, gigabit networks, and accelerating the maturation of systems software at NSF Supercomputer Centers. Joint projects with NASA include an Internet software exchange, system software maturation, and ground stations for the ACTS gigabit satellite system. ARPA, NSF, and NASA also have a joint program in digital libraries. Joint projects with DOE include scalable software libraries and networking applications. ARPA is working with NIST to develop performance measurement technologies and techniques, privacy and trusted systems technologies, and the computer emergency response team system for the Internet. A joint project with NSA is developing gigabit network security technology and other secure and trusted systems technologies. In addition, a variety of early evaluation and experimental use projects involve different kinds of scalable parallel computing systems.

The ARPA program focuses on the advanced technology aspects of all five components of the HPCC Program as follows:

HPCS



A multi-year Caltech project developed high performance interconnect and fine-grained parallel systems including boards consisting of 64 single nodes, scalable to thousands of nodes. Several commercial systems have adopted architectures based on this research, and are beginning to demonstrate how low cost modules can be configured to meet a broad range of applications.

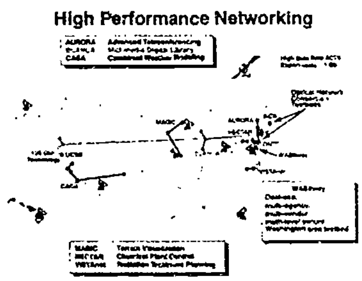
ARPA projects stimulate the development of scalable computing technologies that are capable of being configured as networks of workstations and large scale parallel computing systems capable of sustaining trillions of operations per second. Systems can be configured over a wide performance range. The systems will be balanced to provide the processor-to-memory, scalable interconnection, and input/output bandwidth needed to sustain high internal and external system performance. The modular design of the system units of replication will enable them to cover the full range from workstations to the largest scale distributed and parallel systems. Scalable systems with vector accelerators may be configured as parallel vector systems. Other kinds of accelerators such as field programmable logic arrays may be added for specialized applications. The largest scale parallel systems with hundreds to thousands of processors or more, are sometimes referred to as massively parallel systems. The input/output interfaces of these systems may be used to configure heterogeneous systems with high performance networks.

Scalable microkernel operating systems with a full complement of servers will enable software and system developers to work with a uniform set of application interfaces over the scalable computing base. Through the use of multiple servers, different application interfaces can be supported to enable the transition from legacy systems such as those available today. The system software may be configured as needed for particular applications including trusted and real-time systems.

Advanced components and packaging technologies including the associated design, prototyping, and support tools will enable higher performance and more compact systems to be developed. These technologies also enable the development of embedded systems so that computing can be put in specialized physical and environmental settings (such as airplanes, spacecraft, land vehicles, or ships).

Early evaluation and experimental use of new computing systems is an integral part of the overall development process. Policies and mechanisms have been developed that enable the timely purchase of new small to medium scale computing systems for the purpose of early evaluation and experimental use. As these technologies mature, larger scale systems are purchased and deployed by other parts of the HPCC Program in consultation with their user communities.

NREN



Experimental gigabit networks are overlaying and enhancing the Internet. ARPA's Networking Systems program develops and evaluates these technologies as foundations for a global scale, ubiquitous information infrastructure supporting Grand Challenge. National Challenge. and Defense needs.

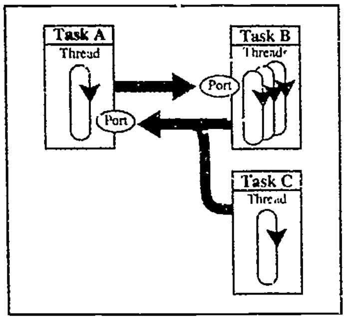
ARPA projects develop scalable networking technologies to support the full range of applications from local networks, to regional networks, to national and global scale networks including their wireless and mobile extensions. Different kinds of communication channels, or "bitways", will be integrated to enable network connectivity to be achieved between users and their applications.

Internet technologies will be developed to enable continued scaling of the networks to every individual and system needing access. Scalable high performance networking technologies will be developed to enable gigabit speeds to be delivered to the end users. A variety of networking testbeds developed in cooperation with other agencies are used to develop, demonstrate, and experimentally deploy new networking technologies. As these technologies mature, larger scale systems are deployed by other parts of the Program in consultation with their user communities.

ASTA

ARPA projects develop the advanced algorithms and software technologies needed for computationally intensive applications. Scalable software libraries will enable optimal performance for common problem domains. Extensions to existing programming languages and the development of new ones will make it easier to program scalable systems. Optimizing compilers for individual processors and their scalable configurations including support for memory and process allocation will enable maximal performance for a given algorithm to be achieved. Design, analysis, visualization, and debugging tools will enable the development and support of scalable parallel and distributed computing systems through the use of common integrated frameworks. This will provide the foundation for the software and system development and support environments needed for computationally intensive problems such as the Grand Challenges. The application of HPCC technologies to specific Defense problems is performed by other ARPA or Defense programs.

Modern Operating Systems



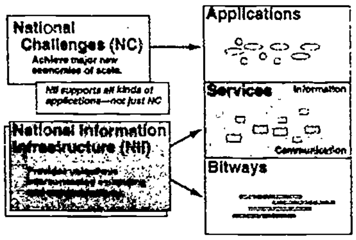
Message passing, transparent access to user services, advanced memory management, and real time response are key to a new generation of portable microkernel-based operating systems. Coupled with scalable computing systems, these operating systems enable the introduction of multilevel caches, scalable device interfaces, and scalability to thousands of processors.

IITA

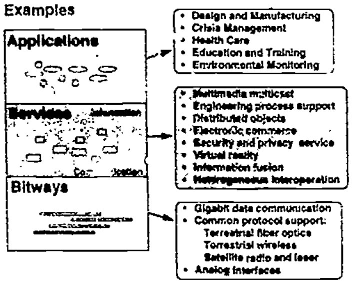
ARPA projects will develop the information infrastructure technologies needed for information and functionally challenging applications that emphasize information processing. Advanced Internet-based services will be developed to enable the effective deployment of distributed Internet-based systems including multiple media with access to high performance computing services and individual access points. Mobile and wireless technologies will enable users and their networks to access the information infrastructure with the appropriate authentication, privacy, and security.

Distributed data and object bases with transparent replication will enable the development and deployment of data, information, and knowledge repositories and services. A variety of access tools and interfaces will be developed to enable interactive access to the infrastructure. These will include intelligent functions such as knowledge based searching and alerting, planning and learning systems, and natural language understanding systems for speech, text, and images. This will provide the foundation for solving the large scale distributed, information and functionally intensive problems such as the National Challenges.

The Program Layers



Examples



The three interconnected layers of the NII are a variety of "Bitways" for communication, services for information-based and computation-based resources, and National Challenge applications. Each layer focuses on protocols for common delivery mechanisms in order to insulate users from the details of the underlying technologies. In so doing, each layer sustains a diverse base of technologies and supports a broad base of suppliers. Communications protocols, for example, ensure delivery in a uniform manner; protocols for multimedia multicast and other services provide similar uniformity to applications developers; and applications developers use scalable open protocols to make rapid advances without major services reengineering.

BRHR

The BRHR component is generally structured as an integral part of the other components. The ARPA program, along with NSF, provides the majority of Federal support to universities in computer and computational science. Fundamental limitations of HPCC technologies are identified and alternatives developed that will provide the basis for more advanced technologies in the future. An Historically Black Colleges and Universities program identifies individuals and groups that have the potential to contribute to the program. Small Business Innovation Research topics are formulated to provide another opportunity for small businesses to participate in the program. University projects provide an extraordinary opportunity to attract new people into the program, benefit from their talent, and prepare them for the new technologies. Basic research programs are often established in cooperation with NSF.

Major FY 1994 Milestones

Develop system architectures for affordable computing systems, scalable from workstations to teraops systems.

Demonstrate scalable libraries, components, and tools to support software development on a scalable parallel system.

Develop and deploy gigabit networking protocols based on experimental testbed results.

Develop information infrastructure system architectures and early prototype services.

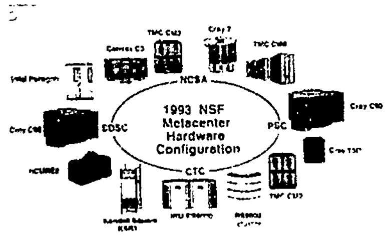
77

National Science Foundation (NSF)

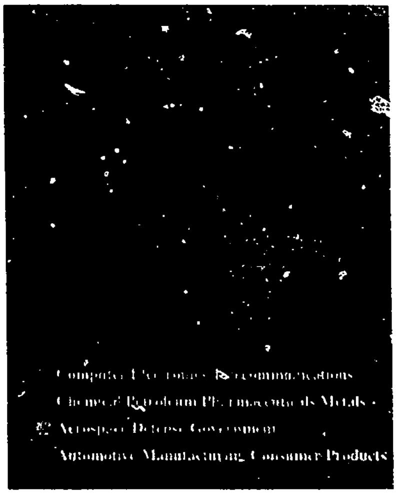
NSF's HPCC program provides support for basic research in HPCC technologies and applications, national HPCC facilities and services, computational infrastructure, and education and training. The program attempts to strategically guide broad currents of scientific inquiry and discovery into critical application problem domains. In FY 1994, NSF will continue to highlight interdisciplinary research, increased collaboration between computer scientists and application scientists, and cross-sector partnerships. Areas that address pressing interests of science and the broader society will be targeted.

Major NSF activities are planned in all five components of the HPCC Program. Program activities are implemented and executed through the existing NSF disciplinary program structure. An HPCC Coordinating Group with a representative from each NSF research directorate is responsible for assuring program development and management consistent with the goals and objectives of the HPCC Program. The result is a shared-management, NSF-wide process touching on almost all science and engineering disciplines. In FY 1992 over 1,200 individual awards were made by 75 separate disciplinary program offices. All HPCC awards are subject to the Foundation's merit-based peer review process.

NSF will continue to pursue substantial industry participation and collaboration. Cooperation with industry is achieved programmatically both through the natural alignment of academic basic research interests in key HPCC areas and through the deliberate structuring of selected NSF programs to foster collaboration. These interactions stimulate the growth of shared knowledge and capabilities, improve the rate of technology transfer, and identify new technologies and products of commercial value.



Two national NSF programs, NSFNET and the NSF Supercomputer Centers, could not be conducted at their present scope without significant industry interest and collaboration. Extensive industry involvement with the National Supercomputer Centers includes partnerships and affiliate relationships, cooperative efforts in technology development, and use of the centers' computing resources and training facilities. Affiliation with a center offers an innovative low risk method of exploring and ultimately exploiting the usefulness of high performance computing technologies in a diverse intellectual and



The approximately 130 industriai affiliates and partners at the NSF Supercomputer Centers fall into the broad groups shown above.

interdisciplinary environment. Currently about 130 non-academic affiliates or partners represent many areas including aerospace, automotive, financial, chemical products, and telecommunications. The centers are also engaged in software and systems development with most of the major U.S. hign performance computing hardware vendors.

Coordination of NSF activities with other agencies is an integral feature of the program. This entails multiagency planning activities leading to joint sponsorship of large scale research and infrastructure projects, scientific conferences and workshops, and development of shared resources.

NSF Activities by HPCC Program Component

HPCS

The HPCS component focuses on the design and development of heterogeneous and distributed scalable parallel computing systems to advance computational science and engineering research capabilities. Included is research on: systems architecture, including application-specific systems and memory hierarchy; early prototyping and evaluation of hardware and software systems; new tools for systems-level automated design and prototyping; distributed design and intelligent manufacturing capabilities; and optical communications systems and devices.

NREN

NSF is responsible for overall coordination of the NREN component and the broad deployment of network infrastructure and services. NREN component activities support all aspects of the HPCC Program through connection of other agency networks, upgrading of the NSFNET backbone network services, providing for enhanced network security, increasing the number of educational institutions connected to the network, and serving as a primary source of information about access to and use of the network. Activities focus on network infrastructure and services and on gigabit research and development. These activities are designed to develop and implement a wide range of networking infrastructure components and services for the Nation's community of researchers, scholars, and students that will foster interaction and collaboration; and to give rapid access to researchers and students to facilities and other resources for use in scholarly endeavors. These facilities and resources include computation centers, laboratories, scientific instruments, and databases and libraries. NSF also participates in the support (under ARPA coordination) of the networking and communications research and development effort leading to early deployment of advanced HPCC systems. This coordinated research effort includes work on very high speed switches and several wide-area research gigabit testbed networks in support of Grand Challenge collaborations and other applications.

ASTA

NSF centers directly supported within the ASTA component include:

- Cornell Theory Center, Ithaca, NY

- National Center for Supercomputing Applications, Champaign-Urbana, IL (NCSA)

- San Diego Supercomputer Center, San Diego, CA

- Pittsburgh Supercomputing Center, Pittsburgh, PA

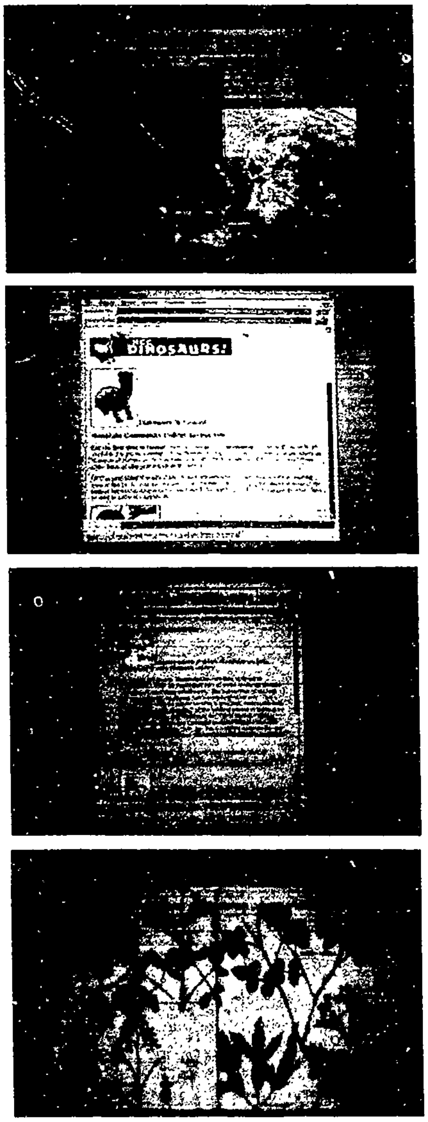
- Center in Computer Graphics and Scientific Visualization, University of Utah

- Center for Research on Parallel Computation, Rice University, Houston, TX

The ASTA component emphasizes the development of algorithms and software technologies and the establishment of research capability to address Grand Challenge scale computational problems. Included in ASTA are: research demonstrating the applicability of HPCS products to critical problems; building interdisciplinary collaboration and exchange; collaborative Grand Challenge and large scale computational research; computational techniques and software technologies for scalable parallel and distributed heterogeneous computing systems; scientific databases; and deploying and configuring new parallel systems in research centers and testbed facilities accessible to researchers, students, and educators nationwide.

NSF's National Supercomputer Centers and Science and Technology Centers are hubs of intellectual and educational activity, serving researchers, students, and teachers from many disciplines and educational levels. They seek to define and provide a premier environment for coordinated approaches to Grand Challenge problems. The centers are currently implementing the concept of a "metacenter:" planning and working closely together, the individual centers pool their technological resources, support services and administrative functions so that the collective resources are viewed and made available to users as a single, powerful computing environment. With the metacenter concept, new systems and services can be incorporated into the overall environment without disturbing the stable computing environment that is in place. The concept can also offer additional flexibility in managing and optimizing the centers' operations.

81



Four sample NCSA Mosaic (described on page 44 in the ASTA section) screens showing access to images contained in public Internet digital data libraries. From top to bottom:

- *The Mosaic Demo document directly accesses a gopher server at the University of Illinois' Weather Machine containing current weather data (here the Flood of '93)*
- *The Dinosaur fossil exhibit developed by Honolulu Community College The exhibit uses audio clips, images, and text to educate and entertain Users can take a narrated tour by selecting the audio links throughout the documents*
- *Two images from "Rome Reborn The Vatican Library and Renaissance Culture," an exhibit at the Library of Congress*
  - *One screen from this hypertext exhibit*
  - *Botanical pages from Materia Medica a 14th or 15th century Italian manuscript*

IITA

IITA component activities are designed to expedite the application of research products and services developed in the original four component areas to National Challenges. Efforts in IITA will build on and integrate fundamental HPCC technologies in order to enable widespread progress on information-intensive applications in such areas as the civil infrastructure, digital libraries, education and lifelong learning, the environment, health care, and manufacturing. NSF will focus primarily on building and promulgating intelligent user interfaces, distributed software systems and support environments, standardized data management and communications processes, and enhanced information infrastructure services. IITA activities include:

- Support for development of tools and mechanisms for enhancing human communication and collaboration across the network.
- Development of multimedia software and educational materials.
- Support for basic and strategic research on development of advanced data capture, storage, representation, and retrieval technologies for "digital libraries."
- Training support for users in the generation and use of digital library resources, fostering "lifelong learning" opportunities and development of new network-based training modalities.
- Pilot projects for information-intensive applications.

BRHR

BRHR component activities are designed to encourage research projects and infrastructure and educational activities that ensure the flow of innovative ideas and talented people into high performance computing technologies and application areas. Activities include support for research in fundamental areas of algorithms, data structures, programming languages, operating systems, software engineering, performance evaluation, databases, digital libraries, information science, knowledge processing, language and speech understanding, computer vision and image understanding, and multimedia computing. BRHR also supports postdoctoral awards in high performance computational science and engineering and experimental computer science; undergraduate research experiences, high school teacher and student training experiences at high performance computing centers; and infrastructure acquisition for computational research and education.

Major FY 1993 Accomplishments

- Six to eight new Grand Challenge Applications Groups funded (listed in Side Bar).

  This program activity was initiated in FY 1992, with awards made to multidisciplinary groups of scientists and engineers whose activities integrated disciplinary research with the design of models, algorithms, and software for high performance computing systems. See Side Bar for areas in which awards were made.

- Research projects on techniques and environments for scalable parallel and distributed heterogeneous computing, collaboration technologies, evaluation of High Performance Fortran, programming language standards, and user environment components.

- Metacenter implementation in selected areas at the NSF high performance computing centers to provide new opportunities for the combined use and shared management of these systems and requisite supporting technologies.

- Deployment and configuration of scalable parallel systems for mainstream scientific computing at the NSF centers.

- NREN awards for Network Information Services; recompetition of the NSFNET Backbone; promulgation of computer security measures; evaluation of the Network Access Point (NAP) concept; regional networks; and testbeds for teaching and education.

On-Going Grand Challenge Applications Groups

Biomolecular Design

Coupled Field Problems and GAFD (Geophysical and Astrophysical Fluid Dynamics) Turbulence

High Capacity Atomic-Level Simulations for Design of Materials

Imaging in Biological Research

Large Scale Environmental Modeling

Machine Learning

Radio Synthesis Imaging

Grand Challenge Application Groups Initiated in FY 1993

Adaptive Coordination Predictive Models with Experimental Observations

Advanced Computational Approaches to Biomolecular Modeling and Structure Determination

Black Hole Binaries: Coalescence and Gravitational Radiation

Earthquake Ground Motion Modeling in Large Basins

High Performance Computing for Land Cover Dynamics

Massively Parallel Simulation of Large Scale, High Resolution Ecosystem Models

Parallel I/0 Methodologies for I/O-Intensive Grand Challenge Applications

The Formation of Galaxies and Large-Scale Structure

Understanding Human Joint Mechanics Through Advanced Computational Models

FY 1994 Plans

NSF will continue to support HPCC basic research activities, metacenter implementation, Grand Challenge Application Groups, postdoctoral research associateships, and broader access to high performance computing research resources for researchers, students, and educators at all levels. Support for NREN development will continue to meet special NSF obligations and responsibilities set forth in the HPCC Program plan, including the development of gigabit testbeds. New activities will support the migration of technologies and capabilities developed in the original HPCC Program component areas to address specific applications associated with such areas as health care, lifelong learning, digital and electronic libraries, and manufacturing.

HPCS support will focus on:

- Developing new systems level design tools that have increased functionality, are highly automated, and accept high level specifications and algorithms as input.
- Novel computing system design approaches drawing on expertise in architecture, systems software, storage technology, I/O subsystems, and networking technologies.
- Optical communications devices and systems.
- Advanced computer and information technologies to enable distributed design and intelligent manufacturing of objects.
- Systems-level issues in understanding, modeling, and integrating component manufacturing technologies.

NREN activities will include:

- Development of a very high speed backbone.
- Increasing the number of network access points (NAPs).
- Funding, improving, and providing new services for mid-level networks.

- Funding a routing authority.
- Increasing connections for post-secondary education and related organizations.

ASTA activities will include:

- Additional Grand Challenge Application Groups awards.
- Research on computational techniques for distributed and heterogeneous computing environments.
- Continued metacenter development and the adoption of common standards for hardware, software, and networking technologies.
- Deployment of prototype scalable systems in Grand Challenge research settings.
- Collaborative activities on software systems, architectural modifications, new programming language standards, and user environment components.
- Projects demonstrating the usability of High Performance Fortran and other high level languages for solving large scale scientific and engineering problems on massively parallel distributed memory systems.

IITA activities will focus on:

- Application of HPCC technologies to such National Challenge areas as health care, lifelong learning, digital and electronic libraries, and manufacturing.
- Software research relevant to scientific databases, digital libraries, collaboration technologies, computer graphics, and scientific visualization.
- Exploring new applications of virtual reality systems.
- Improvements in accessibility and ease of use of Internet-based data and computing.
- Research on intelligent knowledge capture, representation, extraction, and dissemination.

- Projects aimed at simplifying human interaction with digital information resources through development of user friendly, graphical interfaces for knowledge representation and manipulation.

- Activities addressing vital issues of standards.

- Development of prototype libraries designed to fill new roles in research and education.

BRHR support will focus on:

- Basic research on scalable parallel systems, distributed computing environments, and tools.

- Research in fundamental areas of algorithms, databases, visualization, architecture, performance evaluation and modeling, and computer resource management.

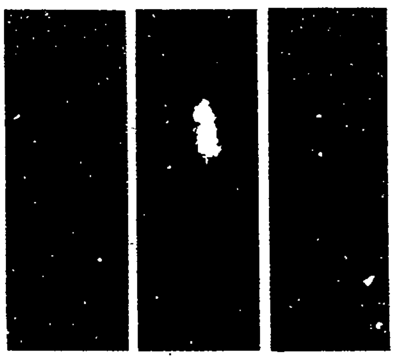
- Increased number of postdoctoral awards in high performance computational science and engineering and computer science, and expansion of the current program to include industry.

- Continued support for HPCC training and education in the form of research experiences for undergraduates, supplements to faculty research awards, high school teacher and student training, and equipment and infrastructure awards.

- Continued acquisition of infrastructure to improve computer science, computer engineering, computational science and engineering, and Grand Challenge research.

Department of Energy (DOE)

DOE's goals in the HPCC Program are to enable effective applications of HPCC technologies and the emerging National Information Infrastructure (NII) to scientific problems that are critical to implementing the Energy Policy Act (PL 102-486), other DOE mission programs, and the national interest. DOE's missions encompass such diverse activities as energy technologies, studies of energy supply and usage in the U.S., environmental remediation, fundamental research, investigations in the health and environmental sciences, and national security.



To better understand combustion processes and the effluent gases produced by these processes, hydrogen/air turbulent reacting flame streams are studied. From left tc right, false color renderings of the temperature. OH radical concentration, and NO concentrations are shown. Low values are in blue and high values in red. (Temperatures in the left image range from 30 C to 2,800 C and the visible flame is in the red band.) The research and object oriented parallel software are the products of a collaboration among Sandia National Laboratory, Lawrence Berkeley Laboratory, and the University of California at Berkeley

All five components of the HPCC Program are important to achieving DOE's goals. The DOE program for FY 1994 will continue to build on the foundation of joint interagency, interdisciplinary, and private sector collaborations established during the earlier phases of the HPCC Program.

With its focus on applying HPCC technologies, it is critical that the DOE effectively encourage technology transfer and collaborations among researchers in different disciplines and among researchers and their counterparts in U.S. commercial enterprises. In order to ensure the effectiveness of this approach, the program:

- Funds activities in cooperation with other DOE mission programs that will develop and use HPCC technology.

- Involves end users of the technology to the maximum extent possible in the initial and ongoing evaluation of projects.

- Carries out work with industrial partners, generally through Cooperative Research and Development Agreements (CRADAs).

The Office of Scientific Computing has been assigned the lead responsibility for Department-wide participation in the HPCC Program.

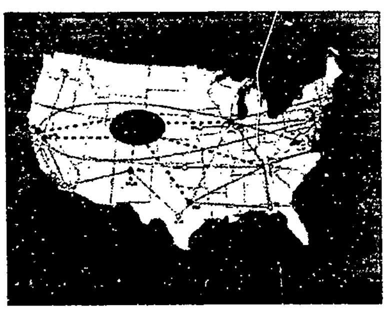
FY 1993 Accomplishments and FY 1994 Plans

HPCS

A research program has been initiated to investigate the performance of computer systems based on an integrated analysis of algorithms used in Grand Challenge applications. In addition, a researcher at DOE's Ames Laboratory was awarded a patent for developing SLALOM, an innovative benchmarking system for any type of computer including parallel computers. Two large university projects in parallel systems research at the University of Illinois and at New York University were extended. Several projects have been initiated to evaluate the effectiveness of early prototype HPCC systems on scientific problems important to the DOE at the Department's new High Performance Computing Research Centers.

NREN

Internet Support



ESnet provides access to Energy Research facilities and resources for DOE and DOE-supported researchers and educators.

Efforts continue to build on the activities initiated in FY 1992 and FY 1993. The ESnet FY 1992 fast packet services procurement will provide advanced network capabilities and services to DOE sites and researchers, in addition to accelerating the availability of these public telecommunications vendor-based services to other Federal, state, and research and education organizations. The ESnet is an integral part of the Internet; it provides multiprotocol seamless connectivity to scientific resources and facilities, for international collaborative science, and supports DOE sponsored education activities. Other activities and accomplishments include collaborative work in packet video and voice, collaborative distributed work environment tools, multiprotocol support, GOSIP/OSI to TCP/IP gateways for e-mail and other services, and other precompetitive and short term research and development efforts.

Gigabit Research

DOE participates in the existing gigabit testbeds, the development of high speed local area networks (e.g., HIPPI), tools needed to support remote experimentation and distributed computing, multimedia data manipulation, high speed network management, and the standards associated with each of these technologies.

ASTA

Support for Grand Challenges

DOE set up an interagency panel, comprised of DOE program managers and other HPCC agency participants and chaired by the NASA HPCC program director, to evaluate proposals for the DOE Grand Challenge computational research projects. The selected projects are cofunded by other DOE programs and by industrial partners and include DOE laboratory, university, industry, and other HPCC agency participation. The six multiyear Grand Challenge projects selected in FY 1992 are:

- Computational Chemistry – to effectively parallelize chemistry codes important to the study of several critical environmental problems. This project is cofunded by four industrial partners and the DOE Chemical Sciences program.

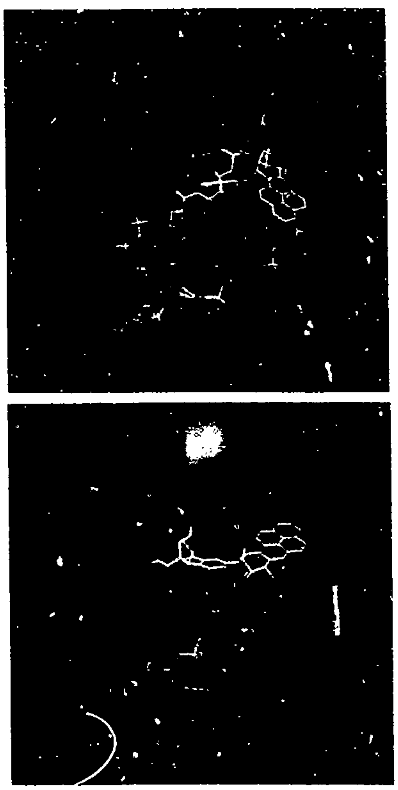
- Computational Structural Biology – for research in computational methods important to understanding the components of genomes and to develop a parallel programming environment for structural biology. This project is cofunded by the DOE Health and Environmental Research program, an NSF Science and Technology Center, a biomedical industrial firm, and a university foundation.

- Mathematical Combustion Modeling – to develop adaptive parallel algorithms for computational fluid dynamics and apply these methods to combustion models important to private sector and government scientists and engineers. This project is cofunded with the DOE Applied Mathematics and the Chemical Sciences programs.

- Quantum Chromodynamics Calculations – to develop lattice gauge algorithms on massively parallel machines. This project is cofunded by the DOE and NSF high energy physics programs and involves particle physicists from a dozen universities.

- Oil Reservoir Modeling – to construct efficient algorithms for parallel systems to simulate fluid flow through permeable media. The work is based on reservoir models from the University of Texas that are widely used by the oil industry and is cofunded by DOE Environmental Sciences, an industry consortium of 26 companies, two computer vendors, NSF, and the State of Texas.

- The Numerical Tokamak Project – to develop and integrate particle and fluid plasma models on massively parallel machines as part of a multidisciplinary study of Tokamak fusion reactors. The



In work jointly supported by NIH, DOE, and NSF, researchers at New York University, Sloan Kettering Cancer Center, and Oak Ridge National Laboratory are elucidating the effects of biochemically activated environmental chemicals, such as benzo-a-pyrene, a material present in automobile exhaust gases, bound to DNA. The figures illustrate calculations of two such derivatives that are mirror images of one another, (+) and (-) anti-benzo-a-pyrene-diol-epoxide [BPDE], bound to a segment of normal right-handed DNA. There is a profound difference in the health effects of these two cases, the (+) BPDE case (red figure) is tumorigenic while the (-) BPDE is benign.

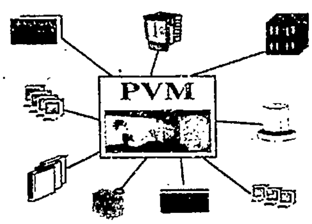
project is cofunded by DOE's Fusion Energy program and is an important component of the National Energy Strategy.

By the end of FY 1992 DOE supported 14 computational projects. No new computational projects or Grand Challenges were started in FY 1993 due to lack of funds. In FY 1994, the DOE expects to evaluate the existing projects and may initiate new projects to balance the program in response to the goals and objectives of the DOE HPCC Program.

Software Tools

Projects were initiated in: object oriented data base technology (applied to and cofunded with the Superconducting SuperCollider), large spatial databases, distributed computing technologies, and software environments.

Heterogeneous Distributed Computing



The PVM (Parallel Virtual Machine) software permits a heterogeneous collection of networked Unix machines (serial, vector, or parallel) to be used as a single large parallel computer. This public-domain software is now used at hundreds of sites worldwide.

Computational Techniques

The DOE initiated projects on computational techniques applied to catalysis, drug design, and in chemistry and the materials sciences in cooperation with private sector firms.

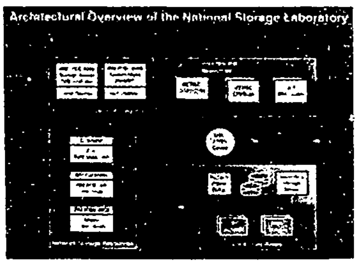
High Performance Computing Research Centers (HPCRC) and Advanced Computing Resources

Two HPCRCs were created in FY 1992, at Los Alamos National Laboratory (LANL) and at Oak Ridge National Laboratory (ORNL), involving industrial and university partnerships as well as cofunding from other DOE programs. These HPCRCs conduct research in all HPCC Program component areas; operate massively parallel high performance computing prototypes (a Thinking Machines Inc. CM5 at LANL and an Intel Corp. Paragon at ORNL); perform computational investigations in global climate research, environmental groundwater transport modeling, and in materials sciences calculations; and provide HPCS resources for the other computational research projects described above. The HPCRC at LANL is also a partner in an NSF-sponsored Science and Technology Consortium. The Serial #1 Kendall Square system was installed at the ORNL HPCRC where it will be evaluated on energy applications.

A Cray Research Inc. C-90 supercomputer system was installed in FY 1993 in the National Energy Research Supercomputer Center (NERSC) at Lawrence Livermore National Laboratory to provide high performance computer technology in a full production environment for DOE applications. The C-90 has 16 processors and is being used in a mode that encourages the use of parallel programming techniques.

National Storage Laboratory (NSL)

In concert with NASA, the NSL was established by a CRADA among six industrial firms and NERSC. This collaboration currently involves several DOE Laboratories, about a dozen U.S. storage vendors, and a university. The NSL will help serve the data storage needs of researchers at the DOE and at other Federal agencies and will help participating vendors develop new mass storage products.



Overview of the architecture of the National Storage Laboratory located at NERSC, in which different networks are used for control of the attached devices and for data transmission. This architecture overcomes the bottleneck in earlier hierarchical storage designs in which all of the data had to pass through the control computer. The technology will first be used in fusion energy and climate modeling applications. The NSL architecture is being extended in a new High Performance Storage System to support scalable parallel I/O and integration with massively parallel computers.

The NSL will advance the state of the art in high performance data storage systems that are capable of storing terabytes of data. Such mass storage is required by applications such as the Grand Challenges and those requiring storage of large amounts of experimental data.

Candidate IITA Activities

DOE brings expertise in information technologies to the task of developing the NII. DOE has developed communication networks and has applied computing technology to a broad range of applications and technical problems addressed in collaborations involving scientists and engineers located across the country as well as around the world The many DOE CRADAs are evidence of industry's interest in working with the Department. DOE Laboratories have a long history of collaborating with research universities.

The early deployment of the NII is important to the successful implementation of the Energy Policy Act. The NII is critical to the Act's goals of:

- Managing energy supply and demand.
- Developing telecommuting to reduce energy usage and provide more efficient use of the Nation's human resources.

- Training technical personnel in the use of information technologies.
- Providing energy related information to citizens.
- DOE initiatives for energy efficient manufacturing and materials processing.

In FY 1994 a candidate DOE activity is to begin developing prototype applications to demonstrate national scale applications and use of the NII in selected energy-related areas. It would work closely with corresponding programs of other agencies in these efforts.

"Superkids," participants in the High-School Science Students Honors Program, with the National Education Supercomputer at NERSC. Each summer this program brings one student from each state, the District of Columbia, Puerto Rico, American Samoa, the DOD Dependents Schools, and eight foreign students to Lawrence Livermore National Laboratory.

BRHR

DOE supports several educational programs including the National Education Supercomputer Program (NESP) at NERSC; the Southwest Indian Polytechnic Institute in which 30 Native Americans participate; and workshops to train teachers in HPCC technology and use, and to conduct such workshops themselves. FY 1993 highlights include:

- The NESP completed its eighth High School Science Students Honors in Supercomputing and an associated teachers curriculum development workshop, and operated the National Education Supercomputer (NES), a Cray X-MP/18 donated by Cray Research, Inc. This program was recognized in a DOE Public Service Award to the NERSC Education Coordinator.
- A syllabus for teaching computational science at the graduate level has been developed and made available on the Internet. This work involved 24 authors from 10 disciplines. It incorporates graphics using Apple MacIntosh and IBM RS 6000 technologies and allows the user to execute examples on Cray Research, IBM, Intel, Kendall Square, and Thinking Machines systems. The syllabus is maintained at Vanderbilt University.
- A Computational Science Graduate Fellowship Program was begun in FY 1992. The fellowships are awarded on one-year renewable terms to support full-time graduate study and thesis research in the U.S. The fellowships are in the applied science and engineering disciplines with applications in high performance computing. In FY 1993 the

number of fellowships was increased to 42 at 35 universities and institutes. This program and the syllabus described above have been instrumental in defining the discipline of computational science.

Fundamental Research Programs

Clockwise from lower left, volume renderings display the density of a bubble being hit by a shock wave. New adaptive gridding and fast, accurate, and robust numerical algorithms are used to simulate the behavior of such highly nonlinear fluid flows in complicated three-dimensional geometries. These techniques are implemented on the latest vector and massively parallel architectures. Applied mathematicians and computational scientists performed this research at Lawrence Livermore National Laboratory, Los Alamos National Laboratory, New York University, and the University of California at Berkeley.

The Applied Mathematics program supports a broad range of activities at universities and DOE Laboratories in modeling, analysis, and numerical simulation of physical and biological phenomena that arise in energy and environmental systems. Most of the projects involve applications-driven studies of the mathematical and numerical tools required to understand the behavior of complex discrete and continuous systems with an emphasis on algorithms for parallel computing. It has cofunded with NSF the Geometry Science and Technology Center at the University of Minnesota to apply new HPCC technology to traditional mathematics problems. It has initiated new efforts in complex nonlinear behavior that underlies most natural phenomena, and in graph and group theories related to discrete phenomena and topology for application to genome sequencing and protein structure.

In computer science the focus is on understanding how parallel and distributed computer systems can be applied more effectively to large-scale scientific problems. Supported projects include research in programming models and tools, improved software libraries for parallel computers, scientific visualization of large data sets, software performance analysis techniques, and message-passing utilities to facilitate distributed computing.

FY 1994 Milestones

HPCS

Begin evaluating one or two additional early prototype systems in cooperation with vendors.

Expand computer systems performance analysis project, including top-down algorithmic analysis of Grand Challenge codes at HPCRCs.

Continue development of high performance hierarchical data storage system at the NSL in NERSC.

NREN

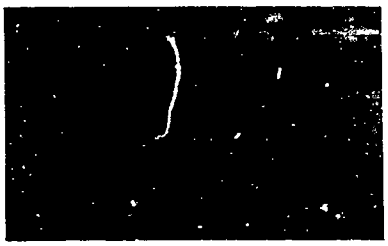
Begin prototype project in telecommuting access to DOE experimental facilities.

Complete ESnet upgrades to 45 Mb/s at 20 sites.

Initiate ESnet upgrades to 144-622 Mb/s at selected sites.

Continue gigabit research projects, emphasizing distributed applications.

Implement production-quality packetized workstation-based video over ESnet and other Internet components.



A single frame from an eight month simulation for the Pacific Northwest using the Penn State/NCAR Mesoscale Meteorological (MM) model. The simulation was executed on the Intel Touchstone Delta at Caltech. MM model results are used in high resolution regional forecasts, allowing more efficient energy use and helping reduce storm damage. Argonne National Laboratory (ANL) and NSF-supported NCAR researchers used fast fine-grained massively parallel software and PCN, a parallel programming language developed at ANL and Caltech with support from the NSF Science and Technology Center for Parallel Computation.

ASTA

Review, evaluate, and adjust program of Grand Challenge and computational research projects begun in 1992.

Implement fully three-dimensional algorithms based on adaptive mesh refinement techniques for computational fluid dynamics.

Deliver prototypes of discipline-oriented computing environments for Grand Challenge applications.

Conduct Grand Challenge workshop together with other HPCC participants, emphasizing progress in software tools and computing environments.

Develop policies, standards, and software engineering methodologies to ease the transition of research tools into commercial products.

Develop software to integrate the high performance hierarchical data storage system at the NSL into advanced computing environments.

Provide detailed feedback to computer vendors on experience gained with prototype computers installed at HPCRCs.

Clockwise from the top, a slice of a three-dimensional model of an 8 kilometer/second impact on a planet the size of the Earth by a planet the size of Mars. at 0. 60, 120. and 200 seconds. The gravity of the larger planet later traps mantle material and forms a moon. Using multidimensional shock physics software developed in a collaboration among several DOE Laboratories and others under DOE and ARPA sponsorship. this is an example of materials modeling under extreme conditions. The same software is used to model other high velocity impact phenomena, including explosives. resulting in fewer experiments, reduced cost. and increased safety.

IITA Candidates

Publish white paper reviewing applicability of NII to telecommuting and defining the enabling technologies required for and the barriers to widespread implementation of telecommuting and other forms of telework.

Evaluate scope of DOE generated information for inclusion in NII digital libraries.

Conduct cooperative evaluation of potential for energy supply and demand management using the NII.

Begin collaborative development of NII technologies for applications testbeds.

BRHR

Refine and make widely available the graduate-level computational science electronic textbook.

Initiate college-level computational science electronic textbook.

Review and evaluate ongoing high school education programs and associated teacher workshops.

Initiate junior high school educational projects.

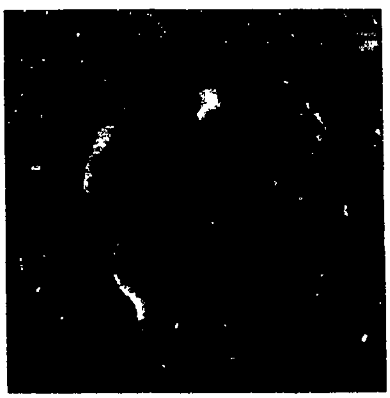
Rebalance applied math research program based on comprehensive review conducted in 1993.

National Aeronautics and Space Administration (NASA)

Improved design and simulation of aerospace vehicles and increased ability of scientists to model the Earth's climate and to forecast global environmental trends are NASA's strategic goals within the HPCC Program. These goals also include broadening the information applications of HPCC in areas complementary to NASA's expertise, and in areas important to the development of the National Information Infrastructure. To reach these goals, NASA participates in the development and use of the advanced high performance computing and communications systems and tools piloted in each of the five components of the HPCC Program.

HPCS and ASTA: The Grand Challenge in Aeronautics

Improved design and simulation of advanced aerospace vehicles at reduced cost will enable the U.S. to enhance its leadership in aerospace trade, especially as future aerospace vehicles become more difficult and expensive to simulate. In the HPCS component, NASA procures and evaluates prototype, scalable parallel computing systems used to develop advanced algorithms and software tools. This systematic approach is important for developing improved methodologies and software tools to model aerodynamics, aerobraking, heat transfer, combustion, and other engine elements through the full flight envelope of the aerospace vehicle (from vehicle takeoff, to flight at different altitudes and conditions, to landing).



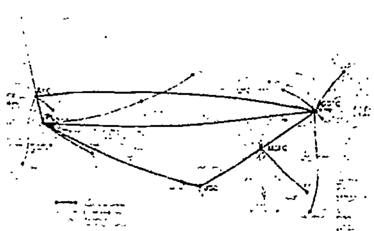
Simulation of three-dimensional distribution of the Earth's ozone layer. Latitude and longitude slices illustrate the varying ozone content

HPCS and ASTA: The Grand Challenges in the Earth's Environment

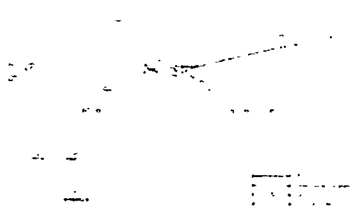
NASA's Earth and Space Sciences (ESS) project is developing multidisciplinary models of physical, chemical, and biological phenomena that will lead to more accurate environmental simulations. Challenges range from analysis of the interactions among Earth's atmosphere, oceans, and land masses, to the reconstruction of planetary evolution. A crucial aspect of this program is the development of software management systems to handle the volumes of scientific data that will be produced late in this decade; multiple terabytes of data will be transmitted and stored each day and will rely on the high performance data management, storage systems, and communications systems developed through the HPCC Program.

*The NASA Aeronautics Network **(AERONet)** provides computer networking facilities to the aerospace community. NASA centers are linked via high speed communications lines, and lower-speed tail circuits connect other members of the aerospace community to NASA. In addition, AERONet also provides Internet access with many other networks, thereby increasing national and international network connectivity.*

NASA Science Internet



NASA Science Internet



*The NASA Science Internet **(NSINet)** provides connectivity among NASA, industry, and academic researchers to facilitate collaboration in Earth and space science research. This network builds on the Internet and includes technology to support the evolution of the Internet*

NREN and BRHR: Support for Grand Challenge Teams

NASA has established Grand Challenge applications interdisciplinary teams in Computational Aerosciences (CAS) and Earth and Space Sciences (ESS). The CAS teams are examining coupled aerodynamics, structures/materials, controls and combustion modeling for high-speed civil transport, high-performance aircraft, subsonic aircraft, and rotorcraft. Through a NASA Research Announcement, ESS teams have been established in Earth system science, space and solar-terrestrial physics, astronomy and astrophysics, biochemical life cycles, planetary evolutionary processors, and massive data analysis.

The CAS teams use the NASA Aeronautics Network (AERONet) while the ESS teams use the NASA Science Internet (NSINet) for high speed network connectivity among NASA, industry, and academic researchers (see map). In the future, these high performance network highways will be constructed from fast-packet network technology. NASA supports the architecture evolving in the NREN component and participates in the development and deployment of gigabit network technologies and architectures.

Pilot programs with the K-12 educational community have been established to discover the best mechanisms for using space and aeronautics assets to support educational communities in the U.S., and to provide practical models for using sophisticated computational and networking resources in these communities. This is to be accomplished in collaboration with Grand Challenge scientists from the CAS and ESS projects, thereby achieving direct involvement between NASA scientists and the K-12 community.

ASTA: Software Sharing

As lead agency in ASTA's software sharing activities, NASA coordinates the collection of and access to the software base developed in the HPCC Program. The goal of the HPCC Software Exchange is to facilitate the exchange and reuse of software. Its specific objectives are to 1) develop and demonstrate distributed architectures and enabling technologies that support software exchange, 2) implement an initial distributed HPCC Software Exchange that satisfies the needs of the

HPCC Program, and 3) specify an open non-proprietary architecture that will facilitate the emergence of a national software exchange.

IITA: Toward a National Information Infrastructure

NASA is committed to broadening its participation in key areas of research and development of prototype solutions to National Challenge applications. NASA is selectively augmenting on-going programs that can contribute to an early start-up of these prototype systems in applications such as education and lifelong learning, digital library technology, manufacturing and industrial design, and access to widely available databases of Earth and space science data. NASA participates in each of the four IITA elements.

BRHR: Fostering High Performance Computing Research

NASA's BRHR component fosters research into new theory and application of high performance computing. These activities will leverage current research efforts in high performance computing at NASA research institutes and universities.

Simulated temperature profile for a three-dimensional hypersonic re-entry body traveling at an altitude of 100 kilometers, a speed of Mach 24, and an angle of attack of 10 degrees. Performed on the Intel Touchstone Delta at CSCC, the simulation employed 70 million particles and all 512 of the Delta's processors running for 50 minutes.

Major FY 1993 Accomplishments and Activities

Concurrent SuperComputing Consortium (CSCC)

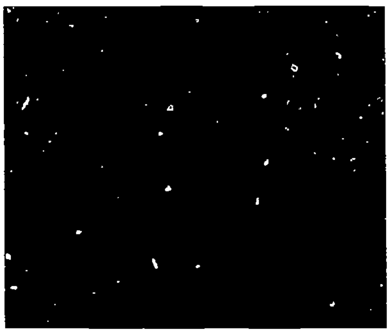
NASA played an active role in the Consortium during its first year of operation. Use of the Consortium's Intel Touchstone Delta supercomputer has enabled major advances in several scientific and engineering applications including the processing of three-dimensional images of Venus taken by the Magellan satellite at the rate of four frames per second – thus putting real-time image processing within reach. Other accomplishments on the Delta include the largest direct numerical simulation of the time-dependent compressible Navier-Stokes equations and the largest three-dimensional compressible turbulence simulations for high Reynolds numbers.

High Performance Networking

NASA continues to develop and deploy advanced networking technologies to allow researchers and educators to carry out collaborative research and educational activities. NASA has entered into a cooperative effort with DOE to procure network services that will operate at 45 Mb/s using synchronous optical fiber network standards. The two agencies are working closely to provide a nationwide fiber optic network that will meet HPCC research communications needs and serve as the foundation for even faster networks ranging from 155 Mb/s to eventual gigabit speeds.

Software on Parallel Computers

NASA's Ames Research Center has ported single discipline computational fluid dynamics code to a number of scalable parallel computers. The objective is to perform multidisciplinary optimization of a High Speed Civil Transport vehicle and the takeoff and landing of a simple powered lift vehicle (described on pages 126-128 in the Case Studies section). The optimization will consider aerodynamic efficiency, structural weight, and propulsion system performance. The multidisciplinary analysis will be performed by solving the governing equations for each discipline concurrently on a parallel computer. Developing scalable algorithms for the solution of this problem will also be central to demonstrating the potential for teraflops execution speed on a massively parallel computer.



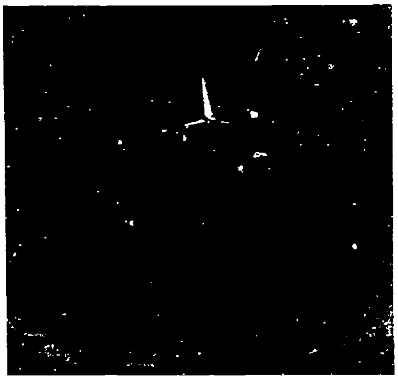
The surface pressure distribution on a High Speed Civil Transport aircraft is simulated using an Intel iPSC/860 high performance computer with 32 processors. Areas of highest pressure are in red, areas of lowest pressure are dark violet.

High Speed Civil Transport (HSCT)

NASA's Langley Research Center is working to improve the processes for the design and analysis of HSCT using advanced computational fluid dynamics and structural analyses. A set of highly optimized software tools has been created that can be used to implement irregular computations on massively parallel machines. These tools can be used both manually by users and by distributed memory compilers to automatically parallelize irregular codes.

Benchmarking Multidisciplinary Codes on Parallel Supercomputers

NASA's Ames Research Center has developed and



Computational fluid dynamics reveals the interaction between the freestream airflow and engine exhaust of a YAV-8B Harrier Vertical Takeoff and Landing (VTOL) aircraft. The simulation shows how hot gas from the nozzles is ingested by the engines, reducing their effectiveness.

demonstrated the ENSAERO computer algorithm, which couples aerodynamics and structures analysis modules on the Intel iPSC/860 massively parallel computer.

Results involve the coupling of a low-fidelity structures code to a high-fidelity aerodynamics analysis routine. This multidisciplinary research, in conjunction with the increased performance offered by massively parallel computers, will enhance the ability of aircraft manufacturers to quickly analyze different design options and accelerate the prototyping process, and thereby reduce design cycle costs in addition to producing vehicles with improved performance.

Computational Aerosciences Consortium

This consortium of U.S. industrial firms works with NASA to facilitate the development of precompetitive software for the implementation of integrated multidisciplinary design, analysis, and optimization systems on heterogeneous computer networks. Incorporation of these systems into the product development process can increase U.S. competitiveness through reduced design cycle time and life cycle costs and increased quality of technology-driven products. Consortium members include:

- Allison Gas Turbines
- Boeing Helicopter
- Ford
- General Electric
- General Motors
- Grumman
- McDonnell Douglas
- Northrop
- Rockwell
- United Technologies
- Vought

High Performance Computing Software Exchange

The inherent size and complexity of HPCC problems will increase the cost and difficulty of developing and maintaining robust applications software. The HPCC Software Exchange Experiment encourages the sharing and reuse of software modules by providing an infrastructure of interconnected software repositories. The HPCC Software Exchange Experiment System represents the first time that a number of different agencies with different philosophies for software management have collaborated to develop an integrated, albeit experimental, system.

The second phase of development, the HPCC Software Exchange Prototype System, was begun in FY 1993. This phase includes the development of 1) an HPCC Software Standards database with initial software submission standards, 2) an HPCC Repository Directory with additional software repositories, 3) an HPCC Union Catalogue populated with mathematical and statistical software packages and programs, 4) a Mathematical and Statistical Software Cross-index, and 5) a Unix-to-Unix client/server system for uniform repository access.

FY 1994 Milestones

HPCS

Install 20- to 50-gigaops testbed (scalable to 100-gigaops) for Computational Aerosciences (CAS) project Grand Challenge teams.

NREN

Complete T3 (45 Mb/s) Level 3 interconnects to NASA HPCC Centers and T3 Level 2 interconnects among NASA HPCC Centers.

BRHR

Evaluate initial K-12 digital educational material.

National Institutes of Health (NIH)

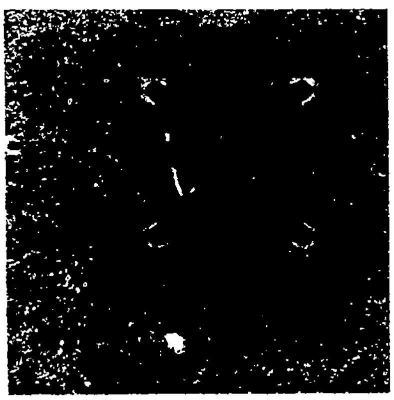
The National Institutes of Health (NIH) contribution to the HPCC Program focuses on biomedical applications of computing and digital communications. Four NIH units have HPCC programs:

NLM – National Library of Medicine
NCRR – National Center for Research Resources
DCRT – Division of Computer Research and Technology
NCI – National Cancer Institute

NIH participates in each of the five components of the overall Federal HPCC Program, as follows:

HPCS: Evaluation for Biomedical Applications

Through DCRT, NCI, and NCRR, NIH applies and evaluates scalable parallel computing systems to problems of biomedical significance. This includes adapting existing molecular analysis algorithms to new computing architectures, and developing entirely new approaches that take advantage of computational parallelism.



A cross-section of the three-dimensional reconstruction of a herpes simplex virus capsid is shown in the inset with the monoclonal antibody fragments bound to the protruding tips of hexons color-coded red and the original electron micrograph shown in the background.

NREN: Increasing Access and Transmitting Biomedical Images

DCRT and NLM have complementary investments in extending gigabit speed communications for high data volume scientific applications, as well as lower speed connections for a broad community of research and education institutions.

DCRT's efforts focus on high speed networking to support the intramural research program of the NIH; NLM serves as a national resource, providing support for medical centers to connect to the Internet, and developing prototype biomedical digital image libraries that use the Internet as a high speed distribution channel. An Intelligent Gateways project, co-sponsored by NLM and NSF, is developing methods to link dissimilar databanks over the Internet, using automated "Knowbots" (Knowledge Robots).

ASTA: Molecular Biology and Biomedical Imagery

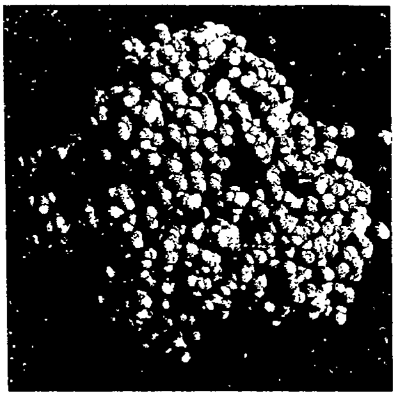
Each of the four NIH organizational units is developing algorithms and software for advanced, high performance computing environments. The two major themes of this work are molecular biology and biomedical imaging. Molecular biology computing includes comparison of genetic and protein sequences, and early development of algorithms to predict molecular structure and function. Biomedical imaging includes imaging of molecules and new methods for correlating and displaying clinical images in three dimensions.

IITA: Applications For Improved Health Care Delivery

In FY 1994 NIH will support the development of HPCC technologies for health care, with particular emphasis on the following applications:

- Testbed networks for linking hospitals, clinics, doctors' offices, medical schools, medical libraries, and universities to enable health care providers and researchers to share medical data and imagery.
- Software and visualization technology for visualizing the human anatomy and analyzing imagery from X-rays, CAT scans, PET scans, and other diagnostic tools.
- Virtual reality technology for simulating operations and other medical procedures.
- Collaborative technology to allow several health care providers in remote locations to provide real-time treatment to patients.
- Database technology to provide health care providers with access to relevant medical information and literature.
- Database technology for storing, accessing, and transmitting patients' medical records while protecting the accuracy and privacy of those records.

Two to five research and development projects in each of these areas will be supported via an NIH Broad Agency Announcement contract mechanism issued in FY 1993.



The myoglobin protein with a hydration of 350 water molecules.

BRHR: Medical Informatics Fellowships and Biomedical Science Education

NIH provides training and an ongoing investment in fundamental technology development. NCRR and NLM both sponsor formal degree-granting fellowship training in medical informatics, and cross-disciplinary training of established investigators in the use of advanced computing systems and methods. NCRR also supports a pilot project to educate high school science teachers and their students about new computing technologies for biomedical science.

FY 1993 Accomplishments

HPCS: Evaluation of Parallel Systems

DCRT evaluated a 128 processor Intel Touchstone Gamma Prototype Parallel Computer and associated system software on biomedical problems in structural biology, computational chemistry, image processing, and genetic database searching.

DCRT designed and implemented remote file access and communication systems software for this computer.

NCRR funded a Kendall Square Research KSR-1 computer jointly with NSF and ARPA at the Cornell Theory Center, and is evaluating its efficiency and effectiveness for biological problems.

NREN: Medical Connections Program

NLM initiated a Medical Connections program in collaboration with NSF, to provide support for academic medical centers to connect to the Internet. In addition, NLM has provided support for a pilot project to connect rural hospitals in the Northwest to the Internet. This project is coordinated by the University of Washington Health Sciences Center Library, a Regional Medical Library in the National Network of Libraries of Medicine.

NREN: Access to X-ray Images

NLM began developing a system to provide Internet-mediated access to an electronic archive of 20,000 digital X-ray images for medical research purposes.

NREN: UMLS Research

NLM created and published on CD-ROM a prototype Metathesaurus, Semantic Network, and Information Sources Map as part of the Unified Medical Language Systems research effort to link together computer-based biomedical information resources. These data files support experimentation on advanced systems that translate user information requests into multiple access vocabularies.

NREN: Digital Anatomy Demonstration Project

NLM initiated a demonstration project that will create a digital anatomy workstation that allows a health professions student to browse three-dimensional images generated by a specialized graphics computer at a distant site. The prototype links the NLM's Learning Center for Interactive Technology in Bethesda to the University of Washington at Seattle via the Internet.

NREN: Access to Cancer Treatment Guidelines

NCI implemented remote file access to treatment guidelines from its PDQ (Physician Data Query) database via the Internet.

High performance systems under evaluation for biomedical applications

Location	*Vendor*	*Machine(s)*
Cornell Theory Center	*Kendall Square Research*	*KSR-1 (64 nodes)*
	IBM	*SP1*
Pittsburgh Supercomputer Center	*Cray*	*C-90 & T3D*
University of Illinois at Champaign-Urbana	*Thinking Machines*	*CM200*
University of Illinois at Champaign-Urbana and Columbia University	*Thinking Machines*	*CM5*
Columbia University	*Convex*	*C210*

ASTA: Parallel Methods for Biomedical Applications

NCRR-supported High Performance Computing Resource Centers are evaluating high performance systems for biomedical research applications (see Side Bar). Investigators at these Centers have developed applications including:

- Modeling the molecular dynamics of proteins in membranes, in order to enable the study of drug permeability across membranes. This requires simulating approximately 10,000 atoms in a membrane patch.

- Simulating the complex interactions between DNA and protein (specifically Eco RI, a restriction enzyme). This work won the 1993 Computerworld Smithsonian Award.

- Developing a "dynamical model" for the structure of proteins essential for the replication of the human immunodeficiency virus 1 (HIV-1), which may help in generating inhibitors of the protein.

- Determining the three-dimensional structure of proteins by developing algorithms that model how proteins fold based on their amino acid sequences.

- Modeling the ability of molecules to bind to target receptors in solution by calculating their "solvation energies." This is a crucial component of rational drug design.

DCRT developed parallel methods for the following biomedical applications:

- The three-dimensional reconstruction of herpes virus images from two-dimensional projections of the virus obtained from electron micrographs.

- The calculation of the solvent accessible surface area of proteins that is used to predict the conformation of these molecules.

- The generation of primary sequence patterns from sets of related protein sequences contained in a database.

- The within-subject registration of PET (Positron Emission Tomography) images for the correction of roll, yaw, pitch, and translation.

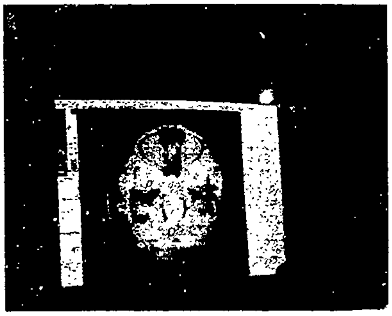
- The processing of NMR (Nuclear Magnetic Resonance) spectroscopy data using the Maximum Entropy Method for determining the structure of proteins.

- Ab initio quantum mechanical calculations for molecules of biological interest.

- Molecular dynamics calculations with a parallel implementation of CHARMM, a molecular modeling program that takes atomic coordinates of molecules and renders graphical images of those molecules that can be manipulated in various ways for experimental chemistry.

Calculation of the binding of small molecules to DNA and other biological macromolecules depends upon a complex interaction of electrostatic fields, solvent-accessible surface, and other physicochemical properties. Structure-based drug design is a key biomedical HPCC Grand Challenge.

- Human genetic linkage analysis to determine the likely position of a disease gene using LINKAGE, a gene linkage analysis program that calculates the probability of the association between a pattern of gene inheritance and a disease condition. It is used to analyze family pedigrees based on data obtained from gene probes.



Two-dimensional clinical images from computed tomography and magnetic resonance imaging underpin modern medical diagnosis. Transmission of these diagnostic images over wide area networks and their reconstruction to form three-dimensional views are important health care applications of HPCC technologies.

ASTA: "Visible Human"

NLM initiated a two-year project to acquire the three-dimensional digital representation of entire human beings at millimeter-level resolution, derived from computed tomography, magnetic resonance imaging, and digitized cryosections. This "Visible Human" research data set will become available nationally via the Internet in 1994.

ASTA: Prototype Program for Retrieving Molecular Biology Information

NLM created a prototype advanced molecular biology information retrieval program that provides integrated access to genetic and protein molecular sequences, and the biomedical literature linked to those sequences. Field testing of the system has begun.

ASTA: Faster Molecular Analysis and Imaging Algorithms

NCRR and NLM achieved order of magnitude speedups in several existing molecular analysis algorithms.

DCRT and NCRR developed new algorithms for registration and rendering of three-dimensional images from two-dimensional clinical images and micrographs.

ASTA: HIV Research

Research conducted at NCI's Biomedical Supercomputer Center is increasing the understanding of the human immunodeficiency virus (HIV) that causes AIDS and is helping to design and develop new drugs to combat the deadly disease. NCI researchers have successfully predicted the secondary structure of the entire 9,000 unit HIV virus RNA. NCI and NCRR supercomputing applications have assisted in the design of new drugs to inhibit HIV replication.

BRHR: Medical Informatics Grants and Other Training

NLM competed the award of 10 Medical Informatics Training Grant programs at academic medical centers. The program supports cross-disciplinary training of health professionals in the use of advanced computing technologies.

NCRR conducted a pilot project to introduce scientific computing methods to high school science teachers and their students.

DCRT and NCRR sponsored "hands on" training of biomedical researchers in the use of new computational biology tools at NSF Supercomputer Centers and on the NIH campus in Bethesda.

BRHR: Basic Research Through Long Distance Microscopy

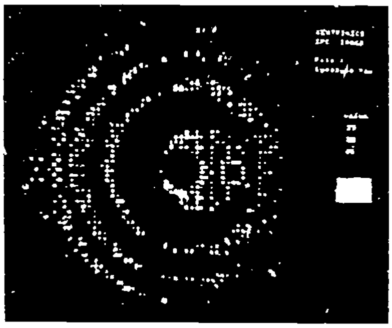
In the first demonstration of its kind, scientists at a workstation in Chicago viewed high-resolution images of nerve cells in a high-voltage electron microscope located 1,700 miles away at the San Diego Microscopy and Imaging Resource (SDMIR), which is supported by NCRR. Further details are given on pages 123-125 in the Case Studies section.

FY 1994 Milestones

NIH will accelerate the pace of molecular and genetic discovery by enabling the solution of currently intractable problems in molecular structure prediction, drug design, and human genome database analysis.

The program will apply and evaluate new computer architectures to key problems of human health and disease, in a manner that gives early feedback to computer designers on the strengths and limitations of their systems for medical applications. The HPCC Program will rapidly build an electronic community among life science researchers by connecting academic medical centers to the Internet. It will create prototype medical imaging applications that use the Internet and provide a model for distance-independent medical consultation. It will double the pool of computationally trained investigators in biomedicine.

Digital radiology techniques require high speed networks: a single X-ray film represented by a 2K-by-2K-by-10 bit gray scale generates a 4 megabyte image file.



X-ray diffraction spectroscopy is a laboratory method for determining the folded structure of proteins and other biological macromolecules. Parallel computing systems can be used to automate interpretation of X-ray diffraction patterns acquired via two-dimensional array sensors.

HPCS: Evaluation of Parallel Systems

DCRT will obtain a next generation parallel computer that will allow the solution of new computationally intensive problems in biomedicine.

NCRR will assess the efficiency and scalability of emerging massively parallel architectures for Grand Challenge problems.

NREN: More Connections to the Internet

NLM will establish Internet connections to 70 to 100 additional medical centers.

NREN: Digital Anatomy Databases

NLM will create advanced three-dimensional imaging databases for digital anatomy on the Internet by the nation's health professions schools. Workstations to access and display those images will be developed.

NREN: "Knowbots" for Natural Language Queries

An operational Knowbot-based database retrieval system will be deployed at NLM to provide integrated access to over 50 computerized knowledge sources in biomedicine, underpinned by a fully operational Unified Medical Language System that allows users to state scientific questions in their own language, and have the answer retrieved and synthesized automatically from multiple databases at multiple sites on the Internet.

ASTA: Software Development and Hardware Placement

NCI will expand its advanced software development program to allow research on a broader range of molecular dynamics, structure-function problems, and structure-assisted drug design.

NLM will develop and deploy advanced molecular biology workstations to approximately 1,000 molecular biology laboratories.

Through funding of five biomedical High Performance Computing Resource Center (HPCRC) programs at

NSF and ARPA-sponsored high performance computing centers, NCRR will support development of algorithms to compare molecular sequences, predict molecular structure from genetic and protein sequences, simulate protein folding, and model complex biological systems such as proteins interacting with membranes in an aqueous solution. Computed images of biological structure, from molecular to whole body, will be an additional area of ASTA software development.

IITA: Health-Related Research and Development

NIH will support the development of HPCC technologies through a Broad Agency Announcement to support two to five research and development projects in each of the six health-related areas listed above.

NCI will provide Xconf, a prototype multimedia (including medical images) group conferencing tool that uses the Internet. Contingent on funding, NCI will 1) connect to the Internet both medical research centers (to facilitate multicenter environmental epidemiology studies) and the Cancer Information Service offices at its cancer centers, and 2) begin developing telemammography over high speed networks.

BRHR: New and Upgraded Centers and Additional Fellowships

In order to meet the needs of biomedical researchers, NCRR will establish at least one additional high performance computing resource center and upgrade the five existing centers. Enhanced cross training of biomedical research scientists will also be possible.

NLM will fund an additional 50 medical informatics training fellowships nationwide.

National Security Agency (NSA)

The goal of NSA's HPCC Program is to accelerate the development and application of the highest performance computing and communications technologies to meet national security requirements and to contribute to collective progress in the Federal HPCC Program.

In support of this goal, NSA:

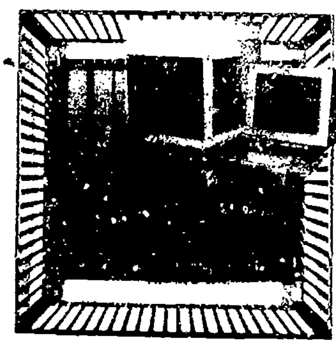
- Develops algorithms and architectural simulators and testbeds that contribute to a balanced environment of workstations. vector supercomputers, massively parallel computer architectures, and high speed networks.

- Sponsors and participates in basic and applied research and development of gigabit networking technology.

- Develops network security and information security techniques and testbeds appropriate for high speed in-house networking and for interconnection with public networks.

- Develops software and hardware technology for highly parallel architectures scalable to sustained teraops performance.

- Investigates or develops new technologies in materials science, superconductivity, ultra-high-speed switching and interconnection techniques, networking, and mass storage systems fundamental to increased performance objectives of high performance computing programs.



By integrating processing directly into otherwise standard memory chips fabricated at its Special Processing Laboratory, the Supercomputing Research Center has developed the ***Terasys*** *workstation, which outperforms one Cray Y-MP processor by 5 to 48 times on a set of nine NSA applications. A mature software environment, available for the workstation, makes the system easy to use.*

NSA participates in all five components of the HPCC Program as follows.

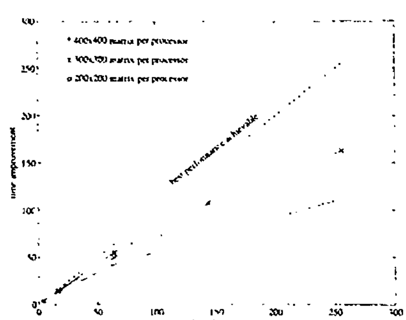
HPCS: Heterogeneous High Performance Computing; Balanced Architectures

NSA deploys experimental scalable computer capabilities, emphasizing interoperability of massively parallel machines in a highly heterogeneous environment of workstations, vector supercomputers, and mass storage devices. NSA's HPCS program also emphasizes an open systems approach to development and maintenance of the HPCC environment and the integration of specialized high speed hardware within this environment.

NREN: Network and Security Technology

NSA uses the Internet to provide high speed network connection among NSA, industry, and academic researchers. NSA takes the lead in developing network security and other information system security technology and products for high speed networks. It establishes high speed network testbeds to explore network and security interface technology issues.

ASTA: High Performance Systems Software, Tools, and Algorithm Research



Time improvement using multiple processors vs. a single processor for a parallel eigensolver algorithm for large dense symmetric matrices, which occur frequently in models of physical phenomena. A parallel algorithm "scales well" when it effectively uses almost all available resources as the problem size and the number of processors increase. The achievable performance improvement is proportional to the increase in the number of processors. Parallel algorithm design and implementation is an iterative process with the goal of achieving the maximum performance possible. SRC research continues seeking ever-faster eigensolvers. These data are from the Intel Touchstone Delta at Caltech. Results are being collected for the IBM SP1, the Thinking Machines CM5, and the Intel Paragon.

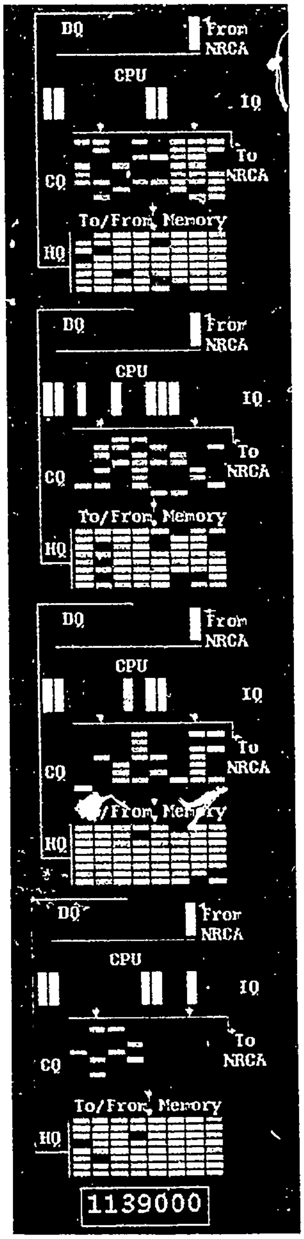
NSA develops systems software to enhance productivity of users of high performance systems. This software includes compilers, simulators, performance monitoring tools, software that is portable among a variety of high performance systems, software for distributing jobs across a network, and visualization tools. New algorithms are developed to map problems to high performance architectures.

IITA: Development of Dual-Use Technology

NSA proposes to investigate developmental technologies to support information infrastructure applications in manufacturing, education, public health, and digital libraries. Research and development in transaction processing, database management systems, and digital data storage and retrieval systems, are expected to make strong dual-use technology contributions to both the Federal and the IITA communities.

BRHR: Fundamental High Performance System Research and Education

NSA supports basic research into new technologies, theory and concepts of high performance computing and networking, and promotes research in high performance computing at the Institute for Defense Analyses' (IDA) Supercomputing Research Center (SRC) and at univer-



Memory traffic flow from CPU memory ports through a memory arbitration network, and back to CPU input ports. Each colored rectangle represents one physical queue organized in groups. The colors represent the queue state: for example an empty queue is white and a full queue is red. The animation of this display shows memory message traffic through the memory network. Its purpose is to point out possible clogs and busy spots within the network.

sities. NSA's National Cryptologic School and the SRC develop courses suitable for users of high performance computing environments.

Management

The overall coordination of NSA's HPCC activities resides in the office of the NSA Chief Scientist, who reports to the Director of NSA. Responsibilities for execution of major elements of the plan lie with the Technology and Systems Directorate and the Information Systems Security Directorate. NSA sponsors the SRC in exploring massively parallel computer architectures and in developing algorithms and systems software for parallel and distributed systems. The SRC heavily supports NSA's HPCC activities directly and by collaborative efforts with other HPCC participants, especially industry and academia.

FY 1994 Plans

HPCS

Extend existing high performance system simulators and testbeds at NSA and the SRC.

Develop mass archival storage – 10^{15} to 10^{16} bits.

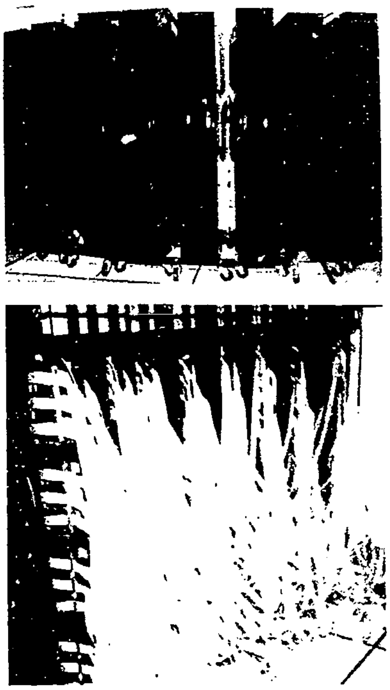
Define techniques for the interoperability of distributed operating systems spanning large sets of heterogeneous computer assets, including supercomputers.

Integrate major new vector/scalar massively parallel processor as a research system.

Demonstrate a terabit/second (10^{12}-bit-operations-per-second) deskside SIMD system.

Continue technical cooperation with major vector/massively parallel processor developers.

Investigate, cooperatively with industry, processing-in-memory (PIM) technology in established systems architectures.



Front and rear views of High Speed Network (HNET) Testbed.

These 64 Unix processors attached to 64 custom-chip switch nodes, each with seven ports, serve as a high speed network routing protocol testbed. The system can be configured to be an Asynchronous Transfer Mode (ATM) switch, for example, to allow experimentation with the efficiency of currently evolving commercial offerings.

NREN

Install in-house gigabit network testbed to explore high speed network architectures and techniques.

Explore network security issues for 622 Mb/s and 2.4 Gb/s networks.

Install a gigabit network testbed to explore compatibility of DOD networks with vendor-provided public networks.

Initiate development of a bulk link encryptor for high speed networks.

Develop proof of concept for cell encryption in very high speed switched network products.

ASTA

Develop parallel extensions to high-level programming languages suitable for parallel architectures.

Develop visualization techniques for performance analysis of parallel and vector high-performance architectures.

Develop system modeling tools for heterogeneous computing environments, including high performance systems.

Develop algorithms and implementations for solving eigensystems and manipulating sparse matrices on parallel systems.

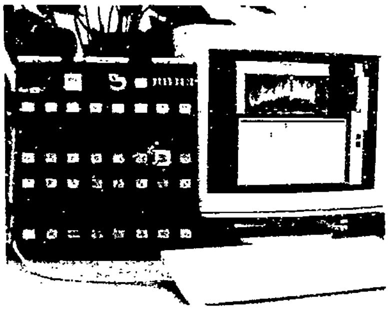
Develop evaluation testbed for exploring routing algorithms and topologies for computer interconnects.

IITA Candidates

Study applicability of heterogeneous data base technology program to information infrastructure application needs.

Investigate transferability of digital library technology to the private sector.

Research public sector security technology issues unique to the IITA as well as dual-use technologies



NSA has used parallel processing to greatly accelerate workstation performance. Designed and fabricated at the SRC, this board employs 32 Field Programmable Gate Arrays which are custom programmed for each application. This customization and parallelization yields speedups of 100 to 1,000 times the already impressive workstation performance. (Field Programmable Gate Arrays are also described on pages 154-155 in the Case Studies section.)

applicable to both national security and public sector communities.

BRHR

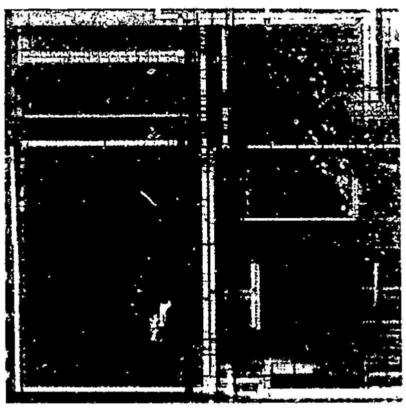
Initiate collaborative and university research efforts in performance modeling of high performance computing systems for generic problem domains.

Explore innovative parallel computer and network architectures.

Investigate new technologies in material sciences, superconductivity, optoelectronics, ultra-high-speed interconnection, and switching.

National Institute of Standards and Technology (NIST)

High performance computing and communications technology is an essential enabling component of NIST's mission to promote U.S. industrial leadership and international competitiveness and to provide measurement, calibration, and quality assurance techniques to support U.S. commercial and technological progress. The objectives of NIST's HPCC program are: to accelerate the development and deployment of high performance computing and networking technologies required for the National Information Infrastructure; to apply and test these technologies in a manufacturing environment; and to serve as coordinating agency for the manufacturing component of the Federal Program.



NIST's MultiKron chip provides low perturbation measurements for performance evaluation of computer and communication systems.

Specific goals of NIST's program are:

- To apply high performance communications and networking technology to promote improved U.S. product quality and manufacturing performance, to reduce production costs and time-to-market, and to increase competitiveness in international markets.
- To promote the development and deployment of advanced communications technology to support the education, research, and manufacturing communities and to increase the availability of scientific and engineering data via the National Information Infrastructure.
- To advance instrumentation and performance measurement methodologies for high performance computing and networking systems and components to achieve improved system and application performance.
- To develop efficient algorithms and portable, scalable software for the application of high performance computing systems to industrial problems, and to develop improved methods for the public dissemination of advanced software and documentation.
- To support, promote, and coordinate the development of voluntary standards that provide interoperability and common user interfaces among systems.

NIST participates in four components of the HPCC Program:

HPCS: Performance Measurement of Scalable Systems

NIST develops instrumentation and methodology for performance measurement of high performance networks and massively parallel computer systems. Emphasis is on the use of low-perturbation data capture hardware and simplified software-based approaches to performance characterization.

NREN: Networking and Information Infrastructure

NIST supports and coordinates the development of standards within the Federal government to provide interoperability, common user interfaces to systems, and enhanced security. NIST works with other agencies to promote open system standards to aid in the commercialization of technology by U.S. industry. NIST promotes the development of communications infrastructure and the use of the Internet through information technology research, development, and related activities to enhance basic communications capabilities.

ASTA: Application Systems and Software Technology

NIST develops algorithms and generic software for advanced scientific, engineering, and manufacturing applications. Common elements and techniques are encapsulated in software libraries to promote ease of use and application portability. NIST's Guide to Available Mathematical Software (GAMS) provides industry and the public with improved electronic access to reusable software.

A computer-controlled coordinate measuring machine determines the exact dimensions of a precision-machined stainless steel part. NIST researchers use this and similar instruments to devise ways to improve machine tool performance.

IITA: Systems Integration for Manufacturing Applications

NIST will build upon its experience in information technology and manufacturing engineering to accelerate the application of high performance computing and communications technology to manufacturing environments.

NIST will support expanded programs in advanced manufacturing systems integration technologies; development and testing of prototype components and interface specifications for manufacturing systems; application of high performance computing and networking technologies to integrate design and production processes; and testbeds for achieving cost-effective application of advanced manufacturing systems and networks.

FY 1993 Accomplishments and FY 1994 Plans

A machinist monitors a machine tool retrofitted with a personal computer controller. The machine is located in NIST's Shop of the 90s where manufacturers can learn how to use open system integration technology and low-cost automation techniques to improve productivity and product quality.

Systems Integration for Manufacturing Applications

Beginning in FY 1994, NIST will establish an Advanced Manufacturing Systems and Networking Testbed to support research and development in high performance manufacturing systems and to test high performance computer and networking hardware and software in a manufacturing environment. The testbed will serve as a demonstration site for use by industrial technology suppliers and users, and to assist industry in the development and implementation of voluntary consensus standards. Research and testing will be conducted at the NIST testbed as well as at testbeds funded through the NIST Advanced Technology Program. A manufacturing systems environment will be developed to support the integration of advanced manufacturing systems and networking software and products. A standards-based data exchange effort for computer integrated manufacturing will focus on improving data exchange among computer aided design, process, and manufacturing activities. Prototype systems and interface specifications will be communicated to appropriate standards organizations. Results will be made available to U.S. industry through workshops, training materials, electronic data repositories, and pre-commercial prototype systems that can be installed by potential vendors for test and evaluation. One role of advanced computing technology in manufacturing process modeling and simulation is described on pages 150-151 in the Case Studies section.

Networking and Information Infrastructure

NIST performance evaluation activities include measurement and characterization of the impact of software protocols on communication performance in order to minimize communication bottlenecks between application programs. NIST researchers have implemented a high speed communications testbed that provides a HIPPI link to a performance instrumented workstation. Testbed performance instrumentation includes a modified NIST MultiKron performance data capture chip interfaced to the test workstation. The workstation HIPPI interface was designed, implemented and obtained through a collaborative effort with the VISTANet project, one of the NREN gigabit testbeds.

In FY 1994, NIST will employ systems instrumented with the MultiKron chip to investigate the behavior and performance of communications protocols suitable for multi-gigabit/second transmission rates. For high performance networking, NIST will support transition planning and deployment of ISDN and OSI-based protocols in the Internet and undertake research and development activities to support emerging Broadband ISDN standards. NIST will continue to interact with industry by sponsoring and hosting the North American ISDN Users Forum.

Since the early 1970s, NIST has been developing cost-effective ways to help protect computerized data. NIST has devised a prototype system for controlling access to a computer system that uses a password, a smart card, a fingerprint reader, and cryptography.

NIST activities in the areas of networking and communications technology will be expanded to address applications of information infrastructure including electronic commerce, distributed multimedia environments, and adaptive systems. Support for electronic commerce will focus on enabling electronic exchange technology and protocols to support business transactions and manufacturing techniques. NIST will develop improved electronic data interchange methodology to describe, access, and update data. In FY 1994, an integration facility will be created for manufacturing applications of electronic commerce.

NIST activities in distributed multimedia environments will include the development of methodology to acquire and administer the large amounts of text and multimedia material associated with electronic libraries and the implementation of information retrieval mechanisms for easy access to information by untrained end users. NIST activities in adaptive systems will include research in wireless communication compression and encoding technology. NIST will also conduct workshops to assess network security requirements and develop network security technology suitable for Internet and other networking technologies.

Performance Measurement

NIST developed the MultiKron chip to achieve low-perturbation data capture while assessing the performance of scalable high performance computer systems. This technology was transferred to Intel Corporation for application in its Paragon systems. NIST also devised a software-based technique for portable sensitivity measurements of parallel programs. The new method is based upon statistical experimental design principles and applies to any MIMD system. A low-cost setup approach makes the technique practical, and promising preliminary results on shared-memory and distributed memory systems have been obtained.

In FY 1994, NIST will extend its hardware and software tools for system performance monitoring through the acquisition of a MultiKron instrumented scalable system. Evaluation activities will focus on the performance characterization of application programs running on this system and on other scalable systems.

Software Development, Distribution, and Reuse

Statistical analysis of video microscope images provides improved precision in determining optical fiber geometry.

The NIST Guide to Available Mathematical Software (GAMS) project develops techniques and tools to help scientists and engineers locate and use computer software to solve mathematical and statistical problems. The GAMS system has been incorporated in the Software Exchange Experiment coordinated by the NASA Goddard Space Flight Center. The status of GAMS and future plans are described on pages 148-149 in Case Studies section.

In FY 1994, NIST algorithm and software development activities will address needs for improved computational performance and visualization capability in application areas drawn from biotechnology and from chemical and materials process design and simulation.

National Oceanic and Atmospheric Administration (NOAA)

NOAA's Grand Challenge research in climate prediction and weather forecasting is critical to its mission to describe and predict changes in the Earth's environment, manage the Nation's ocean and coastal resources, and promote global stewardship of the world's oceans and atmosphere. This research depends on advances in high-end computing and on the collection and dissemination of environmental data.

NOAA Grand Challenges Requiring HPCC Resources

- *Environmental assessment and prediction*
  - *Global and regional modeling to support short term forecasting and warnings*
    - *Weather*
    - *River flow and water resources*
    - *Solar and space*
    - *Living marine resource forecasts*
  - *Assessment of seasonal, interannual, and long-term global environmental change*
    - *Global change modeling*
    - *Ocean modeling*
    - *Coastal analysis and assessment*
- *Environmental data and information dissemination*
  - *Climate and weather*
  - *Oceanographic*
  - *Geophysical*
  - *Navigational, including charts of U.S. waters and airspace*

Increased computing power will enable higher resolution in the current models of the Earth's atmosphere-ocean system. Increased resolution will enable accurate representation of key features such as weather fronts and ocean eddies, and eliminate distortions due to clouds. More accurate NOAA models will improve the understanding of the behavior of climate and weather systems, making possible better decision making by government and industry on issues that affect both the environment and the economy.

NOAA is the Federal agency responsible for archiving and disseminating the Nation's environmental data, and is the principal repository for the Nation's climatic data for the U.S. Global Change Research Program. By increasing NOAA connectivity to the Internet, researchers and other users will have better access to this growing collection of large data sets located at more than a score of sites.

Through HPCC efforts in climate modeling, NOAA will provide better simulations of atmosphere-ocean coupling and a first-ever direct attack on the regional climate change problem. More accurate and more timely assessment of the future impact of climate change will make it possible to avoid "false choices" between the economy and the global environment. In weather forecasting, finer resolution in global and regional models will result in better weather forecasting and warning services, especially for hazardous weather and flight safety.

NOAA participates in the NREN, ASTA, IITA, and BRHR components of the HPCC Program as follows.

NOAA Laboratories Involved in HPCC

- *Geophysical Fluid Dynamics Laboratory (GFDL) Princeton, New Jersey*

- *Forecast Systems Laboratory (FSL) Boulder, Colorado*

- *National Meteorological Center (NMC) Camp Springs, Maryland*

NREN: Improving Access to Systems and Data

By using the Internet, NOAA researchers can easily access massively parallel systems at distant locations. Geographically distributed researchers can collaborate and easily share NOAA computing resources and data by accessing these systems remotely from high performance workstations.

NOAA plans to make its vast environmental data archives more accessible to the scientific community, while preserving the integrity of operational NOAA systems. Greater NOAA Internet connectivity will substantially improve access to these data. To support this increased Internet use, a NOAA Internet Network Information Center has been established. The Center will provide information and assistance to all NOAA Internet users and to scientists at other agencies and in academia with whom they collaborate.

ASTA: Environmental Grand Challenges

NOAA is developing advanced algorithms and redesigning climate prediction and weather forecasting models to use new parallel programming paradigms.

NOAA will conduct a phased acquisition of scalable parallel systems for use in climate prediction and weather forecast modeling. The new models will be installed and evaluated on these systems.

IITA: Environmental Applications and Improved Data Accessibility

NOAA proposes to investigate environmental monitoring, prediction, and assessment applications, and to expand efforts to make its environmental data more accessible.

BRHR: Increasing the Base of Researchers and Users

The number of NOAA users who can effectively use scalable systems will grow considerably. In addition, substantially more visiting scientists will work with NOAA researchers on Grand Challenge problems, encouraging both evolutionary improvements and creative breakthroughs in climate prediction and weather forecasting.

FY 1993 Activities and Accomplishments and FY 1994 Plans

NOAA National Data Centers

- *National Climatic Data Center Asheville, North Carolina*
- *National Oceanographic Data Center Washington, D.C.*
- *National Geophysical Data Center Boulder, Colorado*

Network Access

NOAA's pilot Internet Network Information Center has begun to provide assistance in using the Internet, network management for connected NOAA facilities, mail and other server capabilities, and protected access to selected NOAA operational data sets, such as NMC model output data.

In FY 1994, at least 10 more NOAA computational and data archiving facilities will be connected to the Internet, and connection bandwidth will be increased.

Global Atmosphere and Global Ocean Modeling

Over the past two years, NOAA's Geophysical Fluid Dynamics Laboratory (GFDL) has begun redesigning its most important atmospheric and oceanic models to make them modular and parallel in design. This activity includes a collaborative effort by GFDL with DOE/Los Alamos National Laboratory (LANL) scientists that has led to the successful design of a model shell version of the GFDL Modular Ocean Model (MOM) to a form that is amenable to highly parallel computers and that maintains its modular coding structure. High resolution experiments using this new code were performed on the 1,024-node Thinking Machines CM-5 at LANL.

GFDL's complete SKYHI global atmospheric grid-point model, which is used to study climate, middle atmosphere dynamics, and atmospheric ozone, has been redesigned for parallel execution and ported to the LANL CM-5, as part of the GFDL-LANL collaboration.

The GFDL radiation physics package, which is a critical part of all GFDL atmospheric climate models, was redesigned to a modular form that is suited to highly parallel computer systems.

In FY 1994, high resolution atmospheric modeling experiments using a simplified-physics version of a parallel atmospheric grid model will be performed to test model sensitivity to grid resolution.

In FY 1994, a parallel version of the GFDL limited-area non-hydrostatic model will be developed for use in investigating the interaction between radiation and clouds.

A computer model representation of a blizzard over the Colorado Front Range area, produced jointly by NOAA's Forecast Systems Laboratory (FSL) and the Colorado State University. The image represents a six hour forecast, looking over Colorado from the southeast. The white/gray area depicts the forecast region for significant clouds, and the red area within the clouds represents the region in which aircraft icing is likely to occur.

Weather Forecast Modeling

An adiabatic shell for the NOAA National Meteorological Center (NMC) spectral model has been developed and executed on a Thinking Machines CM-200 at DOD/Naval Research Laboratories (NRL) in collaboration with NRL.

A regional atmospheric model has been restructured for execution on highly parallel systems by NMC, in collaboration with NOAA's Forecast Systems Laboratory (FSL).

The initial restructuring and recoding of the regional/mesoscale Optimum Interpolation (OI) analysis code for execution on the CM-200 at DOD/NRL has been completed.

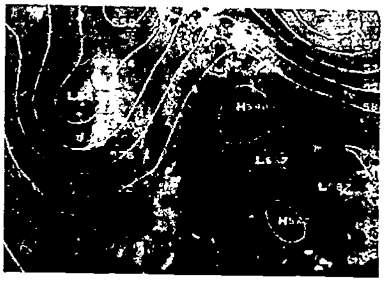
FSL has developed a parallel version of the Mesoscale Analysis and Prediction System (MAPS) as a functional prototype system for both the Federal Aviation Administration and the National Weather Service.

MAPS is the first of several strategic weather models to be parallelized for the Aviation Weather Program. These models will provide high resolution forecasts on both the national and regional levels to support operational and aviation meteorology. The MAPS model was engineered using an FSL-specified layered software approach to provide portability among scalable parallel systems.

NOAA HPCC Systems

A scalable architecture system, a 208-node Intel Paragon, is being installed at NOAA/FSL in conjunction with the ARPA-sponsored National Consortium for High Performance Computing and as a part of the Boulder Front Range Consortium. Consortium members include NCAR (under NSF sponsorship), the University of Colorado (under ARPA sponsorship), and FSL.

FSL has developed a benchmark suite to evaluate parallel processors for use in weather analysis and prediction applications. FSL has also specified and is imple-



Contours representing the wind flow in the middle atmosphere, produced by a weather forecast model running on the Cray Y-MP8 at NOAA's National Meteorological Center, superimposed on a NOAA GOES satellite infrared cloud image.

menting a layered software approach to support the portability of such applications between various scalable systems and networks of workstations, and has taken the initial steps toward implementing a meteorological software library using this approach.

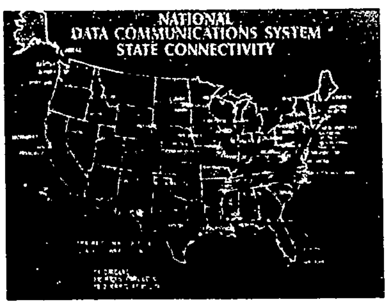
GFDL has solicited technical input from massively parallel computer vendors through a public comment process as part of its effort to redesign its primary atmospheric climate model to a suitable parallel design. The resulting technical collaborations are leading to a more effective code design for this and other models that are being rewritten for parallel systems.

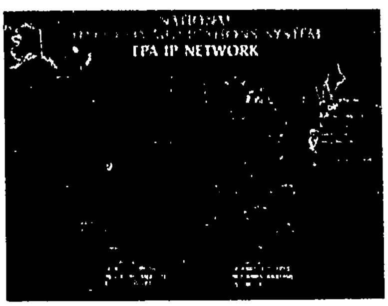
Environmental Protection Agency (EPA)

Central to EPA's mission are its Grand Challenges in air and water pollution management and in ecological assessment. The complex nature of these Grand Challenge problems demands use of the most comprehensive and integrated computational assessment tools available. By the mid-1990s, the achievement of EPA's HPCC Program goals will provide more reliable and useful tools to develop improved national pollution control and prevention strategies involving billions of dollars in control costs.

EPA has three main goals for its HPCC Program activities:

- Advance the capability of environmental assessment tools by adapting them to a distributed heterogeneous computing environment that includes scalable massively parallel architectures.
- Provide more effective solutions to complex environmental problems by developing the capability to perform multipollutant and multimedia pollutant assessments.
- Provide a computational and decision support environment that is easy to use and responsive to environmental problem solving needs to key Federal, state, and industrial policy-making organizations.





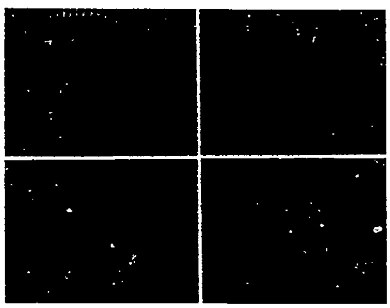
EPA participates in the NREN, ASTA, IITA, and BRHR components of the HPCC Program as follows.

NREN: Increasing Access to a Heterogeneous Computing Environment

EPA will expand Internet connectivity to a critical mass of environmental problem-solving groups at the Federal, state, and industrial levels and to environmental Grand Challenge research teams, enabling distributed computing approaches for linking complex ecological models or multimedia environmental models. EPA is establishing a research network to support geographically distributed collaborative development of prototype environmental assessment frameworks using a distributed, heterogeneous computing environment that includes massively parallel components and distributed data access. EPA will also provide environmental applications software for use in performance and reliability testing of Internet technology.

ASTA: Environmental Assessment Grand Challenges

EPA is integrating advanced environmental assessment tools into distributed, heterogeneous computing environments. A major focus is on the development of parallel algorithms for atmospheric chemistry, transport, and molecular models to enable more effective study of multipollutant air quality issues, the impact of man-made chemicals on the environment, and the relationship between air and water pollution. In the latter case, the linkage of distinct air and water models will enhance the knowledge of complex interactions of pollutants crossing media boundaries, enable integrated assessment of pollutant impacts, and form the foundation for further integration with ecological process models. Such work will ensure that the interaction of multiple pollutants in several media (such as air and water) are considered when assessing pollutant abatement strategies. These assessments are not feasible without HPCC technology.



Researchers evaluate molecular properties of carcinogens from pollution sources. The image shows electric field vectors for benzopyrene.

Environmental scientists and regulatory analysts lack adequate assessment tools because computational complexities and slow response time inhibit effective and extensive use of the most advanced environmental assessment tools. Development of a user friendly framework that puts the power of HPCC technology and advanced multipollutant environmental models directly in the hands of Federal, state, and industry groups charged with solving environmental pollution problems is a primary EPA goal.

IITA: Enhancing User Access to Environmental Data and Systems

EPA, in concert with NASA and NOAA efforts to disseminate environmental information, proposes to expand public access to a variety of environmental databases, such as ecological measurements, air and water quality model predictions, and population exposure to pollutants. Development of intelligent user interfaces for interactive browsing, retrieval, analysis, and on-line multimedia tutorials for inexperienced users will facilitate routine use of environmental models for regulatory analysis and environmental education. In addition to these access and training systems, EPA proposes to support curriculum development for environmental education for grades 9-12 as well as university undergraduate and graduate programs.

BRHR: Broadening the User Community

EPA will develop and evaluate methods and materials for training Federal, state, and industrial users of advanced environmental assessment tools.

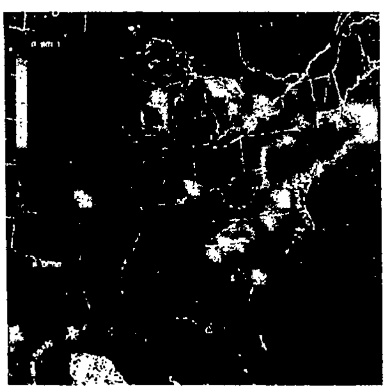
It has initiated a program to train analysts who support environmental decision making in advanced computing technology. The scope of this program ranges from making it as easy as possible to use the models to developing interactive software for on-line training.

EPA also sponsors fellowships, graduate student programs, and high school computational science programs to develop interdisciplinary skills required for air and water quality modeling.

Major FY 1993 Activities and Accomplishments and FY 1994 Plans

Multipollutant Air Quality Management

Working with the North Carolina Supercomputing Center and several universities, EPA has initiated a Grand Challenge air quality management project in which a user friendly, multipollutant, multiscale air quality modeling system is being developed for modeling research, evaluation, and regulatory assessment.



Regional Acid Deposition Model predictions of nitric acid over the eastern U.S.

The first prototype of the system will address regional and urban-scale ozone and regional acid deposition issues using generic scalable algorithms for key physical and chemical processes, and advanced technology for data management, interprocess communications, analyses, and visualization.

EPA has also begun to evaluate the performance of both low and high end massively parallel computers on atmospheric chemistry and transport algorithms and molecular modeling codes critical in environmental assessment. The Regional Oxidant Model has been ported to a 1,000 processor MasPar for benchmarking. Molecular mechanics and dynamics algorithms are also being evaluated on a 32-node Kendall Square Research KSR-1 and an Intel iPSC/860 at the National Institutes of Health. A variety of partial differential equation solvers are also being implemented for performance evaluation on both the MasPar and KSR machines.

To foster communication among scientists currently performing molecular modeling research on highly parallel machines, EPA and the Office of Naval Research jointly sponsored the "Molecular Modeling on Parallel Computers" workshop session at the 1993 Sanibel Symposium.

Multimedia Modeling

A multimedia project has been initiated by porting three single-medium (or single-disciplinary) models to supercomputers – specifically a comprehensive regional air quality model, a watershed-water quality model, and a three-dimensional, bay hydrologic-water quality model. These models are used to address the nitrogen eutrophication of Chesapeake Bay – an environmental problem involving two major pollutant pathways of major concern to the Nation.

Three-dimensional rendition of salinity in Chesapeake Bay.

The three models have been ported to a Cray Y-MP, and the air quality model has been optimized there. Current research focuses on linking these models together when there is only minor interaction among the different media, thereby maintaining the full disciplinary complexity of known and tested individual models while demonstrating the added benefit of multimedia modeling for environmental decision making.

Visualization

As computer capability increases, interaction among the different media and synergisms involving scores of pollutants are being introduced into environmental models. As a result, these models are becoming increasingly complex. Visualization groups supported by EPA's HPCC Program have developed high-quality videos of a regional acidic deposition model and an air quality model, demonstrating the power that visualization has to open up the workings of these models to both scientists and non-scientists.

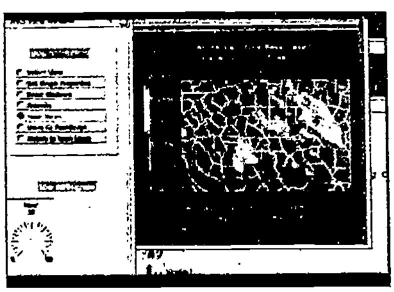
This effort was recently expanded as visualization experts began working with water quality scientists to display the output of increasingly complex sediment transport water quality models. As part of these efforts, a new volume-rendering algorithm has greatly enhanced the realism of the three-dimensional model data, and more visualization algorithms are under development.

Networking

EPA established a 100 Mb/s (FDDI) research network and has provided T1 connectivity to the Environmental Research Center in Georgia and to the Chesapeake Bay Office in Maryland. This connectivity provides the bandwidth needed to satisfy EPA's heterogeneous computing, data management, and collaborative visualization requirements. In FY 1994 EPA will establish a T3 Internet interconnect to EPA's National Environmental Supercomputer Center.

Technology Transfer

EPA is funding the establishment of a prototype training center and an on-site training program for Federal, state, and industrial personnel, initially to support air quality regulatory decision-making policies. Hands-on training with state-of-the-art environmental assessment techniques will foster competency in high performance computing; users will observe the increasing speed, accuracy, scope, and improved management of environmental assessments available through high performance computing.



Ozone concentrations seen through an AVS interface.

EPA is evaluating the usefulness of multimedia electronic tutor and help approaches, and collaborative computing and visualization to provide expert advice to remote locations. A prototype interactive analysis and visualization system for a workstation environment has been developed to support regulatory decision making at the state level. It uses the Urban Airshed photochemical model and the Applications Visualization System (AVS) software. The pilot user group, the North Carolina Department of Environment, Health, and Natural Resources, accesses the North Carolina Supercomputing Center at Research Triangle Park via a newly established T1 link. In FY 1994 the pilot program will be extended to additional Federal, state, and industrial environmental organizations.

Computational Education

EPA's HPCC Program helps support EarthVision, EPA's computational science educational program for high school students and teachers. Students prepare proposals during Saturday tutorials at EPA's National Environmental Supercomputing Center in Bay City, Michigan. Teams are selected to participate in a three-week summer educational program and to conduct environmental research using scientific workstations and EPA's supercomputer during the academic year.

EPA Milestones

NREN – FY 1994

- Complete T3 Internet interconnect to EPA HPCC Center.
- Increase bandwidth and number of state connections to the Internet.

ASTA – FY 1993 - 1994

- Establish Grand Challenge collaborations.
- Demonstrate key environmental chemistry algorithms on scalable massively parallel testbeds.

ASTA – Beginning in FY 1994

- Demonstrate linked air and water models in a distributed computing environment.
- Install 10 gigaops testbed for Grand Challenge teams.

BRHR – FY 1992 - 1994

- Establish a key set of collaborations for technology transfer of advanced HPCC environmental assessment capabilities to Federal, state, and industrial users.

BRHR – Beginning in FY 1993

- Establish fellowships and graduate level support for Grand Challenge teams.

Getting user feedback on a prototype user interface as part of EPA's technology transfer pilot projects.

Department of Education (ED)

The Department joined the HPCC Program in FY 1992. Through program initiatives and through activities sponsored by the regional laboratories and the research and development centers, educators and students will have enhanced access to information and communications resources that will lead to improved teaching and learning.

ED has supporting activities in the NREN component of the Program.

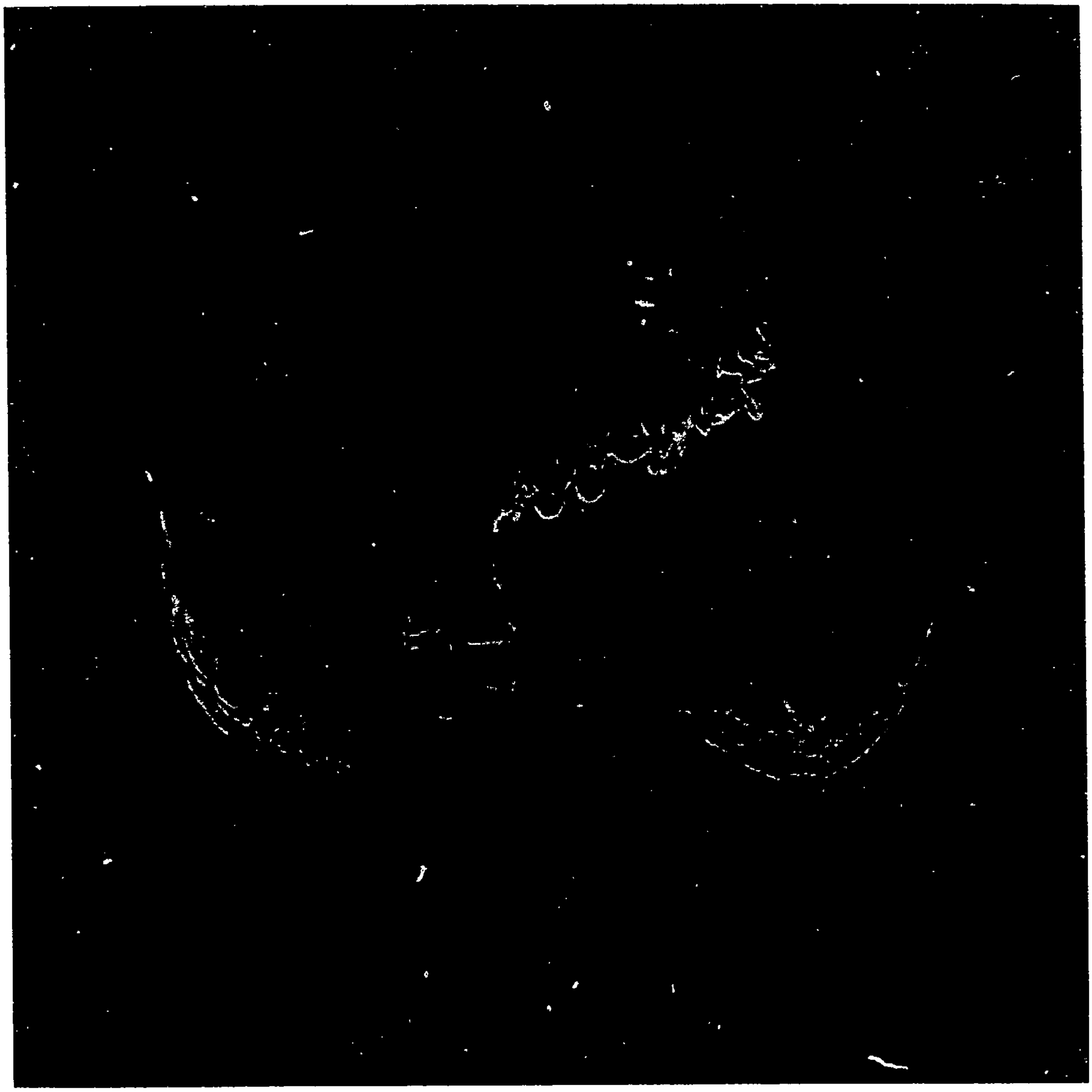
Fairfax County Public School teacher and students use PATHWAYS database and STAR SCHOOLS distance learning materials.

NREN

The Department's on-going activities include:

- Development of PATHWAYS, a computer-based information system designed to use the Internet to provide easy access by teachers, administrators, parents, students, and community members to educational information.
- Expansion of research and development initiatives in networked applications for curricula, instruction, and administration.
- Support for educational technology activities through:
  - -Grants for equitable access to information resources
  - -Star Schools distance learning grants for student instruction and teacher staff development

ED will work with other agencies to define and coordinate educational restructuring initiatives to take advantage of HPCC resources.



The speed of currents at the ocean surface simulated with a three-dimensional global ocean model developed for the Thinking Machines CM-2 and CM-5 at Los Alamos National Laboratory. High speeds are in red, low in blue. Eddies are evident in the Gulf Stream along the east coast of North America and in the tropics. The spatial resolution of the computer model is 0.5 degrees in latitude and longitude with 20 vertical levels; realistic ocean bottom topography is used. The work was jointly sponsored by the HPCC and the DOE CHAMMP (Computer Hardware, Advanced Mathematics, and Model Physics) Program.

Introduction to Case Studies

This section presents examples of on-going Grand Challenge applications research and emerging HPCC hardware and software technologies being developed in support of real-world problems and the HPCC Program.

Only within the past several decades have advances in high performance computing and networking technologies merged with long standing theoretical and experimental methods to yield a powerful new computational approach to scientific inquiry. Thi approach has enabled scientists and engineers to transcend many of the limitations inherent in the more traditional practices. The demonstrated success of computational science in a wide variety of problem areas has led to an infusion of high performance computing technologies and techniques into mainstream scientific practice. An increasing number of researchers, representing almost every discipline, continues to probe new and complex problem areas – many of pressing societal concern.

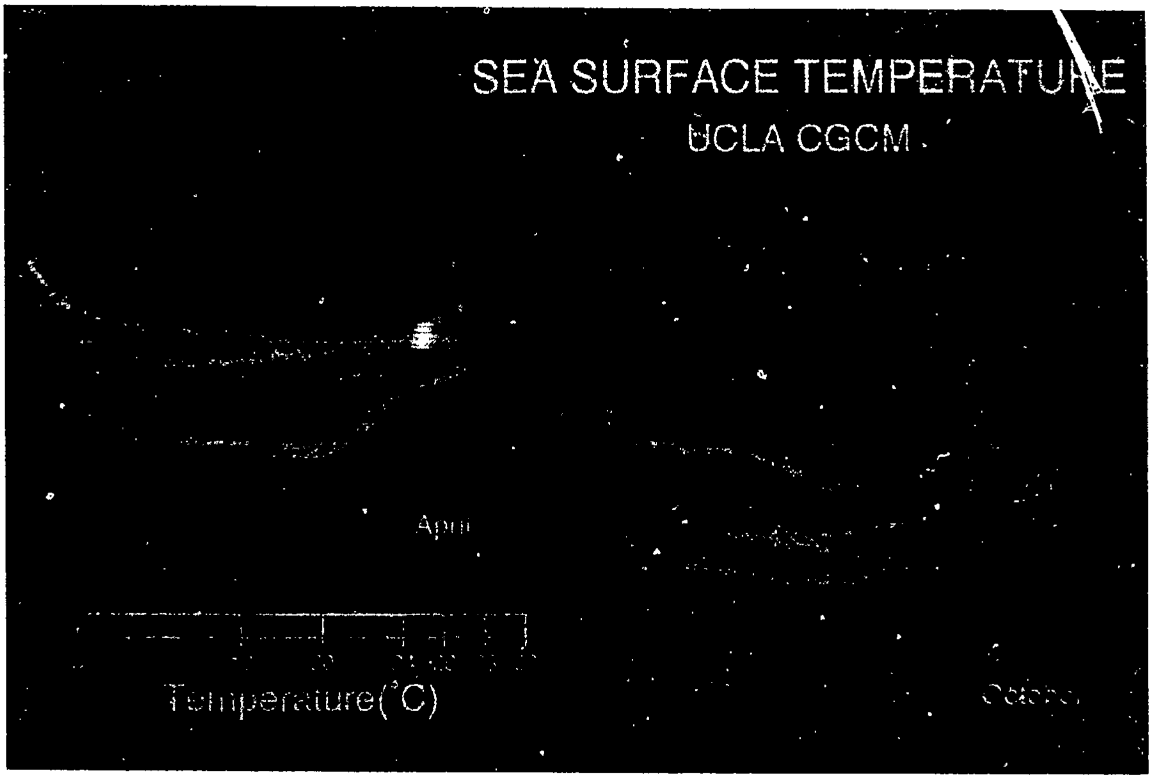
The computational approach adds another dimension to research methods by allowing the researcher to create a mathematical model of some aspect of reality. Solving the model entails translating its equations into a form capable of being programmed and executed on a high performance computing system. Algorithms that structure input data and specify the manner in which the calculations are to be performed are the primary building blocks of the computational model. By exercising the model over broad data ranges and parameter spaces, a picture– albeit a simulated one – of the real phenomena emerges. To the extent that it is complete and accurate, this picture is useful in describing reality as it exists and perhaps more importantly, predicting change. In many cases such as those described in this section, scientists are able to reach and extend the understanding of phenomena far beyond what is possible through pure reasoning and observation.

The success of this approach directly depends upon computing and data management capability. Input data ranges and parameter spaces may be mathematical continua – infinite in dimension. It may require billions of calculations to produce a single solution point, and some phenomena require billions of solution points to approach a useful level of completeness and accuracy. Thus the computational requirements of many applications are potentially boundless. Similarly, the "answers" associated with a simulation may comprise terabytes of data. Scientific visualization, with its ability to store and interpret data, has come to play an indispensable role in the overall success of computational research.

The strengths and limitations of the computational approach are readily evident: a simulation can represent reality only to the degree that it both holds to the physical laws of nature and captures the inherent complexity and detail of that which it attempts to represent. However, for many problems this is the only approach available. Scientific instruments for observation and measurement confront limits – either those imposed by the nature or location of the object of interest, or by the economics of producing the instrument or conducting the experiment. There are few alternatives to the computational approach for studying aspects of nature lying at the extremes of measurability – those that are very small, very large, very fast, very slow, very close, very far, and so on. Yet the universe is filled with such phenomena that affect our daily lives.

The examples presented here are only a small subset of the early achievements to come out of the High Performance Computing and Communications Program – a hint of the promise of HPCC methods and technologies applied to long standing problems of science and engineering – and humanity.

Case Study 1
Climate Modeling



Sea surface temperature from a multi-year simulation with the UCLA coupled atmosphere-ocean model. The atmospheric component of the model is the UCLA General Circulation Model and the oceanic component is the GFDL Modular Ocean Model. The simulation was performed on the Cray Y-MP at the San Diego Supercomputer Center.

Understanding the Earth's climate system and its trends is one of the most challenging problems facing the scientific community today. Better understanding of our climate system is critical for the Nation as it prepares for the 21st century. The Earth's atmosphere-ocean system and the physical laws that control its behavior are complex and contain subtle details. This system is only crudely represented by the most comprehensive of present-day climate models. Improvements in the computational modeling of many of the component physical processes, such as cloud-radiation interaction, will require long-term effort. Climate model improvements will necessitate a hundred to thousand-fold increase in computing, communications, and data management capabilities before these goals can be met. In addition, large increases in model detail and sophistication are necessary for regional climate change forecasting.

Under the sponsorship of the Federal HPCC Initiative and other programs, new computing and communications resources are being developed for climate research. Current state-of-the-art climate models are being redesigned to execute efficiently on promising new scalable architecture systems. Scientists are also investigating distributed computing strategies that will use gigabit-per-second data transfer between distant supercomputers.

The coupled atmosphere-ocean model is the primary tool by which climate scientists simulate the behavior of the Earth's climate system. As a first step in porting a fully coupled model to massively parallel processor systems, scientists at several sites are redesigning separate atmosphere and ocean model elements for execution on scalable parallel systems.

A parallel ocean model, referred to as the Parallel Ocean Program (POP), has demonstrated gigaflop performance during early data-parallel experiments on the 1,024-node Thinking Machines CM-5 at DOE's High Performance Computing Research Center of Los Alamos National Laboratory (LANL). This research effort is being funded under DOE's Computer Hardware Advanced Mathematics and Model Physics (CHAMMP) program. LANL scientists collaborate on the project with scientists at NOAA's Geophysical Fluid Dynamics Laboratory (GFDL), the Naval Postgraduate School (NPS), and NSF's National Center for Atmospheric Research (NCAR).

Over the past decade, the U.S. academic community has made extensive use of NCAR's Community Climate Model (CCM) for global climate research. Recently, researchers at two DOE laboratories, Argonne National Laboratory and Oak Ridge National Laboratory, also funded under the CHAMMP program, have been collaborating with NCAR scientists to develop parallel versions of the latest CCM code to run on next generation massively parallel systems. This model, referred to as the Parallel Community Climate Model 2 (PCCM2), is now demonstrating near gigaflop performance on the 512-processor Intel Touchstone Delta. A data parallel version for the Thinking Machines CM-5 is also being prepared. In ongoing investigations, researchers are exploring new parallel algorithms designed to improve performance on message passing systems and incorporating new numerical methods that will improve climate simulations.

Scientists at the Lawrence Livermore National Laboratory (LLNL) are simultaneously developing a message passing version of the models described above. These versions treat the inherent parallelism of climate problems in a different fashion and can be executed on the class of parallel systems built around the distributed memory architectural approach.

A high-resolution atmospheric general circulation model, known as SKYHI, has been developed and used by scientists at NOAA's GFDL to investigate stratospheric circulation and ozone depletion for a number of years. The ozone depletion study was presented in "Grand Challenges 1993" and is an example of some of the scientific results obtained from this model. GFDL scientists are now working with LANL scientists to reconstruct the programs that make up the SKYHI model in order to execute it on massively parallel systems that support data parallel and message passing programming paradigms.

These various model development efforts all have the same ultimate objective – to develop a coupled atmospheric-ocean climate model that will execute efficiently on scalable parallel computers.

Just as demand for more computational resources is growing within the climate research community, so also is the need to distribute the computing load among different computers, both at one site and across the country when appropriate. With this objective in mind, researchers at the University of California, Los Angeles (UCLA) are collaborating with scientists at NSF's San Diego Supercomputer Center (SDSC), the California Institute of Technology (Caltech), and NASA's Jet Propulsion Laboratory (JPL) to demonstrate the feasibility of distributed supercomputing of a coupled climate model. In this research, separate components of the atmosphere and ocean codes in a single climate computation are distributed over multiple supercomputers connected via a high speed network. Initial prototype experiments have been performed between Cray Research Y-MP systems at SDSC and NCAR, connected by a 1.5- megabit-per-second T-1 data link. The imminent availability of a gigabit-per-second network under the CASA gigabit testbed project will soon allow for a much more advanced communications capability. Plans then call for computations to be distributed between a Cray Y-MP at either SDSC or JPL and an Intel supercomputer at either SDSC or Caltech. Such distributed computing experiments provide realistic tests for the gigabit-per-second connections that are anticipated on the NREN of the future.

SPONSORING AGENCIES AND PROGRAMS

ARPA
DOE
NASA
NOAA
NSF

PERFORMING ORGANIZATIONS

Argonne National Laboratory
California Institute of Technology
Geophysical Fluid Dynamics Laboratory
Jet Propulsion Laboratory
Los Alamos National Laboratory
National Center for Atmospheric Research
Naval Postgraduate School
Oak Ridge National Laboratory
San Diego Supercomputer Center
University of California-Los Angeles Climate Dynamics Center

Case Study 2
Sharing Remote Instruments

Specimens of chick cerebellum stained with osmium and reduced with potassium ferrocyanide. Image B shows a higher magnification of a portion of Image A. The images were obtained in a study of the internal membrane structure of the Purkinje cell.

HPCC-supported advances in computer networks and visualization software are allowing scientists to control and share remote microscopes, telescopes, and other scientific instruments in an interactive, real-time fashion. These types of projects have opened the doors of the laboratory to a new concept – the "distributed laboratory" – which integrates laboratory equipment, high performance computing systems and data visualization tools over a high speed network, resulting in a more comprehensive and scientifically valuable investigative environment.

Remote Microscopy

Researchers at the University of California-San Diego Microscopy and Imaging Resource have implemented a sophisticated, computer-controlled high-voltage transmission electron microscope (HVEM). Working in collaboration with staff scientists at the San Diego Supercomputer Center and the Scripps Research Institute, the research group has coupled the HVEM via a high speed network to high performance computing systems and interactive visualization software running on scientists' workstations. The microscope is a unique resource, one of only a few such microscopes in the United States in use in biological science. More powerful than ordinary electron microscopes, it can accommodate much thicker laboratory specimens, yielding greater amounts of biological information. Using computer tomography and other visualization techniques, the images collected can be used to produce three-dimensional animations, allowing scientists to look at many previously uninterpreted areas of biomedical science relating biological function with structure.

The subjects of such study include the disruption of nerve cell components resulting from Alzheimer's disease, the structural relations of protein molecules involved in the release of calcium inside neurons, and the three-dimensional form of the Golgi apparatus, where sugars are added to proteins.

This project is an example of the application of

Scientists no longer need to be in the same room with their laboratory equipment. Instead, they can control it over the network from their desktop computers. Pictured is a scientist controlling a high-voltage electron microscope (useful for imaging thick three-dimensional biological tissues) from her Sun workstation.

diverse technologies to a single class of scientific problems. The capabilities of the microscope for scientific investigation will soon be extended, combining image data acquisition with computing resources to render, view, and animate the images for real-time analysis. By connecting the microscope to a high speed network, the instrument will someday be made available to investigators located in any geographic region, extending the accessibility of the resource to a broader scientific community in a collaborative environment.

Future plans call for extending this environment to the Apple Macintosh; implementing automatic focusing and calibration of the microscope (now handled by a human operator); developing remote image analysis tools; optimizing the tomography reconstruction code currently running on a Cray Y-MP system; and implementing the code on a parallel computing platform.

Real-time Radio Telescope Observatory

The application of the operational concepts described above is not limited to microscopes. A research group affiliated with the National Center for Supercomputing Applications (NCSA) in Champaign-Urbana, IL is investigating how to apply this model to real-time radio telescope observation. This group is seeking to demonstrate the feasibility of connecting the Berkeley-Illinois-Maryland Array (BIMA) radio telescope array, located at the Hat Creek Observatory in northern California, to NCSA supercomputers via high speed networks.

One of their needs calls for transferring a gigabyte-size observed visibility dataset from the telescope to NCSA for processing, and then returning the processed data to Berkeley for analysis on a workstation. In this way, the astronomer can judge the quality of the data, see if the signal is strong enough to proceed with the observations, judge whether the area of the sky being mapped is correct, and experiment with processing parameters. By connecting BIMA directly to BLANCA, one of five HPCC supported gigabit testbeds, and increasing the number of present antennae from three to six, the group expects to achieve a data transfer rate 10 to 100 times higher than current rates. This could allow real-time steering of the telescope, thereby greatly enhancing the value of an observational session and revolutionizing the way in which astronomy research is performed.

REMOTE MICROSCOPY

SPONSORING AGENCIES AND ORGANIZATIONS

Datacube, Inc.
Network Systems Corporation
NIH
NSF
Photometrics LTD
State of and University of California
Sun Microsystems Computer Corporation

PERFORMING ORGANIZATIONS

San Diego Supercomputer Center
The Scripps Research Institute
University of California- San Diego
- San Diego Microscopy and Imaging Resource
- Computer Science and Engineering Department

REAL-TIME RADIO TELESCOPE OBSERVATORY

SPONSORING AGENCIES AND ORGANIZATIONS

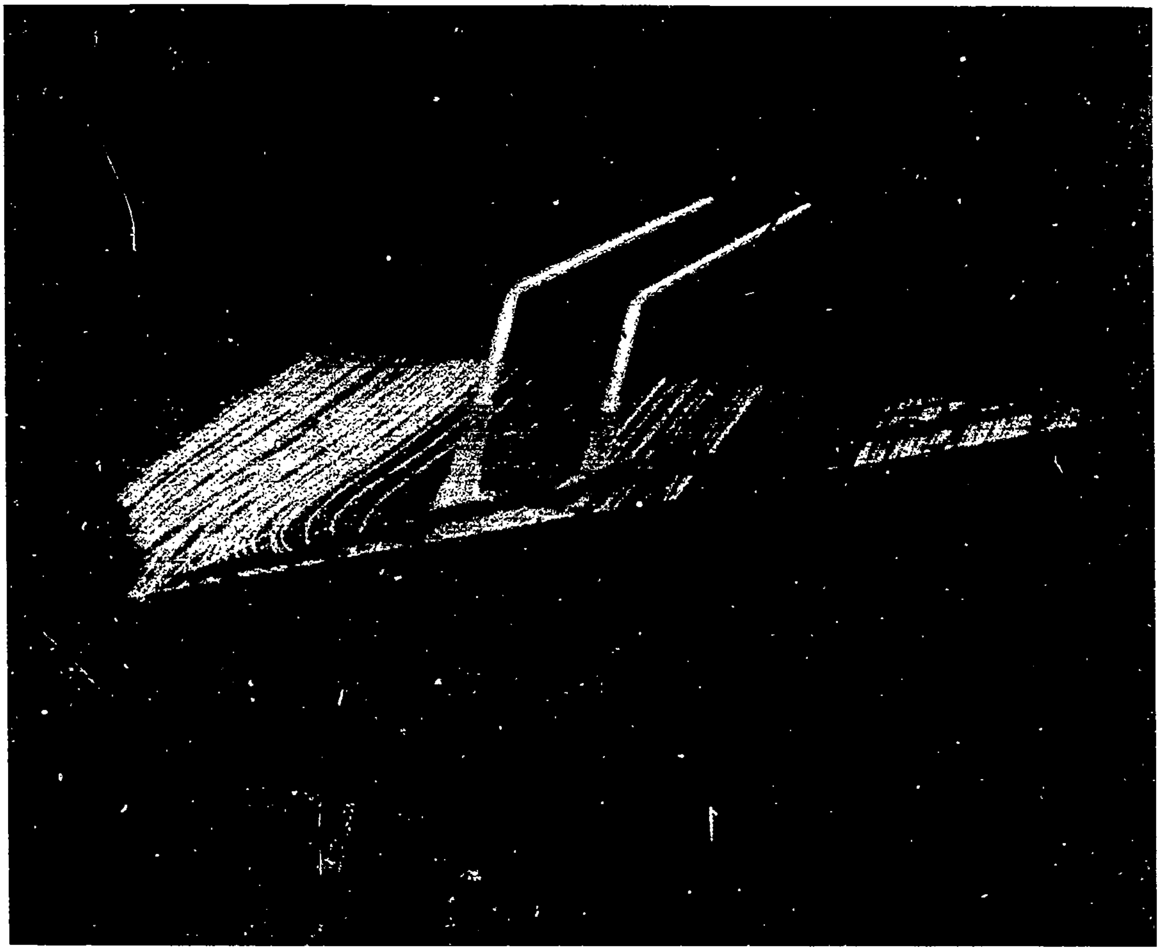
NSF
State of and University of California
State of and University of Illinois
State of and University of Maryland

PERFORMING ORGANIZATIONS

National Center for Supercomputing Applications
University of California-Berkeley
University of Illinois-Champaign-Urbana
University of Maryland

Case Study 3

Design and Simulation of Aerospace Vehicles

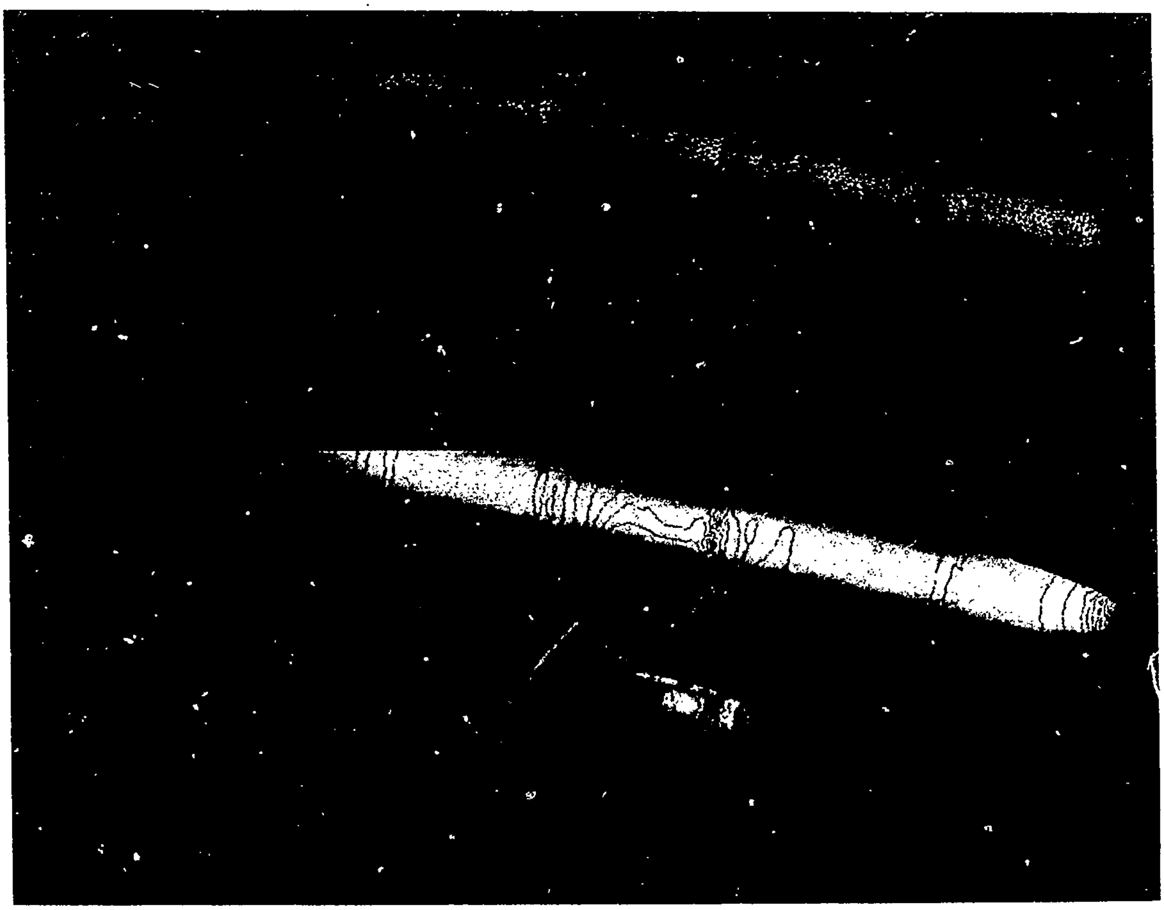


Airflow is simulated over and past the wing of a high performance aircraft that is using vectored thrust while descending to a few feet above the ground (in ground effect).

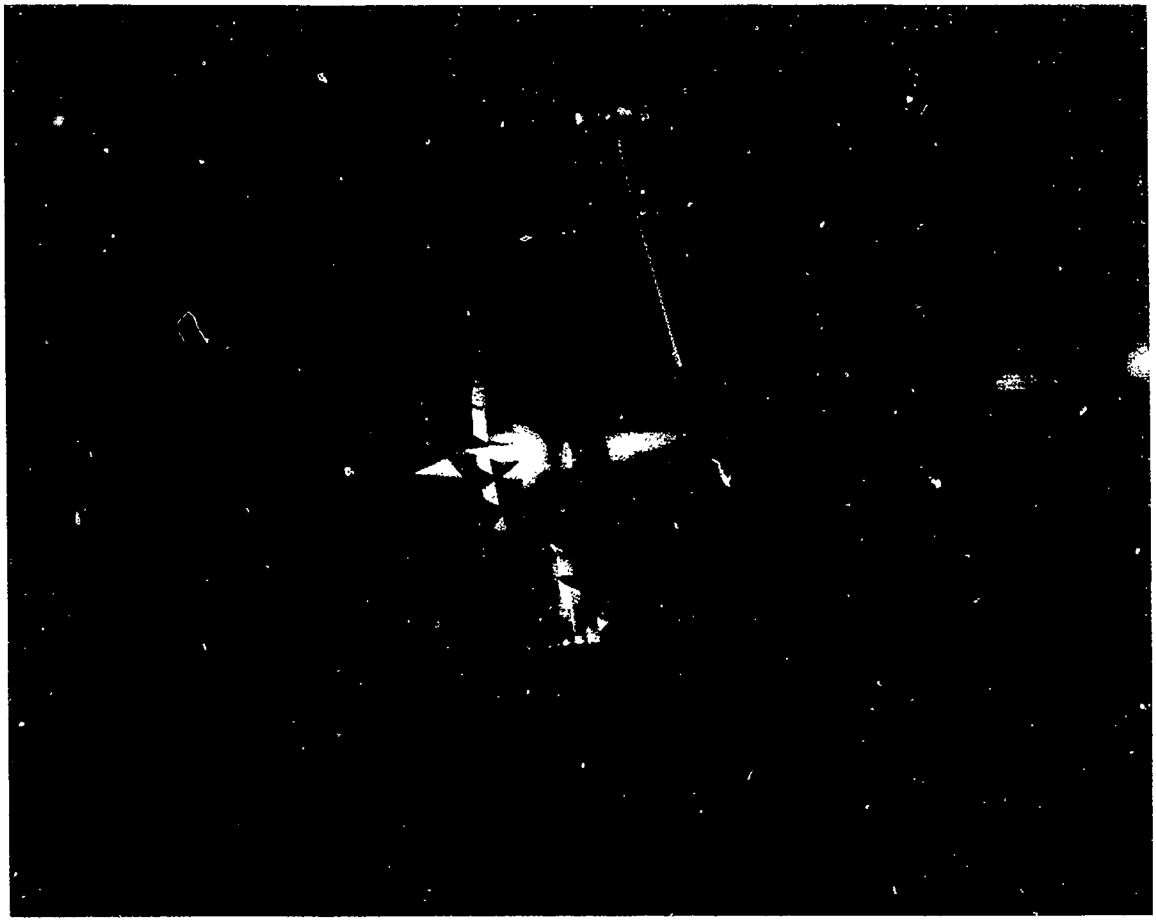
Improved design and simulation of advanced aerospace vehicles is a primary goal of the NASA HPCC Program. Massively parallel processing and related high performance computing technologies are enabling new simulation and optimization methodologies for flight vehicle design. These methodologies combine multiple physical disciplines and integrate vehicle components such as the airframe, propulsion systems, and controls. Improved aircraft design resulting from high performance computing promises to give the U.S. aerospace industry a critical competitive edge. Early and continuous coordination between NASA research centers, other Federal agencies, academia and industry is crucial to the success of this endeavor.

Scientists at NASA's Ames Research Center have developed computer simulations of air flow past a delta wing at takeoff and landing that will help design more efficient and cost-effective aircraft. By porting a single discipline computational fluid dynamics code to a number of scalable parallel computers, the scientists simulated applications involving moderately complex geometries. For example, a simple powered lift vehicle that includes a wing and two lifting jetson is shown on the previous page. Navier Stokes computations were performed to simulate flow past a delta wing with thrust reverser jets in a ground effect environment (takeoff and landing). These preliminary simulations will provide insight into the computational interaction of thrust elements, and upwash and reingestion of hot air and gases, leading to significant improvements in aerodynamic performance.

Improved design processes for advanced aircraft and spacecraft through the use of advanced computational fluid dynamics and structural analyses is also the subject of research by scientists at Langley Research Center. Scientists are focusing on the design and modeling of a High Speed Civil Transport (HSCT). One of the greatest challenges in modeling the fluid dynamics of these vehicles is developing mathematical equations that accurately simulate turbulent airflow. Turbulent flow simulations currently used lose accuracy at high speeds. Development of an optimal airframe design for a HSCT involves thousands of design iterations, each requiring a



Airspeed contours around a three-dimensional aircraft traveling at Mach 0.77 are revealed by using an advanced simulation method (lower figure). Blue areas represent slower moving air, yellow and red the fastest moving air. Flow calculations for this complex geometry were based on the unstructured mesh shown in the upper figure.



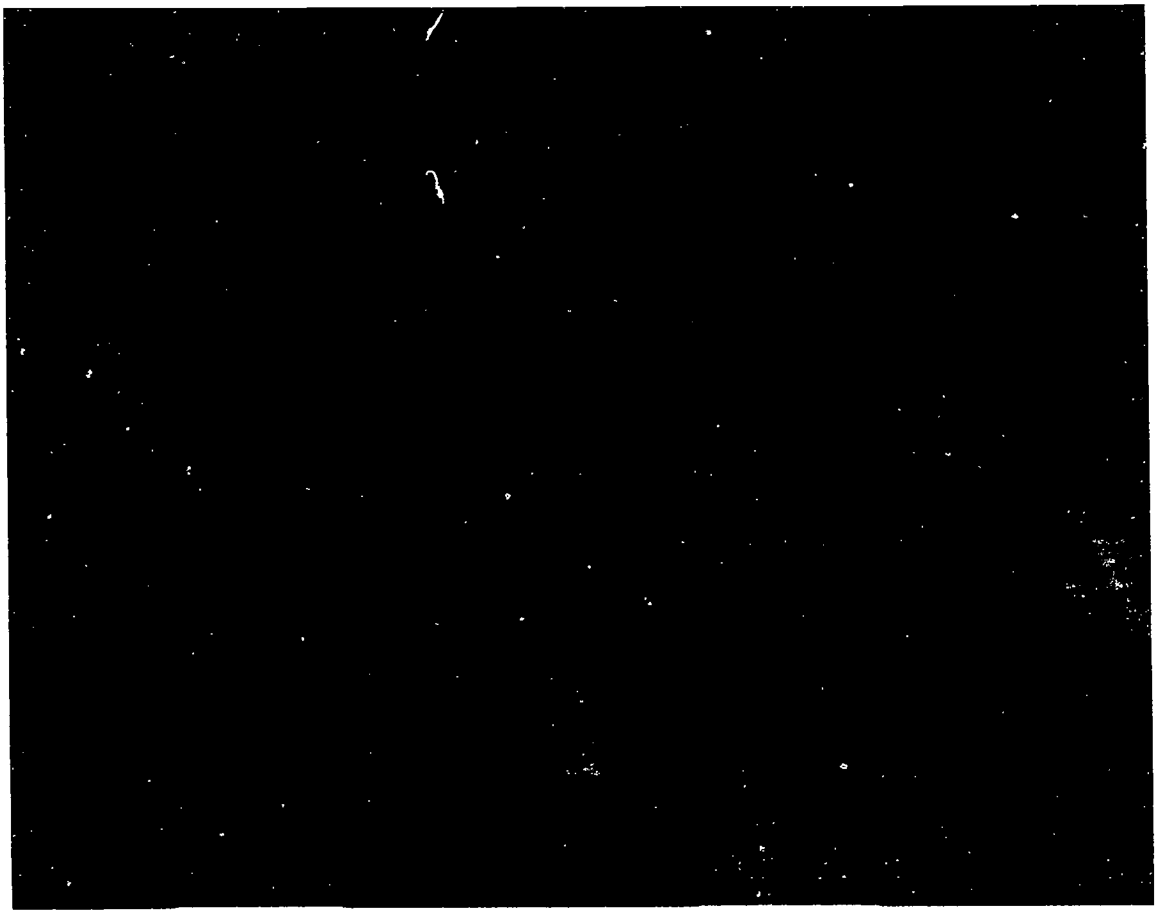
Airflow through several components of a jet engine is simulated on a cluster of IBM RISC 6000 workstations.

large amount of computing time to recalculate how changes in one part of the design affect the rest of the structure. Langley scientists have developed a method to streamline this part of the overall process and reduce the amount of computing time needed for each iteration.

Researchers at NASA Lewis Research Center are developing a prototype Numerical Propulsion System Simulator (NPSS) that will allow dynamic numerical and visual simulation of engine components. The prototype is intended to simplify the process of dynamically connecting engine component software across various machine architectures. This will allow industry to design, engineer, and fabricate tomorrow's propulsion systems. This multidisciplinary research, coupled with the improved performance offered by massively parallel computers, will greatly enhance the ability of aircraft manufacturers to analyze different design options rapidly and then efficiently produce vehicle propulsion systems with optimal performance and reduced design cycle costs.

Case Study 4

High Performance Life Science: From Molecules to MRI



Computer modeling of the Purkinje neuron helps explain its branching network of nearly 200,000 connections to other brain cells.

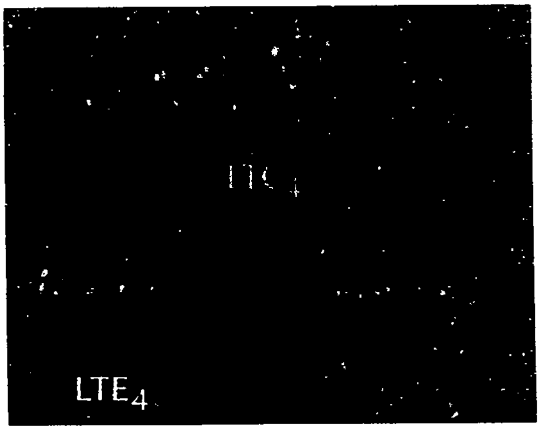
High performance computing and communications is providing valuable insights into the mysteries of human health and disease. These sophisticated new techniques allow for unprecedented study and understanding of the remarkable self-assembly of atoms into molecules, molecules into cells, cells into tissues, and organ systems into the complete organisms that constitute humanity itself. Important new medical knowledge and new diagnostic tools have resulted from the HPCC Program.

At the molecular level, supercomputing and scientific visualization have provided valuable insights into the molecular basis of asthma. Corporate scientists at Eli Lilly collaborated with visualization specialists at the National Center for Supercomputer Applications to model the three-dimensional molecular motions of leukotrienes, naturally occurring "messenger molecules" that induce the lungs to become stiff and inflamed. With the goal of designing new medicines to block the leukotriene receptor in the lungs, these researchers used supercomputers to calculate the position of each atom in several related leukotriene molecules; three-dimensional scientific visualization then displayed the subtle differences in atom positioning that are crucial to biological activity. More effective medications with fewer side effects are the goal of this high performance molecular dynamics computing.

Molecules assemble themselves into cells and tissues, and at this level also computerized analysis has expanded our knowledge. One of the most complicated cells in the brain is a neuron called the Purkinje cell; each one of these cells may have up to 200,000 electrical connections with other brain cells. Researchers at the California Institute of Technology used an experimental massively parallel computer to model the response of Purkinje cells to chemical and electrical stimuli. Their model emulates the observed behavior of the living cell, and provides evidence that although the connections to other neurons are voluminous, there is an elegant simplicity in the spatial arrangement of these brain interconnections.

Cells and tissues organize into body systems, which can also be better understood using HPCC technologies. Magnetic Resonance Imaging (MRI) is a widely used method for imaging internal structure of the human body that produces two-dimensional cross section views. Researchers at Sandia National Laboratories, in collaboration with Baylor University Medical Center in Dallas and the Department of Veterans Affairs Medical Center in Albuquerque, have used massively parallel supercomputers to turn two-dimensional MRI images into three-dimensional views, and revealed previously hidden

Massively parallel supercomputing turns two-dimensional magnetic resonance images into three-dimensional maps that can identify early breast cancers.



The molecular dynamics of three different leukotriene (LTC4, LTD4, and LTE4) molecules portrayed with a new supercomputing visualization technique called "ghosting." Dense shadows represent the configurations in which the molecules spend a greater amount of time, while the range of molecular movement is shown by the extent of the dot patterns.

information. Communicating their data via the Internet, these groups compared three-dimensional MRI to standard X-ray mammography for the detection of early breast cancer, and found that the new technique revealed early tumors that could not be detected by mammography.

SPONSORING AGENCIES AND PROGRAMS

- Department of Veterans Affairs
- DOE
- Eli Lilly & Co.
- NIH
- NSF

PERFORMING ORGANIZATIONS

- Baylor University Medical Center
- California Institute of Technology
- Department of Veterans Affairs Medical Center, Albuquerque
- Eli Lilly & Co.
- National Center for Supercomputing Applications
- Sandia National Laboratory

Case Study 5

Energy Resource Recovery

A numerically intense geological database that describes the physical and chemical conditions of the Gulf Coast Basin is assisting the scientists of the Global Basins Research Network to understand the processes that control the movement of oil and gas in sedimentary basins. By visualizing the data scientists are able to observe sedimentary conditions that would otherwise take much longer to investigate. This image depicts the various structures beneath the ocean flow in an area off the shore of Louisiana. Two faults are shown, in addition to three geographic layers: (from bottom) top-of-salt, shale, and sand and shale.

High performance computing plays an important role in the recovery of non-renewable energy resources. About two-thirds of the U.S. energy supply comes from oil and gas, and although much oil continues to be imported, there are profound advantages to domestic production: for example, ensuring a stable supply and price to the consumer. While most large U.S. oil reservoirs have already been discovered, two-thirds of the oil still remains in old fields after conventional recovery technology has been applied. Enhanced oil recovery (EOR) using advanced technologies has the potential to recover another 100 billion barrels worth about two trillion dollars at today's prices.

Domestic oil and gas producers are working with scientists from other sectors to optimize recovery methods for existing petroleum reservoirs. Petroleum industry scientists use reservoir simulations run on high performance computers to determine the production potential of reservoirs and the most efficient methods to extract petroleum resources before investment in field operations is made. The simulations model large complex field problems quickly, accurately, and efficiently, leading ultimately to reduced recovery costs. Research projects combining the expertise of researchers from universities, industry, and government focus on ways to improve reservoir simulation methods for flow through porous media, pore-scale multiphase flow, and hydrocarbon migration. As a technological spinoff, EOR simulation can be applied to remediation strategies for underground, contaminated sites.

Parallel Algorithms for Modeling Flow in Porous Media

New parallel algorithms for computational models describing the flow of oil and other organic chemicals in porous media are an important technical development. Models employ stochastic and conditional simulation and take into account numerous variables associated with reservoir behavior caused by complex physical and chemical phenomena. Parallel algorithms help researchers utilize the computational power of massively parallel computers to perform the simulations. Researchers are also writing parallel versions of two widely used reservoir simulation codes, UTCHEM and UTCOMP. UTCHEM has been applied to study both surfactant EOR and surfactant remediation of ground water contaminated by dense nonaqueous phase liquids found at weapons sites and other locations.

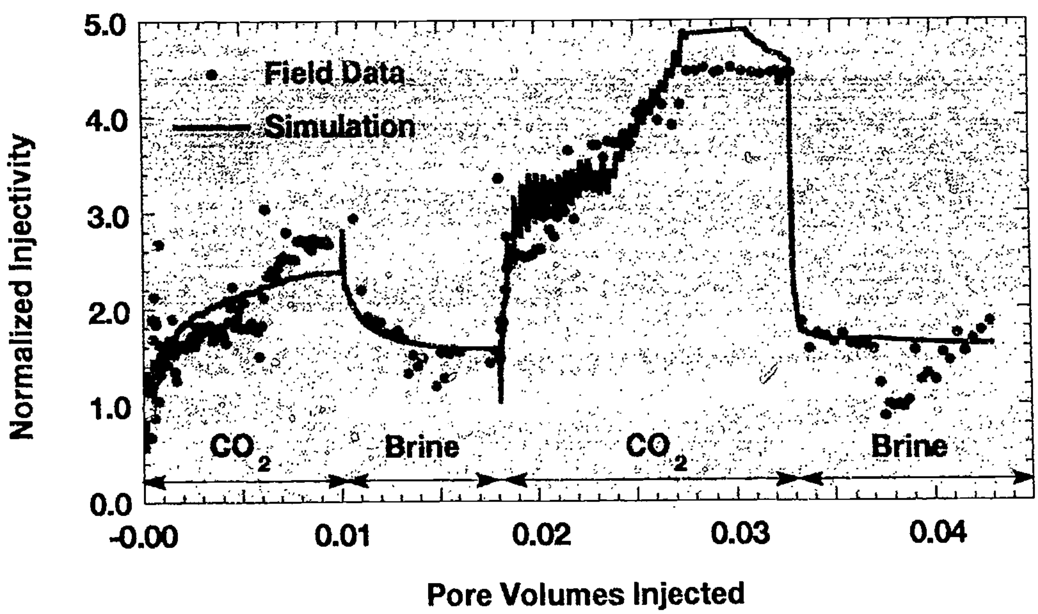
Pore-scale Multiphase Flow Modeling

To better understand and predict the interactions between fluids and rock, field simulation programs must extrapolate results from laboratory samples. To assist in interpreting experimental results, researchers are developing a basic theoretical framework for multiphase flow systems that incorporates current knowledge of displacement processes and the understanding of oil and rock chemistry. One technique in this framework, the Lattice-Boltzmann method, is essential for solving Navier-Stokes equations, the basic equations of fluid flow. These equations are the basis of the mathematical models that simulate the movement of a multiphase system of organic materials through rock pores.

Imaging of Present-day Hydrocarbon Migration

Scientists are developing computer visualizations to chart fluid movements below the ocean floor of the Louisiana coast. Observations in

INJECTIVITY HISTORY DURING CARBON DIOXIDE FLOODING OF WEST TEXAS FIELD TEST



This figure shows a comparison between the computed injectivity of carbon dioxide and field data from an oil reservoir in Texas operated by Texaco, Inc. The University of Texas compositional reservoir simulator UTCOMP was used to make this calculation. The capability to predict the rate at which carbon dioxide can be injected into these old wells to increase the oil production from the field is crucial for the economic success of such enhanced oil recovery operations. This comparison shows close agreement with the field data from a geologically complex San Andres formation.

this area are providing new insights into how over-pressured oil and gas move in sedimentary basins. Because the study area encompasses the largest oil field in the continental United States, these insights may have important economic consequences. Researchers are developing visual models that reflect changes in temperature, pressure, and strata of sand and shale to allow qualitative evaluation of the rate of fluid flow required to produce these changes.

SPONSORING AGENCIES AND ORGANIZATIONS

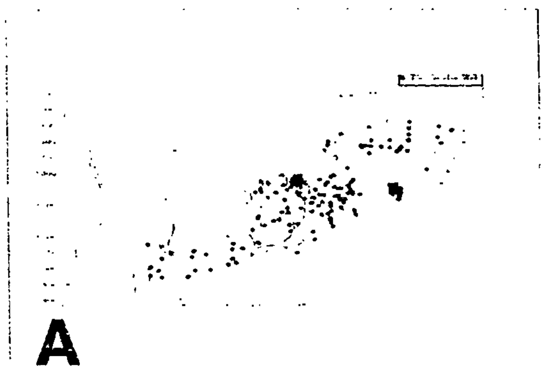
- ARPA
- Cornell Theory Center
- DOE
- EPA
- IBM Corporation
- New York State
- NIH
- NSF
- 41 Affiliates at Rice University
- State of Texas
- University of Texas at Austin

PERFORMING ORGANIZATIONS

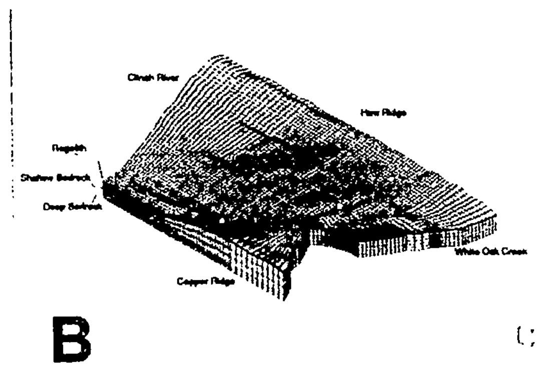
- AGIP
- Chevron
- Computational Mechanics Corporation
- Conoco
- Cornell University
- Elf Aquitaine
- Exxon
- HyperMedia Corporation
- IBM Corporation
- Lamont-Doherty Geological Observatory
- Landmark Graphics
- Lawrence Livermore National Laboratory
- Los Alamos National Laboratory
- Louisiana State University
- Michigan Technological University
- Mobil Exploration and Production Technology and Engineering Center
- Penzoil
- Rice University
- Texaco
- University of Houston
- University of Texas
- Woods Hole Oceanographic Institute

Case Study 6

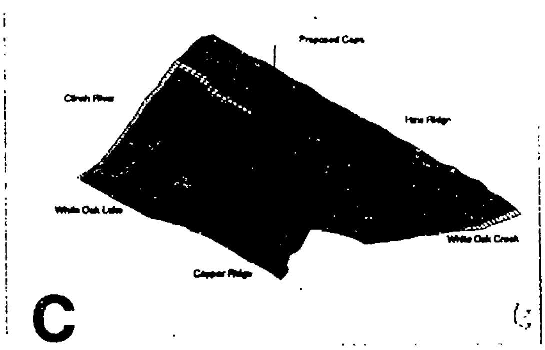
Groundwater Remediation



Map of Oak Ridge Waste Storage Facility



Computational Grid for Waste Area Grouping 6



Modeling the Effect of Capping WAG 6 Waste Burial Trenches

Computed Water Table Elevations after Capping

Contamination of groundwater by potentially toxic organic liquids is one of the more serious environmental problems facing the Nation. The release of such liquids has been widespread, originating from sources ranging from gasoline service stations to large industrial facilities. Once trapped in the subsurface these liquids serve as long-term sources of aquifer contamination. Remediation schemes include optimal liquid recovery pumping schemes, solvent and surfactant flushing, and steam injection.

Hierarchical models of groundwater flow and associated chemical transport are needed at the local, basin, and regional scales to assess waste transport and remediation schemes. Current models are limited computationally in two ways: first, models can only include a limited number of physical and chemical processes before computing runtimes become prohibitive; second, models lack adequate field data to characterize geologic and chemical factors that influence contaminant behavior. High performance computing systems are needed both to model complex transport processes and to overcome data limitations by using computationally intensive parameter estimation techniques.

The Partnership in Computational Science (PICS) consortia project funded by the DOE HPCC program is exploring new approaches to modeling groundwater flow at waste sites in extremely complex hydrogeologic settings. The consortia consists of members from Brookhaven

National Lab, Oak Ridge National Lab, Rice University, SUNY-Stony Brook, Texas A&M, and the University of South Carolina. A map of the site most extensively studied (WAG 6) is given in Figure A. Researchers at ORNL have adapted an existing finite element code to run on an Intel scalable parallel distributed-memory system. By using a preconditioned conjugate gradient solver, users were able to distribute portions of the model grid to different processors and greatly reduce solution time. Using a 40,000 node grid (Figure B), they are obtaining anisotropy factors and boundary conditions that can be used to model smaller areas and test the effects of remediation schemes on models of the waste areas. One such remediation scheme is to cover areas with a water-impermeable membrane (cap) to eliminate surface recharge (Figure C). A calculation of the water table following a capping is given in Figure D. Such results point to the specific piezometric wells for which experimental indications of capping effects should be sought.

Pacific Northwest Laboratory is also exploring the use of scalable massively parallel quantum chemistry algorithms to improve methods for redesigning enzymes to better degrade pollutants, for extracting contaminants from soils, and for burning halohydrocarbons.

Researchers at the EPA's Robert S. Kerr Environmental Research Laboratory and collaborators at various universities have focused efforts on developing process understanding using complementary numerical models and laboratory studies. Implementation of numerical models has followed approaches originally developed in the petroleum industry and advanced multiprocessor techniques. A large scale physical model is currently being used to evaluate the numerical model. Release of an organic liquid is planned for the physical model, followed by testing of remedial technologies. A remediation approach involving surfactant flood, followed by vapor extraction above the water table, is being designed by bench scale laboratory studies and adaptation of a sophisticated numerical model developed by the University of Texas at Austin. Other remediation models are also under development. Joint field and modeling studies in cooperation with the USDA are underway to assess the impact of agricultural chemicals on water resources in the midwest.

Groundwater models will not eliminate the need for adequate field data but will continue to assist scientists in understanding these complex systems, thus contributing to reducing the risks to public water supplies.

SPONSORING AGENCIES AND ORGANIZATIONS

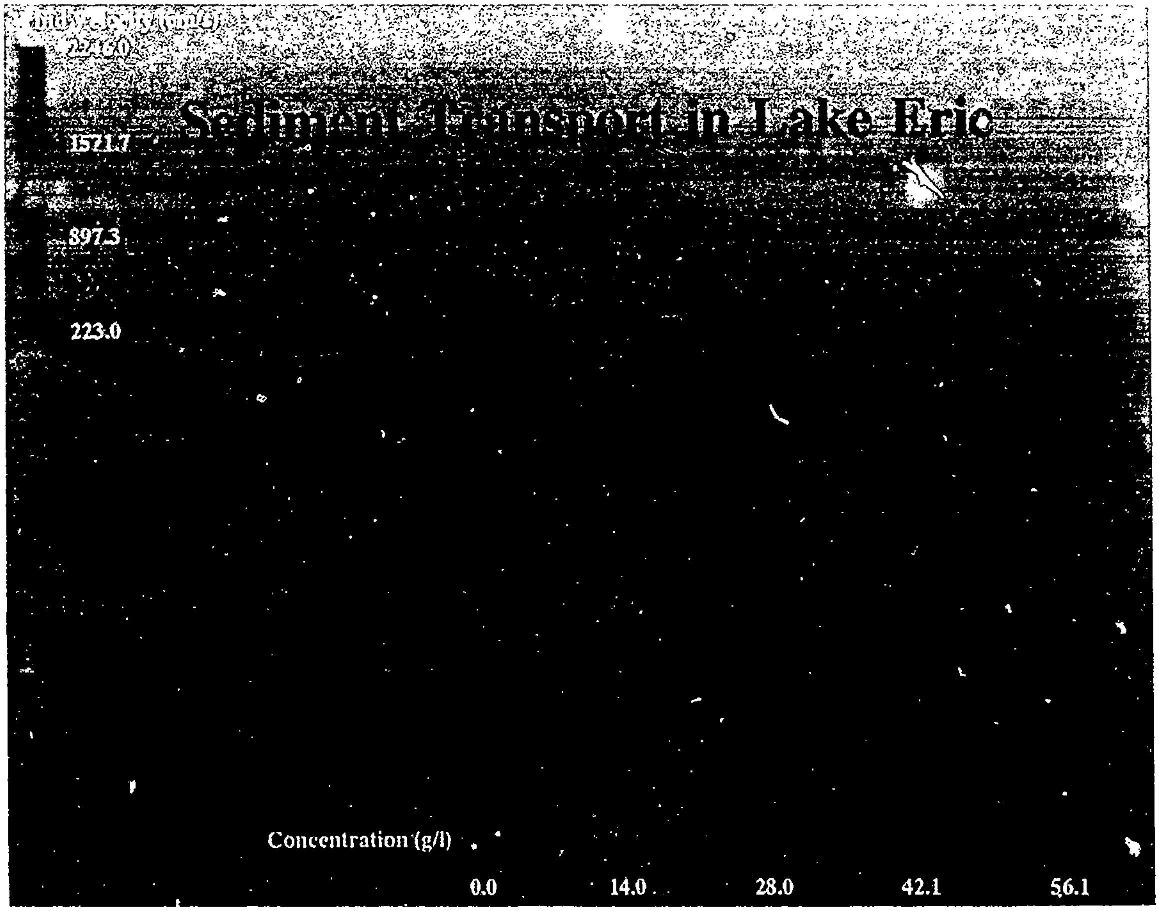
- DOE
- EPA

PERFORMING ORGANIZATIONS

- EPA's Robert S. Kerr Environmental Research Laboratory
- Pacific Northwest Laboratory
- Partnership in Computational Science (PICS) Consortia
  - Brookhaven National Lab
  - Oak Ridge National Lab
  - Rice University
  - SUNY-Stony Brook
  - Texas A&M
  - University of South Carolina
- University of Texas at Austin

Case Study 7

Improving Environmental Decision Making



Concentration of sediments in Lake Erie resulting from a large storm. Concentrations are color coded, with the highest shown in red. Winds from storms produce waves, which in turn create bottom currents and turbulence that can cause contaminated bottom sediments to be resuspended in the water column. The shallower the water, the stronger the effects of the wave action. Consequently, sediment concentrations tend to be largest in shallow sections.

High performance computing is crucial to improve decision making on environmental issues. Improved numerical computer models provide more accurate predictions of pollution so that our leaders will be in a better position to make policy decisions involving the environment and economic growth. The accuracy of these models depends upon descriptions of physical, chemical and biological processes that adequately incorporate important causal interactions, nonlinearities, synergies, and feedbacks, as well as capture nonintuitive interactions. The linkage of single media models and the development of megamodels are essential for more reliable management of the environment.

Three areas of environmental decision support have been established as part of EPA's HPCC

Program: air pollution, water pollution, and the combined effects of air and water pollution on coastal estuaries that lead to de-oxygenation.

Air Pollution

The Clean Air Act amendments mandate State use of air quality models to demonstrate the effectiveness of proposed approaches for reducing air pollution in highly polluted areas. Researchers at EPA and the North Carolina Supercomputing Center in Research Triangle Park, NC are developing a decision support system to satisfy a wide range of related air quality modeling, information and analysis needs. New computational capabilities enable improved air quality models to address multipollutant interactions among acidifying species, oxidants and aerosol, and interactions with meteorology. These complex, three-dimensional, time-dependent models are used to evaluate alternative pollution control strategies where interactions at different scales, from urban to regional, are critical and require new approaches made possible only through high performance computing. The North Carolina Department of Environment, Health and Natural Resources is testing an early prototype of an urban airshed modeling system.

Water Pollution

Contaminated bottom sediments and the associated decrease in water quality are a major problem in hundreds of rivers, lakes, harbors, estuaries, and near-shore areas of oceans. In order to remediate this problem and evaluate possible management alternatives for the disposition of these toxic sediments, EPA-supported researchers at the University of California- Santa Barbara are developing computer intensive three-dimensional, time-dependent models of the hydrodynamics, particle transport, and sediment bed dynamics coupled with meteorology. Significant biochemical reactions that affect the fate of contaminants are also included in the models. Work currently focuses on the Great Lakes region of the United States. Figure 1 shows that redistribution of toxic sediments in Lake Erie due to a large storm is highly dependent on wind direction and speed, and water depth.

Linked Air and Water Pollution

Interactions between pollutant media, such as air and water, often significantly affect the environment. Air and water pollution both contribute to an overabundance of nutrients in coastal estuaries causing lack of oxygen and eventual decline of commercial productivity of the estuaries. Researchers at EPA's Environmental Research Laboratories in North Carolina and Georgia, and the Chesapeake Bay Program Office in Maryland are developing loose linkages between air and water quality models to study the impact of nitrogen deposition by air and water sources on the decline of coastal estuaries of the Chesapeake Bay.

In some cases, only minor feedbacks exist between the different media. Loosely linking current single-medium models retains the known features and complexity of each model. This linkage provides additional information for more effective policy determination by individual pollutant media decision making communities.

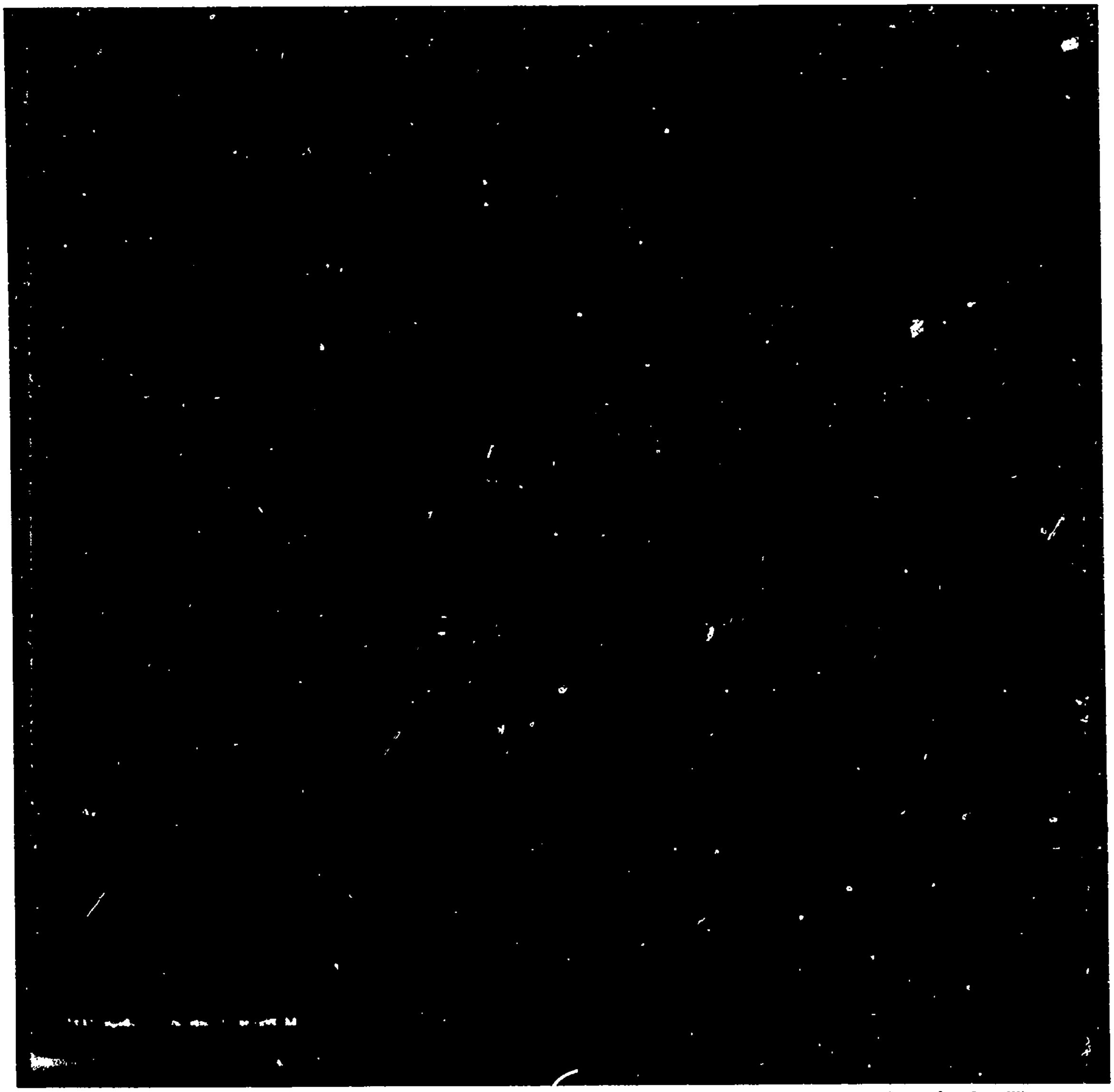
SPONSORING AGENCIES AND ORGANIZATIONS

EPA

PERFORMING ORGANIZATIONS

EPA's Chesapeake Bay Program Office
EPA's Environmental Research Laboratories
North Carolina Department of Environment, Health and Natural Resources
North Carolina Supercomputing Center
University of California - Santa Barbara

Case Study 8
Galaxy Formation



Visualization of a small portion of the computational results of an astrophysical simulation of 8.8 million gravitating bodies run on 512 nodes of the Concurrent Supercomputing Consortium's Intel Touchstone Delta System. The image shows about 137,000 gravitating bodies that form a single halo about 400 megaparsecs across. Color represents projected density along the line of sight, i.e., different colors correspond to pixels that "contain" different numbers of bodies. Densities are much higher at the center of the object. The entire simulation contained about 700 halos like the one depicted here.

The process by which galaxies form is among the most important unsolved problems in physics. There is a wealth of observational data – modern observations span the electromagnetic spectrum from radio frequency to gamma rays. However, a firm theoretical understanding is yet to be acquired. Questions as simple as, "Why are there two families of galaxies, spiral and elliptical?" remain unanswered.

The first step in understanding how galaxies and stars form is to understand the environment in which their formation occurs. Scientists infer that approximately 90 percent of the mass of the Universe is in the form of dark matter, that which can be observed only indirectly, through its gravitational influence on observable matter, such as stars, gas clouds, and nebulae. Simulations are used to study the shapes and dynamics of dark matter halos that are known to surround observed galaxies.

One obstacle to the development of efficient codes to simulate the galaxy formation process is the sheer size of the problem. To obtain useful information, the software must simulate the interactions of millions of bodies. Improved algorithms for N-body calculations have been developed and adapted to this problem. Even with the new algorithms, however, conventional supercomputers lack the computational power to perform simulations in a reasonable period of time at an acceptable level of resolution. A few hundred thousand bodies is the most that conventional systems can model at a time, about two orders of magnitude too few to obtain new scientific results. Hence, scientists must turn to parallel computers – and very large ones at that – to perform the computations.

Parallel versions of simulation code are generally more difficult to develop than those for sequential computing systems because of the inherent complexity of parallel systems. In this particular case, code development was further complicated by the requirements for internal data organization in the program that needed to adapt as the computation progressed to reflect changes in the structure of the evolving universe.

Concurrent Supercomputing Consortium (CSCC) scientists overcame these programming difficulties and recently developed a parallel simulation code for the 512 node Intel Touchstone Delta System, which achieved speedups in excess of 400 over the single processor speed.

In March 1992, researchers ran a simulation with almost 8.8 million bodies for 780 timesteps on 512 processors of the Touchstone Delta system. The simulation was of a spherical region of space of diameter 10 megaparsecs (or, about 30 million light years); a region large enough to contain several hundred typical galaxies. The simulation ran continuously for 16.7 hours, and carried out 3.24 x 1,014 floating point operations, for a sustained rate of 5.4 gigaflops per second. Algorithm improvements accounted for nearly a 3,000-fold improvement in execution time over traditional calculation methods. Subsequently, the research team ran two 17.15 million body simulations using the Cold Dark Matter model of the Universe. These simulations, the largest N-body simulations ever run, took into account recently acquired microwave background measurements gathered by the COBE satellite.

SPONSORING AGENCIES AND ORGANIZATIONS

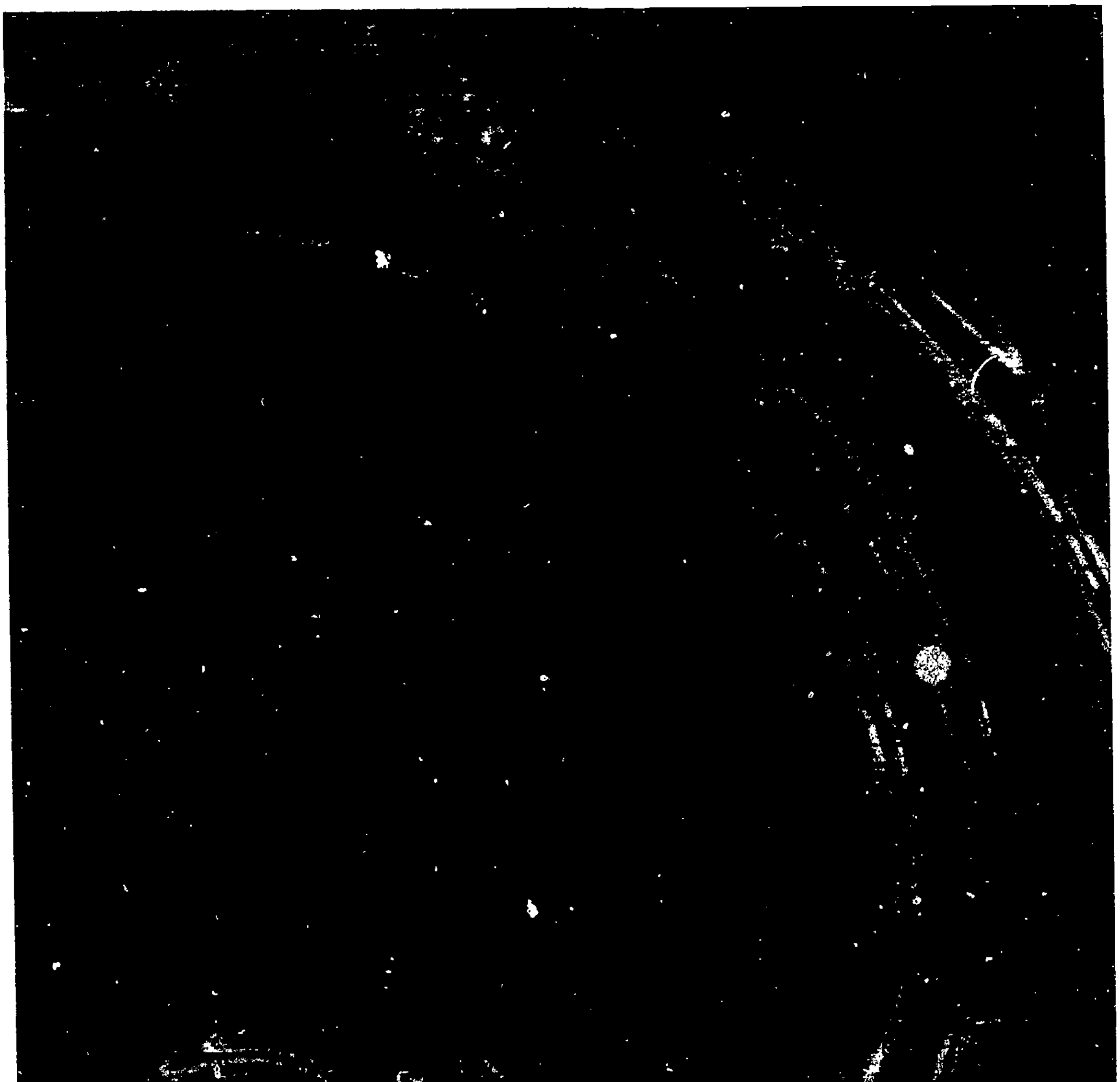
Concurrent Supercomputing Consortium
DOE
NASA
NSF

PERFORMING ORGANIZATIONS

California Institute of Technology
Center for Research on Parallel Computation
Los Alamos National Laboratory
University of California-Santa Barbara

Case Study 9

Chaos Research and Applications



Simple visual example of the patterns that Chaos finds in complexity.

"Chaos" is used here as a term that refers to a variety of new methods for dealing with nonlinear interdependent systems. This class of complex systems, noted for sensitivity to small effects and seemingly erratic behavior, has been difficult to comprehend using traditional methods and conceptualizations. The methods of chaos have proven to be remarkably effective in revealing new forms of simplicity hidden within complexity. This in turn has led to new understandings of how such systems operate. Nonlinear systems phenomena are present in many fields – from physics, engineering and meteorology to medicine, psychology and economics. Consequently there are many opportunities for application of the methods of chaos.

Computing capability makes the new approaches possible. The methods of chaos focus on the geometry of behavior of a system as a whole and are computationally intensive. Underlying patterns of behavior are best revealed through the use of graphical representation of data. It is only with the availability of high speed processing and improved technologies for graphical display of data that it has become possible to explore a range of behavior of nonlinear systems sufficiently large to permit observation of patterns within complexity. The need for geometric approaches to such problems has been known for many years – it was first demonstrated by Poincare in 1892 in his proof that the three-body problem could not be solved by linear or simple curve approximation methods. However, computational complexity of the problems prevented Poincare's methods from being widely applied until the advent of high performance computing.

As a field chaos is less than 30 years old, but its methods are now being applied to problems in many fields of science and engineering. For example, in physics chaos has been used to refine the understanding of planetary orbits, to reconceptualize quantum level processes, and to forecast the intensity of solar activity. In engineering, chaos has been used in the building of better digital filters, the control of sensitive mechanisms such as ink-jet printers and lasers, and to model the structural dynamics in such structures as buckling columns. In medicine it has been used to study cardiac arrhythmias, the efficiency of lung operations, EEG patterns in epilepsy, and patterns of disease communication. In psychology it has been used to study mood fluctuations, the operation of the olfactory lobe during perception, and patterns of innovation in organizations. In economics it is being used to find patterns and develop new types of econometric models for everything from the stock market to variations in cotton prices.

As an example of the multi-applicability of chaos research products, a research group at the University of Maryland is developing control methods for highly sensitive chaotic processes. The control methods are to be used in such diverse applications as laser control, arrhythmical cardiac tissue, buckling magnetoelastic ribbon, and, in conjunction with Oak Ridge National Laboratory, fluidized bed devices used in chemical and energy applications.

Because chaos is new, growing and deeply interdisciplinary, it benefits greatly from the emerging NREN services for access to remote systems, for sustaining collaborative research activities, and for initiating and maintaining scientific dialogue.

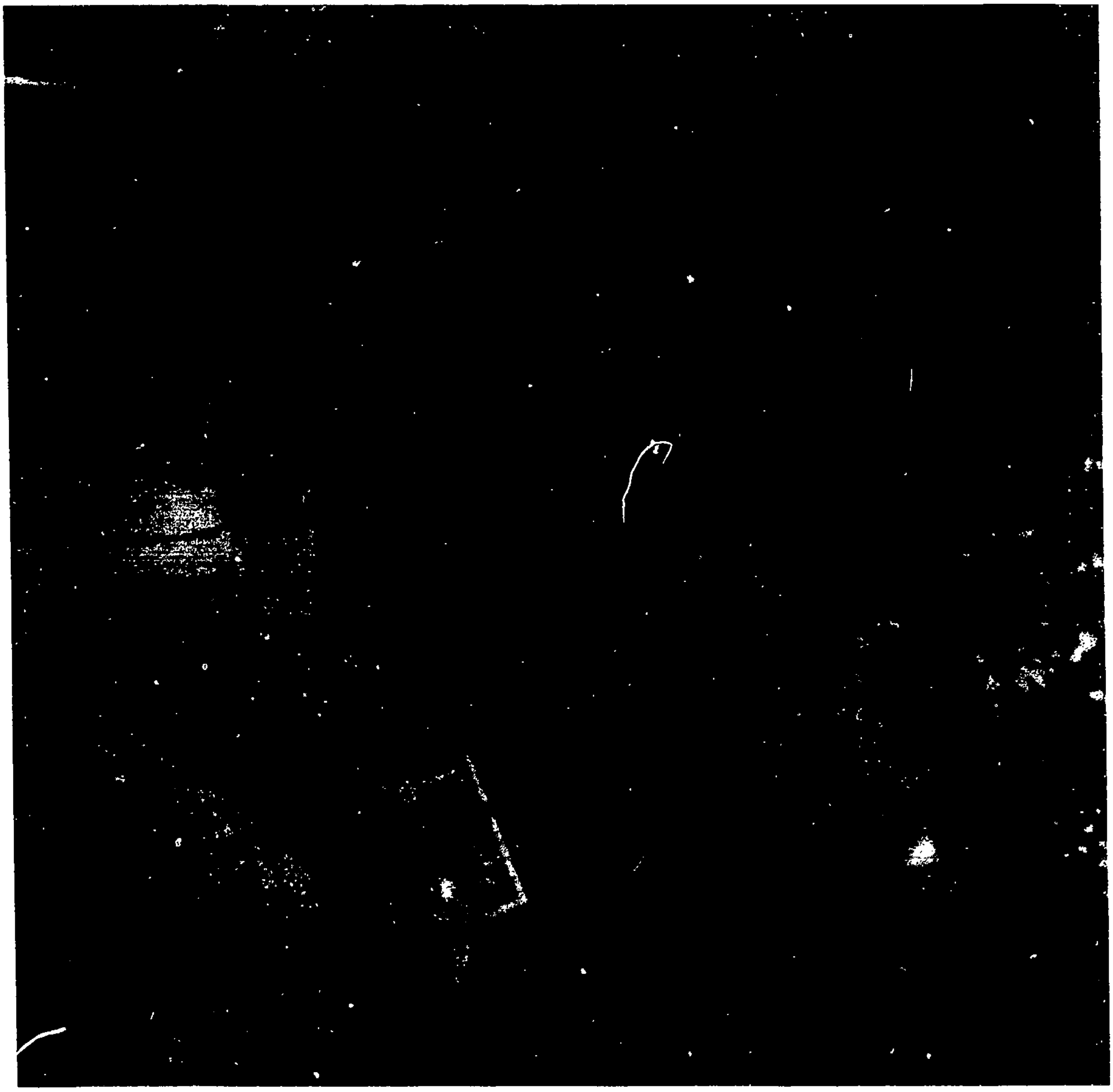
SPONSORING AGENCIES AND ORGANIZATIONS

DOE

PERFORMING ORGANIZATIONS

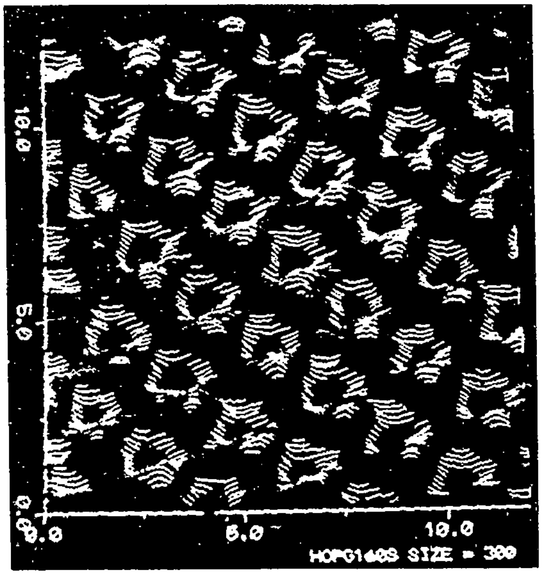
Oak Ridge National Laboratory
Triangle Center for the Study of Complex Systems
University of Maryland at College Park

Case Study 10
Virtual Reality Technology



The photograph shows an investigator, with a head-mounted display and a force-feedback manipulator controlling the tip of a scanning tunneling microscope (STM) probe, exploring the surface of a sample of material. The image projected on the wall indicates the view that the investigator sees and changes as his head position and orientation change. The region of the surface shown is 25 Angstroms on a side.

Virtual reality (VR) is the human experience of perceiving and interacting through sensors and effectors with a synthetic (simulated) environment containing simulated objects as if it were real. It is supported by advances in simulation technology that allow linking human capabilities and computational resources, sensor systems and robotic devices for real-time tasks. VR technol-



In this closeup of the sample on the preceding page, atomic-level structures of carbon rings at the surface are seen. A button permits energy to be supplied on demand through the STM tip to modify the surface by breaking bonds, displacing various structures and enabling physical and chemical reactions.

ogy can be applied to many tasks that would be more difficult to do by other methods.

The view of the simulation that a person sees is an encompassing "immersion" perspective that appears to be from a position inside the simulation itself. By changing the view with head motion and orientation, VR "engages" a human's visual perception, vestibular system, sense of balance, cognition and motor reactions to such an extent that the immersion scenario is experienced as real.

With VR technology, a human and a computer operate together as a combined system, one that is far more capable than either taken individually when performing a large variety of tasks. For example, the environment is particularly well-suited to support reaction to unanticipated scenarios, ones that can exploit the corrective judgment of a person or of skilled teams of experts.

This enabling technology can be used to improve productivity in many important areas. There are numerous applications in the domains of health care, education and lifelong learning, manufacturing and other areas where this technology shows great promise. Early results have shown increased productivity and a dramatic reduction of resource requirements in many instances. Examples of current use include:

- Manipulation of molecules for development of nanotechnology devices and chemical systems.
- Shared surgical interventions.
- Exploration of networked databases and digital libraries for learning and research.
- Modeling, simulation, and analyses.
- Scientific and technical visualization applications.
- Prototyping and planning.
- Training for and monitoring of complex human-computer tasks.

In order to keep pace with real-time interaction, VR technology must be supported by high performance computers, associated software and high bandwidth network capabilities. VR also requires developing new technologies in displays that update in real-time with head motion; advances in sensory feedback such as force, touch, texture, temperature, and smell; and, intelligent models of environments.

With future increases in technological capability VR holds the promise to provide significant improvements in many new application areas.

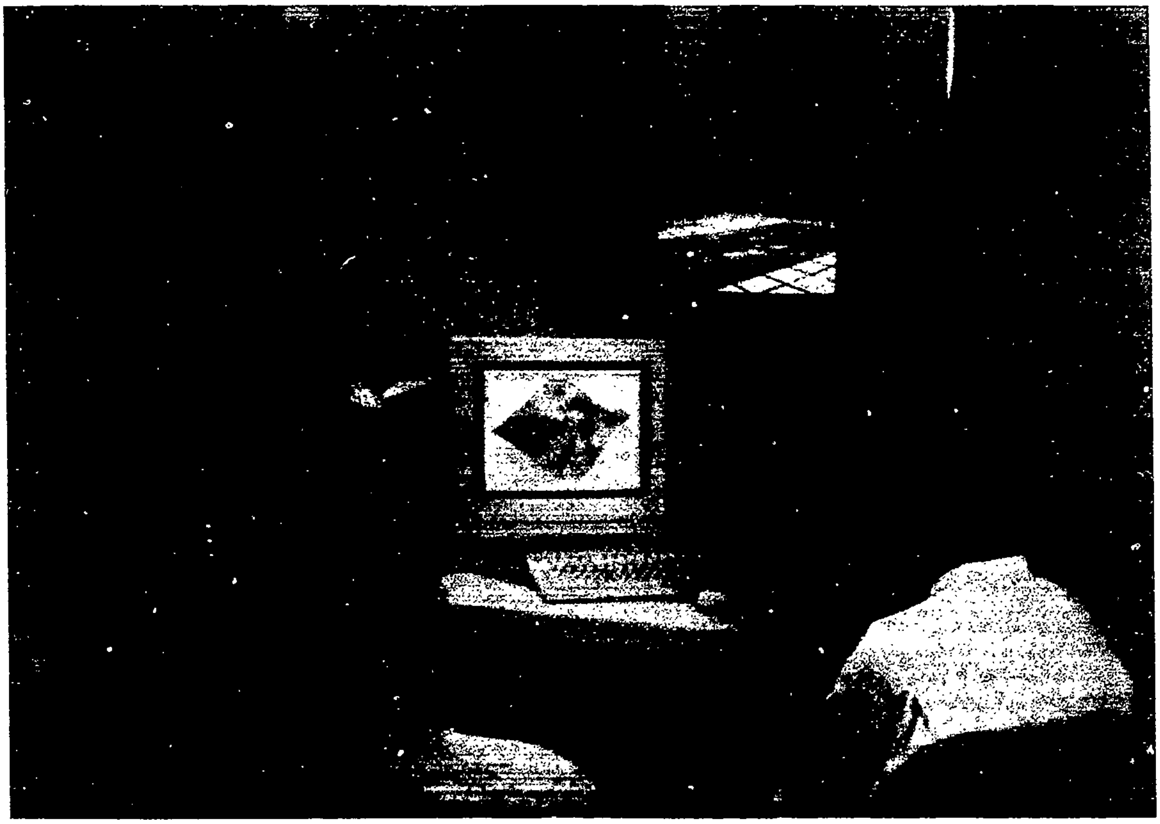
SPONSORING AGENCIES AND ORGANIZATIONS

NASA
NIH
NSF

PERFORMING ORGANIZATIONS

University of California - Los Angeles
University of North Carolina

Case Study 11
HPCC and Education



HPCC, science education and science research are ushering in a new era of learning. Through hands-on activities, state-of-the-art experiments, and novel interactive advanced computing and communications technologies, high school students immerse themselves in educational materials based on cutting-edge, truly modern science research. Closer linkages and shared methods between students and research scientists help attract students of many backgrounds to science as a subject and as a profession.

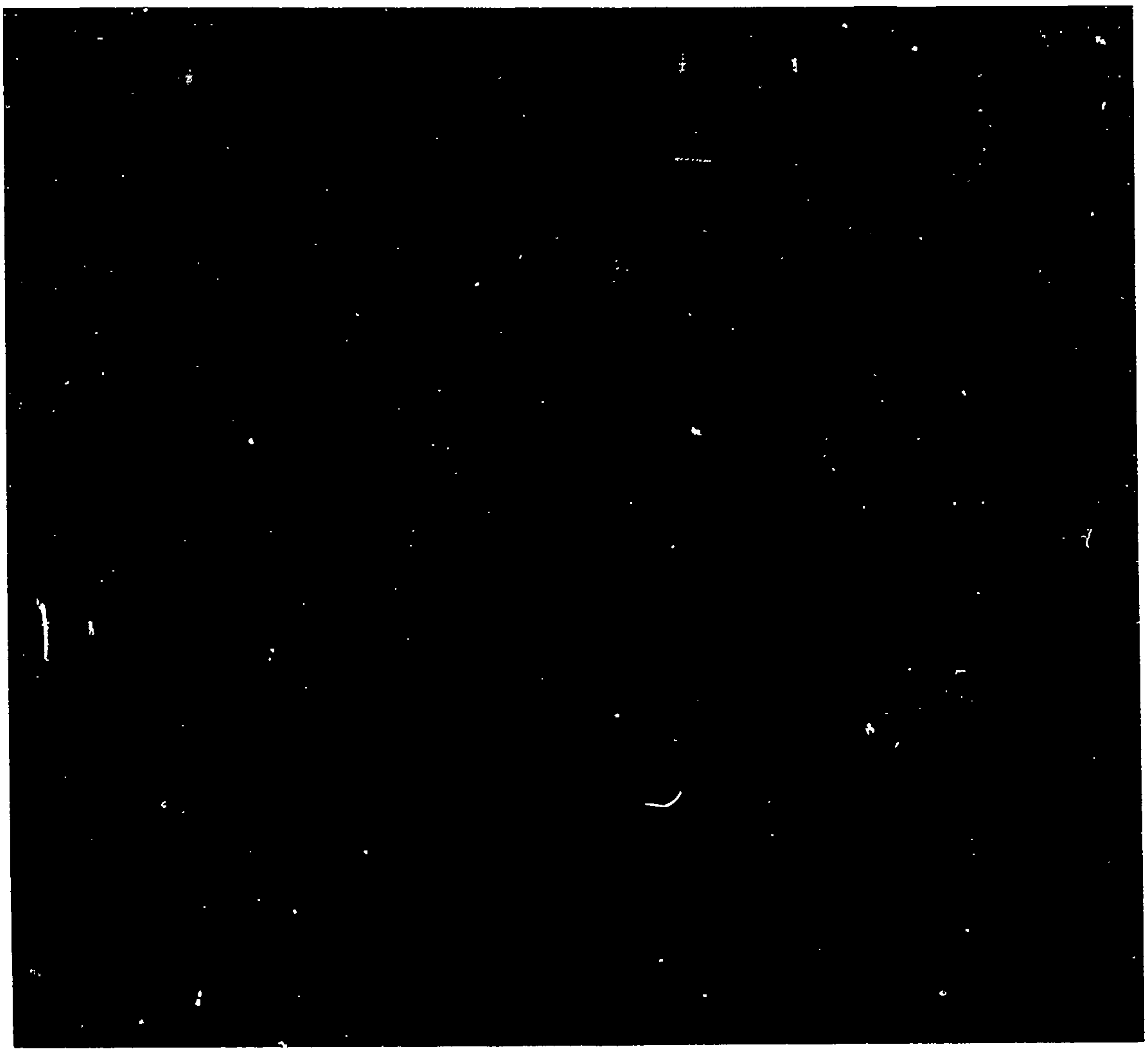
The major impact of HPCC technologies on education will follow the growing integration of HPCC technologies with inquiry-based science learning in the classroom. Direct collaborations with professional scientists and engineers, carried over high bandwidth networks, allow teachers and students to dynamically integrate new scientific advances and methodologies with their own education. In addition, a growing number of HPCC funded programs such as SuperQuest provide individual students with access to high performance computers for student-directed research projects. The two research projects described below illustrate the usefulness of both approaches. In many cases, these activities lead to science and technology careers for students who previously considered these careers outside of their reach.

Learning Through Collaborative Visualization

The Learning Through Collaborative Visualization testbed explores educational uses of scientific visualization and collaborative software tools through shared computer workspaces and two-way audio/video connections. In this project, students and teachers near Chicago *(shown on previous page)* are working on meteorology with atmospheric scientists and graduate students in Champaign-Urbana, and will be joined in 1993 by students and teachers in Michigan. Together, they will study the modeling of unusual weather patterns in their area, utilizing real-time weather data from satellites and other sources, such as clouds, temperature and moisture. Any current information on weather, such as the impending conditions giving rise to a developing storm, can be accessed using networks and displayed using local visualization tools, by researchers and students alike.

Diffusion Limited Aggregation

Diffusion limited aggregation (DLA) structures arise naturally in fields of science, ranging from electrochemical deposition and dendritic solidification to various breakdown phenomena such as dielectric breakdown, viscous fingering, chemical dissolution, and the rapid crystallization of



Color-coded electric field in the vicinity of an electrical conductor molded in the shape of a recently discovered DLA fractal object. The high school student performing the research was a finalist in the 1993 Westinghouse Science Talent Search.

lava. It is not surprising then that there is much interest in the computer modeling of DLA and its variants. Using the computer simulation of DLA, the student can discover new structures that are aesthetically beautiful and at the same time have physical meaning. Students model structures ranging from snowflakes to cancer cells. In the illustration, a student models straight DLA and ascertains its potential energy distribution along the perimeter. As a consequence, the student discovers that the most active regions on the cluster are near the tips.

SPONSORING AGENCIES AND ORGANIZATIONS
ED
NSF

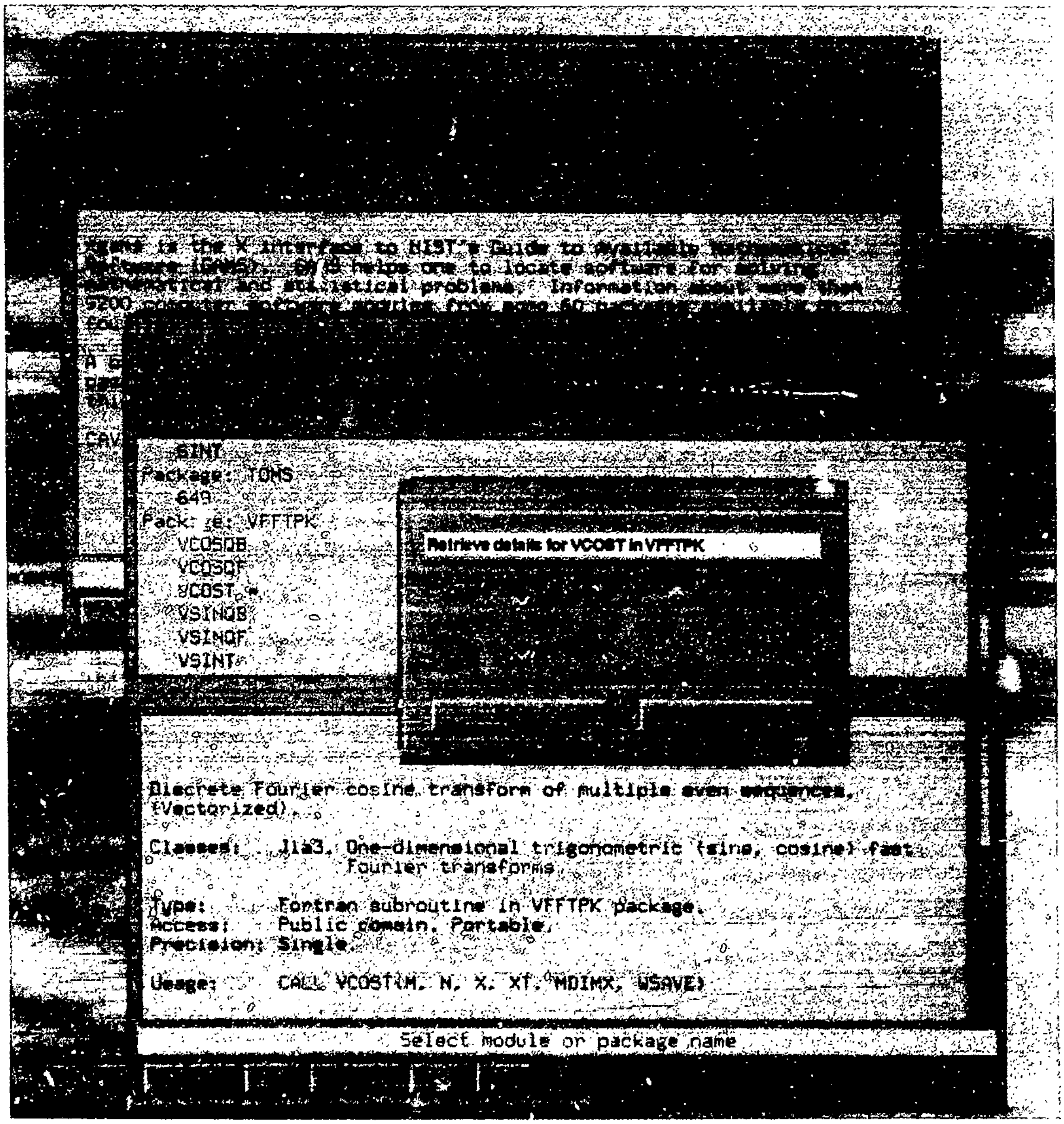
PERFORMING ORGANIZATIONS
Bellcore
High Schools in Illinois and Michigan
Learning through Collaborative Visualization
National Center for Supercomputer Applications
Northwestern University, Department of Psychology
San Francisco Exploratorium
University of Illinois at Champaign-Urbana
University of Michigan, Computer Science

Technology-Dissemination Participants
Ameritech
ScienceKit Publishers
Technical Education Research Center and affiliated (national) LabNet teachers

Diffusion Limited Aggregation
Akron High School, Akron, OH
Boston University, Boston, MA

Case Study 12

HPCC Community Resources



Making effective use of high performance computers requires considerable expertise in diverse areas such as computer architecture, algorithms, numerical analysis, and applied mathematics. The need for insights in all of these arenas is a serious stumbling block to a researcher whose primary interest is in a particular physical application. An essential key to making high performance computers generally usable in today's fast-paced research environment lies in the development and dissemination of reusable software that encapsulates the expertise of computer scientists and numerical analysts.

A wealth of reusable software for solving recurring mathematical and statistical problems is already emerging. Unfortunately, the volume of this software and its distribution over labyrinthine computer networks has made it quite difficult for the average user to locate appropriate tools.

The Guide to Available Mathematical Software (GAMS) is a centralized software resource maintained for scientists at the National Institute of Standards and Technology (NIST). GAMS functions like a management and retrieval system for a large central software repository. However, instead of maintaining the cataloged software itself, it provides a seamless cross-index to multiple repositories managed by others. The term virtual software repository has been used to describe systems of this type.

The steadily increasing collection of software indexed by GAMS now numbers more than 9,200 software modules in about 60 mathematical and statistical software packages available from four physically-distributed software repositories, including the well-known netlib collection of public-domain research software maintained jointly by Oak Ridge National Laboratory, AT&T Bell Laboratories, and the University of Tennessee at Knoxville. Users access GAMS anonymously by logging in to a server system on the Internet. Searches based upon problem classification, keyword, or name may be used to locate appropriate software, after which items such as abstracts, documentation, examples and source code can be retrieved. Both generic and networked (X11) graphical user interfaces to GAMS have been developed.

The classification-based retrieval technology used in GAMS is not unique to mathematics and statistics. The system itself can be easily reconfigured to foster the reuse of software in any application field.

Future directions for the GAMS system include support of the system as a public Internet resource, development of networked GAMS clients and servers, development of improved classification schemes, and research in alternate search mechanisms, including domain-dependent expert system methodologies.

GAMS is a component of an experimental software exchange being coordinated for the HPCC by the NASA Goddard Space Flight Center. The experiment facilitates the development and evaluation of multiple approaches to the problem of software reuse in pursuit of an open non-proprietary architecture to facilitate the emergence of a national software exchange.

SPONSORING AGENCIES AND ORGANIZATIONS

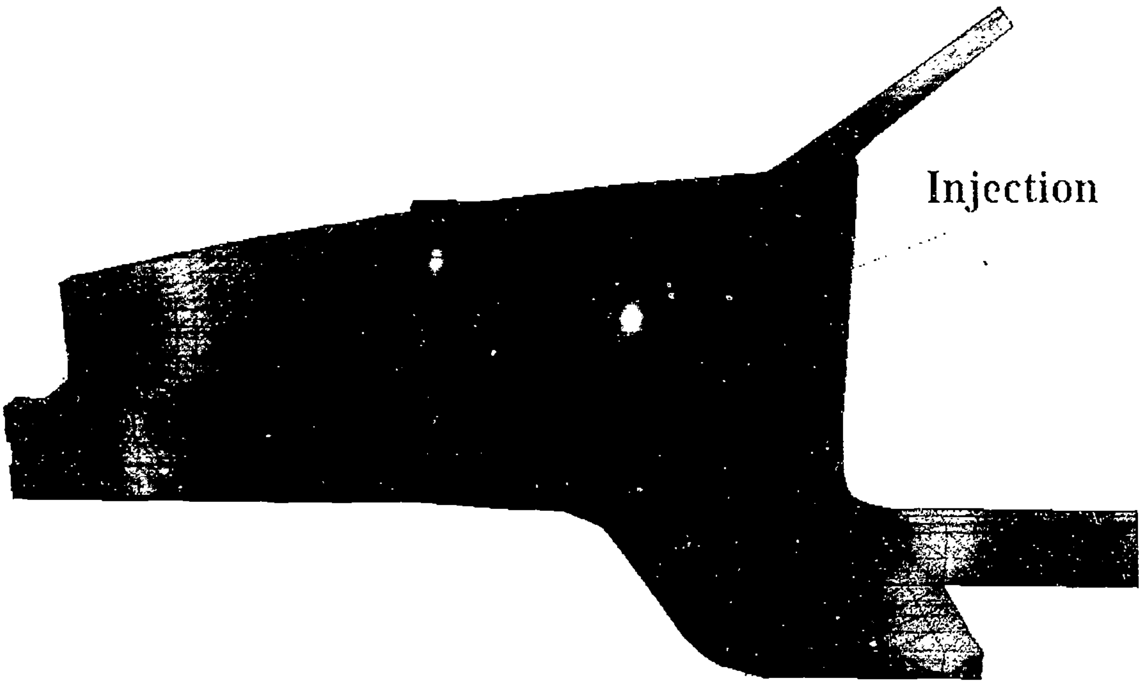
NASA
NIST

PERFORMING ORGANIZATIONS

AT&T Bell Laboratories
NASA
NIST
Oak Ridge National Library
University of Tennessee - Knoxville

Case Study 13

Process Simulation and Modeling



Results of a computer simulation of a liquid molding process for manufacturing high strength polymer composite automotive body panels. The simulation shows how the liquid polymer is injected into the mold and predicts the flow front position as a function of time. The results of the simulation are used to better understand the design of a particular mold and part and to select polymers and preplaced fiber preforms.

In the future, one of the keys to optimizing new product designs and manufacturing processes will be the ability to model and simulate production methods using advanced computer hardware and software. Today, the development of many products, including those that require castings, forgings and injection-molded parts, for example, typically require lengthy and expensive prototyping and experimentation to refine details of the product design, tooling, and manufacturing parameters. The use of computerized process modeling and simulation will eliminate much of this prototyping and dramatically reduce product development times and costs.

The traditional approach of using experimentation and prototyping to refine production processes is indicative of the fact that manufacturing practices have historically been based as much on "art" as on sound scientific understanding of the processes involved. This is changing rapidly as a result of research to achieve better theoretical understanding and predictability of specific processes and materials and their effect on product performance and costs. Improved theoretical frameworks are advancing progress on a wide range of manufacturing applications, processes and materials, including stamping and forming, machining, powder compaction processes and even assembly of parts.

Process modeling and simulation is being pursued by a number of organizations, including the National Institute of Standards and Technology (NIST). For example, as part of a cooperative agreement with Ford, Chrysler and General Motors, NIST is conducting research into the use of polymer composites for structural applications in automobiles. As part of the project, NIST has developed a finite element simulation of the mold filling process used to manufacture high strength body components. In a similar manner, NIST is developing process models and simulations for robotic manufacturing operations and production of powder metal particles.

Much of the ability to apply new theoretical understanding depends on the performance and cost of advanced computing technologies. Process modeling and simulation draws heavily on computationally intensive techniques in such areas as heat transfer and solidification kinetics, coupled fluid flow, finite difference approximations, turbulent viscosity calculations, fracture mechanics, stress analyses, and three-dimensional graphics and visualization techniques. Advanced computing technology directly affects the speed and level of detail that can be incorporated into the models, and therefore the realism and benefits that can be obtained.

SPONSORING AGENCIES AND ORGANIZATIONS

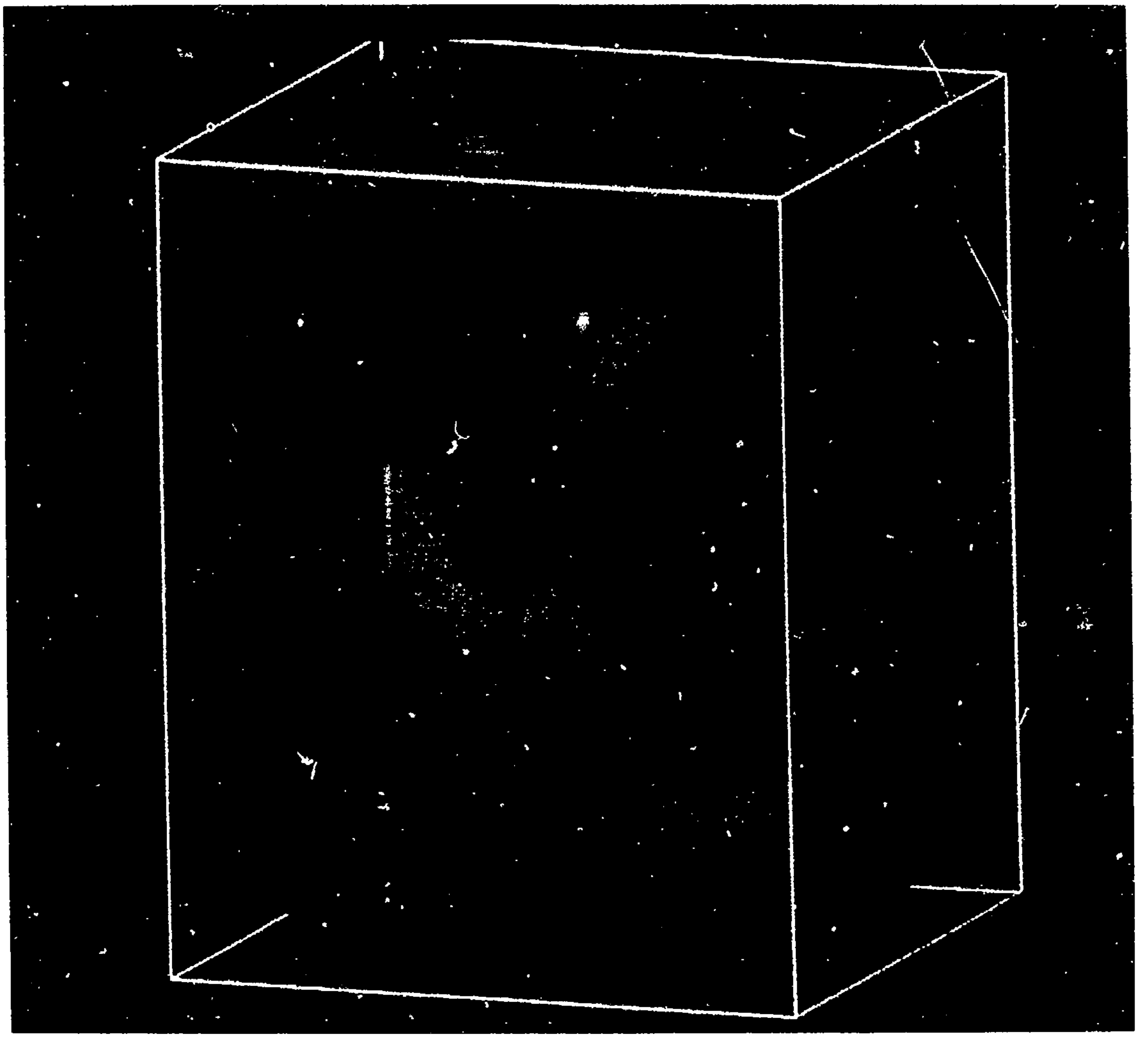
Automotive Composites Consortium
NIST

PERFORMING ORGANIZATIONS

Automotive Composites Consortium
NIST

Case Study 14

Semiconductor Manufacturing for the 21st Century



The results of applying a fully-scalable parallel iterative matrix solver to create a three-dimensional simulation of current flow in a bipolar transistor are shown. The transistor was simulated on the Intel Delta at Caltech using software developed at Stanford University. This simulation is the largest yet in this field and represents the state of the art in three-dimensional device modeling. Such modeling efforts enable process design in the "Virtual Factory."

Semiconductor manufacturing plants today are characterized by high capital costs and low flexibility. They are built to serve one or two generations of technology and usually a narrow product base. New technology development is very expensive and largely empirically done. To be competitive in such an environment requires high capital investment.

In the 1970s, chip designers faced similar investment problems because new designs were largely hand crafted. As a result, few VLSI (Very

Large Scale Integration) applications were viewed as being able to justify design costs. The CAD (Computer-Aided Design) revolution changed this situation by empowering a vastly increased number of creative people with the ability to design chips at low economic risk. The CAD revolution also provided first pass success in most cases on new designs.

A current view, widely held, is that a similar revolution using the same general approach can revolutionize manufacturing. The concept is to build a highly flexible computer controlled manufacturing facility; a "Programmable Factory." In parallel with this factory, a suite of simulation tools would be constructed; a "Virtual Factory" capable of emulating all functions of the real factory. Both of these facilities would be controlled and integrated through a Manufacturing Automation Framework (MAF). The comparison of actual results from the real factory with predicted results from the Virtual Factory provides a means for improving models.

One example of work in progress is ongoing at Stanford University supported by ARPA. Here, the overall aim is to implement a basic version of these constituents in a working IC (Integrated Circuit) facility and to demonstrate the power of this approach in developing and applying new technologies. Much of the work in the first two years of this project has focused on developing specific software tools needed for the Virtual Factory such as simulators and design tools; developing the software needed for the Automation Framework including information storage methods and frameworks for linking software tools; and putting into place the hardware and software tools needed to connect the Virtual Factory to the Programmable Factory.

Recent accomplishments include an integrated demonstration of the concept including operational prototypes of many of these tools. More than 20 different projects in the research program were integrated together to show the power of connecting the Virtual and Real Factories to an audience of more than 100 people from industry, government and universities. Scalable algorithms, developed on a workstation, were used to demonstrate an extremely large three-dimensional device simulation; a major new software tool, SPEEDIE, which simulates thin film etching and deposition steps in IC manufacturing, has been released to industry; and a new multiprocessing reactor designed to perform multiple manufacturing steps in a single machine is being used to test new manufacturing concepts in the real factory.

SPONSORING AGENCIES AND ORGANIZATIONS

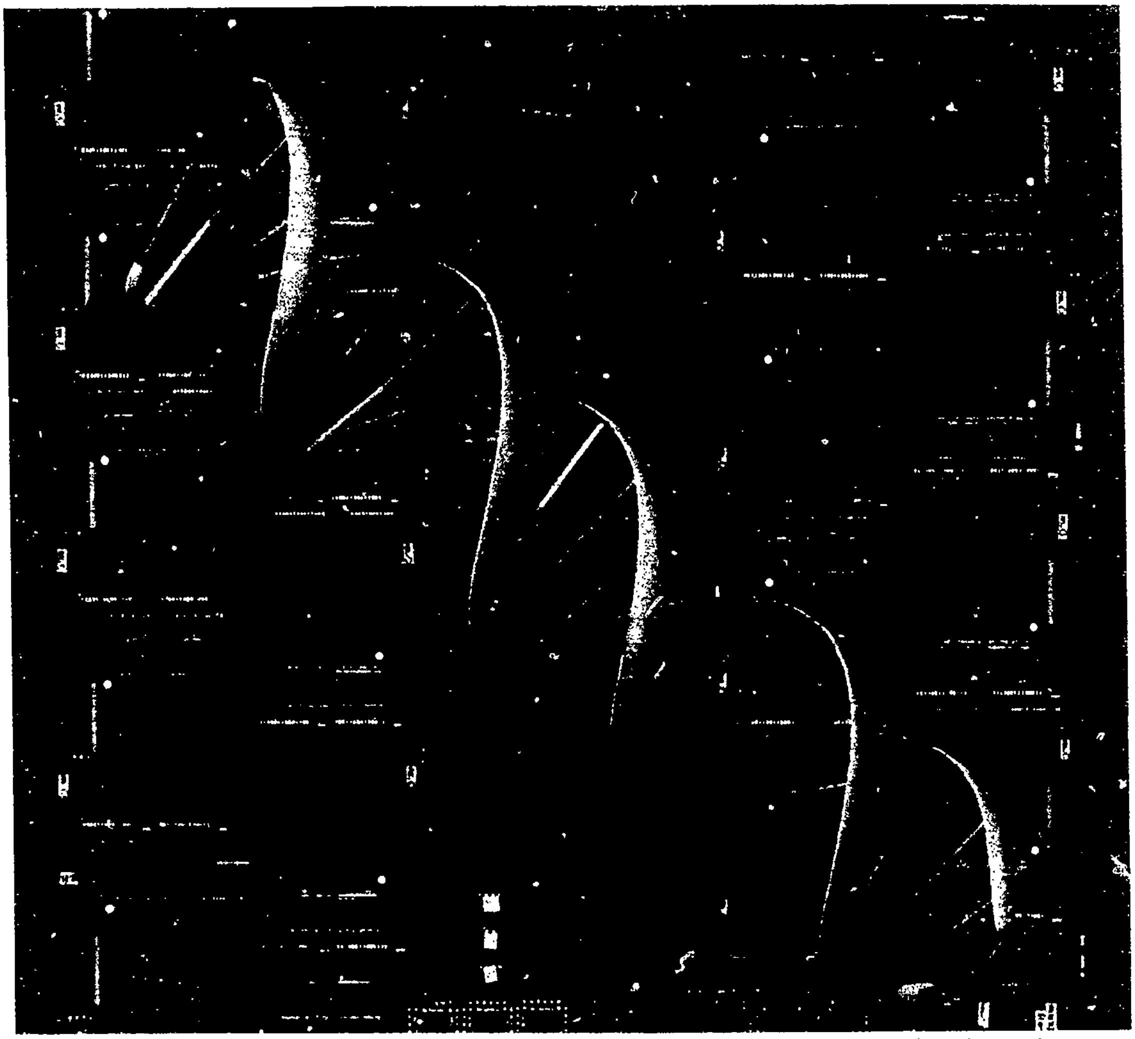
ARPA
Digital Equipment Corporation
Semiconductor Research Corporation
Stanford Center for Integrated Systems industrial sponsors
Texas Instruments

PERFORMING ORGANIZATIONS

Stanford University

Case Study 15

Advances Based on Field Programmable Gate Arrays



Splash 2 board, containing Field Programmable Gate Arrays, which attaches to a workstation. It is programmed in the DOD standard VHDL language and can achieve high performance on problems such as DNA and protein sequence comparison.

Since computers were first built in the 1940s and 1950s, the structure of their processors has been fixed at the time of their construction. As a result, even the fastest computers will perform poorly on a computation whose structure is a bad match for that of the processor.

This basic feature of processor architecture is changing with the use of Field Programmable Gate Arrays (FPGAs) as processor elements. The hardware of an FPGA-based computer is reconfigured for each application. In this way, the structure of the processor can be made to match the structure of the computation, and an FPGA-based computer can achieve supercomputer performance on a range of applications at a fraction of the cost. Importantly, the reconfiguration of the processor structure is done not by designing new computer hardware but by programming the FPGA-based computer in a standard high-level language.

The Splash 1 and Splash 2 attached processor systems, which were designed and built at the Supercomputing Research Center (SRC) using FPGAs from Xilinx, Inc., are just this kind of FPGA-based reconfigurable computer. Together with the National Cancer Institute, the SRC is writing software that will perform DNA and protein sequence comparisons and alignment on existing databases at speeds equal to or greater than those of supercomputers. Applications for Splash 2 are written in VHDL (the DOD standard Very High Speed Integrated Circuit Hardware Description Language). On applications running on Splash 2, performance equal to or as much as 100 times that of some conventional supercomputers has been achieved.

The CM-2X, built by Thinking Machines, Inc., for the SRC, is a CM-2 in which the Weitek floating point accelerator chips have been replaced by Xilinx FPGAs, which can be programmed with the usual CM-2 languages and utilities. It is common on multiprocessor machines to find such floating point accelerators: the use of FPGAs in their place allows computationally intensive kernels of an application to be executed on FPGA-based hardware specifically configured for the application. On complete applications programmed at the SRC, the CM-2X has achieved three to four times the performance of the CM-2.

Similar approaches are being evaluated for other architectures. For example, the CM-5, also built by Thinking Machines, Inc., has from 32 to 1,024 network-connected nodes each consisting of a high performance microprocessor and four custom designed vector processors. Embedding FPGAs into its nodes and network offers the potential for producing dramatic improvement on specific applications and, by altering the FPGA programming, observation of internal system states to assist in program tuning and debugging.

SPONSORING AGENCIES AND ORGANIZATIONS

NSA

Case Study 16

High Performance Fortran and Its Environment

In an effort to provide a machine-independent programming interface, a working group of representatives from industry, academia and government laboratories has defined an informal standard called High Performance Fortran (HPF). HPF consists of language extensions to Fortran 90 that support data-parallel programming by permitting specification of the distribution of data structures across the processors of a scalable parallel computer system. The ideas behind HPF have been taken in large part from earlier research on a language called Fortran D. This work, which was conducted at the Center for Research on Parallel Computation at Rice University and at Syracuse University, was sponsored under the HPCC Program by NSF and ARPA. The Fortran D project designed a scalable language on the basis of analysis of existing parallel applications, prototype compiler implementations and an extensive benchmark suite that are all available to the HPCC community. The HPF draft standard, which was produced over a single year, promises to dramatically reduce the burden of specifying parallel scientific programs for the new generation of massively parallel computer systems. Four companies have already announced HPF compilers that will be available in 1993. HPF will accelerate the use of parallel machines, as now users can use a high level software with the assurance that their applications will be supported by future as well as current systems.

Since HPF hides many of the details of parallel programming for a specific target machine, the programmer will need assistance in preparing, debugging and tuning HPF programs. To this end, ARPA is sponsoring a research project at Rice University and the University of Illinois at Champaign-Urbana to design and build a collection of programming tools for HPF, including an intelligent editor to help construct efficient data distributions, a debugger to locate data races and other programmer errors and a performance analysis and visualization system. This project has a secondary goal of providing a platform for the development of other software tools for HPF. When complete, this software platform will be distributed throughout the computational science community in source form.

The current version of HPF supports data parallel programming for about 50-75 percent of applications but has limited support for "unstructured" or "irregular" problems. Since this class includes many important science and engineering problems, NASA, ARPA, ONR and NSF are jointly sponsoring a research project at Rice, Syracuse and the University of Maryland to design and implement extensions to HPF to support this important problem class. This work will affect the next round of HPF standardization, scheduled for calendar year 1994.

SPONSORING AGENCIES AND ORGANIZATIONS
- ARPA
- NASA
- NSF
- ONR

PERFORMING ORGANIZATIONS
- Center for Research on Parallel Computation
- Syracuse University
- University of Illinois at Champaign-Urbana
- University of Maryland

```
Jacobi iteration, using HPF:

    REAL X(N,N), Y(N,N)
!HPF$ DISTRIBUTE (*,BLOCK)::X,Y

    FORALL ( J = 2 : N-1 )
     FORALL ( I = 2 : N-1 )
      X(I,J) = 0.25 * (Y(I-1,J) + Y(I+1,J) + Y(I,J-1) + Y(I,J+1))
     END FORALL
    END FORALL
```

```
Jacobi iteration, using message passing:

    REAL X(N,N_OVER_P), Y(N,0:N_OVER_P+1)

    IF (MY_ID .NE. 0) THEN
     SEND( Y(1:N,1), MY_ID-1 )
    END IF
    IF (MY_ID .NE. P-1) THEN
     SEND( Y(1:N,N_OVER_P), MY_ID+1 )
    END IF
    IF (MY_ID .NE. P-1) THEN
     RECEIVE( Y(1:N,N_OVER_P+1), MY_ID+1 )
    END IF
    IF (MY_ID .NE. 0) THEN
     RECEIVE( X(1:N,0), MY_ID-1 )
    END IF
    LOW = 1
    IF (MY_ID .EQ. 0) LOW=2
    HIGH = N_OVER_P
    IF (MY_ID .EQ. P-1) HIGH=N_OVER_P-1
    DO J = LOW,HIGH
     DO I = 2, N-1
      X(I,J) = 0.25 * (Y(I-1,J) + Y(I+1,J) + Y(I,J-1) + Y(I,J+1))
     END DO
    END DO
```

These code segments both perform square mesh Jacobi relaxation on a distributed memory parallel computing system.

Contributing Writers

The editorial group wishes to thank the following people who contributed to this section.

Ron Boisvert, NIST
Duncan Buell, Supercomputing Research Center, NSA
Linda Callahan, Cornell Theory Center
Y. T. Chien, NSF
Richard Crutcher, University of Illinois - Champaign-Urbana
Eduardo D'Azevedo, Oak Ridge National Laboratory, DOE
Robin Dennis, EPA
Mark Ellisman, San Diego Microscopy and Imaging Resource
Sally Goerner, Triangle Center for the Study of Complex Systems
John Hestenes, Drexel University
Jim Jenkins, NASA
Chuck Koebel, Rice University
James Lick, University of California - Santa Barbara
Wilbur Lick, University of California - Santa Barbara
Dan Masys, NLM/NIH
Paul Messina, California Institute of Technology
John Meyer, NIST
Joan Novak, EPA
Bob Parker, ARPA
Tina Pauna, Concurrent Supercomputing Consortium
Kathryne Pickens, University of California-Santa Barbara
Fred Phelan, NIST
Cal Ramos, NCO/HPCC, NASA
Ann Redelfs, Center for Research on Parallel Computation
John Riganati, Supercomputing Research Center, NSA
Charles Romine, Oak Ridge National Laboratory, DOE
Bruce Ross, GFDL/NOAA
Nora Sabelli, NSF
John Salmon, Center for Research on Parallel Computation, California Institute of Technology
Paulette Sancken, National Center for Supercomputing Applications
Stephanie Sides, San Diego Supercomputer Center
Kevin Timson, Center for Research on Parallel Computation
John Toole, ARPA
Laura Toran, Oak Ridge National Laboratory, DOE
Robert Voigt, NSF
Michael S. Warren, University of California - Santa Barbara, Los Alamos National Laboratory, DOE
James Weaver, Environmental Research Laboratory, EPA
Gil Weigand, ARPA
Olivia West, Oak Ridge National Laboratory, DOE
Mary Wheeler, Rice University
Patricia N. Williams, NCO/HPCC

Glossary

ACTS
Advanced Communications Technology Satellite, part of a joint NASA-ARPA NREN collaboration that will provide high speed ATM/SONET transmission, and will provide interface and operations experience in mating high speed terrestrial communications systems with high speed satellite communications systems

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

ARPA
Advanced Research Projects Agency, part of DOD (formerly DARPA, Defense Advanced Research Projects Agency)

ASTA
Advanced Software Technology and Algorithms, a component of the HPCC Program

ATDnet
Advanced Technology Demonstrations Network, an ARPA project that will make use of technology developed for and by gigabit testbeds

ATM
Asynchronous Transfer Mode, a new telecommunications technology, also known as cell switching, which is based on 53 byte cells

Backbone Network
A high capacity electronic trunk connecting lower capacity networks, e.g., NSFNET backbone

Bit
Acronym for binary digit

BRHR
Basic Research and Human Resources, a component of the HPCC Program

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

CLNP
ConnectionLess Network Protocol

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

CPMES

Committee on Physical, Mathematical, and Engineering Sciences

DOC

Department of Commerce

DOD

Department of Defense

DOE

Department of Energy

ED

Department of Education

EPA

Environmental Protection Agency

ESnet

Energy Sciences Network

FCCSET

Federal Coordinating Council for Science, Engineering, and Technology

FDDI

Fiber Distributed Data Interface

Flops

Acronym for floating point operations per second. The term "floating point" refers to that format of numbers that 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.

FNC

Federal Networking Council

Giga-

10^9 or billions of ... (e.g., gigabits, gigaflops, gigaops)

GOSIP
Government Open Systems Interconnection Profile

Grand Challenge
A fundamental problem in science and engineering, with broad economic and scientific impact, whose solution can be advanced by applying high performance computing techniques and resources

Heterogeneous system
A distributed system that contains more than one kind of computer

High performance computer
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
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 systems and applications software with all components well integrated and linked over a high speed network

HIPPI
High Performance Parallel Interface

HHS
Department of Health and Human Services

HPCC
High Performance Computing and Communications

HPCCIT
High Performance Computing, Communications, and Information Technology Subcommittee

HPCS
High Performance Computing Systems, a component of the HPCC Program

IINREN – see Interagency Internet

IITA
Information Infrastructure Technology and Applications, a component of the HPCC Program

Interagency Interim NREN – see Interagency Internet

Interagency Internet

The federally funded part of the Internet that is directly used by the HPCC Program. It includes NSFNET, ESnet, and NSI. The Interagency Internet is also the name of an element of the NREN component of the HPCC Program, and is the successor to the Interagency Interim NREN (IINREN).

Internet

The global collection of interconnected, multiprotocol computer networks including Federal, mid-level, private, and international networks. The Interagency Internet is a part of the Internet, and NSFNET, ESnet, and NSI are part of the Interagency Internet.

InterNIC

Internet Network Information Center

Interoperability

The effective interconnection of two or more different computer systems, databases, or networks in order to support distributed computing and/or data exchange

ISDN

Integrated Services Digital Network

Kb/s

Kilobits per second or thousands of bits per second

LAN

Local Area Network

Mb/s

Megabits per second or millions of bits per second

MIMD

Multiple Input Multiple Data

MPP

Massively parallel processor

NAPs

Network Access Points, a set of nodes connecting mid-level or regional networks and other service providers to NSFNET

NASA

National Aeronautics and Space Administration

National Challenge

A fundamental application that has broad and direct impact on the Nation's competitiveness and the well-being of its citizens, and that can benefit by the application of HPCC technology and resources

National Information Infrastructure (NII)

The integration of hardware, software, and skills that will make it easy and affordable to connect people with each other, with computers, and with a vast array of services and information resources.

NCO

National Coordination Office for High Performance Computing and Communications

Network

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

NIH

National Institutes of Health, part of HHS

NIST

National Institute of Standards and Technology, part of DOC

NLM

National Library of Medicine, part of NIH

NOAA

National Oceanic and Atmospheric Administration, part of DOC

NREN

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

NSA

National Security Agency, part of DOD

NSF

National Science Foundation

NSFNET

NSF computer network

NSI

NASA Science Internet

NTIA

National Telecommunications and Information Administration

OC-3

Network transmission speed of 155 Mb/s

OC-12

Network transmission speed of 622 Mb/s

OMB
Office of Management and Budget

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.

OSI
Open Systems interconnection

OSTP
Office of Science and Technology Policy

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

Peta-
10^{15} of ... (e.g., petabits)

Port
Transport a computer program from one computer system to another

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.

R&D
Research and development

Scalable
A system is scalable if it can be made to have more (or less) computational power by configuring it with a larger (or smaller) number of processors, amount of memory, interconnection bandwidth, input/output bandwidth, and amount of mass storage.

SIMD
Single Instruction Multiple Data

SONET
Synchronous Optical Network

Tera-
10^{12} or trillions of ... (e.g., terabits, teraflops)

T1
Network transmission of a DS1 formatted digital signal at a rate of 1.5 Mb/s

T3
Network transmission of a DS3 formatted digital signal at a rate of 45 Mb/s

vBNS
Very high speed Backbone Network Services

WAN
Wide Area Network

Contacts

NCO (National Coordination Office for High Performance Computing and Communications)

Bldg. 38A/B1-N30
8600 Rockville Pike
Bethesda, MD 20894
301-402-4100
fax: 301-402-4080
nco@hpcc.gov

Donald A. B. Lindberg, M.D.
Director
lindberg@hpcc.gov

Patricia R. Carson
Executive Assistant
carson@hpcc.gov

Sally E. Howe, Ph.D.
Assistant Director for Technical Programs
NIST 1992-93 Commerce Science and Technology (ComSci) Fellow
howe@hpcc.gov

Charles R. Kalina
Executive Secretary, HPCCIT Subcommittee
kalina@hpcc.gov

Calvin T. Ramos
Program Associate
NASA 1992-93 Professional Development Program

Shannon N. Uncangco
Staff Assistant
uncangco@hpcc.gov

Patricia N. Williams
Information Officer
pwilliams@hpcc.gov

HPCC Internet Server

www.hpcc.gov

InterNIC Information Services

General Atomics
PO Box 85608
San Diego, CA 92186-9784
800-444-4345
619-455-4600
fax: 619-455-3990
info@internic.net

ARPA

Lt Col John C. Toole
Acting Director, Computing Systems Technology Office
Advanced Research Projects Agency
3701 N. Fairfax Dr.
Arlington, VA 22203
703-696-2264
toole@arpa.mil

Stephen L. Squires
Special Assistant to the Director
Advanced Research Projects Agency
3701 N. Fairfax Dr.
Arlington, VA 22203
703-696-2226
squires@arpa.mil

Randy Katz, Ph.D.
Program Manager for Information Infrastructure Technology
Advanced Research Projects Agency
3701 N. Fairfax Dr
Arlington, VA 22203
703-696-2228
katz@arpa.mil

181

DOE

HPCC Program Coordinator
Office of Scientific Computing
ER-30, GTN
Department of Energy
Washington, DC 20585
301-903-5800
hpcc@er.doe.gov

ED

James A. Mitchell, Ph.D.
Special Assistant for Technology
Office of the Assistant Secretary
Department of Education
555 New Jersey Ave., NW #604B
Washington, DC 20208
202-219-2053
jmitchel@inet.ed.gov

EPA

General information:

Joan H. Novak
HPCC Program Manager, MD-80
Environmental Protection Agency
Research Triangle Park, NC 27711
919-541-4545
novak.joan@epamail.epa.gov

Robin L. Dennis, Ph.D.
Senior Science Program Manager
Environmental Protection Agency
Research Triangle Park, NC 27711
919-541-2870
rdennis@flyer.ncsc.org

Internet use:

Bruce P. Almich
EPA Federal Network Council Representative, MD-90
Environmental Protection Agency
Research Triangle Park, NC 27711
919-541-3306
almich.bruce@epamail.epa.gov

NASA

Lee B. Holcomb
Director, High Performance Computing and Communications Office
Office of Aeronautics
National Aeronautics and Space Administration
Code RC
Washington, D.C. 20546
202-358-2747
l_holcomb@aeromail.hq.nasa.gov

Paul H. Smith, Ph.D.
Program Manager, HPCC
National Aeronautics and Space Administration
Code RC
Washington, D.C. 20546
202-358-4617
ph_smith@aeromail.hq.nasa.gov

Paul E. Hunter
HPCC Program Manager for Basic Research and Education
National Aeronautics and Space Administration
Code RC
Washington, D.C. 20546
202-358-4618
P_Hunter@aeromail.hq.nasa.gov

NIH

NIH HPCC Coordinator and NLM contact:

Daniel R. Masys, M.D.
Director, Lister Hill National Center for Biomedical Communications
National Library of Medicine
Bldg. 38A/7N707
Bethesda, MD 20894
301-496-4441
masys@lhc.nlm.nih.gov

NCRR program:

Judith L. Vaitukaitis, M.D.
Director, National Center for Research Resources
National Institutes of Health
Bldg. 12A/4007
Bethesda, MD 20892
301-496-5793
judyv@nihrr12a.bitnet@cu.nih.gov

Caroline Holloway, Ph.D.
Director, Office of Science Policy and Communication
National Center for Research Resources
National Institutes of Health
Bldg. 12A/4057
Bethesda, MD 20892
301-496-2992
CarolineH@nihrr12a.bitnet

DCRT program:

David Rodbard, M.D.
Director, Division of Computer Research and Technology
National Institutes of Health
Bldg. 12A/3033
Bethesda, MD 20892
301-496-5703
dar@cu.nih.gov

William L. Risso
Deputy Director, Division of Computer Research and Technology
National Institutes of Health
Bldg. 12A/3033
Bethesda, MD 20892
301-496-8277
risso@nih.gov

Robert L. Martino, Ph.D.
Chief, Computational Bioscience and Engineering Laboratory
Division of Computer Research and Technology
National Institutes of Health
Bldg. 12A/2033
Bethesda, MD 20892
301-496-1111
martino@alw.nih.gov

NCI/BSC program:

Jacob V. Maizel, Ph.D.
Biomedical Supercomputer Center
National Cancer Institute
Frederick Cancer Research Facility
301-846-5532
jmaizel@fcrfv1.ncifcrf.gov

NIST

General information:

James H. Burrows
Director, Computer Systems Laboratory
National Institute of Standards and Technology
Technology Building, Room B154
Gaithersburg, MD 20899
301-975-2822
burrows@ecf.ncsl.nist.gov

NOAA

General information:

Thomas N. Pyke, Jr.
Director for High Performance Computing and Communications
National Oceanic and Atmospheric Administration
1315 East-West Highway
Room 15300
National Oceanic and Atmospheric Administration
Silver Spring, MD 20910
301-713-3573
tpyke@hpcc.noaa.gov

Directors of high performance computing centers:

Jerry Mahlman, Ph.D.
Director, Geophysical Fluid Dynamics Laboratory
National Oceanic and Atmospheric Administration
P.O. Box 308
Princeton, NJ 08542
609-452-6502
jm@gfdl.gov

Ron MacPherson, Ph.D.
Director, National Meteorological Center
National Oceanic and Atmospheric Administration
5200 Auth Road, Room 101
Camp Springs, MD 20746
301-763-8016
ronmcp@nic.fb4.noaa.gov

Sandy MacDonald, Ph.D.
Director, Forecasting Systems Laboratory
National Oceanic and Atmospheric Administration
325 Broadway
Boulder, CO 80303
303-497-6378
macdonald@fsl.noaa.gov

Internet Use:

Robert L. Mairs
Chief, Information Processing Division
Office of Satellite Data Processing and Distribution
National Oceanic and Atmospheric Administration
FB4 - Room 0301
Suitland and Silver Hill Roads
Suitland, MD 20746
301-763-5687
rmairs@nesdis.noaa.gov

NSA

George Cotter
Chief Scientist
National Security Agency
9800 Savage Rd.
Ft. Meade, MD 20755-6000
301-688-6434
grcotte@afterlife.ncsc.mil

NSF

Merrell Patrick, Ph.D.
HPCC Coordinator
National Science Foundation
Directorate for Computer and Information Science and Engineering
4201 Wilson Blvd., Room 1105
Arlington, VA 22230
703-306-1900
mpatrick@nsf.gov

Index

This is an index to major sections and selected key words and phrases that appear in this document. It complements the Table of Contents (pages v-vi), List of Tables (page vii), List of Illustrations (pages viii-ix), and Glossary (pages 159-165).