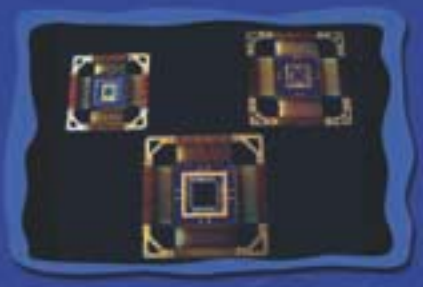
NITRD Networking and Information Technology Research and Development

2005

GUIDE TO THE NITRD PROGRAM

FY 2004 - FY 2005

Supplement to the President's Budget



NITRD Guide

About the National Science and Technology Council

The National Science and Technology Council (NSTC) was established by Executive Order on November 23, 1993. This Cabinet-level council is the principal means by which the President coordinates science, space, and technology policies across the Federal government. NSTC coordinates the diverse parts of the Federal research and development enterprise. An important objective of the NSTC is the establishment of clear national goals for Federal science and technology investments in areas ranging from information technologies and health research to improving transportation systems and strengthening fundamental research. The Council prepares research and development strategies that are coordinated across Federal agencies to form a comprehensive investment package that is aimed at accomplishing multiple national goals.

The NSTC Web site is: www.nstc.gov. To obtain additional information regarding the NSTC, please contact the NSTC Executive Secretariat at (202) 456-6100.

About the Office of Science and Technology Policy

The Office of Science and Technology Policy (OSTP) was established by the National Science and Technology Policy, Organization, and Priorities Act of 1976. OSTP's responsibilities include advising the President in policy formulation and budget development on all questions in which science and technology are important elements; articulating the President's science and technology policies and programs; and fostering strong partnerships among Federal, state, and local governments, and the scientific communities in industry and academe. The Assistant to the President for Science and Technology serves as the Director of the OSTP and directs the NSTC on behalf of the President.

The OSTP Web site is: www.ostp.gov. For additional information about OSTP, please call (202) 456-7116.

Copyright Information

This is a work of the U.S. Government and is in the public domain. It may be freely distributed and copied, but it is requested that the National Coordination Office for Information Technology Research and Development (NCO/IT R&D) be acknowledged.

To Request Additional Copies

To request additional copies of this Supplement to the President's FY 2005 Budget or other publications of the National Coordination Office for Information Technology Research and Development, please contact: National Coordination Office for IT R&D, Suite II-405, 4201 Wilson Boulevard, Arlington, Virginia 22230; (703) 292-4873; fax: (703) 292-9097; e-mail: nco@nitrd.gov.

Buy American Report

Congress requires information concerning non-U.S. high-performance computing and communications funding activities. In FY 2004, DARPA was the only NITRD agency that entered into grants, contracts, cooperative agreements, or cooperative research and development agreements for information technology research and development with either 1) a company other than a company that is either incorporated or located in the U.S. and that has majority ownership by individuals who are citizens of the U.S., or 2) an educational institution or nonprofit institution located outside the U.S. DARPA funded an IT research-related award of $1.076 million to Cambridge University, Cambridge (UK). In FY 2004, no NITRD procurement exceeds $1 million for unmanufactured articles, materials, or supplies mined or produced outside the U.S., or for manufactured articles, materials, or supplies other than those manufactured in the U.S. substantially all from articles, materials, or supplies mined, produced, or manufactured in the U.S.

NETWORKING AND INFORMATION TECHNOLOGY
RESEARCH AND DEVELOPMENT

GUIDE

TO THE

NITRD PROGRAM

FY 2004 - FY 2005

SUPPLEMENT
TO THE PRESIDENT'S BUDGET
FOR FY 2005

A Report by the Interagency Working Group
on Information Technology Research and Development

National Science and Technology Council

December 2004

EXECUTIVE OFFICE OF THE PRESIDENT
OFFICE OF SCIENCE AND TECHNOLOGY POLICY
WASHINGTON, D.C. 20502

MEMBERS OF CONGRESS:

I am pleased to forward with this letter the annual report on the multi-agency Networking and Information Technology Research and Development (NITRD) Program. This Supplement to the President s Budget for Fiscal Year 2005 describes activities funded by Federal NITRD agencies in the areas of advanced networking and information technologies. Investments in these fundamental technologies are continuing to fuel the engine of innovation in science and technology innovation that is essential to the Nation s current and future security and economic prosperity.

The impact of the NITRD Program is felt not only in the research communities directly supported by NITRD funds, but also in a wide range of scientific and engineering research efforts that span both Federal programs and the private sector. Multi-agency programs like NITRD provide the Federal government the opportunity to maximize this impact through a coordinated investment strategy, and this report provides a clear picture of the value of such coordination.

The NITRD Program helps to assure that the United States continues to lead the world in science and engineering by supporting fundamental research, education, and the development of new information technologies that continue to transform our economy and enhance our standard of living. I am pleased to provide to you this report.

Sincerely,

John H. Marburger, III
Director

Table of Contents

Introduction 1

The NITRD Agencies 2

NITRD Research Interests of the Agencies 3

The NITRD Program Today 4

IWG: Coordination and Activities 5
Grand Challenges: IT Goals in the National Interest 6
16 Illustrative Grand Challenges 6
IT 'Hard Problem' Areas 6

President's Information Technology Advisory Committee 7
Members of PITAC 7

High End Computing – Infrastructure and Applications 8

Definition of HEC I&A PCA 8
Benefits of HEC to U.S. Science and Engineering 9

High End Computing – Research and Development 10

Definition of HEC R&D PCA 10
Roadmaps for Federal HEC R&D in the Years Ahead 11
HEC PCAs: Coordination and Activities 12
Multiagency ESMF Effort 13
New HEC-URA 13

High End Computing Revitalization Task Force (HECRTF) 14
Excerpts From the HECRTF Report 14
Key Recommendations of the HECRTF Report 15
HECRTF Proposed R&D Priorities to FY 2010 (table) 16

HEC I&A and R&D Programs by Agency:
Selected FY 2004 Activities and FY 2005 Plans 17

Human-Computer Interaction and Information Management 32

Definition of HCI&IM PCA 32
HCI&IM PCA: Coordination and Activities 33
HCI&IM R&D Programs by Agency: Selected FY 2004 Activities and FY 2005 Plans . 34

Large Scale Networking **42**

Definition of LSN PCA 42
LSN PCA: Coordination and Activities 43
Agencies' Current Networking R&D Interests 43
LSN R&D Programs by Agency: Selected FY 2004 Activities and FY 2005 Plans 45

Software Design and Productivity **54**

Definition of SDP PCA 54
SDP PCA: Coordination and Activities 55
Scope of SDP R&D Topics 55
SDP R&D Programs by Agency: Selected FY 2004 Activities and FY 2005 Plans 57

High Confidence Software and Systems **64**

Definition of HCSS PCA 64
HCSS PCA: Coordination and Activities 65
HCSS R&D Programs by Agency: Selected FY 2004 Activities and FY 2005 Plans . . . 66

Social, Economic, and Workforce Implications of IT and IT Workforce Development **74**

Definition of SEW PCA 74
SEW PCA: Coordination and Activities 75
SEW R&D Programs by Agency: Selected FY 2004 Activities and FY 2005 Plans . . . 76

Interagency Working Group on IT R&D **81**

IWG Members 81
PCA Coordinating Groups and Team Chairs 81
Participation in Federal NITRD Activities 82

Agency NITRD Budgets by Program Component Area **83**

Agency Contacts **84**

Glossary 92
Acknowledgements 105
Buy American Report . . . . . inside front cover
Abstract . . . . . inside back cover
Cover Design and Imagery . . . . . inside back cover

Dedication

This Blue Book is dedicated to the memory of Frank D. Anger, an esteemed research manager in the National Science Foundation. Frank died in an auto accident on July 7, 2004. He was an active and dedicated participant in the NITRD Program who, as co-chair of the Software Design and Productivity Coordinating Group, vigorously championed the need for new engineering paradigms to improve the quality, reliability, and ease of use of contemporary software. His grace, wisdom, kindness, and unfailing good humor will be greatly missed by all who knew him.

Introduction

The Supplement to the President's FY 2005 Budget reports on the FY 2004 research and development (R&D) activities and FY 2005 plans of the multiagency Networking and Information Technology Research and Development (NITRD) Program. A collaborative effort of many Federal agencies (listed on pages 2-3), the NITRD Program is the Nation's principal source of long-term, fundamental information technology (IT) R&D, including advanced technologies in high-end computing systems and software, high-speed networking, software assurance and reliability, human-computer interaction, and information management, as well as research in the socioeconomic and workforce development implications of these new technologies.

Each year, the NITRD Supplement to the President's Budget, also known as "the Blue Book," seeks to illuminate the breadth of the NITRD portfolio and the impact of NITRD research advances on U.S. leadership in national defense and national security, cutting-edge science and technology, and economic prosperity, and on improving the quality of life for all Americans.

This year's Blue Book highlights the technical domains, called Program Component Areas (PCAs), in which the NITRD agencies conduct IT research and collaborate to achieve common goals. The report, based on information provided by the agencies, is structured to serve as a detailed guide to the program, including both collaborative and agency-by-agency activities in FY 2004 and plans for FY 2005. The document begins with an overview of the NITRD Program, followed by sections on each NITRD PCA. The NITRD budget request for FY 2005, by agency and by PCA, appears on page 83, along with FY 2004 estimates.

The NITRD Agencies

National Science Foundation (NSF)

National Institutes of Health (NIH)

Department of Energy/Office of Science (DOE/SC)

National Aeronautics and Space Administration (NASA)

Defense Advanced Research Projects Agency (DARPA)

National Security Agency (NSA)

Agency for Healthcare Research and Quality (AHRQ)

National Institute of Standards and Technology (NIST)

National Oceanic and Atmospheric Administration (NOAA)

Environmental Protection Agency (EPA)

Department of Energy/National Nuclear Security Administration (DOE/NNSA)

NITRD Research Interests of the Agencies

NSF – supports basic research in all NITRD areas, incorporates IT advances in science and engineering applications, supports computing and networking infrastructure for research, and educates world-class scientists, engineers, and IT workforce professionals.

NIH – is applying the power of computing, both to manage and analyze biomedical data and to model biological processes, in its goal to develop the basic knowledge for the understanding, diagnosis, treatment, and prevention of human disease.

DOE/SC – is exploring, developing, and deploying computational and networking tools that enable researchers in the scientific disciplines to model, simulate, analyze, and predict complex physical, chemical, and biological phenomena important to DOE. The Office also provides support for the geographically distributed research teams and remote users of experimental facilities whose work is critical to DOE missions. FY 2004 is the fourth year of the Office's Scientific Discovery through Advanced Computing (SciDAC) initiative, which is focused on the next generation of scientific simulation and collaboration tools for the scientific areas that are the focus of DOE research.

NASA – is extending U.S. technological leadership to benefit the U.S. aeronautics, Earth and space science, and spaceborne research communities.

DARPA – is focused on future-generations computing, communications, and networking as well as embedded software and control technologies, and human use of information technologies in national defense applications such as battlefield awareness.

NSA – is addressing some of the most challenging problems in the country in computing, storage, communications, networking, and information assurance in order to help ensure our national security.

AHRQ – focuses on research into state-of-the-art IT for use in health care applications such as computer-based patient records, clinical decision support systems, and standards for patient care data, information access, and telehealth.

NIST – is working with industry and with educational and government organizations to make IT systems more useable, secure, scalable, and interoperable; to apply IT in specialized areas such as manufacturing and biotechnology; and to encourage private-sector companies to accelerate development of IT innovations. It also conducts fundamental research that facilitates measurement, testing, and the adoption of industry standards.

NOAA – is an early adopter of emerging computing technologies for improved climate modeling and weather forecasting, and of emerging communications technologies for disseminating weather forecasts, warnings, and environmental information to users such as policymakers, emergency managers, industry, and the general public.

EPA – has the IT research goal of facilitating multidisciplinary ecosystem modeling, risk assessment, and environmental decision making at the Federal, state, and local levels, and by other interested parties, through advanced use of computing and other information technologies.

DOE/NNSA – is developing new means of assessing the performance of nuclear weapon systems, predicting their safety and reliability, and certifying their functionality through high-fidelity computer models and simulations.

Other Participating Agencies

Other Federal agencies participate in networking and information technology research and development, and/or coordinate with NITRD activities, using funds that are not budgeted under the program. These agencies include:

- Air Force Research Laboratory (AFRL)
- Department of Defense/High Performance Computing Modernization Program Office (DoD/HPCMPO)
- Federal Aviation Administration (FAA)
- Food and Drug Administration (FDA)
- General Services Administration (GSA)
- Office of Naval Research (ONR)

The NITRD Program Today

One of the few formal multiagency enterprises in the Federal government, the Networking and Information Technology Research and Development (NITRD) Program was chartered by Congress under the High-Performance Computing Act of 1991 (P.L. 102-194) and the Next Generation Internet Act of 1998 (P.L. 105-305) to improve coordination and cooperation among agencies engaged in IT R&D.

Over the past 13 years, the scope of the agencies' collaborative activities has evolved and expanded to encompass emerging technological fields – such as wireless and optical networking, cybersecurity, high-assurance software and systems, and embedded systems – that did not yet exist when the program began. The successful NITRD collaboration has come to be viewed as a model Federal research effort that leverages agency strengths, avoids duplication, and fosters interoperable systems that maximize the benefits of Federal IT R&D investments to both agency missions and private-sector innovation.

Today, the NITRD Program remains the Nation's primary source of not only fundamental technological breakthroughs but also highly skilled human resources in the advanced computing, networking, software, and information management technologies that underpin our 21st century infrastructure and quality of life. NITRD's broad impact derives in part from its highly diversified and multidisciplinary research strategy, which funds fundamental scientific investigations across Federal laboratories and centers, research universities, nonprofit organizations, and partnerships with industry.

Goals of the NITRD Program

The NITRD Program provides agencies that perform IT R&D with a framework to plan, budget, coordinate, implement, and assess their research agendas. The program is an R&D priority of the Administration that is a distinct feature in the President's annual budget. NITRD's signal role in supporting the national interest is reflected in the program's goals, which are to:

- Assure continued U.S. leadership in computing, networking, and information technologies to meet Federal goals and to support U.S. 21st century academic, industrial, and government interests
- Accelerate deployment of advanced and experimental information technologies to maintain world leadership in science, engineering, and mathematics; improve the quality of life; promote long-term economic growth; increase lifelong learning; protect the environment; harness IT; and enhance national security
- Advance U.S. productivity and industrial competitiveness through long-term scientific and engineering research in computing, networking, and related information technologies

Management and Structure

The NITRD Program functions under the aegis of the National Science and Technology Council (NSTC). Overall program coordination at the operational level is provided by the Interagency Working Group on IT R&D (IWG/IT R&D), made up of senior managers from each of the NITRD agencies and representatives of the Office of Science and Technology Policy (OSTP) and the Office of Management and Budget (OMB). (See roster on page 81.) The IWG, which is co-chaired by an agency representative (currently, NSF's Assistant Director for the Computer and Information Science and Engineering Directorate) and the Director of the National Coordination Office (NCO) for IT R&D, meets quarterly to coordinate NITRD policies, programs, and budget planning.

Program Component Areas (PCAs)

At the core of the NITRD enterprise are seven research domains called Program Component Areas (PCAs). The PCAs, which have evolved over time as the ambit of information technology has expanded, encompass

the principal areas of mission-related IT research engaged in by the NITRD agencies. The PCAs are:

- High-end Computing (HEC), with two PCAs:
  - HEC Infrastructure and Applications (I&A)
  - HEC Research and Development (R&D)
- Human Computer Interaction and Information Management (HCI&IM)
- Large Scale Networking (LSN)
- Software Design and Productivity (SDP)
- High Confidence Software and Systems (HCSS)
- Social, Economic, and Workforce Implications of IT and IT Workforce Development (SEW)

In each PCA, program managers from NITRD agencies with R&D in that area and from other agencies interested in the topic participate in a Coordinating Group (CG). The CGs meet monthly to share information, develop collaborative activities, and coordinate their PCA's overall research agenda. The CGs report to the IWG, serving as the link between the coordinating body and research activities at the agency program level. CG chairs are listed on page 81.

National Coordination Office for IT R&D

The NCO provides technical and administrative support for the interagency NITRD Program, including extensive activities on behalf of the IWG and planning, budget, and assessment tasks for the program as a whole. The NCO also supports the President's Information Technology Advisory Committee (PITAC), an external, non-government panel appointed by the President to provide independent reviews and guidance on IT R&D-related topics (story on page 7). The NCO supported the work of the High End Computing Revitalization Task Force and is helping coordinate implementation of the HECRTF plan (details on pages 14-16). The NSF serves as the NCO's host agency. Information about NITRD and PITAC, and copies of NITRD documents, are posted on the NCO web site: *www.nitrd.gov*.

IWG: Coordination and Activities

IWG Highlights

In FY 2004, the IWG is leading several major coordination initiatives in addition to its regular program oversight activities. As part of its ongoing responsibility to review the NITRD structure periodically and evaluate the need for program updating, the working group:

- Charged the NITRD agencies to report on their FY 2004 program activities, highlighting interagency collaborations and plans, and to present their views on research gaps and/or issues in the NITRD structure at Special Meetings of all the CGs
- Charged each CG to review the results of its Special Meeting and, using those results, to develop a definition of the PCA that reflects participating agencies' current research investments and interests
- Established a PCA/CG Task Group to coordinate review of the new PCA definitions, evaluate the need for changes in the PCA/CG structure based on agency comments, and make recommendations to the IWG
- In coordination with OMB, authorized preparation of an Interagency Coordination Report (ICR) that would provide technical details of NITRD Program collaborative activities on an annual basis

Also in FY 2004, the IWG Issued "Grand Challenges: Science, Engineering, and Societal Advances Requiring Networking and Information Technology Research and Development," a report by a multiagency NITRD task force that spent a year on the effort (story on page 6).

Research Linkages

The PCA/CG Task Group, based on its review of agency perspectives, identified two areas – cybersecurity issues and interoperability standards – that cross PCA boundaries in the current coordination structure. The task group recommended, and the IWG approved, setting up "temporary linkage groups" enabling agencies across PCAs with an interest in these or future crosscutting subjects to meet and plan activities on an ad hoc basis.

PCA Definitions and FY 2004 Programs

This Supplement to the President's Budget provides the new one-page PCA definitions developed by the CGs as well as FY 2004 program activities reported by each agency at the Special Meetings. The document also summarizes coordinated activities in each PCA reported by the CG and participating agencies. The IWG's full Interagency Coordination Report is available on the NCO web site, at *www.nitrd.gov/pubs/icr/*.

Grand Challenges: IT Goals in the National Interest

In FY 2003, the IWG established a Grand Challenges Task Force and charged it with identifying a set of science, engineering, and societal challenges that require innovations in IT R&D. The charge to formulate a new set of grand challenges was a direct response to update the list of grand challenges in IT called for in the HPC Act of 1991 and documented in the FY 1994 Blue Book. The Task Force comprised volunteers from ten NITRD agencies, FAA, the NCO, and OSTP. The group completed its work in FY 2004.

The Task Force began by revising the HPC Act's definition of "grand challenge," taking into account the impact of current technological advances. The new definition of a NITRD grand challenge is as follows:

> *A long-term science, engineering, or societal advance, whose realization requires innovative breakthroughs in information technology research and development and which will help address our country's priorities.*

Key criteria for defining the grand challenges included:

- Description of the multi-decade nature of the challenge
- Focus of the grand challenge in the next ten years
- Benefits of the grand challenge to the social, economic, political, scientific, and technological well-being of mankind
- Relationship of the grand challenge to vital national priorities. In consultation with OSTP, the Task Force identified six national priority areas that reflect the country's broad-based scientific, military, social, economic, and political values and goals. The national priorities identified are: Leadership in Science and Technology; Homeland and National Security; Health and Environment; Economic Prosperity; A Well-Educated Populace; and A Vibrant Civil Society.

The Task Force formulated 16 new illustrative grand challenges (below) mapped to these criteria and designed to stimulate multidisciplinary thinking. The goal was to create a set of visionary research attainments that would test the intellectual aspirations of the Nation's researchers, asking them to reach for possibilities beyond their understanding today or in the next decade.

The Task Force then considered what fundamental IT research will be required to enable realization of each of the 16 grand challenge goals. Out of this discussion came a list of 14 "IT hard problem areas" – broad topical categories of research in which solutions or advances are required to achieve progress toward the grand challenge goals. Each IT hard problem area was mapped to the relevant challenges. The resulting Task Force document is the first NITRD report of its kind. It is available on the NCO web site, at: *www.nitrd.gov/GCs/*.

16 Illustrative NITRD Grand Challenges

- **Knowledge Environments for Science and Engineering**
- **Clean Energy Production Through Improved Combustion**
- **High Confidence Infrastructure Control Systems**
- **Improved Patient Safety and Health Quality**
- **Informed Strategic Planning for Long-Term Regional Climate Change**
- **Nanoscale Science and Technology: Explore and Exploit the Behavior of Ensembles of Atoms and Molecules**
- **Predicting Pathways and Health Effects of Pollutants**
- **Real-Time Detection, Assessment, and Response to Natural or Man-Made Threats**
- **Safer, More Secure, More Efficient, Higher-Capacity, Multi-Modal Transportation System**
- **Anticipate Consequences of Universal Participation in a Digital Society**
- **Collaborative Intelligence: Integrating Humans with Intelligent Technologies**
- **Generating Insights From Information at Your Fingertips**
- **Managing Knowledge-Intensive Organizations in Dynamic Environments**
- **Rapidly Acquiring Proficiency in Natural Languages**
- **SimUniverse: Learning by Exploring**
- **Virtual Lifetime Tutor for All**

IT 'Hard Problem' Areas

- **Algorithms and Applications**
- **Complex Heterogeneous Systems**
- **Hardware Technologies**
- **High Confidence IT**
- **High-End Computing**
- **Human Augmentation IT**
- **Information Management**
- **Intelligent Systems**
- **IT System Design**
- **IT Usability**
- **IT Workforce**
- **Management of IT**
- **Networks**
- **Software Technologies**

President's Information technology Advisory Committee

PITAC Highlights

The President's Information Technology Advisory Committee (PITAC) is appointed by the President to provide independent expert advice on maintaining America's preeminence in advanced information technology. PITAC members are IT leaders in industry and academe with expertise relevant to critical elements of the national information infrastructure such as high-performance computing, large-scale networking, and high-assurance software and systems design. The Committee's studies help guide the Administration's efforts to accelerate the development and adoption of information technologies vital for American prosperity.

Chartered by Congress under the High-Performance Computing Act of 1991 and the Next Generation Internet Act of 1998, the PITAC is a Federal Advisory Committee formally renewed through Presidential Executive Orders. The current Committee held its first meeting November 12, 2003.

In June 2004, the PITAC delivered its first report – "Revolutionizing Health Care Through Information Technology" – to the President, following eight months of study by its Subcommittee on Health and IT. The report concludes that although the potential of IT to improve the delivery of care while reducing costs is enormous, concerted national leadership is essential to achieving this objective. The PITAC study focuses on specific barriers to the nationwide implementation of health IT that can only be addressed by the Federal government.

The report offers 12 specific recommendations for Federal research and actions to enable development of 21st century electronic medical records systems. At the core of such systems is the concept of a secure, patient-centered electronic health record (EHR) that: 1) safeguards personal privacy; 2) uses standardized medical terminology that can be correctly read by any care provider and incorporated into computerized tools to support medical decision making; 3) eliminates today's dangers of illegible handwriting and missing patient information; and 4) can be transferred as a patient's care requires over a secure communications infrastructure for electronic information exchange.

Two other PITAC subcommittees – on Cyber Security and Computational Science – have been charged by OSTP to review and evaluate Federal R&D investments in these areas to assess whether they are appropriately balanced between long-term and short-term efforts and effectively focused to achieve maximum benefits. Reports on these topics are to be presented to the Administration in FY 2005.

Members of PITAC

Co-Chairs

Marc R. Benioff
Chairman and CEO
Salesforce.com, Inc.

Edward D. Lazowska, Ph.D.
Bill and Melinda Gates Chair in Computer Science
Department of Computer Science and Engineering
University of Washington

Members

Ruzena Bajcsy, Ph.D.
Director, Center for Information Technology Research in the Interest of Society (CITRIS) and Professor
University of California, Berkeley

J. Carter Beese, Jr.
President
Riggs Capital Partners

Pedro Celis, Ph.D.
Software Architect
Microsoft Corporation

Patricia Thomas Evans
President and CEO
Global Systems Consulting Corporation

Manuel A. Fernandez
Managing Director
SI Ventures/Gartner

Luis E. Fiallo
President
Fiallo and Associates, LLC

José-Marie Griffiths, Ph.D.
Professor and Dean
School of Information and Library Science
University of North Carolina at Chapel Hill

William J. Hannigan
President
AT&T

Jonathan C. Javitt, M.D., M.P.H.
Senior Fellow
Potomac Institute for Policy Studies

Judith L. Klavans, Ph.D.
Director of Research, Center for the Advanced Study of Language, and Research Professor
College of Library and Information Science, University of Maryland

F. Thomson Leighton, Ph.D.
Chief Scientist
Akamai Technologies, and
Professor of Applied Mathematics
Massachusetts Institute of Technology

Harold Mortazavian, Ph.D.
President and CEO
Advanced Scientific Research, Inc.

Randall D. Mott
Senior Vice President and CIO
Dell Computer Corporation

Peter M. Neupert
Chairman of the Board
Drugstore.com, Inc.

Eli M. Noam, Ph.D.
Professor and Director of the Columbia Institute for Tele-Information
Columbia University

David A. Patterson, Ph.D.
Professor and E.H. and M.E. Pardee Chair of Computer Science
University of California, Berkeley

Alice G. Quintanilla
President and CEO
Information Assets Management,Inc.

Daniel A. Reed, Ph.D.
Chancellor's Eminent Professor, Vice Chancellor for Information Technology and CIO, and Director, Institute for Renaissance Computing
Department of Computer Science
University of North Carolina at Chapel Hill

Eugene H. Spafford, Ph.D.
Professor and Director, Center for Education and Research in Information Assurance and Security (CERIAS)
Purdue University

David H. Staelin, Sc.D.
Professor of Electrical Engineering
Massachusetts Institute of Technology

Peter S. Tippett, M.D., Ph.D.
CTO and Vice-Chairman
TruSecure Corporation

Geoffrey Yang
Managing Director
Redpoint Ventures

High-End Computing

Infrastructure and Applications

Definition of HEC I&A PCA

The facilities and activities funded under the NITRD Program's HEC I&A PCA include R&D infrastructure and research and development to extend the state of the art in computing systems, science and engineering applications, and data management, keeping the U.S. at the forefront of 21st century science and engineering discoveries. HEC researchers develop, deploy, and apply the most advanced hardware, systems, and applications software to model and simulate objects and processes in biology, chemistry, environmental sciences, materials science, nanoscale science and technology, and physics; address complex and computationally intensive national security applications; and perform large-scale data fusion and knowledge engineering. For scientific researchers in every field, these advanced computing capabilities have become a prerequisite for discovery.

The R&D that produces these capabilities requires collaborations across Federal and academic institutions, industry, and the international research community. Interdisciplinary teams of scientists, engineers, and software specialists design and maintain the large, complex body of applications software. The largest and fastest computational platforms available are required because of the great range of space scales (from sub-atomic to supernova) and time scales (such as nanosecond to multi-century climate shifts) in the models and simulations. Modeling and simulation produce vast amounts of data that require leading-edge storage and visualization technologies.

Even with skilled teams and leading-edge technologies, however, today's HEC systems remain for the most part fragile and difficult to use. Specialized systems and applications software are needed to distribute calculations across hundreds or thousands of processors in a variety of massively parallel systems. Computational scientists are faced with a proliferation of architectures and variety of programming paradigms, resulting in a multitude of questions that must be addressed and tasks that must be performed in order to implement a modeling or simulation algorithm on any specific system architecture. But while this work progresses, advances in the size and speed of computing systems open up opportunities to increase the size, scale, complexity, and even nature of the modeling and simulation problems that can be addressed. The result is that a new cycle of systems and applications software R&D is required to enable scientists to take advantage of the increased computing power.

To maintain or accelerate the pace of scientific discovery, HEC I&A efforts are needed to develop breakthroughs in algorithm R&D, advances in systems software, improved programming environments, and computing infrastructure for development of the next-generation applications that will serve Federal agency missions. Focus areas include, but are not limited to, cyberinfrastructure tools to facilitate high-end computation, storage, and visualization of large data sets encountered in the biomedical sciences, climate modeling and weather forecasting, crisis management, computational aerosciences, Earth and space sciences, and a wide range of other human activities.

Broad Areas of HEC I&A Concern

- Algorithm R&D
- Data management and understanding

HEC I&A Agencies

			Participating Agency
NSF	NASA	NIST	
NIH	DOE/NNSA	EPA	
DOE/SC	NOAA		DoD/HPCMPO

HEC I&A PCA Budget Crosscut

FY 2004 Estimate	FY 2005 Request
$516.2 M	$505.7 M

- Programming environments
- Scientific applications
- Computational facilities

Technical Goals

- Understand the interaction between applications and architectures
- Provide mathematical and computational methods needed for critical mission agency applications
- Provide technology base for next-generation data management and visualization
- Enable new generations of mission agency computational applications
- Provide measures of progress
- Reduce time and cost in HEC procurements
- Reduce cost of ownership of HEC systems

Illustrative Technical Thrusts

- Scientific discovery through advanced computing
- Cyberinfrastructure
- Leadership computing
- System monitoring and evaluation
- Integrated end-to-end data management
- Visualization clusters
- Common procurement methodology

Benefits of HEC Applications to U.S. Science and Engineering

Application Area	Science Challenge	Potential at 100-1,000 Times Current HEC Capability
Astrophysics	Simulation of astrophysical environments such as stellar interiors and supernovae.	Yield understanding of the conditions leading to the origin of the heavy elements in the universe.
Nuclear Physics	Realistic simulations of the characteristics of the quark-gluon plasma.	By developing a quantitative understanding of the behavior of this new phase of nuclear matter, facilitate its experimental discovery in heavy ion collisions.
Catalyst Science/ Nanoscale Science and Technology	Calculations of homogeneous and heterogeneous catalyst models in solution.	Reduce energy costs and emissions associated with chemicals manufacturing and processing. Meet Federally mandated NOx levels in automotive emissions.
Nanoscale Science and Technology	Simulate and predict mechanical and magnetic properties of simple nanostructured materials.	Enable the discovery and design of new advanced materials for a wide variety of applications impacting many industries.
Simulation of Aerospace Vehicle in Flight	Simulate a full aerospace vehicle mission, such as a full aircraft in maneuver or an RLV in ascent or descent.	Reduce aerospace vehicle development time and improve performance, safety, and reliability.
Structural and Systems Biology	Simulations of enzyme catalysis, protein folding, and transport of ions through cell membranes.	Provide ability to discover, design, and test pharmaceuticals for specific targets and to design and produce hydrogen and other energy feedstock more efficiently.
Signal Transduction Pathways	Develop atomic-level computational models and simulations of complex biomolecules to explain and predict cell signal pathways and their disrupters.	Yield understanding of initiation of cancer and other diseases and treatments on a molecular level; prediction of changes in ability of microorganisms to influence natural biogeochemical cycles such as carbon cycling and global change.
Signal & Image Processing & Automatic Target Recognition	Replace electromagnetic scattering field tests of actual targets with numerical simulations of virtual targets.	Design more stealthy aircraft, ships, and ground systems and create the ability to rapidly model new targets, enabling more rapid adaptation of fielded weapon systems' ability to target new enemy weapon systems.
Climate Science	Resolve additional physical processes such as ocean eddies, land use patterns, and clouds in climate and weather prediction models.	Provide U.S. policymakers with leading-edge scientific data to support policy decisions. Improve understanding of climate change mechanisms and reduce uncertainty in the projections of climate change.
Solid Earth Science	Improved statistical forecasts of earthquake hazards (fault-rupture probabilities and ground motion).	Provide prioritized retrofit strategies. Reduced loss of life and property. Damage mitigation.
Magnetic Fusion Energy	Optimize balance between self-heating of plasma and heat leakage caused by electromagnetic turbulence.	Support U.S. decisions about future international fusion collaborations. Integrated simulations of burning plasma crucial for quantifying prospects for commercial fusion.
Combustion Science	Understand interactions between combustion and turbulent fluctuations in burning fluid.	Understand detonation dynamics (for example, engine knock) in combustion systems. Solve the "soot" problem in diesel engines.

Excerpt from a table in the "Federal Plan for High-End Computing" illustrating the critical role of HEC I&A in future scientific advances.

High-End Computing

Research and Development

Definition of HEC R&D PCA

The activities funded under the HEC R&D PCA include research and development to optimize the performance of today's high-end computing systems and to develop future generations of high-end computing systems. These systems are critical to addressing Federal agency mission needs and in turn many of society's most challenging large-scale computational problems.

Most high-end systems are clusters of processor nodes developed for the commercial computing market for use in personal computers and Web servers, and are not specifically targeted for HEC computing. Although the "peak" performance of these processors has followed Moore's Law, increasing more than five orders of magnitude in speed in the last decade, the sustained performance of scientific applications measured as a fraction of the peak value has continued to decline. Two reasons for this decline are: (1) these HEC systems are hundreds of times larger than those used in commercial applications and require a highly specialized software infrastructure to use effectively, and (2) the current HEC systems are unbalanced in that they do not have the optimal ratios of system parameters such as computation rate versus memory bandwidth.

To remedy this situation, HEC R&D supports both hardware and software R&D specifically aimed at scientific and national security applications. HEC R&D focuses on teraflop- through petaflop-scale systems and computation. Research activities in this area seek fundamental, long-term advances in technologies to maintain and extend the U.S. lead in computing for generations to come. Current research focuses on advanced computing architectures, software technologies and tools for high-end computing, mass storage technologies, and molecular, nano-, optical, quantum, and superconducting technologies.

HEC R&D engages collaborative research teams from academic institutions, national laboratories, and industrial partners in the development of new architectures that are well suited for algorithms used in scientific applications. HEC R&D supports fundamental investigations in memory, interconnect, and storage technologies in order to improve system balance. HEC R&D involves research across the entire spectrum of software issues – operating systems, languages, compilers, libraries, development environments, and algorithms – necessary to allow scientific applications to use the hardware effectively and efficiently. In addition, HEC R&D supports system modeling and performance analysis, which enable researchers to better understand the interaction between the computational requirements of applications and the performance characteristics of any proposed new high-end architectures.

Broad Areas of HEC R&D Concern

- Hardware architecture, memory, and interconnects
- Power, cooling, and packaging
- I/O and storage
- Comprehensive system software environment
- Programming models and languages
- System modeling and performance analysis
- Reliability, availability, serviceability, and security

HEC R&D Agencies

NSF	DOE/NNSA	NOAA	Participating Agency
DOE/SC	NSA	NASA	
DARPA	NIH	NIST	DoD/HPCMPO

HEC R&D PCA Budget Crosscut

FY 2004 Estimate	FY 2005 Request
$355.1 M	$317.5 M

Technical goals

- Parallel architecture designed for scientific application requirements
- Parallel I/O and file systems for sustained high data throughput
- Systems scalable to hundreds of thousands of processors
- Reliable and fault-tolerant systems
- Improved programming model expressability and parallel compiler support
- Effective performance measurement, optimization tools
- Improved ease of use and time to solution

Illustrative technical Thrusts

- Parallel component architecture
- Parallel I/O and high-performance file systems
- Next-generation architecture
- Extreme-scale operating and runtime systems
- Global address space language
- Software infrastructure tools
- Development and runtime performance tools
- High-productivity computer systems
- Quantum computing

Roadmaps for Federal HEC R&D in the Years Ahead

Hardware

Component	Near- to Mid-Term	Long-Term
Microarchitecture	Prototype microprocessors developed for HEC systems available	Innovative post-silicon technology optimized for HEC
Interconnect technologies	Interconnect technology based upon optical interconnect and electrical switches	All-optical interconnect technology for HEC
Memory	Memory systems developed for HEC needs. Accelerated introduction of PIMs	Revolutionary high-bandwidth memory at petaflop scale
Power, Cooling, and Packaging	Stacked 3-D memory and advanced cooling technologies address critical design limitations	Ability to address high-density packaging throughout the entire system
I/O and storage	Petaflop-scale file systems with RAS focused on HEC requirements	Revolutionary approaches to exascale "file systems"

Software

Component	Near- to Mid-Term	Long-Term
Operating systems (OSs)	New research-quality HEC OSs that address scalability and reliability	Production-quality, fault-tolerant, scalable OSs
Languages, compilers, and libraries	Optimized for ease of development on selected HEC systems. Research-quality implementations of new HEC languages, supporting multiple levels of abstraction for optimization.	High-level algorithm-aware languages and compilers for automated portability across all classes of HEC systems
Software tools and development environments	Interoperable tools with improved ease of use across a wide range of systems. First research-quality IDEs available for HEC systems.	IDEs that support seamless transition from desktop to largest HEC systems
Algorithms	New multiscale algorithms suitable for HEC systems. Initial prototypes of architecture-independent parallel computations.	Automatic parallelization of algorithms for irregular and unbalanced scientific problems. Scaling up of parallel algorithms to enable detailed simulations of physical systems.

Systems

Component	Near- to Mid-Term	Long-Term
System architecture	Three or more systems capable of sustained petaflops (up to 100,000 processors or more) on wider range of applications. Programming much simpler at large scale. Emergence of adaptable self-tuning systems.	High-end systems capable of sustained 10 to 100 petaflops on majority of applications. Programmable by majority of scientists and engineers. Adaptable self-tuning systems commonplace.
System modeling and performance analysis	Accurate models/tools for HEC systems and applications. Tools and benchmarks provide better understanding of architecture/application interactions.	Models enable analysis, prediction of software behavior. Automated, intelligent performance and analysis tools and benchmarks widely available, easy to use.
Programming models	Research implementations of novel parallel computing models. "Non-heroic" programming: MPI follow-on for 1,,024 processors and robust DSM implementations (UPC, CAF, …) widespread and available for 1,024 processors.	Parallel computing models that effectively and efficiently match new or planned architectures with applications. Novel parallel computation paradigms foster new architectures and new programming language features.
Reliability, availability, and serviceability (RAS) + Security	Semi-automatic ability to run through faults. Enhanced prevention of intrusion and insider attack.	Self-awareness: reliability no longer requires user assistance. Systems will have verifiable multilevel secure environments.

"Federal Plan for High-End Computing" roadmaps outline mid-term and long-term R&D advances required to revitalize U.S. HEC leadership.

HEC PCAs: Coordination and Activities

HEC Highlights

The HEC Coordinating Group (CG) coordinates activities funded under both the HEC I&A and HEC R&D PCAs. The CG provides a forum in which Federal agency program managers coordinate and collaborate on HEC research programs and on implementation of Federal high-end computing activities. The HEC CG is charged with:

- Encouraging and facilitating interagency coordination and collaborations in HEC I&A and R&D programs
- Addressing requirements for high-end computing technology, software, infrastructure, and management by fostering Federal R&D efforts
- Providing mechanisms for cooperation in HEC R&D and user access among Federal agencies, government laboratories, academia, industry, application researchers, and others

The HEC CG held a Special Meeting on March 19, 2004, with presentations and program descriptions from ten Federal agencies: DARPA, DoD/HPCMPO, DOE/NNSA, DOE/SC, EPA, NASA, NIST, NOAA, NSA, and NSF. The meeting enabled the agencies to identify opportunities for collaboration and coordination, areas for HEC CG focus, and areas – such as open source software – whose scope is broader than the HEC CG's.

Multiagency Interests and Collaborations

The HEC agencies reported the following common interests, collaborations, and commitments across a wide range of topics, including:

- Acquisition coordination: DoD/HPCMPO, DOE/NNSA, NASA
- Air-quality modeling applications: DOE/SC, EPA, NOAA
- Applied research for end-to-end systems development: NASA, NOAA, NSF
- Benchmarking and performance modeling: DoD/HPCMPO, DOE/SC, NASA, NSA, NSF
- Connectivity and technology delivery to universities: NOAA, NSF
- Climate and weather applications: DOE/SC, NASA, NOAA, NSF
- Earth System Modeling Framework (ESMF): DOE/SC, NASA, NOAA, NSF/NCAR (details on next page)
- Grid demonstrations: DOE/SC, NOAA, NSF
- Hardware development: DARPA, DOE/NNSA, NSA
- HECRTF: DoD/HPCMPO, DoD/OSD, DOE/NNSA, DOE/SC, EPA, NASA, NIST, NOAA, NSA, NSF
- HEC-URA: DARPA, DOE/SC, NSA, NSF (details on next page)
- HPCS Phase II: DARPA, DOE/NNSA, DOE/SC, NASA, NSA, NSF
- HPCS productivity metrics: DARPA, DoD/HPCMPO, DOE/NNSA, DOE/SC, NASA, NSA, NSF
- Joint memo expecting open source software and Service Oriented Data Access (SODA) for work funded at DOE labs: DOE/NNSA, DOE/SC
- MOU among HPCS mission partners for joint planning, coordination of directions, and leveraging each other's activities: DARPA, DoD/OSD, DOE/NNSA, DOE/SC, NSA
- National Research Council study on the "Future of Supercomputing": DOE/SC, DOE/NNSA (report expected in 2004)
- Optical switches and interconnects: DARPA, DOE/NNSA, NSA
- Quantum information science: DARPA, NIST, NSA
- Reviews of SV2 [procurements?]: DOE/NNSA, NSA [and DOE/SC?]
- Quarterly reviews of Cray X1e/Black Widow R&D programs: DoD/HPCMPO, DOE/NNSA, DOE/SC, NRO, NASA, NSA
- Reviews of ASC White and ASC Q software environments: DOE/NNSA, DOE/SC
- Single photon sources: DARPA, NIST
- Spray cooling: DOE/NNSA, DOE/SC
- Standards: DoD/HPCMPO, NIST, NOAA, NSA
- Technology transfer from universities: DoD/HPCMPO, DOE/NNSA
- Testbeds: DoD/HPCMPO, DARPA, NIST, NOAA, NSA
- Unified Parallel Compiler (UPC): DOE/SC, NSA
- Weather research and forecast (WRF): NOAA, NSF/NCAR

Multiagency ESMF Effort

Another broad-based FY 2004 NITRD activity that involves collaboration among multiple Federal agencies and coordination across multiple CGs, including HEC and SDP, is the building of the Earth System Modeling Framework (ESMF). Initiated by NASA researchers to develop tools enabling broader scientific use of the agency's extraordinary Earth data collections, the effort has become a key R&D component of Federal climate research across many organizations.

The ESMF is a high-performance, flexible software infrastructure designed to increase ease of use, performance portability, interoperability, and reuse in climate, numerical weather prediction, data assimilation, and other Earth science applications. The ESMF is an architecture for composing multi-component applications and includes data structures and utilities for developing model components. The goal is to create a framework usable by individual researchers as well as major operational and research centers, and to engage the weather research community in its development.

The ESMF addresses the challenge of building increasingly interdisciplinary Earth systems models and the need to maximize the performance of the models on a variety of computer architectures, especially those using upwards of thousands of processors. The new structure allows physical, chemical, and biological scientists to focus on implementing their specific model components. Software engineers design and implement the associated infrastructure and superstructure, allowing for a seamless linkage of the various scientific components.

The following agency organizations collaborate:

- NSF-supported National Center for Atmospheric Research (NCAR)
- NOAA's Geophysical Fluid Dynamics Laboratory (GFDL)
- NOAA's National Center for Environmental Prediction (NCEP)
- DoD's Air Force Weather Agency (AFWA)
- DoD's High Performance Computing Modernization Program Office (HPCMPO)
- DoD's Naval Research Laboratory (NRL)
- DOE's Argonne National Laboratory (ANL) and Los Alamos National Laboratory (LANL)
- NASA's Goddard Space Flight Center (GSFC)
- NASA Goddard Global Modeling and Assimilation Office (GGMAO)
- NASA Goddard Institute for Space Studies (GISS)
- NASA Goddard Land Information Systems project (GLIS)

ESMF Version 2.0, the first version of the software usable in real applications, was released in June 2004. It includes software for (1) setting up hierarchical applications, (2) representing and manipulating modeling components, fields, bundles of fields, and grids, and (3) services such as time management and logging messages.

New HEC-URA

DARPA, DOE/SC, NSA, and NSF collaboratively developed the HEC-University Research Activity (HEC-URA) launched in FY 2004 as an outgrowth of the HECRTF effort. The initiative funds universities to conduct research in software specifically for high-end computing. The strategy is to invest in:

- Basic and applied research to refill the academic pipeline with new ideas and people
- Advanced development of component and subsystem technologies
- Engineering and prototype development
  - Integration at the system level
  - Development of new technologies
- Test and evaluation to reduce the risk for development, engineering, and government procurement

Research emphases include DOE/SC and DARPA efforts on runtime operating systems for extreme-scale scientific computations and the NSF and DARPA focus on languages, compilers, and libraries for high-end computing. The agencies will participate in annual topic selections, solicitations, peer review selections, PI meetings, and the overall execution of this activity. Selections will be in accordance with individual agency requirements. Both university and industry researchers will be invited to make proposals.

In its first year, the HEC-URA is focused on software R&D to build a critical mass in research teams, advance the field toward 2010/2015 HEC software, avoid duplication, share lessons learned, and develop links between basic research and advanced development and engineering.

High-End Computing Revitalization Task Force (HECRTF)

In FY 2004, the NITRD agencies and others are engaged in a major planning initiative to reinvigorate Federal leadership in the development of leading-edge supercomputing capabilities to serve the national interest.

The activity was launched in April 2003 with a charge from OSTP to all government agencies with a major stake in high-end computing. The charge directed the agencies, under the auspices of the NSTC, to develop "a plan to guide future Federal investments in high-end computing" that would lay out an overall investment strategy based on "the needs of important Federally funded applications of HEC." The plan was to include five-year roadmaps for investment in R&D in core HEC technologies, as well as approaches to addressing Federal HEC capability, capacity, and accessibility issues, and Federal HEC procurement.

The 63-member High-End Computing Revitalization Task Force (HECRTF), representing 12 departments and agencies, spent 15 months gathering information, soliciting expert opinion from scientists, IT researchers, industry, and Federal HEC users, and developing a comprehensive report addressing the OSTP charge.

On May 13, 2004, OSTP Director John H. Marburger released the "Federal Plan for High-End Computing" at a hearing of the House Science Committee on H.R. 4218, "the High Performance Computing Revitalization Act of 2004." In his testimony, Dr. Marburger commended the task force and called its report "an important first step" in the "sustained and coordinated effort" required to address the issues facing the Nation's high-end computing enterprise. He noted that results of the HECRTF effort were already beginning to emerge – for example, the new High-End Computing University Research Activity (HEC-URA) jointly developed by DARPA, DOE/SC, NSA, and NSF to fund academic investigations in key HEC areas. FY 2004 solicitations focus on HEC software tools and operating/runtime systems for extreme-scale scientific computation.

Beginning in June 2004, the HECRTF Implementation Committee, which was created to activate the Federal Plan's recommendations, began a series of activities that include: 1) a meeting to coordinate FY 2006 agency HEC budget requests; 2) a meeting to plan cooperative development and use of HEC benchmarks and metrics; and 3) planning for workshops on testbeds for computer science research and on file systems.

The Need for HEC Revitalization: Excerpts from Report

The full "Federal Plan for High-End Computing" is available online at: http://www.nitrd.gov/pubs/2004_hecrtf/20040510_hecrtf.pdf

The United States has repeatedly demonstrated that leadership in science and technology is vital to leadership in national defense and national security, economic prosperity, and our overall standard of living. Today, progress in many branches of science and technology is highly dependent on breakthroughs made possible by high-end computing, and it follows that leadership in high-end computing is increasingly crucial to the nation.

***

In the past decade, computer modeling and simulation of physical phenomena and engineered systems have become widely recognized as the cost-effective "third pillar" of science and technology – sharing equal billing with theory and experiment. Simulations are performed on computing platforms ranging from simple workstations to very large and powerful systems known as high-end computers. High-end computers enable investigations heretofore impossible, which in turn have enabled scientific and technological advances of vast breadth and depth. High-end computing (HEC) thus has become an indispensable tool for carrying out Federal agency missions in science and technology.

***

[Recent agency studies] have revealed that current high-end computing resources, architectures, and software tools and environments do not meet current needs. Of equal concern, they observe that sustained investigations of alternative high-end systems have largely stopped, curtailing the supply of ideas and experts needed to design and develop future generations of high-end computing systems.

***

Revitalization of high-end computing is needed to refill the research pipeline with new ideas and highly trained people, support the development of robust and innovative systems, and lower industry and end-user risk by undertaking the test and evaluation of HEC systems and software technologies. This revitalization must support advancement across the innovation spectrum – from basic and applied research, to advanced development, to engineering and prototyping, to test and evaluation. Such a comprehensive approach is essential to the establishment of a sustainable R&D process that encourages the generation of competing innovations from the basic research phase, the development of early prototype HEC systems, and their evaluation on mission-critical test applications.

Key Recommendations of the HECRTF Report

Overarching Goals

The HECRTF recommends that the Federal government and its private-sector partners carry out comprehensive, complementary, and synchronized actions over the next five years with these goals: 1) Make high-end computing easier and more productive to use; 2) Foster the development and innovation of new generations of HEC systems and technologies; 3) Effectively manage and coordinate Federal high-end computing; and 4) Make high-end computing readily available to Federal agencies that need it to fulfill their missions.

HEC R&D

To achieve goals 1 and 2, the plan's roadmaps propose an integrated, balanced, and robust program of basic and applied research, advanced development, engineering and prototype development, and test and evaluation activities in each of three core HEC technologies:

- **Hardware** (microarchitecture; memory; interconnect; power, cooling, and packaging; and I/O and storage)
- **Software** (operating systems; languages, compilers, and libraries; software tools and development environments; and algorithms)
- **Systems technology** (system architecture; system modeling and performance analysis; reliability, availability, serviceability, and security; and programming models)

As part of this research effort, the task force recommends that the government return to a strategy – the development of HEC systems specifically for HEC R&D – that had been used in the early 1990's. These "early access" systems, called Research and Evaluation Systems by the HECRTF, will enable testing of early prototypes and provide development platforms for new algorithms and computational techniques, software functionality and scalability testing, and system and application testbeds. "Evaluation projects that identify failed approaches save the government from acquiring systems that simply do not perform as expected," the report states. "They may also suggest more fruitful approaches to remove the sources of failure."

The HECRTF also recommends that its proposed R&D agenda be balanced and prioritized to provide a continuum of efforts across the four developmental stages "required to sustain a vigorous and healthy high-end computing environment." The table on page 16 provides an overview of the proposed R&D priorities over the next five years.

HEC Resources

To achieve goal 4, the plan offers strategies to address three distinct aspects of Federal HEC resources:

- **Access –** Federal agencies with limited or no HEC resources could partner with large agencies under mutually beneficial multiyear HEC utilization agreements. The plan proposes establishing pilot partnerships with industry to work on HEC problems, with metrics to evaluate the success of these arrangements.
- **Availability –** Agencies such as DoD, NSA, DOE/NNSA, DOE/SC, EPA, NASA, NOAA, and NSF all have programs that provide HEC resources to their respective user communities. The plan calls for an increase in the resources available to Federal mission agencies to acquire and operate HEC resources.
- **Leadership –** To provide world-leading high-end computing capabilities for U.S. scientific research, the plan proposes that the Government develop what the HECRTF calls Leadership Systems, HEC systems that will offer 100 times the computational capability of what is available in the marketplace. These very advanced resources would be made available for a limited number of cutting-edge scientific applications, enabling the U.S. to be "first to market" with new science and technology capabilities. The Leadership machines, managed as a shared Federal research resource, would not only support research but drive HEC commercial innovation.

HEC Procurement

A significant barrier to achieving all of the plan's goals is presented by the technical complexity and demands of Federal procurement processes, which burden both agencies and industry vendors. The plan proposes testing more cost-effective, efficient approaches in interagency pilot projects on:

- **Performance measurement, or benchmarking –** agencies with similar applications will develop a single suite of benchmarks for use in a pilot joint procurement process.
- **Accurate assessment of the total cost of ownership (TCO) –** a multiagency team will evaluate TCO (acquisition, maintenance, personnel, external networking, grid and distributed computing) for several similar systems and develop "best practices" for determining TCO for comparing possible procurements.
- **Procurement consolidation –** Using these new tools, agencies will test a combined procurement solicitation and evaluate its effectiveness.

Plan Implementation

The HECRTF offers one example of an interagency governance structure for the plan but notes that it is only a starting point for discussion.

HECRTF Plan's Proposed R&D Priorities to FY 2010

		Current Programs*	Increment compared to HEC R&D Current Program				
		FY2004 ($ in millions)	FY2006	FY2007	FY2008	FY2009	FY2010
Hardware	Basic and Applied Research	$5					
	Advanced Development	$5					
	Engineering and Prototypes	$0					
	Test and Evaluation	$2					
	TOTAL	$12					
Software	Basic and Applied Research	$33					
	Advanced Development	$21					
	Engineering and Prototypes	$15					
	Test and Evaluation	$2					
	Long-term Evolution and Support	$0					
	TOTAL	$71					
Systems	Basic and Applied Research	$4					
	Advanced Development	$40					
	Engineering and Prototypes	$1					
	Test and Evaluation	$30					
	TOTAL	$75					
TOTALS		$158[1]					

Modest redirection
Moderate funding increment
Modest funding increment
Robust funding increment

*Assumes no planning changes from FY 2004

This table provides an overview of the Federal HEC R&D investments recommended by the HECRTF to address the needs of Federally funded R&D over the next five years. The level of effort called for in the major HEC components – hardware, software, and systems – is shown for each of the following categories of R&D work, which the Task Force deemed vital to a robust HEC R&D program:

- **Basic and Applied Research:** Focus on the development of fundamental concepts in high-end computing, creating a continuous pipeline of new ideas and expertise.
- **Advanced Development:** Select and refine innovative technologies and architectures for potential integration into high-end systems.
- **Engineering and Prototype Development:** Perform integration and engineering required to build HEC systems and components.
- **Test and Evaluation:** Conduct testing and evaluation of HEC software and current and new generations of HEC systems at appropriate scale.

The HECRTF report notes that innovations in hardware and systems have a natural transition from research into industrial manufacturing but that software research and development will require a different strategy. Since major advances in HEC system software have generally occurred only when academe, industry, and government research laboratories have teamed to solve common problems (such as by developing the message passing standard, MPI, to resolve internodal communications issues), the Task Force concluded that the software revitalization strategy should also include significant government involvement to ensure long-term maintenance of critical HEC software infrastructure components. The table thus shows a Long-Term Evolution and Support category under software.

NOTE 1: The FY 2004 funding levels shown in the table represent aggregate investments by all the agencies that participated in developing the "Federal Plan for High-End Computing," not just those of NITRD agencies. In addition, the figures represent only HEC R&D activities as defined by the HECRTF plan. They do not include many related activities (such as networking, visualization, and application-specific software development) considered outside the scope of the plan.

Columns four through eight present investment levels recommended by the Task Force for FY 2006 through FY 2010 in each of the categories and activities. Modest funding increments above the FY 2004 level are shown in medium orange. Moderate funding increments above the FY 2004 level are shown in lined orange. Robust funding increments above the FY 2004 level are shown in solid orange. Modest funding redirections from the FY 2004 level are shown in pale orange.

HEC I&A and R&D Programs By Agency

Selected FY 2004 Activities and FY 2005 Plans

HEC **NSF** HEC

High-end computing capabilities are important to work supported by all the NSF research directorates and offices including: Biological Sciences (BIO); Computer and Information Science and Engineering (CISE); Education and Human Resources (EHR); Engineering (ENG); Geosciences (GEO); Mathematical and Physical Sciences (MPS); and Social, Behavioral and Economic Sciences (SBE); and the Office of Polar Programs (OPP).

Directorates have complementary HEC investments. For example: R&D on computer architecture, networking, software, and cyberinfrastructure is funded by CISE; HEC devices are funded by MPS and ENG; mathematical algorithms are funded MPS and CISE; computational algorithms and libraries are funded by CISE, with some funding from MPS; science and engineering applications are developed primarily with funding from MPS, ENG, GEO, and BIO.

NSF HEC R&D Programs

NSF's largest investments in HEC resources are in the CISE Directorate. All CISE Divisions are responsible for HEC activities, as follows:

Computing and Communication Foundations (CCF) Division – has responsibility for:

- Formal and Mathematical Foundations – in particular algorithmic and computational science
- Foundations of Computing Processes and Artifacts – in particular high-end software, architecture, and design
- Emerging Models for Technology and Computation – includes biologically motivated computing models, quantum computing and communication, and nanotechnology-based computing and communication systems

Computer and Network Systems (CNS) Division – supports programs in computer and network systems, in particular distributed systems and next-generation software systems.

Information and Intelligent Systems (IIS) Division – supports programs in data-driven science including bioinformatics, geoinformatics, cognitive neuroscience, and other areas.

Shared Cyberinfrastructure (SCI) Division – supports programs in:

- Infrastructure development – creating, testing, and hardening next-generation deployed systems
- Infrastructure deployment – planning, construction, commissioning, and operations

The following illustrate ongoing HEC R&D:

- Chip Multiprocessor Architectures: on-chip shared-cache architectures, arithmetic logic unit (ALU) networks, shared memory architectures, and networks
- Scalable Multiprocessing Systems: system architecture (communication substrate) and (heterogeneous) ensembles of chip multiprocessors
- System-on-a-Chip building blocks and integrated functionality
- Dynamic and Static (compiler) Optimizations: memory systems (SMT/CMP support latency reduction, prefetching, and speculative execution [load/value prediction]), and verification and runtime fault tolerance
- Networking and storage

NSF also funds the following speculative research that could provide breakthrough HEC technologies:

- Nanotechnology that promises unprecedented scale-up in computing and communication through research in nano-architecture and design methodologies
- Post-silicon computing: development and application of technologies such as quantum, DNA, chemical, and other computing concepts to provide radically new models of computation, algorithms, and programming techniques.

NSF Cyberinfrastructure Framework

A key NSF-wide interest in FY 2004 is development of an overarching IT framework to support the communities performing research and using computing resources. Components include hardware (distributed resources for computation, communication, and storage), shared software

cybertools (grid services and middleware, development tools, and libraries), domain-specific cybertools, and applications to make possible discovery and innovative education and training. The framework will consist of:

- Computation engines (supercomputers, clusters, workstations), including both capability and capacity hardware
- Mass storage (disk drives, tapes, and related technologies) with persistence
- Networking (including wireless, distributed, ubiquitous)
- Digital libraries and data bases
- Sensors and effectors
- Software (operating systems, middleware, domain-specific tools, and platforms for building applications)
- Services (education, training, consulting, user assistance)

R&D in every component of this infrastructure will enable HEC capabilities.

Support for the Supercomputing Needs of the Broad Scientific Community

NSF plays a unique role nationally in supporting supercomputing needs across the entire academic spectrum of scientific research and education. NSF's SCI Division provides support for and access to high-end computing infrastructure and research. NSF supports more than 22 high-performance computing systems that include both capacity and capability systems. In FY 2004, NSF supports supercomputing resources at the following national partnerships and leading-edge sites:

- The National Computational Science Alliance (the Alliance) led by the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Urbana-Champaign
- The National Partnership for Advanced Computational Infrastructure (NPACI) led by the San Diego Supercomputer Center (SDSC) at the University of California-San Diego
- The Pittsburgh Supercomputing Center (PSC) led by Carnegie Mellon University and the University of Pittsburgh, with its Terascale Computing System (TCS)

In FY 2003, NSF had 4,450 HEC users including 1,800 NPACI users, 1,200 NCSA users, and 1,450 PSC users – for an estimated total of 3,000 unique users across all the centers. These organizations participate in the following ongoing NSF infrastructure initiative:

Extensible Terascale Facility (ETF) – multiyear effort to build and deploy the world's largest, most comprehensive, distributed infrastructure (also known as the TeraGrid) for open scientific research. Four new TeraGrid sites, announced in September 2003, will add more scientific instruments, large datasets, computing power, and storage capacity to the system. The new sites are ORNL, Purdue University, Indiana University, and the University of Texas at Austin. They join NCSA, SDSC, ANL, CalTech, and PSC.

With the addition of the new sites, the TeraGrid will have more than 20 teraflops of computing power, be able to store and manage nearly one petabyte of data, and have high-resolution visualization environments and toolkits for grid computing. All components will be tightly integrated and connected through a network that operates at 40 gigabits per second.

NSF FY 2005 plans include the following:

HEC R&D:

- ***CISE emphasis area focused on high-end computing –*** *new initiative intended to stimulate research and education on a broad range of computing and communication topics, including architecture, design, software, and algorithms and systems for computational science and engineering. While the range of topics is broad, the research and education activity is to focus strictly on extreme-scale computing. One component of this emphasis area is:*
  - *The special NSF/DARPA activity "Software & Tools for High-End Computing" (ST-HEC), developed as a High End Computing University Research Activity (HEC-URA) – supporting high-end software tools for extreme-scale scientific computation, which are highly computation- and data-intensive and cannot be satisfied in today's typical cluster environment. The target hosts for these tools are high-end architectures, including systems comprising thousands to tens of thousands of processors.*
- ***Activities in planning or evaluation in all NSF directorates on cyberinfrastructure development –*** *will include research and education areas in computational science and engineering, such as high-performance computational algorithms and simulation-based engineering*

HEC I&A:

- *Expansion of the Extensible Terascale Facility (ETF) will continue, providing high-end computing and networking for colleges and universities*

- ***Cyber-Infrastructure Training, Education, Advancement, and Mentoring (CI-TEAM)*** *– new FY 2005 activity will support the preparation of a workforce with the appropriate knowledge and skills to inform the development, deployment, support, and widespread use of cyberinfrastructure. Goal is to support efforts that will prepare current and future generations of scientists, engineers, and educators to use, develop, and support cyberinfrastructure, and broaden participation of underrepresented groups and organizations in cyberinfrastructure activities.*
- ***Linked Environments for Atmospheric Discovery (LEAD)*** *– will create grid-computing environments for on-demand prediction of high-impact local weather such as thunderstorms and lake-effect snows.*
- ***International Virtual Data Grid Laboratory (iVDGL)*** *– collaboration of 16 institutions and more than 80 scientists is deploying a computational and data grid for four international high-energy physics experiments, including the Large Hadron Collider.*

HEC DOE/SC HEC

DOE/SC supports basic research that underpins DOE missions and constructs and operates for the U.S. scientific community large scientific facilities such as accelerators, synchrotron light sources, and neutron sources. Six offices oversee DOE/SC research: Basic Energy Sciences (materials, chemistry, and nanoscale science); Biological and Environmental Research (global climate, genomics research); Fusion Energy Sciences (plasma fusion energy); High Energy Physics (accelerator design, petascale experimental data analysis); Nuclear Physics (quantum chromodynamics [QCD], astrophysics); and Advanced Scientific Computing Research.

Role of Computational Science

Computational science is critical to the DOE/SC mission in energy production, novel materials, climate science, and biological systems, where:

- Systems are too complex for direct calculation and descriptive laws are absent
- Physical scales range over many orders of magnitude
- Several scientific disciplines are involved such as in combustion and materials science
- Multidisciplinary teams are required
- Experimental data may be costly to develop, insufficient, inadequate, or unavailable
- Large data files are shared among scientists worldwide

DOE/SC's major HEC activities are integrated in the following large-scale multiyear computer science initiative to accelerate advanced research in scientific fields of interest to DOE:

Scientific Discovery Through Advanced Computing (SciDAC) Program – provides terascale computing and associated information technologies to many scientific areas to foster breakthroughs through the use of simulation. Through collaborations among application scientists, mathematicians, and computer scientists, SciDAC is building community simulation models in plasma physics, climate prediction, combustion, and other application areas. State-of-the-art collaboration tools facilitate access to these simulation tools by the broader scientific community to bring simulation to a level of parity with theory and observation in the scientific enterprise.

DOE/SC HEC R&D Programs

Recent DOE/SC accomplishments in HEC R&D that help address DOE computational science needs include:

- PVM, a widely used early message passing model
- MPI, MPICH – THE standard message passing model in use today (MPICH is the reference implementation used by all commercial vendors to develop tuned versions)
- Global arrays – global memory model used by NWChem, which greatly simplified electronic structure programs in computational chemistry
- Co-Array Fortran – the first and only open source Co-Array Fortran compiler
- UPC – the first open source Unified Parallel C compiler
- SUNMOS/Puma/Cougar/Catamount – microkernel operating system used by ASC Red system and its 40 teraflops replacement being developed by Cray, Inc.
- Science Appliance – pseudo single system image Linux-

based operating system used by large-scale (10 teraflops) clusters

- OSCAR – in a partnership with industry, the most widely used open source toolkit for managing Linux clusters
- Low-cost parallel visualization – use of PC clusters to achieve high-performance visualization

SciDAC Integrated Software Infrastructure Centers (ISICs) – SciDAC interdisciplinary team activities that address specific technical challenges in the overall SciDAC mission. Three Applied Mathematics ISICs are developing new high-performance scalable numerical algorithms for core numerical components of scientific simulation, and will distribute those algorithms through portable scalable high-performance numerical libraries. These ISICs are:

- Algorithmic and Software Framework for Applied Partial Differential Equations (APDEC) – to provide new tools for (nearly) optimal performance solvers for nonlinear partial differential equations (PDEs) based on multilevel methods
- Terascale Optimal PDE Solvers (TOPS) – hybrid and adaptive mesh generation and high-order discretization techniques for representing complex, evolving domains
- Terascale Simulation Tools and Technologies (TSTT) – tools for the efficient solution of PDEs based on locally structured grids, hybrid particle/mesh simulations, and problems with multiple- length scales

Four Computer Science ISICs are working closely with SciDAC application teams and the math ISICs to develop a comprehensive, portable, and fully integrated suite of systems software and tools for the effective management and use of terascale computational resources by SciDAC applications. The Computer Science ISICs are:

- Center for Component Technology for Terascale Simulation Software ISIC – to address critical issues in high-performance component software technology
- High-End Computer System Performance: Science and Engineering ISIC – to understand relationships between application and architecture for improved sustained performance
- Scalable Systems Software ISIC – for scalable system software tools for improved management and use of systems with thousands of processors
- Scientific Data Management ISIC – for large-scale scientific data management

SciDAC National Collaboratory Software Environment Development Centers and Networking Research – DOE's investment in National Collaboratories includes SciDAC projects focused on creating collaboratory software environments to enable geographically separated scientists to effectively work together as a team and to facilitate remote access to both facilities and data.

DOE/SC Workshops

Workshops – some general, others highly focused – are a standard DOE/SC planning tool. FY 2004 workshops include:

- SciDAC
- Fast Operating Systems
- Multiscale Mathematics Workshop – to bridge the wide scale between modeling on continuum or atomistic basis; expected output includes a roadmap
- High Performance File Systems – such systems should support parallelism, be efficient, and be reasonably easy to use
- Distributed Visualization Architecture Workshops (DiVA)
- Data Management Workshop series – to address where file system ends and data-management system begins
- Raising the level of parallel programming abstraction
- Parallel tool infrastructure

DOE/SC HEC I&A Programs

National Energy Research Supercomputing Center (NERSC) – delivers high-end capability computing services and support to the entire DOE/SC research community including the other DOE laboratories and major universities performing work relevant to DOE missions. Provides the majority of resources and services supporting SciDAC. Serves 2,000 users working on about 700 projects – 35 percent university-based, 61 percent in national laboratories, 3 percent in industry, and 1 percent in other government laboratories.

NERSC's major computational resource is a10-teraflops IBM SP computer. Overall NERSC resources are integrated by a common high-performance file storage system that facilitates interdisciplinary collaborations.

Advanced Computing Research Testbeds (ACRT) – activity supports the acquisition, testing and evaluation of advanced computer hardware testbeds to assess the prospects

for meeting future computational needs such as SciDAC and special-purpose applications. The ACRT provide two types of computer testbeds for evaluation – early systems and experimental systems. These research and evaluation (R&E) prototypes have been identified as a critical element in the HECRTF plan because they enable early partnerships with vendors to tailor architectures to scientific requirements. The results from these partnerships also play a key role in the choice of both high-performance production systems and potential leadership class systems governmentwide.

Scientific Application Pilot Programs (SAPP) – support targeted efforts to integrate new applied mathematics and computer science algorithms into SciDAC applications projects through strong liaisons between applications projects and ISICs. SAPP provides expertise in adapting high-performance algorithms to terascale computers and in adapting numerical modules to include application-specific knowledge.

Innovative and Novel Computational Impact on Theory and Experiment (INCITE) Projects – new FY 2004 competitive program that supports a small number of computationally intensive, large-scale research projects that can make high-impact scientific advances through the use of a substantial allocation of computer time and data storage at NERSC. INCITE encourages proposals from universities and other research institutions.

In December 2003, DOE/SC selected three projects to receive a total of 4.9 million hours of supercomputing time – 10 percent of the total computing time available this year on NERSC's Seaborg system. The projects are expected to significantly advance understanding of the makeup of the universe, the chemical process by which plants convert sunlight to energy while removing carbon dioxide from the atmosphere, and the turbulent forces that affect everything from weather to industrial processes. The projects are:

- Thermonuclear Supernovae: Stellar Explosions in Three Dimensions
- Quantum Monte Carlo Study of Photoprotection via Carotenoids in Photosynthetic Centers
- Fluid Turbulence and Mixing at High Reynolds Number

DOE/SC FY 2005 plans include:

HEC I&A

- *Installation of a Linux Cluster with more than 5 Tflops at NERSC to support high-end capacity needs of science applications that do not scale to more than 256 processors*
- *Partnership with DOE Office of Fusion Energy to develop next generation of Tokamak simulation tools*
- *Initiative in multiscale mathematics to develop the mathematical underpinnings for the simulation of complex physical and biological systems*

HEC R&D

- *Continuation of partnership with DARPA in productivity metrics for high-end computers*
- *Continuation of HEC-URA efforts to develop software base for future high-end systems*
- *Complete evaluation of Cray X1 computer*

HEC NIH HEC

The scientific drivers behind requirements for supercomputers at NIH are recent and are increasing rapidly. Although biomedical supercomputing is still in an early stage of development, NIH grantees, taking the individual initiative to apply for computer time, already consume approximately one-third of the cycles provided by the NSF supercomputing centers at Illinois, Pittsburgh, and San Diego. NSF's planned TeraGrid will most likely receive similar use by the biomedical research community.

The largest portion of the current biomedical supercomputing use is for biomolecular modeling and simulation. This class of application will continue to grow rapidly. The areas of high-throughput bioinformatics, data integration, and modeling of complex biosystems will grow explosively as software linking high-throughput data gathering with automated analysis and modeling tools becomes widely available. This software is currently in early development, with NIH support through new grant programs instituted in the last few years. To realize the potential of imaging and medical informatics will also require increased computational resources.

NIH has in place a planning process for the next stage of growth by preparing an eight- to ten-year plan for bioinformatics and computational biology that outlines the software needs and challenges for the next decade of biomedical computing.

Principal new investment areas for NIH include:

- Expanded development of molecular and cellular modeling software
- Applied computer science, such as algorithm creation and optimization, creation of appropriate languages and software architectures applicable to the solution of biomedical problems
- R&D in biomedical computational science, by developing and deploying software designed to solve particular biomedical problems
- Providing infrastructure to serve the needs of the broad community of biomedical and behavioral researchers
- Enhancing the training for a new generation of biomedical researchers in appropriate computational tools and techniques
- Disseminating newly developed tools and techniques to the broader biomedical research community

In FY 2005, NIH plans to:

- *Continue developing long-term plan for bioinformatics and computational biology, including listed areas of new investment*

HEC DARPA HEC

In FY 2004, DARPA supports the following HEC programs:

High Productivity Computing Systems (HPCS) program – addresses all aspects of HEC systems (packaging, processor/memory interfaces, networks, operating systems, compilers, languages, and runtime systems). Goal of multiyear effort is to develop a new generation of economically viable high- productivity computational systems for the national security and industrial user community in the time frame of FY 2009 to FY 2010. Sustained petaflop performance and shortened development time are key targets. DARPA is working with DOE/SC and other agencies to understand how long it takes to develop solutions today. These systems will be designed to have the following impact:

- Performance (time to solution) – provide a 10-to-40-times performance speedup on critical national security applications
- Programmability (idea to first solution) – reduce the cost and time of developing applications solutions
- Portability (transparency) – insulate the research and operational application software from the system on which it is running
- Robustness (reliability) – apply all known techniques to protect against outside attacks, hardware faults, and programming errors

Target applications are intelligence and surveillance, reconnaissance, cryptanalysis, weapons analysis, airborne contaminant modeling, and biotechnology.

The HPCS program is structured to enable ongoing fielding of new high-performance computer technology through phased concept, research, development, and demonstration approach. The program is in its second phase, with productivity goals including:

- Execution (sustained performance) of 1 petaflop/second (scalable to over 4 petaflop/second)
- Development productivity gains of 10 times over today's systems

In phase 2, a productivity framework baselined to today's systems is used to evaluate vendors' emerging productivity techniques and will provide a reference for phase 3 evaluations of vendors' proposed designs. Phase 2 subsystem performance indicators are a 3.2 petabytes/second bisection bandwidth, 64,000 gigabyte updates per second (GUPS), 6.5 petabytes data streams bandwidth, and 2+ petaflops Linpack Fortran benchmark (an HPC Challenge) that augments the Top 500 benchmarks.

The NITRD agencies that provide additional funding for HPCS activities are DOE/SC, DOE/NNSA, NASA, NSA, and NSF; DoD's HPCMPO provides complementary support. Cray, Inc., IBM, and Sun Microsystems are phase 2 industry partners.

Council on Competitiveness HPC Initiative – HPCS-related co-funded by DARPA, DOE/SC, and DOE/NNSA has the goal of raising high productivity computing to be a major economic driver in the United States. Effort includes: an HPC Advisory Committee with representatives from the private sector; an annual private-sector user survey; and annual user meetings. The first conference, entitled "HPC: Supercharging U.S. Innovation and Competitiveness," was held July 13, 2004, in Washington, D.C.

Polymorphous Computing Architectures (PCA) Program– aims to develop the computing foundation for agile systems by establishing computing systems (chips, networks, and software) that will morph to changing missions, sensor configurations, and operational constraints during a mission or over the life of the platform. Response (morph) times may vary from seconds to months. The PCA objective is to provide:

- Processing diversity with a breadth of processing to address signal/data processing, information, knowledge, and intelligence, and uniform performance competitive with best-in-class capabilities
- Mission agility to provide retargeting within a mission in millisecond, mission-to-mission adaptation in days, and new threat scenario adaptation in months
- Architectural flexibility to provide high-efficiency, selectable virtual machines with two or more morph states, mission adaptation with N minor morph states, and portability/evolvability with two-layer morphware

The PCA hardware being developed includes:

- Monarch/MCHIP – 2005 prototype is projected to demonstrate 333Mhz, 64 GFLOPS, 64 GOPS, and 12 Mbytes EDRAM.
- Smart Memories – 2005 prototype is projected to demonstrate Brook Language based on modified commercial Tensilica CPU cores.
- TRIPS – 2005 prototype is projected to demonstrate 500 MHz, 16 Gflops peak floating point, and 16 GIPS peak integer. Production technology is projected to provide 4 Tflops peak floating point and 512 GIPS peak integer.
- RAW (early prototype) to provide chips and development boards (available now) and kernel evaluations with test results in 2004. The technology provides early morphware, software, and kernel performance evaluation. The morphware will be demonstrated in a development environment in which APIs, including Brook, will use high-level compilers to provide stream and thread virtual machine APIs to low-level compilers using PCA technology such as TRIPS, Monarch, Smart Memories, and RAW.

Networked Embedded Systems Technology (NEST) – will enable fine-grain fusion of physical and information processes. The quantitative target is to build dependable, real-time, distributed, embedded applications comprising 100 to 100,000 simple compute nodes. The nodes include physical and information system components coupled by sensors and actuators.

In FY 2005, work will continue on the following DARPA efforts:

- *High Productivity Computing Systems (HPCS)*
- *Polymorphous Computing Architectures (PCA)*

HEC **NASA** HEC

FY 2004 NASA accomplishments in high-end computing:

- Testbeds
  - Built the first SSI Linux base supercomputer, the world's fastest machine on the Triad Benchmark
- Advanced Architectures
  - Invested in new architecture streams and PIM working on languages and compilers and key application kernels for stream technology (this is tied into DARPA's PCA and morphware efforts)
- Grid Technology
  - Continued work on NASA's Information Power Grid (IPG) demonstration in cooperation with Earth Science and Shuttle payload management
- Applications
  - Developed advanced applications for aerospace, nanotechnology, and radiation biology
  - Provided major advance in ocean modeling in the Estimating the Circulation and Climate of the Ocean (ECCO) code
  - Major advance in aerospace vehicle modeling
- Programming environments
  - Improved scalability of multilevel parallelism (MLP)
- Tools
  - Workflow and database management programming environments
  - Aero database/IPG Virtual Laboratory for workflow management and ensemble calculation database-management prototypes
  - Hyperwall (for viewing large datasets) used in nanotechnology research, the Sloan Digital Sky Survey, and the Columbia Space Shuttle Mission accident investigation. This technology is now being transferred within NASA for general use.

Estimating the Circulation and Climate of the Ocean (ECCO) – project goal is to use NASA data and community code to improve Earth oceans modeling. Starting in August 2003, the project has:

- Developed serial #1 512 Altix supercomputer
- Ported and scaled ECCO code from 64 to 128 to 256 to 512 processors
- Determined correct environment variables, optimized code
- Assisted in revamping grid technique to improve prediction at poles
- Developed Kalman filter techniques at high resolution for data assimilation
- Developed archive for community access to data
- Improved resolution from 1 degree to 1/4 degree
- Developed unique visualization approaches

As a result of this work:

- Time to solution dropped from approximately one year to five days (about a 50-fold speedup)
- Resolution improved by a factor of four
- Code accuracy and physics improved
- Data assimilation at high resolution was enabled

Shuttle return-to-flight problems are now being ported to the Altix supercomputer for speedup by a factor of four or more for some turbo-machinery.

Supercomputing support for the Columbia Accident Investigation Board (CAIB) – NASA Ames supercomputing-related R&D assets and tools produced time-critical analyses for the CAIB. R&D and engineering teams modeled a free object in the flow around an airframe with automatic regridding at each time step. The environment in which this was accomplished entailed:

- State-of-the-art codes
  - Codes honed by R&D experts in large-scale simulation, with their environment of programmer tools developed to minimize the effort of executing efficiently
- A modeling environment that included experts and tools (compilers, scaling, porting, and parallelization tools)
- Supercomputers, storage, and networks
  - Codes and tools tuned to exploit the advanced hardware that was developed for large tightly-coupled computations
- Computation management by components of the AeroDB and the iLab
  - An environment and tools for efficient executions of hundreds or thousands of large simulations and handling of results databases

- Advanced data analysis and visualization technologies to explore and understand the vast amounts of data produced by this simulation

NASA HEC Resources

NASA high-end computing resources include:

- Goddard Space Flight Center, where the foci are in Earth Science and Space Science
  - 1,392 processing element (PE) Compaq (2.2 Teraflops)
  - 640 PE SGI O3K (410 Gigaflops)
  - Sun QFS (340 Terabytes)
  - SGI DMF (370 Terabytes)
- Ames, where the foci are in Aeronautics and Earth Science
  - 512 PE SGI Altix (2.3 teraflops)
  - 1,024 PE SGI O3K (850 gigaflops)
  - SGI DMF (1,600 terabytes)

NASA FY 2005-FY 2007 HEC R&D plans include:

- *Collaborative decision systems to improve decision making*
- *Discovery systems to accelerate scientific progress*
- *Advanced networking and communications*
- *Advanced computing architectures and technologies*
- *Reliable software*
- *Adaptive embedded information systems*

HEC DOE/NNSA HEC

DOE/NNSA's Office of Advanced Simulation and Computing (ASC) provides the means to shift promptly from nuclear test-based methods to computer-based methods and thereby predict with confidence the behavior of nuclear weapons through comprehensive science-based simulations. To achieve this capability, ASC has programs in:

- Weapons codes and the science in those codes (advanced applications, materials and physics modeling, and verification and validation)
- Computational platforms and the centers that operate them
- Software infrastructure and supporting infrastructure (Distance computing [DisCom], Pathforward, Problem Solving Environments, and VIEWS)

ASC HEC R&D focuses on:

- Software-quality engineering oriented to processes to provide for longevity of codes
- Verification to ensure numerical solutions are accurate
- Validation to ensure that a problem is understood correctly
- Certification methodology – improvements through scientific methods
- Capability computing to meet the most demanding computational needs
- Capacity computing to meet current stockpile workloads
- Problem Solving Environments (PSEs) to create usable execution environments for ASC-scale applications
- Industry collaboration to accelerate key technologies for future systems
- Tracking requirements

ASC activities – Pathforward work includes:

- Optical switch (with NSA)
- Memory usage
- Lustre file system
- Spray cooling (with DOE/SC)

ASC HEC I&A platforms – HEC systems and software from Cray, HP, IBM, Intel, SGI, and Linux NetworX. ASC procures machines with differing technologies to support diverse requirements. These types include:

- Capability systems: Vendor-integrated SMP clusters used for critical Stockpile Stewardship Program (SSP) deliverables and for particularly demanding calculations. (These systems run a small number of large jobs.)
- Capacity systems: Linux HPC clusters that carry the largest workloads in terms of numbers of jobs supporting current weapons activities and as development vehicles for next-generation capability systems. (These systems run a large number of smaller jobs.)
- Disruptive technologies: The Blue Gene/L system, for example, is being procured as a highly scalable, low-

power, low-cost computer. It is the first high-performance computing system to incorporate all three attributes in a single system. ASC works with vendors and universities to push production environments to petaflop scales. ASC works with industry to develop new technologies such as optical switches, scalable visualization, memory correctness, spray cooling, and the Lustre file system.

DOE/NNSA FY 2005 HEC plans include:

- *Continue R&D in software-quality engineering; validation, verification, and certification of software; PSEs; and Pathforward topics*
- *Continue development of Blue Gene/L*

HEC NSA HEC

NSA has unclassified HEC R&D programs in the areas of:

- Architectures and systems
- High-speed switches and interconnects
- Programming environments
- Quantum information sciences
- Vendor partnerships

Cray X1e/Black Widow – under multiyear joint development effort, NSA supports Cray Inc. in extending its X1 NSA/ODDR&E/Cray development to build its next-generation scalable hybrid scalar-vector high-performance computer. The technology will result in an X1e upgrade in the market in late 2004. Black Widow system to be introduced in 2006 will provide outstanding global memory bandwidth across all processor configurations and will be scalable to hundreds of teraflops. Should be the most powerful commercially available system in the world at that time. Many agencies have regularly participated in the quarterly development reviews; several have serious plans for acquisition of these systems.

HPC Applications Benchmarking Study – project initiated by HPC users, the HPC User Forum, and IDC to study the performance of applications on leading supercomputers. The project will expand the understanding of the performance of a set of complete, difficult HPC applications on the largest computers around the world. Each application will have a specific data set and will represent an actual scientific problem. The results will be posted at the HPC User Forum Web site. This study's objectives are to compare leading large HPC capability class computers on a small set of large HPC problems, show major science accomplishments on these capability systems, explore the relative performance characteristics of the machines on these problems, and explore the scaling abilities of the computers. The HPC Users Forum effort is funded by NSA, DoD/HPCMPO and several other agencies.

Reconfigurable Computing Research and Development – an approach to general-purpose high-performance computing that takes advantage of field programmable gate array (FPGA) technologies and commercial platforms. This continuing effort is leveraging prototype results from academic, industrial, and government research and the continuing increase in the computational capability of commercial FPGA chips. NSA is conducting experiments and demonstrations of commercial reconfigurable computing machines using its applications and benchmarks. The principle difficulty being addressed is the need for software-level programming tools. There is substantial collaboration with the Air Force, NASA Langley, and the NRO in reconfigurable research and development.

Major FY 2004 achievements:

- Completed a highly successful demonstration on an NSA application of the C Compiler on a particular reconfigurable computer to achieve a 65x performance improvement over a modern 64-bit commodity microprocessor.
- Co-funded an NSA vendor partnership to develop an FPGA addition to a commercial architecture, with a prototype machine to be delivered in FY 2004.

Fine-grain Multithreaded Architectures and Execution Models – ongoing research program in collaboration with the University of Delaware to develop architectures, execution models and compilation techniques that can exploit multithreading to manage memory latency and provide load balancing.

Performance analysis of many real-world applications on today's large scale massively parallel processor systems indicates that often only a relatively small fraction of the potential performance of these machines is actually achieved in practice. The applications that exhibit these problems are generally characterized by memory access patterns that are dynamic and irregular, and hence difficult to partition efficiently, either by programmers or compilers. This situation frequently results in processors being "starved" for data due to the relatively high memory latencies. Fine-grain multithreaded models coupled with effective compilation strategies provide a promising and easy to use approach to reducing the impact of long, unpredictable memory latencies.

Work in the past year has resulted in an FPGA-based implementation of a hardware emulation platform and a compiler infrastructure that is enabling experimentation with thread scheduling and synchronization units.

Optical Technologies – Columbia University and Georgia Tech University commissioned by NSA to investigate high-end computer optical interconnect systems. Columbia's comparison of electrical and optical technologies showed conclusively that specifically designed all-optical switching can significantly outperform current and future electronic switches. The collaborative effort demonstrated that optical technologies can significantly improve bandwidth if implementation difficulties of all-optical switching can be overcome. Future efforts will investigate how current optical components might be integrated to form a demonstration system.

Support to DARPA HPCS Program, Phase 2 – NSA participation includes funding, technical guidance, and participation in program reviews.

Quantum Computing – major research program being conducted through NSA's ARDA affiliate. Collaboration involves DARPA and ARL. Efforts focus on quantum information sciences including investigation of selective materials, related quantum properties, and managing decoherence and entanglement issues. Funded research under this multiagency collaboration is underway at many universities, government laboratories, and other centers of excellence.

FY 2005 NSA plans in HEC include:

- *Continue R&D in all listed FY 2004 HEC R&D areas*
- *Increase funding support for Black Widow development consistent with its calendar-year 2006 introduction*
- *Maintain current research effort in all other efforts, including quantum computing.*

HEC **NOAA** HEC

NOAA uses HEC resources and cutting-edge applications to provide advanced products and services delivered to users. NOAA's HEC strategy is to:

- Develop skills, algorithms, and techniques to fully utilize scalable computing for improved environmental understanding and prediction
- Acquire and use high-performance scalable systems for research
- Optimize the use of HEC resources across all NOAA activities. This reflects NOAA's view of HEC as an enterprise resource.

NOAA's HEC I&A resources are located at the:

- Geophysical Fluid Dynamics Laboratory (GFDL) in Princeton, New Jersey. (In FY 2003, GFDL acquired a 1,408-processor SGI system.)
- Forecast Systems Laboratory (FSL) in Boulder, Colorado
- National Center for Environmental Prediction (NCEP), headquartered in Camp Springs, Maryland

NOAA expedites the development and use of improved weather and climate models by:

- Supporting advanced computing for NOAA environmental modeling research
- Developing software tools to optimize the use of scalable computing
- Infusing new knowledge through new talent such as postdoctoral students and contractors

NOAA Development Test Center (DTC) – establishing a collaborative environment in which university research and NOAA personnel can work together, focusing initially on development of the Weather Research and Forecasting (WRF) model that is designed as both an operational model and a research vehicle for the larger modeling community. Collaborators on the WRF include FNMOC, NAVO, AFWA, NCAR, and others. The goal is for Federal and university researchers working on specific problems to explore models, develop improvements, plug in to that part of the model, help the entire community, and increase speed going from basic research to applications. Potential benefits include frequent high-resolution analyses and forecasts produced in real time that are valuable to commercial aviation, civilian and military weather forecasting, the energy industry, regional air pollution prediction, and emergency preparedness.

NOAA is working cooperatively on the ESMF with NASA and others, and is coordinating with other agencies its research plans including the Climate Change Strategic Plan and the U.S. Weather Research Program. Developing grid technology with GFDL, PPPL, PU, FSL, PMEL, and CSU.

NOAA FY 2005 plans include the following:

- *Continue work on the ESMF for climate model development and extend it to all domains*
- *Explore the use of specific types of grid technology*
- *Develop a common weather research computing environment*
- *Apply the WRF modeling system standards and framework to NCEP's Mesoscale Modeling Systems, including the continental domain Meso Eta and nested Nonhydrostatic Meso Threats runs*

HEC NIST HEC

NIST works with industry and with educational and government organizations to make IT systems more usable, secure, scalable, and interoperable; to apply IT in specialized areas such as manufacturing and biotechnology; and to encourage private-sector companies to accelerate development of IT innovations. NIST also conducts fundamental research that facilitates measurement, testing, and the adoption of industry standards.

NIST HEC I&A

Activities and accomplishments include:

- Interoperable Message Passing Interface (IMPI) to support parallel computing research. The standard enables interoperability among implementations of Message Passing Interface (MPI) with no change to user source code.
- Java/Jini/Javaspaces-based screen-saver science system. Distributed-computing environment leverages java's portability to use otherwise unused cycles on any set of computing resources accessible over a local network (including desktop PCs, scientific workstations, and cluster compute nodes) to run any compute-intensive distributed Java program.
- Immersive visualization infrastructure for exploring scientific data sets drawn from both computer and laboratory experiments. Visualization projects include a three-wall Immersive Virtual Reality System that uses the DIVERSE open source software and locally built visualization software to display results at scale. The local software includes the glyph toolbox and techniques for immersive volume visualization, surface visualization, and interaction. All the software runs unchanged on Linux machines.
- Fundamental mathematical software tools to enable high-end computing applications. Examples include:
  - PHAML, a parallel adaptive grid refinement multigrid code for PDEs
  - Sparse Basic Linear Algebra Subroutines (BLAS)
  - Template Numerical Toolkit (TNT) for object-oriented linear algebra
- HEC applications combining parallel computing with immersive visualization for basic sciences, such as physics, and other uses, such as building structure and material strength. Typical applications to which NIST applies these resources are:
  - Nanostructure modeling and visualization (especially in nano-optics)
  - Modeling cement and concrete, including modeling the flow of suspensions and fluid flow in complex geometries
  - Computation of atomic properties
  - Visualization of "smart gels" (which respond to specific physical properties such as temperature or pressure) to gain insight into the gelling mechanism
  - Visualization of tissue engineering to study and optimize the growth of cells on scaffolding
- Applications-oriented problem-solving environments, such as Object-Oriented Finite Element Modeling of Material Microstructures (OOF) for modeling materials with complex microstructure and Object-Oriented MicroMagnetic Computing Framework (OOMMF) for micromagnetics modeling

NIST HEC R&D

Activities and accomplishments include:

- Research in quantum computing and secure quantum communications, as well as in measurement science aspects of nanotechnology, photonics, optoelectronics, and new chip designs and fabrication methods. Theoretical and experimental investigations include:.
  - Algorithms and architectures for scalable quantum computing
  - Demonstrations of quantum computing within specific physical systems such as ion traps, neutral atoms, and Josephson junctions
  - High-speed testbed for quantum key distribution
  - Single photon sources and detectors
  - Quantum protocols

FY 2005 NIST plans include:

HEC I&A

- *Continue work in activities listed for FY 2004*
- *Pursue additional project in "Virtual Laboratories"*

HEC R&D

- *Continue work in activities listed for FY 2004*

HEC **EPA** HEC

EPA's mission is to protect human health and safeguard the environment through research, regulation, cooperation with state governments and industry, and enforcement. Areas of interest extend from groundwater to the stratosphere.

EPA HEC I&A programs are focused on tools to facilitate sound science using high-end computation, storage, and analysis. These programs enable relevant, high-quality, cutting-edge research in human health, ecology, pollution control and prevention, economics, and decision sciences. This facilitates appropriate characterization of scientific findings in the decision process. The HEC programs are performed in-house and as problem-driven research.

EPA is launching the Center of Excellence (CoE) for Environmental Computational Science to integrate cutting-edge science and emerging IT solutions to facilitate Federal and state-level partnerships and enhance the availability of scientific tools and data for environmental decision making. The CoE will enable collaboration from within and without the organization and will provide a flexible, dynamic computing and information infrastructure to ensure optimized yet secure access to EPA resources.

FY 2004 HEC programs include:

Multimedia Assessments and Applications (MAA) Framework – provides a foundation for research on how to structure compartmental modules and improve model integration and interchangeability. The MAA framework's objective is to provide software that supports composing, configuring, applying, linking, and evaluating complex systems of object-oriented models. It will improve EPA's ability to simulate the interaction between individual environmental media components (e.g., chemical fluxes, water cycle) and will enable distributed computation.

The MAA framework is tailored to multimedia models but is adaptable and generalized. It supports EPA programs such as the Chesapeake Bay Program, the Tampa Bay Region Atmospheric Chemistry Experiment, and the Office of Air Quality Planning and Standards. The framework is currently being tested by a number of clients.

Uncertainty Analysis (UA) Framework Development – developing tools to support the analysis of model sensitivities and the effects of input uncertainties on model predictions. Specific tasks of this EPA work are to:

- Construct a 400-node Intel-based supercomputing cluster called SuperMUSE, for Supercomputer for Model Uncertainty and Sensitivity Evaluation
- Develop platform-independent systems software for managing SuperMUSE
- Conduct uncertainty and sensitivity analyses of the Multimedia, Multipathway, Multireceptor Risk Assessment (3MRA) modeling system
- Develop advanced algorithmic software for advanced statistical sampling methods and global sensitivity analyses

Air Quality Modeling Applications (AQMA) – program aims to advance computational performance of the state-of-the-art Community Multiscale Air Quality (CMAQ) Chemistry-Transport Modeling System while maintaining modularity, portability, and single-source code. Efforts to improve CMAQ take into account both the typical Linux cluster in the States and also HEC deployments. Major areas of effort include algorithmic improvement, microprocessor tuning, and architecture assessment. Involves a phased deployment that enables the states, which are the key stakeholders, to participate in the development.

Grid deployment – goal is to provide phased deployment of an EPA-wide enterprise grid that will identify, develop, and integrate key technologies, align organizational policies such as security and networking, and field grid pilot systems to demonstrate success and benefits. Historically, agency researchers with high-end applications competed for time on EPA's high-performance computing resources located at the National Environmental Scientific Computing Center (NESC2). With the implementation of grid middleware, researchers will be able to tap unused processing capacity on local and remote clusters at the campus level or enterprise level. EPA's compute grid is being implemented in a phased approach with parallel development of both grid infrastructure and security policy. Pilot clusters have now been linked to demonstrate EPA compute grid services both internally and also to external partners.Ultimately, EPA researchers and trusted partners will be able to access a partner (or global) grid extending to organizations outside the agency.

Grid demonstration projects – part of EPA effort to combine state-of-the-science air quality modeling and observations to enable timely communication of meaningful

environmental information, improve emission management decisions, and track progress in achieving air quality and public health goals.Technical collaborators in demonstration project on air quality include DOE/Sandia and NOAA; pilot partners include the state of New York and regional planning organizations. Phase I of the project includes delivery of an optimized CMAQ model and data sets to the client community in summer 2004 and eventually the running of CMAQ over the grid and at client sites.

In FY 2005, HEC I&A work will continue on the following EPA efforts:

- *EPA Center of Excellence for Environmental Computational Science (CoE)*
  - *Grid deployment and grid demonstration projects*
- *Air Quality Modeling Applications (AQMA)*

Participating Agency

HEC DoD/HPCMPO HEC

The mission of DoD's HPCMPO is to deliver world-class commercial, high-end, high-performance computational capability to DoD's science and technology (S&T) and test and evaluation (T&E) communities, thereby facilitating the rapid application of advanced technology in superior warfighting capabilities. Development of future technologies supported by HPC includes: Micro Air Vehicles; Joint Strike Fighter; surveillance systems; smart weapons design; ocean modeling; parachute simulations; Unmanned Air Vehicle; and blast protection.

HPCMPO requirements are categorized in the following ten key Computational Technology Areas (CTAs):

- Computational Structural Mechanics (CSM)
- Computational Fluid Dynamics (CFD)
- Computational Chemistry and Materials Science (CCM)
- Computational Electromagnetics and Acoustics (CEA)
- Climate/Weather/Ocean Modeling and Simulation
- Signal/Image Processing (SIP)
- Forces Modeling and Simulation (FMS)
- Environmental Quality Modeling and Simulation (EQM)
- Computational Electronics and Nanoelectronics (CEN)
- Integrated Modeling and Test Environments (IMT)
- Other

HEC I&A

High Performance Computing Centers – include four major shared resource centers (MSRCs) and 16 distributed centers. MSRCs provide complete networked HPC environments, HPC compute engines, high-end data analysis and scientific visualization support, massive hierarchical storage, and proactive and in-depth user support and computational technology area expertise to nationwide user communities. The MSRCs are the Army Research Laboratory (ARL), Army Signal Command (ASC), ERDC, and NAVO.

Software Applications Support (SAS) – encompasses the following four components:

- Common HPC Software Support Initiative (CHSSI)
- HPC Software Applications Institutes (HSAIs)
- Programming Environment and Training (PET) program to enhance user productivity

Human-Computer Interaction and Information Management

Definition of HCI&IM PCA

The activities funded under the NITRD Program's HCI&IM PCA increase the benefit of computer technology to humans through the development of future user interaction technologies, cognitive systems, information systems, and robotics. Current systems overwhelm the user with information but provide little in the way of adaptable access, necessitating adaptability on the part of the user. The HCI&IM research vision is to provide information that is available everywhere, at any time, and to everyone regardless of their abilities; to increase human use of this information by providing customized access including the ability to interact using a variety of devices and to meet varied needs for manipulation, analysis, and control; and to provide comprehensive management of vast information environments. HCI&IM research focuses on developing systems that understand the needs of the users and adapt accordingly.

Broad Areas of HCI&IM Concern

- Usability and universal accessibility
- Interaction
- Cognitive systems, learning, and perception
- Information management and presentation
- Autonomous agents and systems

Technical Goals

- Foster contributions from different branches of science and engineering that are required to address the problems in this multidisciplinary field and encourage the needed technological convergence through communication and coordination
- Develop a deeper scientific understanding of both human and machine cognition, linking the various parts of human-computer interaction and information management together
- Enable development of machines and systems that employ the human senses of perception to their full potential, maximizing the information-flow bandwidth between people and their tools
- Discover new and better ways to achieve information integration across a wide range of different modalities, media, and distributed resources
- Develop tools that are increasingly able to take on planning, aid in decision making, learn from experience, and develop knowledge from raw data, in order to tie together information management and human needs
- Develop flexible systems of control that manage the degree of autonomy exhibited by machines according to the constantly changing needs of humans, especially in that broad and valuable area where HCI and IM unite for the benefit of people

Illustrative Technical Thrusts

- Fundamental science and engineering on human-computer interfaces for stationary, mobile, and ubiquitous computing and communications environments
- Multi-modal, speech, gesture, visual, language, haptic, and physical interactions
- Efforts to improve the collection, storage, organization, retrieval, summarization, analysis, and presentation of

HCI&IM Agencies

			Participating Agency
NSF	NASA	EPA	
NIH	AHRQ	NOAA	
DARPA	NIST		FAA

HCI&IM PCA Budget Crosscut

FY 2004 Estimate	FY 2005 Request
$469.2 M	$419.5 M

information of all kinds in databases, distributed systems, or digital libraries

- An understanding of the use of multiple modalities and their relation to information content
- Information management and access that adapts to the needs and preferences of a diverse population including young, old, and disabled users as well as expert and novice users and in complex, multi-person, collaborative and distributed systems
- Research and technology development related to perception, cognition, and learning by machines and by human beings interacting with machines
- Mobile autonomous robots
- Remote or autonomous agents
- Collaboratories
- Visualizations
- Web-based repositories
- The semantic web
- Information agents
- Evaluation methodologies and metrics for assessing the progress and impact of HCI&IM research

HCI&IM PCA: Coordination and Activities

HCI&IM Highlights

Early in FY 2004, the HCI&IM Coordinating Group released its "Human-Computer Interaction and Information Management Research Needs" report. This document identifies and illustrates the challenges underlying HCI&IM R&D to achieve benefits such as:

- Changing the way scientific research is conducted
- Expanding the science and engineering knowledge base
- Enabling a more knowledgeable, capable, and productive workforce

The report places agency HCI&IM R&D investments within a conceptual framework that illuminates the broader social context in which the need for "human-computer interaction and information management" arises. The framework's four main categories are:

- Information creation, organization, access, and use
- Managing information as an asset
- Human-computer interaction and interaction devices
- Evaluation methods and metrics

The CG is using the report to identify and assess current agency investments, research gaps, and a wish list of possible FY 2005 activities, including refining plans in new R&D areas such as cognition, robotics, and devices and identifying and coordinating activities in areas of shared interests with other CGs, such as:

- With LSN on moving the bits in distributed data
- With SEW on the impact of distributed data on organizations and groups sharing data
- With LSN in middleware
- With SDP in automatic software design and integration
- With SEW in universal accessibility and the use of multiagent systems

The following is a sampling of FY 2004 activities in which more than one NITRD agency participates (other agencies involved in these efforts are not cited):

DARPA – Improving Warfighter Information Intake Under Stress Program, with NSF; Translingual Information Detection, Extraction, and Summarization (TIDES), with NIST, NSA; Effective, Affordable, Reusable Speech-to-Text (EARS), with NIST, NSA

NASA – High-end modeling and simulation, with DOE/SC; global climate modeling, with DOE/SC and NOAA; automated vehicle cockpits and air traffic management and risk, with FAA

AHRQ – Considering funding proposals submitted to AHRQ's Health Information Technology (HIT), with NIH/NLM; developing standards for improving patient safety, as part of the Patient Safety Task Force, with NIH/NLM; developing a U.S. health data standards landscape model and application, with NIST; developing data standards critical to improving patient safety in the use of prescription drugs, with FDA; invites experts from other agencies to participate in peer reviews of its grant proposals

NIST – Human Language Technology and Interactive Systems Technology Programs, with DARPA and NSA

HCI&IM R&D Programs By Agency

Selected FY 2004 Activities and FY 2005 Plans

HCI&IM NSF HCI&IM

The bulk of NSF's FY 2004 HCI&IM R&D investments are in its Directorate for Computer and Information Sciences and Engineering (CISE), in the Division of Information and Intelligent Systems (IIS). Other HCI&IM activities are funded under the NSF-wide Information Technology Research Program (ITR).

CISE/IIS Division

IIS goals are to:

- Increase the capabilities of human beings and machines to create, discover, and reason with knowledge
- Advance the ability to represent, collect, store, organize, locate, visualize, and communicate information
- Conduct research on how empirical data lead to discovery in the sciences and engineering

HCI&IM-related research is funded across IIS's three clusters:

Systems in Context – research and education on the interaction between information, computation, and communication systems, and users, organizations, government agencies, and the environment. This cluster integrates the following programs:

- Human-Computer Interaction
- Universal Access
- Digital Society and Technologies
- Robotics
- Digital Government

Data, Inference, and Understanding – basic research with the goal of creating general-purpose systems for representing, sorting, accessing, and drawing inferences from data, information, and knowledge. Integrates the following programs:

- Artificial Intelligence and Cognitive Science (AI&CS) – focuses on advancing the state of the art in AI&CS. Includes research fundamental to development of computer systems capable of performing a broad variety of intelligent tasks and computational models of intelligent behavior across the spectrum of human intelligence
- Computer Vision
- Human Language and Communication
- Information and Data Management
- Digital Libraries

Science and Engineering Informatics – includes:

- Collaborative Research in Computational Neuroscience that will continue as a collaborative program
- Science and Engineering Information Integration and Informatics (SEI), which focuses on development of IT to solve a particular science or engineering problem and generalizing the solution to other related problems
- Information Integration, to provide a uniform view to a multitude of heterogeneous, independently developed data sources including reconciling heterogeneous formats, Web semantics, decentralized data-sharing, data-sharing on advanced cyberinfrastructure (CI), on-the-fly integration, and information integration resources

ITR Program

Other NSF HCI&IM investments are in the NSF-wide ITR Program, which involves all NSF directorates and offices including Biological Sciences; Engineering; Geosciences; Mathematical and Physical Sciences; Social, Behavioral, and Economic Sciences; Office of Polar Programs; and Education and Human Resources.

NSF HCI&IM plans for FY 2005 include:

- ***Research to develop computer technology that empowers people*** *with disabilities, young children, seniors, and members of traditionally under-represented groups, so that they are able to participate as first-class citizens in the new information society, including for example work on how blind people can benefit from the information communicated when other people gesture or exhibit facial expressions.*
- ***Information Integration*** *– special emphasis will be placed on domain-specific and general-purpose tools for integrating information from disparate sources, supporting projects that will advance the understanding of technology to enable scientific*

discovery, and that will creatively integrate research and education for the benefit of technical specialists and the general population

- ***Intelligent robots and machine vision technology –*** *research to develop intelligent robots and machine vision technology that will help people do what they cannot or would rather not do; protect critical infrastructure and monitor the environment, and continue to explore extreme environments*
- ***Projects to push the frontiers in human-computer communication –*** *automatic multilingual speech recognition toolkits, systems to recognize spontaneous speech despite disfluencies and multiple speakers, and development of information infrastructure to enable cutting-edge research and development about spoken, written, and multimodal communication*

HCI&IM NIH HCI&IM

Illustrative of NIH HCI&IM activities is the Joint NIH Bioengineering Consortium/NIH Biomedical Information Science and Technology Initiative Consortium (BECON/BISTIC) 2004 Symposium on Biomedical Informatics for Clinical Decision Support: A Vision for the 21st Century held June 21-22, 2004. The symposium involved 15 NIH Institutes seeking to gain consensus about standards for reducing medical errors and variability in patient information by reviewing software tools and approaches to deliver the benefits of biomedical information technologies to patients at the time and place of decision making regarding risk, diagnosis, treatment, and follow-up. Specifically, the meeting provided a scientific vision of the health care information technologies that may be more fully deployed in the workflow to improve efficiency and outcomes.

NIH FY 2005 plans in HCI&IM R&D include:

- ***Curation and analysis of massive data collections*** *– a focus that has emerged as a critical element for advances in biomedical and clinical research in the 21st Century. NIH is making substantial new investments in tools for management and use of the massive new databases that will permit their use by researchers throughout the country. Specific R&D topics include:*

  – Tools for building and integrating ontologies
  – Software tools for visualizing complex data sets
  – Curation tools
  – Support for standard vocabularies, nationwide
  – Information integration tools

HCI&IM DARPA HCI&IM

DARPA's FY 2004 HCI&IM investments involve four programs that fall within the agency's Information Processing Technology Office (IPTO). IPTO's mission is to create a new generation of computational and information systems that possess capabilities far beyond those of current systems. The programs are:

Compact Aids for Speech Translation (CAST) – developing rapid, two-way, natural-language speech translation interfaces and platforms for the warfighter in the field that overcome the technical and engineering challenges limiting current multilingual translation technology; goal is to enable future full-domain, unconstrained dialog translation in multiple environments, replacing the DARPA RMS (one-way) translator. The two-way capability will greatly improve the ability of operators to converse with, extract information from, and give instructions to a foreign-language speaker encountered in the field. Major research thrusts are:

- Core research
- DARPA 1+1 Way
- DARPA Two-Way
- Evaluation and data collection

Improving Warfighter Information Intake Under Stress – (formerly Augmented Cognition) R&D to extend, by an order of magnitude or more, the information management capacity of the human-computer warfighting integral by developing and demonstrating quantifiable enhancements to human performance in diverse, stressful, operational environments. The goal is to empower one person successfully to accomplish the functions currently carried out by three or more people. The research aims to enable computational systems to dynamically provide:

- Real-time assessment of warfighter status (phase one)
- Real-time maximization of warfighter potential (phase two)
- Autonomous adaptation to support warfighter performance under stress (phase three)
- Operational demonstration and test (phase four)

Warfighters are constrained in the amount of information they can manage. Adaptive strategies will mitigate specific information processing roadblocks impeding increased performance and information flow. Strategies include:

- Intelligent interruption to improve limited working memory
- Attention management to improve focus during complex tasks
- Cued memory retrieval to improve situational awareness and context recovery
- Modality switching (i.e., audio, visual) to increase information throughput

Technical challenges and goals for IT capabilities that can improve performance include:

- Demonstrating enhancement of warfighter performance (assess warfighter status in less that two seconds with 90 percent accuracy; adapt information processing strategies in less than one minute, with no performance degradation)
- Overcoming information processing bottlenecks (500 percent increase in working memory throughput; 100 percent improvement in recall and time to reinstate context; 100 percent increase in the number of information processing functions performed simultaneously; 100 percent improvement in successful task completion within critical time duration)

Translingual Information Detection, Extraction, and Summarization (TIDES) – seeks to enable military personnel to operate safely and effectively around the globe by:

- Developing advanced language processing technology to enable English speakers to find and interpret critical information in multiple languages without requiring knowledge of those languages
- Developing sophisticated statistical modeling techniques for human language
- Taking advantage of substantial increases in electronic data and computational power
- Developing and testing technology using speech and text from English, Arabic, and Chinese news sources

TIDES has the following military impact:

- Enhancing the ability of U.S. forces to operate safely and effectively around the globe
- Enabling commanders and policymakers to know what is being said in a region by and to the local population
- Less dependence on scarce linguists
- Potential customers throughout the military and intelligence communities

Effective Affordable, Reusable Speech-To-Text (EARS) – developing automatic transcription technology whose output is substantially richer and more accurate than current methods. Research goals are to:

- Automatically transcribe and extract useful metadata from natural human speech
- Develop powerful statistical techniques for modeling variability of speech
- Take advantage of substantial increases in electronic data and computational power
- Develop and test technology using broadcasts and conversations in English, Arabic, and Chinese

The military impact of the EARS program includes:

- Substantial increases in productivity and situation awareness
- Enabling analysts to read transcripts rapidly (in lieu of listening to audio slowly)
- Many potential customers throughout the military and intelligence communities

In FY 2005, work will continue on the following DARPA efforts:

- *Compact Aids for Speech Translation (CAST)*
- *Improving Warfighter Information Intake under Stress*
- *Translingual Information Detection, Extraction and Summarization (TIDES)*
- *Effective, Affordable Reusable Speech-to-Text (EARS)*

HCI&IM NASA HCI&IM

NASA HCI&IM investments include fundamental research in:

- Human information processing and performance
- Multimodal interaction
- Human-robotic systems
- Automation and autonomy
- Software engineering tools
- Knowledge management and distributed collaboration
- Computational models of human and organizational behavior

In the aviation domain, for example, HCI&IM investments include:

- Human-computer interaction for highly automated vehicle cockpits (aviation/shuttle flight control systems; system health management systems; International Space Station [ISS] interfaces)
- Human-computer interaction for air traffic management applications (air traffic control; cockpit display of traffic information; distributed air-ground collaboration)
- Proactive management of system risk (air traffic management; onboard flight-recorded data; human and organizational risk)
- Human-computer interfaces for mission control and ground control applications (shuttle and ISS mission control; remote interactive automation for near-Earth and planetary missions)

In the information management domain, NASA missions require R&D in:

- Data mining and scientific discovery
- Interactive visualization of complex systems for analysis, design, and evaluation
- High-end modeling and simulation (vehicle geometry, functions, and behavior; climate, weather, and terrain; science planning and operations with automation support; homeland security)

In FY 2005, NASA plans to address the following technical challenges in HCI&IM:

- *Multimodal interaction*
- *Data visualization and understanding*
- *Human-in-the-loop supervisory control (the design of intelligent and intelligible automation; human-robotic interaction)*
- *Knowledge engineering and capture*
- *Risk management and technology (aerospace, Earth science, space flight, biological and physiological research, and space science)*

HCI&IM AHRQ HCI&IM

FY 2004 HCI&IM R&D investments are for improved patient safety and health quality. Specifically, these investments support the Patient Safety Health Information Technology (HIT) Initiative. They will support HCI&IM-related work for transforming health care quality (for example, promoting and accelerating the development, evaluation, adoption, and diffusion of IT in health care) and establishing health care data standards (e.g., development, evaluation, and adoption of information standards and technology to support patient safety). A major goal is to support the development of an NHII that includes an Electronic Health Record System (EHRS). The EHRS will be a longitudinal collection of electronic health information for and about persons. It will allow electronic access to person- and population-level information by authorized users. The system will provide knowledge and decision support that enhances the quality, safety, and efficiency of patient care. An EHRS will support more efficient health care delivery processes.

In the area of transforming health care quality, AHRQ provides:

- Planning grants to rural and small communities for health care systems and partners to implement HIT to promote patient safety and quality of care
- Implementation grants for rural and small hospitals to evaluate the measurable and sustainable effects of HIT on improving patient safety and quality of care
- Assessment grants to increase the knowledge and

understanding of the clinical, safety, quality, financial, organizational, effectiveness, and efficiency value of HIT. Practice-based research networks are encouraged to apply.

AHRQ envisions the establishment of a center that will coordinate the vast amounts of HIT information that results from these grants. Potential duties include coordinating and assisting grantees, supporting outreach and dissemination, and hosting meetings.

In health data standards, AHRQ provides:

- Funding for expert consensus on EHR functions
- Funding for standards development
- Funding for mapping between standards
- Funding for developing a metadata health data registry
- Coordination of U.S. standards developing organizations and standards users
- Consensus of U.S. positions on international health data standards issues

"Making the Health Care System Safer: The 3rd Annual Patient Safety Research Conference" is scheduled for September 26-28, 2004, in Arlington, Virginia. The goal of this conference series is to improve patient safety and health quality by:

- Increasing awareness of Patient Safety Events (PSEs) by supporting information systems that collect and analyze (including classification, root cause analysis, failure mode and effects analysis, probabilistic risk assessment) a large body of medical error data and by disseminating scientific findings that are translated into a useful decision-making context
- Understanding the organizational, cultural, human, and IT capabilities needed to guide the development of a network for the collection and analysis of PSEs and build capacity for patient safety research
- Developing informatics standards for uniform reporting, storage, and retrieval of comparable PSE data

HCI&IM DOE/SC HCI&IM

National Collaboratories Program – goal is to support end-to-end scientific discovery processes. Petabyte-scale experimental and simulation data systems will be increasing to exabyte-scale data systems. The sources of the data, the computational and storage resources, the scientific instruments, and the scientists who are consumers of the data and users of the instruments and computation are seldom collocated. DOE/SC uses planning workshops to develop its programs, including those reported in the HCI&IM PCA.

DOE/SC is developing pilot collaboratories for the high-energy nuclear physics, supernovae, and cell biology research communities. These pilot collaboratories are early implementations of virtual laboratories that:

- Are focused on a problem of national scientific or engineering significance clearly related to DOE mission and having high visibility
- Involve geographically separated groups of personnel and/or facilities that are inherently required to collaborate or be used remotely for success of the project
- Have implementations to test and validate that scientific research programs can integrate unique and expensive DOE research facilities and resources for remote collaboration, experimentation, simulation, and analysis through use of middleware technologies to enable ubiquitous access to remote resources – computation, information, and expertise
- Demonstrate capabilities that make it easier for distributed teams to work together over the short term and long term

Developing such collaboratories involves creating partnerships among researchers in scientific disciplines and R&D personnel in middleware, networking, and computational science. DOE/SC's middleware research has two HCI&IM efforts:

- Information integration and access that address:
  - Mechanisms for seamless interface to multiple storage systems and mechanisms for replication management
  - Open issues such as data provenance, metadata standards, and mechanisms to resolve global names into access mechanisms
- Services to support collaborative work that address:
  - A wide variety of community services such as the Access Grid

– Open issues such as scalable, easy-to-use primitives to support multiple modes of secure wide-area collaboration

A potential benefit is speeding up the multidisciplinary research process. Today data resulting from research in one field must be validated and published, then be discovered and understood by people in other fields. DOE/SC is looking to improve data pedigree and communications between groups of researchers.

The DOE/SC National Collaboratories Program activities reported in the HCI&IM PCA in FY 2004 have been moved to the LSN PCA for FY 2005 because this R&D is more aligned with LSN concerns.

HCI&IM NIST HCI&IM

Information Technology Laboratory (ITL)

The mission of ITL's Information Access Division (IAD) is to accelerate the development of technologies that allow intuitive, efficient access, manipulation, and exchange of complex information by facilitating the creation of measurement methods and standards. IAD achieves the objectives by contributing to R&D in these technologies; enabling faster transition into the commercial marketplace; and enabling faster transition into sponsors' applications (performance metrics, evaluation methods, test suites, and test data; prototypes and testbeds; workshops; and standards and guidelines). IAD works in collaboration with industry, academe, and other government agencies. HCI&IM-related programs and activities are:

Human Language Technology Program – includes the following FY 2004 efforts and evaluations:

- TREC (for Text Retrieval Conference) – work on novelty Web and Q&A, add robust retrieval, High Accuracy Retrieval from Documents (HARD), and genomics (retrieval within that domain)
- AQUAINT – continue dialogue testing, "what is" and "who is" questions
- TIDES/DUC – continue evaluations and add cross-language Arabic summarization
- TIDES/TDT – continue evaluations
- EARS – speech-to-text (STT) and metadata evaluations
- ACE – evaluations for foreign languages
- TIDES/MT – continue evaluations
- Meeting Transcription – implement STT evaluation and work with ARDA's Video Analysis and Content Extraction (VACE) program
- Speaker/Language Recognition – continue evaluations

Biometrics Technology Program – FY 2004 efforts include:

- Tests of FBI's Integrated Automated Fingerprint Identification System (IAFIS) on DHS and other data
- Tests to support conversion of US VISIT to more than two fingers (US VISIT is the entry/exit system being implemented by DHS at airports, seaports, and land border crossings to secure the Nation's borders)
- Integration of face recognition and multi-vendor verification into US VISIT
- Fingerprint vendor technology evaluation (FpVTE) testing and analysis
- Data and performance standards development
- Development of Multimodal Biometric Accuracy Research Kiosk
- ANSI and ISO standardization of the Common Biometric Exchange File Format (CBEFF) and the Biometric Application Programming Interface (BioAPI)
- Chair International Committee for IT Standards (INCITS) Committee on Biometrics (M1) and the ISO Committee on Biometrics (SC37)

Multimedia Technology Program – multimedia standards, video retrieval, and visualization and virtual reality for manufacturing. FY 2004 efforts include:

- Work with industry on use of MPEG-7 (the metadata standard for accessing video information) Interoperability Test Bed
- Development of "light-and-simple" profiles for existing industry-wide formats
- Video-retrieval evaluations, with a new story segmentation task and more high-level features to extract; process and ship 120 hours of new data

- X3D and Human Animation (H-ANIM) standardization
- Seek medical applications within Web3D and for human modeling

Interactive Systems Technology Program – activities include:

- Webmetrics – seek additional opportunities with World Wide Web organizations
- Industry Usability Reporting (IUSR, a standard by which software development companies can test results and provide results to potential customers) – extend Common Industry Format (CIF) to hardware and finalize CIF requirements
- Voting – R&D for developing usability and accessibility tests for certifying voting systems
- Novel Intelligence for Massive Data (NIMD) – experiment with glass box data and develop NIMD roadmap
- Human Robot Interaction (HRI) – develop a user interface for robotic mobility platform
- Digital Library of Mathematical Functions (DLMF) – usability study for DLMF web site and animation of Virtual Reality Modeling Language (VRML) worlds
- Accessibility standards – develop a prototype incorporating new architecture from the INCITS Committee on Accessibility (V2)

Smart Space Technology Program – includes a smart space testbed for human interaction in an environment that includes microphones and video cameras. FY 2004 work includes:

- New version of NIST's Smart Flow (a system that allows large amounts of data from sensors to be transported to recognition algorithms running on a distributed network) that is easier to use, provides better performance, and is more stable, portable, and scalable
- Upgrade Meeting Room with new capabilities (originally for speech, Meeting Room is now being used to analyze videos)
- INCITS V2 collaboration
- Provide data-acquisition services and analysis methods to NIST's Chemical Science and Technology Lab (CSTL) and Physics Lab and upgrade real-time processing systems

Manufacturing Engineering Laboratory (MEL)

MEL Manufacturing Systems Integration Division (MSID) – promotes development of technologies and standards that lead to implementation of information-intensive manufacturing systems that can be integrated into a national network of enterprises working together to make U.S. industry more productive. HCI&IM R&D includes work in system interoperability, based on ontologies (knowledge bases that provide the meaning of information held in databases, whether structured or unstructured, in ways that can be understood across systems and also by humans). Ontologies are important to the development of software systems in general, partly because such systems are often created today by integration of existing systems and partly because they can improve software engineering by carefully and formally defining the requirements in ways that can save money over the software life cycle and avoid hidden mistakes. In commerce, ontologies can address such problems as communication within machine shops, the distribution of higher level information both between factories and between factories and front offices.

MEL Intelligent Systems Division (ISD) – conducts the following HCI&IM activities:

- Development of test arenas for robotics competitions, particularly for urban search and rescue applications. Sites of 2004 competitions are New Orleans, Osaka, Lisbon, and San Jose.
- Study of methods and metrics for evaluating performance of intelligent systems. ISD hosted 5th annual workshop on Performance Metrics for Intelligent Systems (PerMIS) on August 24-26, 2004. Included new "Information and Interpretation and Integration Conference" (I3CON) centering on testing tools for alignment of different but overlapping ontologies.

In FY 2005, NIST plans to continue the following HCI&IM R&D activities:

- *Human Language Technology Program*
- *Biometrics Technology Program, including new work in:*
  - *Biometrics, used singly and in combinations, for forensics and security*
- *Multimedia Technology Program*
- *Interactive Systems Technology Program*
- *Cooperation with NIH in biomedical informatics and related areas*
- *MEL programs in manufacturing systems integration and intelligent systems*

HCI&IM NOAA HCI&IM

NOAA's current HCI&IM work involves digital libraries. The R&D is developing methods of cataloging, searching, viewing, and retrieving NOAA data distributed across the Web. Researchers are exploiting Geographical Information Systems (GIS) and XML technologies to display and describe data and are developing methods for distributing model data. Recent results include:

- Fisheries-oceanography GIS with 3D presentations
- Prototype tools for analyzing and integrating hydrographic data for a multi-server distribution system
- Using Really Simple Syndication (RSS) for service registration and information discovery

In FY 2005, planned NOAA work in HCI&IM will include:

- *Developing improved access to current National Weather Service (NWS) model data subsets*
- *Remote collaborative tools development*

HCI&IM EPA HCI&IM

EPA's HCI&IM programs are focused on tools to facilitate sound science using information management, analysis, and presentation. These programs enable relevant, high-quality, cutting-edge research in human health, ecology, pollution control and prevention, economics, and decision sciences. They facilitate appropriate characterization of scientific findings in the decision process and convey important information in such a way that researchers and policymakers can better understand complex scientific data. These programs are performed in-house and as problem-driven research. FY 2004 activities cover a broad spectrum of areas, including:

- Integrating data analysis and decision making across physical, chemical, biological, social, economic, and regulatory influences and effects
- Finding significant relationships, phenomena, differences, and anomalies with multiple data techniques
- Integrating search and retrieval across multiple digital libraries containing environmental and health information
- Enabling efficient management, mining, distribution, and archiving of large data sets
- Creating services to enable and support collaborative science across geographically distributed laboratories and offices including data and computational grids
- Developing interactive visualizations of complex systems for analysis, design, and evaluation
- Generating knowledge engineering and capture

FY 2004 HCI & IM activities have enabled the following pilot projects:

Metabonomics Pilot – a collaborative framework to enable shared access to data and results, joint use of computational applications and visualization tools, and participatory analysis across scientific boundaries. Metabonomics data will be harmonized with both proteomics and genomics data in order to better understand and predict chemical toxicity and support risk assessment. Activities on this pilot continue during FY 2005.

Air Quality Model Pilot – investigating tools and approaches to explore potential linkages between air quality and human health. Assessments include impacts of regulatory and policy decisions through exploration of relationships between regulations, emission sources, pollutants, and ecosystems.

Tools and techniques developed in this pilot will be transferred to states and applied to solve real problems such as enhancing states' abilities to predict and evaluate various control strategies, expanding state-level forecasting capabilities to include more pollutants, enhancing the states' abilities to explore potential linkages between air quality and human health, and improving states' abilities to assess the impact of regulatory and policy decisions. During FY 2005 HCI & IM activities will continue to enable this pilot.

FY 2005 EPA plans in HCI&IM R&D include:

- *Evaluation of tools and models through the EPA Grid*
- *Evaluation and investigation of distribution, management and archival of large data sets*
- *Investigation of significant relationships, phenomena, differences, and anomalies in data*

Large Scale Networking

Definition of LSN PCA

The activities funded under the NITRD Program's LSN PCA maintain and extend U.S. technological leadership in high-performance networks through R&D in leading-edge networking technologies, services, and techniques to enhance performance, security, and scalability. These advances will support Federal agency networking missions and requirements and will provide the foundation for the continued evolution of global scale networks and the services that depend on them. Today's networks face significant challenges in performance, scalability, manageability, and usability that restrict their usefulness for critical national missions. These challenges are addressed by the LSN PCA, which encompasses Federal agency programs in advanced network components, technologies, security, infrastructure, and middleware; grid and collaboration networking tools and services; and engineering, management, and use of large-scale networks for science and engineering R&D.

Broad areas of LSN concern

- The development of networks that deliver end-to-end performance at least 1,000 times greater than today
- The development of a manageable, scalable security system to support the evolution of the global Internet and associated applications that require trust
- The development of advanced network services and middleware that enable development of new generations of distributed applications
- Specialized large-scale networks to enable applications with special needs, such as sensor networks and networks for real-time control
- International cooperation in the area of optical network coordination and certificate authorities
- Outreach and cooperation with university and commercial communities
- Network science education and workforce issues

Technical Goals

- End-to-end performance measurement including across institutional and international boundaries
- Network trust, security, privacy, availability, and reliability through research on vulnerability analysis, scalable technologies, security testbeds, and promotion of best practices
- Collaboration and distributed computing environments including grid
- Adaptive, dynamic, self-organizing, self-scaling, and smart networking
- Revolutionary networking research including revisiting networking fundamentals

Illustrative Technical Thrusts

- Federal network testbeds and their coordination with non-Federal networks and testbeds
- Infrastructure support such as Federal Internet exchange points
- Network technologies, components, and services for optical, mobile, wireless, satellite, and hybrid communications
- Middleware to enable applications
- Sensornets
- Large-scale network modeling, simulation, and scalability
- Protocols and network services to support high-performance networking applications such as high energy physics data transfers, astronomy community collaboration, and biomedical research
- Technologies for measuring end-to-end performance

LSN Agencies

NIH	DOE/SC	AHRQ	DARPA	NOAA	Participating Agency
NSF	NASA	DOE/NNS	NIST	NSA	DoD/HPCMPO

LSN PCA Budget Crosscut

FY 2004 Estimate	FY 2005 Request
$329.8 M	$332.8 M

LSN PCA: Coordination and Activities

LSN Highlights

Agency representatives meet monthly in the LSN Coordinating Group to plan collaborative activities and evaluate Federal R&D in advanced networking technologies. In addition, the CG has established and works with specialized teams to implement agreed-upon policies and to provide interagency coordination. The LSN teams and their roles are:

- Joint Engineering Team (JET) – coordinates connectivity and addresses management and performance of Federal research networks and their interconnections to the commercial Internet
- Middleware and Grid Infrastructure Coordination (MAGIC) Team – coordinates Federal agency grid and middleware research and development and support to grid applications
- Networking Research Team (NRT) – coordinates Federal agency research programs in advanced networking technology and protocols

Annual Planning Meeting

The LSN CG holds a yearly Interagency Coordination Meeting at which each agency describes its programs reported under the LSN PCA. This allows the agencies to identify:

- Opportunities for collaboration and coordination with other agencies on networking R&D activities
- LSN R&D focus areas
- Areas of interest to the LSN agencies that are outside or larger than the stated scope of the LSN CG, such as cybersecurity

At the November 3, 2003, planning meeting, participants identified the following three overarching areas in which most of the LSN agencies have significant efforts and interests that would benefit from cooperation:

- Network testbed infrastructure
- Network performance measurement
- Grid outreach

Agencies also developed a broader list of networking R&D areas that are currently of interest to LSN participants. The list appears below. In optical networking, for example, NSF and DOE/SC are directly coordinating their work in building new experimental optical network testbeds to assure that the networks will be complementary, offering cross-network connectivity, experimentation capability, and advanced infrastructure for each agency's community of researchers.

Network Testbed Infrastructure

Several high-performance testbeds are being built to provide an infrastructure for networking researchers within the Federal agencies, some at the national level and others called Regional Operational Networks (RONs). The national networks include:

- NSF's Experimental Infrastructure Networks (EINs)
- DOE/SC's UltraScience Net
- National LambdaRail (NLR)

The RONs include:

- California Research and Education Network 2 (CalREN 2)
- I-Wire (Illinois)
- Mid-Atlantic Exchange (MAX)

Continued on next page

Agencies' Current Networking R&D Interests

- Networking Research
  - Basic technology
  - Optical networks
  - Services
  - Applications
- Infrastructure
  - Production
  - Experimental
  - Research
- Collaboration Support
  - Middleware and Grid
  - Collaboration
- Security
- Monitoring and Measurement
- Automated Resource Management
- Wireless, Ad Hoc
  - Wireless/nomadic
  - Crisis response, CIP
  - Sensornet
- Standards and Specifications
- Education and Training

There is an opportunity to coordinate connectivity, research, and applications development for these networks to address larger-scale issues such as transparency across the networks. However, controlled-use networks or dual-use networks will be required to assure that researchers doing research for network development (with a high risk of bringing down the network) do not interfere with applications researchers who require highly reliable networks.

The LSN CG tasked the NRT to study the issue of coordination of network testbed infrastructures and report back on its recommendations.

Network Performance Measurement

Many LSN agencies have active network performance measurement programs and others identified the need to measure end-to-end performance for all points of the network. Critical issues include development of standard interfaces and architectures for network performance measurement. This would enable comparisons of performance measurement across network providers to facilitate end-to-end performance measurement, which are needed in identifying and eliminating network bottlenecks and isolating network faults when they occur. The ability to measure and compare performance across network provider boundaries and to isolate network faults is needed for optimizing end-to-end application performance to take advantage of the increasing network bandwidths.

The LSN CG tasked the JET to investigate mechanisms of coordinating network performance measurement, particularly standard interfaces and architectures, and to provide recommendations on how to improve performance measurement capabilities.

Grid Outreach

Many user and application communities are developing grid capabilities, often focused on their particular needs. The LSN agencies are fostering the development of standardized grid middleware resources, tools, architectures, and infrastructure to facilitate such grid-enabled applications and are encouraging outreach to the user and application development communities so these users can:

- Benefit from the extensive grid resource base and tools that currently exist and are being developed
- Avoid balkanizing grid resources by relying on standardized tools, middleware, and architecture rather than unique discipline-specific grid capabilities
- Benefit from the common policies, resource infrastructure, and the Global Grid Forum (GGF). The GGF is a community-initiated forum of thousands of individuals from government, industry, and research leading the global standardization effort for grid computing.

The LSN CG tasked the MAGIC Team to develop a strategy for outreach to grid user communities to inform them of existing and developing tools, architectures, and capabilities, to promote commonality, and to report back.

Possible FY 2005 LSN Coordination

The LSN CG has identified the following three additional topics in which the LSN agencies have a significant commonality of interest and which hold potential for future collaborative activities:

- Autonomic networking
- High-speed transport
- Security

Current High-Speed Networks

Following is a list of all the networks, called "JETnets," across which LSN and its technical teams foster coordination and collaboration. The JETnets are Federal, academic, and private-sector networks supporting networking researchers and advanced applications development. They are:

- Advanced Technology Demonstration Network (ATDnet), established by DARPA and including the Defense Intelligence Agency, NASA, the Naval Research Laboratory, NSA, and Bell Atlantic
- DARPA's BOSSnet and SuperNet
- DoD/HPCMPO's DREN
- DOE/SC's ESnet and UltraScience Net
- NASA's NREN and NISN
- National LambdaRail (NLR)
- Next Generation Internet Exchange Points (NGIXs)
- NSF's EINs (now called DRAGON and CHEETAH) and vBNS+
- StarLight, the international optical peering point in Chicago
- UCAID's Abilene

LSN R&D Programs By Agency

Selected FY 2004 Activities and FY 2005 Plans

LSN **NIH** LSN

Healthcare and the Next Generation Networking program – funds testbeds demonstrating advanced networking capabilities such as Quality of Service (QoS), security and medical data privacy, nomadic computing, network management, and infrastructure technology as a means for collaboration.

Applications of Advanced Network Infrastructure Technology in Healthcare and Disaster Management – funds work on applications that demonstrate self-scaling technology; use self-optimizing end-to-end network-aware real-time technology and/or middleware; depend on wireless technology; involve advanced authentication (biometrics or smartcards); and are nomadic or use GIS technology.

NLM Scalable Information Infrastructure Awards – supporting the following 11 testbed projects in FY 2004:

- Scalable Medical Alert and Response Technology (SMART)
- An Adaptive, Scalable, and Secure Internet2-Based Client Server Architecture for Interactive Analysis and Visualization of Volumetric Time Series Data
- National Multi-Protocol Ensemble for Self-Scaling Systems for Health
- Project Sentinel Collaboratory
- Advanced Health and Disaster Aid Network (AID-N)
- Advanced Network Infrastructure for Distributed Learning and Collaborative Research
- Advanced Network Infrastructure for Health and Disaster Management
- Wireless Internet Information System for Medical Response in Disasters (WIISARD)
- Advanced Biomedical Tele-Collaboration Testbed
- A Tele-Immersive System for Surgical Consultation and Implant Modeling
- 3D Telepresence for Medical Consultation: Extending Medical Expertise Throughout, Between and Beyond Hospitals

In FY 2005, NIH plans to:

- *Increase its investment in datagrid infrastructure including middleware and other software development for support of biomedical and clinical research activities using federated databases*
- *Begin exploring optical networking testbeds for data sharing between sites engaged in clinical research and for testing different approaches to conducting distributed queries against federated medical databases*
- *Expand use of wireless networks in clinical research environments*

Computing and networking form a continuum of support services for applications on desktop to high-end platforms. Elements of the continuum include distributed computational resources, clusters of computers, grid infrastructure and applications, and the Extensible Terascale Facility [ETF].) The support networking ranges from high-performance, end-to-end, and fine-grained to episodic network services using optical, fiber, wireless, satellite, and hybrid network links. Program areas and illustrative awards are listed below.

Networking Research Testbed (NRT) program – supports prototyping and benchmarking as part of networking research. The Computer and Network Systems (CNS) Division manages this program. The NRT program supports research in:

- Disruptive technologies and approaches
- Hybrid and experimental designs
- End-device research
- Core technology development
- New protocol research
- Alternative network architectures
- Testbed implementations

Awards were made in FY 2003, including:

- A unified experimental environment for diverse networks
- Testing and benchmarking methods for future network security mechanisms
- Orbit: Open-access research testbed for next generation wireless networks
- Agile and efficient ultra-wideband wireless network testbed for challenged environments
- Heterogeneous wireless access network testbed
- Scalable testbed for next generation mobile wireless networking technologies
- National radio network research testbed (NRNRT)

Experimental Infrastructure Network (EIN) – supports research applications with high-performance networking. EIN complements the NRT program. Some EIN focus areas are:

- Control/management of the infrastructure end to end
- End-to-end performance/support with dedicated provisioning
- Pre-market technologies, experimental hardware, and alpha software
- Significant collaboration vertically and across sites
- Persistence, with repeatable network experiments and/or reconfigurability
- Experimental protocols, configurations, and approaches for high throughput, low latency, and large bursts

EIN projects include: dynamic resource allocation for Generalized Multi-Protocol Label Switching (GMPLS) optical networks (DRAGON); end-to-end provisioned optical network testbed for large-scale e-science applications (CHEETAH, for Circuit-switched High-speed End-to-End Transport ArcHitecture); cyber defense technology experimental research network (DETER); PlanetLab, an overlay testbed for disruptive network services; and WAN-in-Lab.

NSF Middleware Initiative (NMI) – program's goal is to design, develop, deploy, and support a set of reusable, expandable middleware functions and services that benefit applications in a networked environment. In FY 2003, 20 System Integrator (SI) Awards were made to further develop the integration and support for long-term middleware infrastructure. Ten other awards focused on near-term capabilities and tool development.

Funded projects include: disseminating and supporting middleware infrastructure – engaging and expanding scientific grid communities; designing and building a national middleware infrastructure (NMI-2); integrative testing framework for grid middleware and grid environments; extending integrated middleware for collaborative environments in research and education; instruments and sensors as network services; middleware for grid portal development

Strategic Technologies for the Internet (STI) – FY 2003 theme areas are: complex network monitoring, problem detection, and resolution mechanisms; applications that promote collaborative research and information sharing; networked applications tools or network-based middleware;

development of automated and advanced network tools; and innovative access network technologies. Awards include:

- A security architecture for IP telephony
- Network Startup Resource Center (NSRC)
- Plethora: A wide-area read-write object repository for the Internet
- Marist Grid collaboration in support of advanced Internet and research applications
- Implementation of a handle/DNS server
- Efficient diagnostic strategies for wide-area networks (tools and technologies to detect, diagnose, and correct network faults in local broadcast domains, whether they are wired (Ethernet) or wireless)
- Development of an infrastructure for real-time super media over the Internet
- Viable network defense for scientific research institutions
- The Strategic Technology Astronomy Research Team (START) collaboratory: Broadening participation in authentic astronomy research
- Network measurement, monitoring, and analysis in cluster computing
- Toward more secure inter-domain routing
- eXplicit Congestion Control Protocol (XCP) development – potential transport protocol for high-performance network environments
- Media-aware congestion control
- Self-organizing spectrum allocation

NSF International Connections – include these initiatives receiving support in FY 2004:

- TransPac award supports networking to the Asia-Pacific region with two OC-12 Packet Over SONET (POS) links. It supports a 1 GigE circuit between Chicago and Tokyo
- Starlight project provides a 10 GigE facility in Chicago to serve as a peering point for international high-performance optical research networks including links to CERN and Amsterdam
- NaukaNet provides OC-3 links to Moscow and Hong Kong

* AmPath provides OC-2 service from Miami to NAP (Network Access Point) of the Americas
* TransLight program supports additional optical networking connectivity to Canada (Canarie2 network), Prague, Stockholm, and London

Information Technology Research (ITR) Initiative – supports research in theme areas that cut across science disciplines and NSF division interests. ITR research areas that are part of the LSN PCA include:

- Cybertrust: projects in operating securely and reliably and assuring that computer systems protect information
- Education and workforce

Large FY 2003 awards related to LSN include:

- Sensitive Information in a Wired World
- Responding to the Unexpected
- Networked InfoMechanical Systems (NIMS)
- 100 Megabits per Second to 100 Million Households

Extensible Terascale Facility (ETF) Connectivity – networks connect computation and storage components at each ETF site and connect the ETF sites with one another. The sites currently are, or soon will be: ANL, CalTech, IU, NCSA, ORNL, PSC, Purdue University, SDSC, and TACC.

NSF's LSN plans for FY 2005 include:

- *NSF Middleware Initiative (NMI) – will continue to address middleware deployment challenges and develop common enabling middleware and domain-specific cybertools for grid-computing*
- *International research connections – new investments will be made to enable international science and engineering collaborations*
- *Programmable Wireless Networks – capabilities of programmable radios to make more effective use of the frequency spectrum and to improve wireless network connectivity will be exploited*
- *Networking of Sensor Systems – architectures, tools, algorithms, and systems that will make it easy to assemble and configure a network of sensor systems will be created*
- *Internet Architecture – the core architecture of the Internet will be reexamined, because there are signs that the current IP-based architecture cannot handle expected increases in communication loads. This requires devising means to test and eventually deploy*

LSN **DOE/SC** LSN

Applications-related drivers of the DOE/SC scientific research community's growing needs for high-performance network environments are the following:

- Petabyte-scale experimental and simulation systems will be increasing to exabyte-scale data systems in such areas as bioinformatics, climate, and the Large Hadron Collider
- Computational systems that process or produce data continue to advance with Moore's Law
- Network requirements are projected to double every year.
- The sources of the data, the computational resources, and the scientists are seldom collocated

The result is increasing demands on DOE/SC's network. For example, in five years DOE/SC scientists are expected to need the network to:

- Transfer three petabytes per year
- Move a petabyte of data in 24 hours for the high energy nuclear physics community using secure file movement over a 160 to 200 Gbps "best effort" service
- Provide the cell biology community with shared immersive environments that include multicast and latency and bandwidth guarantees using 2.4 Gbps links with strong QoS guarantees
- Support real-time exploration of remote data sets using secure remote connectivity over 1 to 10 Gbps links to the desktop with modest QoS services
- Support experimental (non-production) networking and grid research using unused wavelengths

DOE/SC works with its users to identify mid- and long-term network requirements. Its Office of Advanced Scientific Computing Research (OASCR) is building partnerships among its networking programs, its researchers, and users of its facilities as well as with other LSN agencies. Other DOE/SC priorities are program integration and priority setting for its facilities and research.

UltraScience Net – new FY 2004 initiative to support a breakable, schedulable, ultra-high-speed, all-optical network for conducting networking research. Its technologies rely on fundamental hardware research results of other LSN agencies. The first UltraScience Net link, 10 GigE from Oak Ridge National Laboratory (ORNL) to StarLight, was put in place in early 2004. Linking to Sunnyvale,California, CERN, and Lawrence Berkeley Laboratory (LBNL) is being done cooperatively with NLR. This network will support end-to-end scientific discovery. Towards that end, DOE/SC is supporting R&D in the following areas:

Networking research – addresses end-to-end network measurement and analysis for applications to enable localization and elimination of bottlenecks, effective tuning of the network path for an application, and resolution of performance issues across provider boundaries. Developing new protocols to support large data transfers; cybersecurity policy and technology advances required for OC-12 to OC-192 speeds; QoS, and dynamic provisioning services.

Middleware research – focuses on: providing transparent and scalable cybersecurity infrastructure; developing services to support collaborative scientific work; integrating remote resources and collaboration capabilities into local experimental and computational environments; and facilitating the discovery and use of scientific data, computers, software, and instruments over the network in a controlled fashion. (The National Collaboratories Program-related research activities that are reported in HCI&IM in FY 2004 have been moved to the LSN PCA for FY 2005 because they are more aligned with LSN concerns).

ESnet – network facilities and testbed activities include maintaining the high-performance production ESnet for DOE/SC mission requirements. DOE/SC has as a goal to transition the UltraScience Net prototype to a production environment and to tighten the coupling of the prototype and ESnet.

Connectivity – developing IPv6 protocol implementation, multiplatform video conferencing services, and implementation of distributed network measurement and analysis tools, in coordination with the NSF cyberinfrastructure program and other LSN agency program managers

DOE/SC plans for LSN efforts in FY 2005 include:

- *Completion of UltraScience Net deployment and initiation of testbed activities*
- *Implementation of fiber ring metropolitan area networks in the Bay Area and Chicago to improve connectivity of ESnet core to DOE laboratories and its reliability*
- *Continue network research, grid middleware, and collaboratory pilot projects in coordination with related efforts at NSF and other agencies*
- *Participate in interagency network measurement collaboration*

LSN NASA LSN

NASA Research and Education Network (NREN) adaptive networking program – developing a tool called Resource Allocation, Measurement and Adjustment System (RAMAS) to reserve bandwidth for applications, either when requested or later. RAMAS features include: passive monitoring and measurement; data capture of OC-3 and OC-12 data streams; graphical user interface (GUI) enabling user-defined specifications to designate traffic of interest and types of information to collect on designated flows; data-collection summarization delivered to a central process/control management system that archives summaries and generates graphical representations; interfaces for Ethernet, ATM, POS, and wireless.

A RAMAS experiment captured low-bandwidth wireless data using secure FTP from a field site connected by wireless to a satellite link to mass storage facilities at Ames Research Center and the University of Cincinnati. The test identified the difficulty of optimizing TCP for satellite transmissions.

In FY 2004, work on RAMAS includes: upgrades to OC-48 and OC-192; enhanced data archive capabilities; enhanced data analysis to reduce massive amounts of data to information; security mechanisms to enable use on the NASA Grid (previously called the Information Power Grid); a distributable RAMAS package, possibly on a CDROM; and identification of performance bottlenecks for selected wireless applications.

Nomadic networking: Portable Satellite Dishes – goal is to enable NASA science and engineering in remote locations. Three components: ad hoc networking, hybrid satellite and terrestrial networking, and mobile IP/IPv6. FY 2004 activities include:

- Implement dynamic source routing, an ad hoc network routing protocol
- Investigate use of directional antennas
- Use portable satellite dishes to demonstrate mobile IP in a hybrid environment
- Evaluate mobile IP/IPv6
- Deploy a portable Ka-band (radio spectrum from 18 GHz to 31 GHz used in satellite communications) satellite dish in collaboration with HP Labs

Mobile Agents Field Experiment – developing techniques for planetary exploration, including command of roving vehicles, wireless on-site networking, and autonomous software testing. Experimental components will include the NREN Transportable Earth Station (TES), a rover, all-terrain vehicles, an astronaut, and a remote mission support team at NASA Johnson Space Center.

Geology ground-truthing experiment – goal is to calibrate satellite-derived geology data with in-situ data taken in real time by on-site geologists. Uses wireless networking and the TES to enable researchers to use mass storage facilities, supercomputers, and the grid to conduct experiments while they are on site.

In one experiment, real-time satellite spectro-radiometer data were transported to a mass storage facility while scientists in Utah uploaded ground spectra data to a second facility. The grid was used to move both data sets to supercomputers at NASA Glenn and NASA Ames Research Centers for analysis. The results were accessed by local scientists and sent to the remote science team, which used the results to locate and explore new critical compositions of interest.

Wide Area Network (WAN) testbed – connects five NASA centers and two exchange points, using Asynchronous Transfer Mode (ATM) and Packet Over SONET (POS) OC-12 and OC-3 circuits. For research and development of technology to enable next-generation NASA missions, demonstrate NASA applications, and transfer technology to enhance the capabilities of NASA operational networks. In FY 2004, MPLS, IPv6, and Differentiated Services (DiffServ) will be deployed on the NREN backbone. The PCMon network monitoring tool will be integrated on the TES platform.

LSN **DARPA** LSN

Network modeling and simulation – Information Processing Technology Office (IPTO) program supports projects to:

- Develop modeling and simulation tools for online measurement and analysis with the goals of designing better protocols and new services such as for dynamic provisioning
- Predict end-to-end performance and vulnerabilities through traffic pattern analysis
- Dynamically optimize performance

In an effort co-funded with NSF, researchers achieved the largest network simulation – of a 1.1 million-node network – to date. The simulation used 1,500 parallel processors, illustrating the use of parallel processing to improve the scale of network simulations and advance the modeling of network behavior. The simulation results are being used to design better, faster protocols.

A DARPA fast-throughput experiment demonstrated 8.6 Gbps throughput using ten simultaneous flows.It attained 34 petabyte meters per second on a path from CERN through StarLight in Chicago to Sunnyvale, California. This metric attempts to capture the interrelationship between bandwidth and delay.

Cognitive Networking – will assess the feasibility of information and communication networks that possess significant degrees of self-reliance and responsibility for their own behavior and survival. Focuses on self-diagnosis, automatic adaptation to changing and hostile environments, reconfiguration in response to changes in the environment, intelligent negotiation for tasks, and resources and robustness under attack. Will also explore the possibility of a virtual "application-private network," whose on-demand protocols are based on specific application requirements and current network conditions.

Self-Aware Collective Systems – this technology thrust will enable heterogeneous teams of individuals (e.g., people, software agents, robots) and/or organizations (e.g., coalition forces) to rapidly form, easily manage, and maintain virtual alliances concerned with a specific task. Thrust involves two efforts:

- Self-Aware Peer-to-Peer Systems – will develop resilient, scalable sensor/computation networks with decentralized control. This technology will support battlespace awareness by enabling the self-formation of large ad hoc networks of sensors and computational elements within the severely resource-constrained environment (power, bandwidth, stealth) of military operations while enabling networks to survive component failure, network intrusion, and the subversion of elements.
- Collective Cognitive Information Processing for Improved Asset Performance – will develop learning and reasoning algorithms that can identify and classify emergent problems and opportunities for proactive maintenance of equipment and use of sensors in a dynamic operational environment. These new self-aware distributed systems will be able to reflect globally on their overall operation (including understanding trends), and make decisions based on the collective disposition of assets connected by networked sensors (e.g., vehicles or other equipment).

In FY 2005, DARPA work will continue on:

- *Cognitive Networking*
- *Self-Aware Collective Systems*

LSN DOE/NNSA LSN

Distance Computing (DisCom) program – a component of NNSA's Advanced Simulation and Computing (ASC) Program, formerly the Accelerated Strategic Computing Initiative (ASCI). DisCom assists in the development of an integrated information, simulation, and modeling capability to support the design, analysis, manufacturing, and certification functions of DOE's defense programs complex through advances in distance computing. Extends the environments required to support high-end computing to remote sites. Provides remote access to major ASC platforms – Red, Blue Mountain, White, and Q.

In FY 2004, DisCom is delivering additional key communications technologies to efficiently integrate the ASC platforms of Red Storm, Purple, Blue Gene/L, and beyond. Supporting tools development for more reliable and persistent file-transfer mechanisms over the WAN.

DisCom cooperates and collaborates with NSA to closely monitor the IP encryptor technology.

In FY 2005, LSN activities of DOE/NNSA will include:

- *Begin deployment of new tools for reliable, persistent file transfer over the WAN*
- *Continue development of integrated distance computing capabilities for ASC*

LSN NIST LSN

NIST supports network-related programs in its Advanced Network Technologies Division (ANTD) and in the Computer Security Division (CSD) of its Information Technology Laboratory (ITL). Their mission is to provide the networking industry with the best in test and measurement research. Their goals are to improve the quality of emerging networking specifications and standards and to improve the quality of networking products based on public specifications. Their core technical contributions are:

- Models and analyses from specifications to assess consistency, completeness, precision, and performance characteristics
- Prototypes and empirical studies from specifications to determine feasibility
- Test and measurement tools, techniques, metrics, and data to assess conformance, interoperability, and performance

ANTD research programs include:

- Pervasive Computing including Wireless Personal Area Networks (WPAN), service discovery, wireless access networking, wired access, core networks, and wireless ad-hoc networking. Focus areas are UltraWide Band (UWB) and grid computing.
- Agile Switching: routing, signaling, protection, restoration, and network management and metrology
- Internet Telephony including session-initiated protocol (SIP) for voice to support nomadic information appliances, particularly for the health care industry
- Infrastructure Protection for securing core services, secure BGP, and survivable control planes
- Wireless Ad Hoc Networks for architecture, routing, and services, particularly standards, and first responder technologies
- Cryptographic: standards and quantum encryption keys

Details of some projects follow.

FY 2004 priority taskings – evaluating the economic consequences of IPv6; planned a January 2004 workshop on "spam" at the request of the White House

Resilient Agile Networking (RAiN) program – developing, testing, and standardizing technologies for fault-tolerant, self-adaptive and ad-hoc networks. RAiN projects include: survivable control planes (work with the Internet Engineering Task Force (IETF) and the Network Processor Forum to define security mechanisms that will scale to performance levels necessary for the core Internet infrastructure); large-scale BGP attack modeling to characterize the performance impact and benefits of various proposed approaches to secure BGP; new measurement platforms; wireless ad hoc networks (work with public safety agencies to develop requirements and standards for emerging public safety communication and networking technologies)

Self-Managing Systems – conducting research and developing the test and measurement basis and standards foundations for future self-managing systems. Working with DARPA, the Global Grid Forum, Distributed Management Task Force (DMTF), IETF, the Institute of Electrical and Electronics Engineers (IEEE), IBM, Sun Microsystems, Cisco, HP, and others, NIST is: exploring information and knowledge management, learning, and reasoning techniques to enable new levels of autonomic operation of increasingly complex distributed systems; establishing the measurement basis to assist industry in evaluating the behavior and performance of emerging self-managing systems.

Information Security – focuses on cryptographic standards and applications, such as: secure encryption, authentication, non-repudiation, key establishment, and pseudo-random number generation algorithms; standards and guidance for e-Gov and e-Authentication; PKI and Domain Name System (DNS) security standards, interoperability, assurance, and scalability; wireless and mobile devices and smart card security; quantum computing and quantum cryptography (in coordination with DARPA)

GLASS: GMPLS/Optical Simulation Tool – developing a modeling tool to evaluate architectures and protocols for routing, signaling, and management of GMPLS for optical networks and to support multilevel, multiprotocol schemes for traffic engineering, QoS, protection, and restoration. Also developing a modular simulation framework to support: optical network restoration; MPLS DiffServ protocol recovery; GMPLS and optical common control and measurement plane (CCAMP); and a protection and restoration tool to provide multilevel recovery schemes.

In FY 2005, NIST plans in LSN R&D include:

- *Continue work in RAiN program, self-managing systems, cryptographic standards for information security, and the GLASS project*

LSN NOAA LSN

Advanced Information Technology Program – R&D strategy to:

- Explore advanced technologies to make NOAA's vast amount of environmental data available easily, quickly, and completely
- Exploit innovative data access technologies including Web-based tools and agents as they evolve
- Explore emerging security technologies to safeguard NOAA data and information
- Transfer new technology to other NOAA users

NOAA is implementing the strategy through:

- Real-time collaborations such as Internet@sea and Ocean Share
- Seasonal-Interannual Climate Collaboration: Distributed collaboration visualizing the environment
- Access to Data: Satellites, radar, aircraft, in situ, and models
- Fisheries Model Analysis such as for Pollack larvae in Shelikof Strait in Alaska
- Hazardous Spill Response: Anytime, anywhere connectivity

Related NOAA R&D includes:

Computer and network security – projects address: encrypted file systems; firewalls and intrusion detection at gigabit speeds; automated, enterprise-wide patching; and wireless security.

Next Generation Internet program – goal is to use advanced networking technologies to enhance NOAA data collection and dissemination and to support the NOAA HPC architecture. Provides connectivity among seven NOAA sites and to radar sites providing near-real-time weather data. Currently developing phased-array radars (NEXRAD) to collect near-real-time weather data. Deployment of these radars will significantly increase data transport requirements.

Discipline-specific user toolsets – being developed to support collaboration and grid applications. Plans to explore the use of grid systems for data handling, computation, and collaboration, and to test deployment of IPv6 at three NOAA sites.

FY 2005 NOAA plans include:

- *Continue developing advanced networking capabilities*
- *Optical networking*

- *Implement additional network-enabled applications as they become practical, including:*
  - *Collaboratories*
  - *Grid computing*
  - *Storm-scale simulations with immersive environments*
  - *Wireless: Data anywhere, anytime*
  - *Remote operations*

LSN NSA LSN

Laboratory for Telecommunications Sciences (LTS) – programs continue to emphasize transmission of quantum communications through optical elements. This work is part of a joint research program with DOE/SC. Demonstrated a quantum channel on the same fiber as a SONET channel.

In FY 2003, NSA demonstrated techniques for regionalized QoS management and developed a model for QoS pricing of Internet services. In FY 2004 NSA is working to join these two efforts to provide a viable pricing model and requisite enforcement mechanisms. Also continuing are programs on Critical infrastructure protection issues for converged networks. Work slowed in nonlinearities and transients in optical networks, regionalized quality of service management, and firewalls in high-speed networks. Work was phasing out in optical burst-switched protocols.

In FY 2005, NSA will continue network research in:

- *Quantum communications, quality of service, and high- speed network interfaces*

Participating Agency

LSN HPCMPO LSN

Defense Research and Education Network (DREN) – maintains a production WAN, using MCI services, providing DS-3 through OC-768 links and deploying GigE and 10GigE services. This network supports IP, IPv6, and multicast with increased security. It supports both Continental U.S. (CONUS) and Outside the Continental U.S. (OCONUS) sites. It provides OC-3 and OC-12 peering to OCONUS. Core network expansion by MCI includes OC-192 in June 2004 and OC-768 when required. Key network services provided:

- IP performance: latency, packet loss, and throughput
- ATM availability
- IPv6 and multicast
- Increased security

FY 2004 activities include:

- Connectivity: to the Advanced Technology Demonstration network (ATDnet), StarLight, and other experimental networks to support development of new technologies and services; to remote sites in Alaska and Hawaii, in collaboration with NASA

- IPv6 pilot program: DREN deploys IPv6 native and supports the Abilene IPv6 Transit Access Point (6TAP) testbed and the IPv6/MPLS end-to-end testbed. Services include: efficient routing using protocols such as BGP+, GMPLS, and OSPF; QoS; priority, preemption, policy, authorization, and audit; voice, video and data support; dynamic network monitoring; scalable speeds, including network test and measurement capabilities and passive monitoring with time-stamping.

- Performance measurement: testing 10 GigE equipment. In FY 2004, DREN is able to achieve near wire-speed performance over 80 percent of the time across its OC-12 links.

- Network security: providing end-to-end WAN encryption; developing IPv6 security with beta-testing of NetScreen IPv6 currently in-place. It is considering NS-500 firewall protection. (NS-500 is a GigE security device that uses a Virtual Private Network [VPN].)

Software Design and Productivity

Definition of SDP PCA

The activities funded under the Software Design and Productivity (SDP) PCA will lead to fundamental advances in concepts, methods, techniques, and tools for software design, development, and maintenance that can address the widening gap between society's need for usable and dependable software-based systems and the ability to produce them in a timely, predictable, and cost-effective manner. The SDP R&D agenda spans both the engineering components of software creation and the economics of software management across all IT domains, including the emerging areas of embedded systems, sensor networks, autonomous software, and highly complex, interconnected systems of systems. Today, software development and maintenance represent by far the most costly, time-consuming, labor-intensive, and frustrating aspects of IT deployment in every economic sector. SDP R&D seeks to foster a new era of innovation in software engineering that addresses these serious design issues.

Broad Areas of SDP Concern

- Overall quality of software
- Overall cost – in time, labor, and money – of software development and maintenance
- Growing complexity of software
- Enabling more people to more easily create software and software-based systems
- Need for expertise in emerging software areas such as embedded systems, large-scale sensor networks, autonomous software, and grid environments
- Workforce development

Technical Goals

- Scientific foundations for creating, maintaining, and improving software that incorporates such qualities as usability, reliability, scalability, and interoperability
- More cost-effective methods for software testing, analysis, and evaluation
- New frameworks for understanding and managing the economics of software
- Methods for developing and evaluating specialized software systems for specific purposes and domains

Illustrative Technical Thrusts

- Development methodologies such as model frameworks; tunable, adaptable processes; component technologies; open source practices for code portability and re-use; integrated environments and tools for development; and programming standards
- Theoretical and technical aspects of programming languages; compilers; software for visualizing and verifying code and data structures; and automatic program synthesis
- Software architectures and component-based methods to incorporate re-usable software resources
- Measuring and managing software quality
- Scalability: Enhancing the ability of software to "grow" along various axes without significant redevelopment
- Interoperability: Software that moves easily across heterogeneous platforms as well as software units that cooperate and exchange information seamlessly
- Fault-tolerant, adaptable software that can continue to function under unexpected conditions

SDP Agencies

			Participating Agency
NASA	DOE/NNSA	NIH	
NSF	NIST	NOAA	FAA

SDP PCA Budget Crosscut

FY 2004 Estimate	FY 2005 Request
$179.3 M	$166.0 M

- Enhancing software performance through code modeling and measurement; understanding how to improve runtime effectiveness of applications
- Techniques and tools - including automated approaches, new metrics, and reference data - for testing, analysis, validation, and verification
- Techniques and tools for assessing engineering risks and tradeoffs and estimating and balancing developmental and operational costs
- Specialized and domain-specific software such as for application frameworks, platform reliability, and modeling or control of embedded systems

SDP PCA: Coordination and Activities

SDP Highlights

Since the inception of the SDP PCA in FY 2000, its Coordinating Group (CG) has worked on developing a common understanding of the SDP R&D terrain. In this process, it has worked with the HCSS CG to make clear the research distinctions between their two domains (see HCSS definition on page 64). These efforts contribute to a broader understanding by the policymaking community and the public of software issues and the critical need for R&D that addresses them.

In December 2001 at the outset of its deliberations, the SDP CG sponsored a workshop at which researchers from academe, industry, and government were encouraged to "think out of the box" about "New Visions for Software Design and Productivity." The participants identified five major facets of the growing problems in contemporary software design and development:

- The large number and inherent complexity of requirements
- Multiple independent sets of requirements and their associated functionalities
- Testing, verification, and certification of large, complex systems
- Composition and integration of software from multiple sources
- Multidisciplinary distributed software development

Building on that workshop's conclusions, in FY 2004 the SDP CG will complete a taxonomy of the PCA's scope of research. The preliminary taxonomy appears

Continued on next page

Scope of SDP R&D Topics

Create, maintain, and improve software

- Software development methods
  - Model development framework
  - Tunable development processes
  - Component technologies
  - Open source
  - Development environments
  - Programming standards
  - Adaptable systems
  - Legacy systems
- Languages and tools
  - Programming languages
  - Cross-platform Fortran 90
  - Compilers
  - Programming environments
  - Software visualization
  - Automatic program synthesis
- Runtime environment

Investigating and achieving the "-ilities"

- Quality management
- Platform reliability
- Scalability
- Interoperability
- Robustness
- Performance tuning

Testing, analysis, and evaluation

- Evaluation of complex integrated systems with real-time characteristics
- Evaluation of models and simulations
- Evaluation of COTS
- Metrics and reference data
- Automatic verification and validation (V&V)
- Testing tools
- Software analysis

Management and economics

- Tools for various aspects of software process management
  - Project management
  - Cost and schedule estimation and prediction
  - Cost of testing
  - Document management systems
  - Quality vs. risk measurement and management

Software systems for specific domains and specialized software

- Application frameworks
  - Domain-specific tools etc.
  - Specialized software products
  - Large-scale data and information access, sharing, security, etc.
  - Scientific data management
  - Visualization and analysis of scientific data sets
- Software for platform reliability
  - Software to model platform architecture
- Embedded systems
  - Software control
  - Modeling and simulation
  - Self-evolving (adaptable) systems

Briefings with IT 'User' Agencies

In FY 2004, the SDP CG is holding briefings with representatives of Federal "IT user" agencies to gather information about software issues from agencies whose missions involve large-scale, critical applications – such as for Social Security records, Medicare transactions, and Internal Revenue systems. The goal is to better understand what IT managers – from their firsthand practical knowledge – perceive to be the key problems of their complex, real-time software environments. The first briefing of the series was provided by the DHS Customs and Border Protection agency, describing development of a new distributed IT system to manage U.S. export-import procedures.

Multiscale Modeling Workshop Planned

The SDP CG also is planning an FY 2005 workshop to explore software issues in multiscale modeling and simulation of complex physical systems. Modeling and simulation software has become a principal research tool across the sciences, and creating models that combine multiple factors at varying scales – such as climate from global to local scales, or interacting systems in the human body through the shape and functions of protein molecules – is both a key opportunity for scientific discovery and a key challenge for high-end computing and domain sciences. The workshop will bring computer and domain scientists together to assess the state of the art and possible synergies among computational tool sets developed for differing scientific domains.

Multiagency SDP Activities

The following is a sampling of FY 2004 activities in which more than one NITRD agency participates (other collaborating agencies are not cited):

NASA – Software design for Earth System Modeling Framework (ESMF); partners include DOE/SC, NOAA, and NSF; problem-solving frameworks, with DOE/SC and NSF; automated software engineering, with NSF; joint work on grid software with DOE/SC and NSF through the Global Grid Forum. (For more details on ESMF, see page 13.)

NSF – collaborates with NASA, DARPA, and others with overlapping missions, such as in the Highly Dependable Computing and Communications Systems Research (HDCCSR) program, with NASA, and ESCHER, with DARPA (see below)

DOE/NNSA – the Advanced Simulation and Computing (ASC) program is open to collaborations and dialogues with other agencies, such as funding R&D in open source software for high-performance computing. ASC researchers collaborate with ASC university alliance members that are also funded by other NITRD agencies, strengthening the collaborative environment in academic research and education.

DARPA – the Model-Based Integration of Embedded Systems (MoBIES) program has an interagency activities aspect, the Embedded Systems Consortium for Hybrid & Embedded Research (ESCHER), with NSF. ESCHER is a repository of technologies and tools from DARPA and NSF programs in hybrid and embedded systems. The prototypes and documentation are intended to promote rapid transition of MoBIES results to DoD and industry.

NIH – holding preliminary discussions with other agencies about an FY 2005 multiagency workshop on informatics associated with microbial science

NIST – collaborates with a wide range of Federal organizations on SDP-related topics, including:

- Data uniformity and Standards Structural Bioinformatics (includes Protein Databank, HIV Structural Database), with DOE/SC, NIH, NSF, and universities
- Digital Library of Mathematical Functions, with NSF
- NeXus Data Exchange Standard (effort to get neutron researchers to use sharable data structures), with DOE/SC
- Numerical Data Markup Language (developing an XML schema of encoding units of measurement, called UnitsML), with DOE/SC
- Product Data Standards for HVAC/R, with DOE
- Product Engineering Program (role of CAD systems; metrics for interoperable representation standards in product engineering), with DARPA, NASA

SDP R&D Programs By Agency

Selected FY 2004 Activities and FY 2005 Plans

SDP NASA SDP

NASA missions are increasingly and critically dependent on complex software systems; in many cases, software cannot be repaired or reloaded after a mission begins. The agency therefore has a critical need for:

- Automated software verification and validation tools
- Automated program synthesis
- Evolvable and adaptable software

Science and engineering applications in general are becoming more complex and multidisciplinary, and require distributed computing including "high-performance" applications, integration and interoperability of independent components, and runtime support for efficient discovery, access, and use of distributed resources including data and software. The agency uses a Technology Readiness Level (TRL) scale that runs from 1 (basic concept) to 9 (embedded in a mission) to categorize R&D.

FY 2004 agencywide SDP R&D activities are located in the Computing, Information, and Communications Technology (CICT) Program, Engineering for Complex Systems (ECS) Program, and Earth Science Technology Office (ESTO), in the following programs:

Information Technology Strategic Research, CICT

Automated Software Engineering Technologies (with NSF) – automated mathematical techniques for the software development process, yielding tools for cost-effective development of high-confidence, highly reliable software systems for aerospace applications. Low TRL. Technology development includes:

- Scalable software model checking, automated program abstraction, state-space search algorithms, and formal method verification of integrated modular avionics design to make possible analytical verification of concurrent advanced aerospace software architectures and code
- Program generation through automated reasoning, product-oriented certification methods, and automated tools that certify automatically synthesized code
- Generate runtime monitors from requirements specifications, automated behavioral verification, machine learning to optimize exploration of potential behaviors, and automated generation of software fault recovery, to enable software that monitors itself and recovers from faults at runtime with minimal computational overhead

Evolvable Systems – aiming to dramatically increase mission survivability and science return through development and application of evolutionary and adaptive algorithms. Includes:

- Advanced evolutionary algorithms and automated design and optimization for evolved fault recovery, to enable programmable logic that automatically rewires following damage
- Evolved sensor suite and evolutionary scheduling algorithms to enable sensor electronics that survives extreme radiation and temperature
- Future developments may include defect-tolerant nanosystems, distributed control, and evolved control algorithms for evolvable robotic control and coordination

Intelligent Systems, CICT

Verification and Validation (V&V) for Autonomy – addresses the complexity in verifying autonomous systems that operate in rich and uncertain environments with no human intervention, and that must adhere to internal correctness constraints involving communication among components, control flow, and resource utilization. This project to develop state-of-the-art benchmarking tools for V&V has successfully used rover software to test experimental versions of V&V tools. TRL 3.

Computing, Networking, and Information Systems, CICT

Problem Solving Frameworks (with DOE/SC, NSF) – objective is to develop infrastructures that support the management of simulation, analysis, and data components in application-specific environments. Challenges:

- Component representations that support the access, use, and composition of multi-component applications

- Methods for automatic and user-based generation of work plans
- Efficient management of the flow of work across distributed resources

Projects include:

– Component and Data Representation Models

– GridWorks: Workflow Management Framework

– CIAPP: CORBA-based Framework

– TAF-J: Agent-based Framework

– AeroDB

Grid Middleware Services (with DOE/SC, NSF) – developing a distributed infrastructure for seamless access to resources that are heterogeneous (computers, data, and instruments), distributed across multiple administrative domains, and dynamically coupled. This involves developing grid services (e.g., application and user-oriented, execution management, grid management, and data management and access) that provide a useful function independent of the underlying resources, are discoverable, encourage the design and use of reusable software components, and are easily incorporated into application frameworks. The runtime aspects of this program also fall under LSN's MAGIC effort.

Parallel Programming Paradigm – seeks to improve parallel programming tools and techniques to increase the performance of HEC systems scaling to thousands of processors, often in a grid environment of multiple distributed platforms. The portability, scalability, and usability of codes are technology challenges. The effort is evaluating (and, if necessary, extending standards for) parallel multilevel programming paradigms. Exploring: Multilevel Parallelism (MLP); MPI + OpenMP Hybrid Parallelism; and Distributed Shared Memory

Automated Programming Tools for Parallelization – aims to reduce time-consuming and error-prone code porting and development times for applications to run on HEC systems and grids by providing an integrated set of automated tools for parallelization that retains good performance. Challenges:

- Straightforward porting process
- Accurate, effective code analysis
- Multi-level parallel code generation
- Correct answers with ported code
- Performance analysis
- Good performance

Projects include:

– CAPO: Computer-Aided Parallelizer and Optimizer

– Performance analysis

– ADAPT: Automatic Data Alignment and Placement Tool

– Automatic debugging

Performance Modeling, Benchmarking, and Optimization – techniques, strategies, and tools to model, predict, and optimize application performance, and to improve applications' maintainability, portability, and performance in parallel and distributed heterogeneous environments. Technical approach:

- Predicting application execution times on various platforms
- Predicting wait times in scheduler queues
- Developing benchmarks to understand the behavior of key NASA applications on single (parallel) and distributed systems
- Understanding application cache usage and optimizing performance of regular and irregular problems

Projects are:

– NAS Parallel Benchmark

– NAS Unstructured Mesh Benchmarks

– NAS Grid Benchmarks

– Performance prediction

– Cache optimization

Resilient Software Engineering, ECS Program

High Dependability Computing – developing tools, case studies, and testbeds to identify and characterize risk precursors in mission software systems. NASA has a user collaboration portal for design and development; computer scientists and mission experts work together on dependability models and metrics. Testing is done on rover software and onboard spacecraft IT systems. The goal is to transfer successful dependability metrics and technologies to missions and industry. High TRL.

Intelligent Software Engineering Software Suite – investigating software algorithms, processes, and development procedures to improve software integrity and reliability. The work begins with the study of critical NASA software risks and prototype tool productization (methodology development, tool enhancement, and beta testing), then moves to a tool maturation and evaluation

stage (methodology integration and tool customization), and finally to integration in mission processes and adoption by mission partners. Low TRL.

Earth System Modeling Framework (ESMF), ESTO

NASA leads this collaboration of many agencies and academic researchers in building high-performance, flexible software infrastructure to increase ease of use, performance portability, interoperability, and reuse in climate, numerical weather prediction, data assimilation, and other Earth science applications. See story on page 13.

SDP NSF SDP

Applications software development is supported across CISE and specific applications (such as in biology, chemistry, and physics) are developed in other directorates. Across all NSF directorates, more emphasis is being placed on cyberinfrastructure, including software to support the science and engineering enterprise. In FY 2004, a new, cross-divisional "Science of Design" theme is being developed in CISE. SDP work is also funded under the ITR program, which involves all NSF directorates. As of December 2003, there were 1,300 active ITR awards, about 10 percent of which involve SDP research.

CISE SDP project areas include:

- Software Engineering and Languages (SEL)
- Advanced Scientific Computing (ASC)
- Embedded and Hybrid Systems (EHS)
- Next Generation Software (NGS)
- Distributed Systems and Compilers (DSC)

The current NSF-funded topics can be mapped into the SDP CG's taxonomy (page 55) as follows:

- Create, maintain, and improve software
  - Software development methods – ASC, EHS, NGS, SEL, ITR
  - Languages and tools – ASC, DSC, NGS, SEL, ITR
  - Runtime environments – DSC, NGS, ITR
- Investigating and achieving the "-ilities"
  - Platform reliability – DSC, ITR
  - Robustness – EHS, SEL
- Testing, analysis, and evaluation
  - Evaluation of models and simulations – ASC, DSC, NGS, ITR
  - Metrics and reference data – SEL, ITR
  - Automatic V&V – SEL, ITR
  - Testing tools – SEL, ITR
  - Software analysis – SEL, ITR
- Management and economics – SEL, ITR
- Software systems for specific domains
  - Application frameworks – ASC, NGS, ITR
  - Embedded systems – EHS, NGS, ITR

NSF FY 2005 plans in SDP areas include:

- *The CISE emphasis area "Science of Design" (SoD) will promote the scientific study of the design of software-intensive systems. Complex interdependencies strain our ability to create, maintain, comprehend and control these systems. The Science of Design activity seeks to rectify this situation by building a foundation for the systematic creation of software-intensive systems. This foundation will consist of a body of theoretical and empirical knowledge on design, computational methods and tools for design, and a new design curriculum for the next generation of designers.*
- *The special NSF/NASA "Highly Dependable Computing and Communication Systems Research" (HDCCSR) program will continue support for R&D to advance the design, test, implementation, and certification of highly dependable software-based systems. Awardees are able to use a NASA testbed facility to experimentally evaluate their research products on significant real hardware/software artifacts.*

SDP DOE/NNSA SDP

The Advanced Simulation and Computing (ASC) Program is responsible for providing designers and analysts with a high-fidelity, validated, 3D predictive simulation capability to certify the safety, reliability, and performance of the Nation's nuclear stockpile in the absence of physical testing. ASC's SDP R&D is focused on the:

Problem Solving Environment (PSE) program – one of four Simulation and Computer Science (SCS) thrusts involved in developing the ASC computational infrastructure. PSE goals:

- Create a common, usable application development environment for ASC platforms
- Produce an end-to-end high-performance I/O and storage infrastructure to enable improved code execution
- Ensure secure, effective access to ASC platforms across NNSA labs for both local and distributed computing

FY 2004 PSE priorities include:

- Focus on Red Storm software development environment, preparing for initial delivery and the final integrated platform
- Plan and prepare software environments for Purple C and Blue/Gene L. Planning for new platforms is a significant multiyear activity that involves developing and testing software configurations on smaller platforms and scaling up toward the final ultrascale integrated system.
- Continue performance modeling and measurement for increased effectiveness and efficiency of applications and platforms
- Develop high-performance open-source, Linux-based computing environments, targeting capacity computing. Requests for information (RFI) and requests for proposals (RFP) for open source software (OSS) projects are under way.
- Track and test development of Lustre file system for clusters and continue to evaluate alternative file systems
- Continue to evaluate alternatives to the distributed computing environment (DCE) technology, which DOE/NNSA labs have used for authentication, security, and administrative services, due to technology phase-out at the end of calendar year 2005

In FY 2005, DOE/NNSA plans to:

- *Continue the SDP-related work in ASC's PSE program*

SDP DARPA SDP

In FY 2004, DARPA has two programs in the SDP PCA, both within the agency's Information Exploitation Office (IXO). They are:

Software Enabled Control (SEC) – focuses on the design of advanced control systems for innovative vehicles (Unmanned Air Vehicle [UAV], Organic Air Vehicle [OAV], rotorcraft, fighters), including controls at both the vehicle and mission level. SDP-related technology areas are:

- Active state models enabling dynamic prediction and assessment
- Adaptive, dynamic customization of online controls
- Hybrid, multimodal controls
- Open-control platform with reusable (CORBA-based) middleware and tool support

Final demonstrations of small rotorcraft and fixed-wing vehicles are scheduled.

Model-Based Integration of Embedded Systems (MoBIES) – automates the design, construction, and testing of complex embedded systems by exploiting mathematical relationships between the physical platform, the computational platform, and real-world constraints – i.e., "software that is too complex to write manually." Mathematical models are used to meet multiple requirements for application-independent design tools, such as real-time control, network connectivity, fault tolerance, environmental resilience, physical constraints, and component libraries. Resulting tools for designing embedded applications include:

- Intelligent programming tools
- Smart process schedulers
- Communications configuration
- Online resource allocation
- User interfaces
- Automatic code generation
- Automatic verification and validation
- COTS integration

The MoBIES tool integration framework is being demonstrated in avionics and automotive design projects.

DARPA reports no SDP programs in FY 2005.

SDP NIH SDP

In FY 2004, the Center for Bioinformatics and Computational Biology of NIH's National Institute of General Medical Sciences (NIGMS) launches several initiatives:

Establishment of four NIH National Centers for Biomedical Computing – each center will serve as a node for developing, disseminating, and providing relevant training for computational tools and user environments in an area of biomedical computing.

Investigator-initiated grants – intended to foster collaboration with the National Centers. The objective is to avoid competition between "big" and "small" science by encouraging collaborations within the Centers.

Program for creation and dissemination of curriculum materials – designed to embed the use of quantitative tools in undergraduate biology education

The National Centers for Biomedical Computing grants solicitation includes a specific software requirement: Source code supported in this initiative must be shared and users may modify it, creating a pathway to commercialization.

In addition, the FY 2004 grant program of NIH's agencywide Biomedical Information Science and Technology Initiative (BISTI) includes a category for computer science grants. BISTI also has a program in software development and maintenance.

SDP **NIST** SDP

SDP-related activities at NIST include:

- Health: Informatics, privacy, ubiquitous aids, device controllers, diagnosis aids, decision support, and computer-aided surgery
- Nanotechnology:Computer control, modeling, and visualization
- Information and Knowledge Management: Data and knowledge bases, formal representations, intelligent access, and manipulation and testing tools
- Cybersecurity and Critical Infrastructure: Computer security, encryption, biometrics, monitoring software, and computer forensics as well as precision engineering and calibration for gun-tracing equipment, bullet-proof vests, and other security-related devices
- Weapons Detection: Monitoring devices such as sensors, threat-detection information systems
- Measurement Science: Conformance testing, scientific visualization and models, intelligent measuring tools, and machine learning tools

Within NIST's Systems Integration for Manufacturing Applications (SIMA) program, SDP-related projects include:

Automating Equipment Information Exchange (AEX) – aims to automate equipment design, procurement, and operation through software interoperability. Interoperability issues focus mainly on data.

Digital Library of Mathematical Functions (with NSF) – FY 2004 activities include usability studies and completion of the Web site.

Electronic Commerce for the Electronics Industry (ECEI) – supply chain software interoperability

IT Infrastructure Conformance Testing – working with industry partners on standards, such as ebXML, a version of XML for electronic business, and is partnering with Korea and Europe on testbeds for trying out b2b solutions

Interoperability of Databases for the Structure, Stability, and Properties of Inorganic Materials

Manufacturing Enterprise Integration – involves a testbed and work with industries on software interoperability, including research on automated methods for integrating systems. Research takes an ontological approach, using formal logic, to enable agent technologies and expert systems to automate the process of integrating systems.

NeXus Data Exchange Standard – developing sharable data structures for neutron researchers

Numerical Data Markup Language – developing UnitsML, an XML schema for encoding measurement units

Open Architecture Control – developing key interface standards and associated conformance tests to achieve interoperability of manufacturing control systems architectures with security taken into consideration. This has CIP applications, such as electrical generating plants or hydroelectric dams.

Product Data Standards for HVAC/R

Product Engineering Program – developing a semantically-based, validated product representation scheme as a standard for seamless interoperability among CAD systems and with systems that use CAD data

Standards for Exchange of Instrument Data and NIST Chemical Reference Data

Standards for Physical and Chemical Property Data Interchange

Anthropometric Data Standards (with U.S. Air Force) – seeking accurate 3-D representation of human measurements; an application is cockpit design.

NIST FY 2005 plans in SDP R&D include:

- *Continue SDP-related work in health informatics, nanotechnology, information and knowledge management, cybersecurity and critical infrastructure, weapons detection, and measurement science*
- *Continue SIMA program activities listed for FY 2004*

SDP NOAA SDP

NOAA SDP activities focus on using two software modeling frameworks:

Flexible Modeling System (FMS) – NOAA's Geophysical Fluids Dynamic Lab (GFDL) uses the FMS programming structure to develop atmospheric, ocean, and coupled climate models for climate projection studies.

Earth System Modeling Framework (ESMF) – ESMF, the next-generation community-wide version of FMS, is being worked on by GFDL in collaboration with NASA, NSF's National Center for Atmospheric Research (NCAR), the university community, and NOAA's National Weather Service (NWS).

The challenges of building increasingly interdisciplinary Earth system models and the need to maximize the performance of the models on a variety of computer architectures, especially those using upwards of thousands of processors, necessitates a new program structure. This structure allows physical, chemical, and biological scientists to focus on implementing their specific model components. Software engineers then design and implement the associated infrastructure and superstructure, allowing for a seamless linkage of the various scientific components.

FY 2005 plans include:

- *Continued cooperative ESMF development with NSF and NASA*
- *Establish Development Test Center (DTC) with NCAR and the university community*
- *Provide modeling framework for weather research and atmospheric science*
- *Improve the transition of research to operations*

Participating Agency

SDP FAA SDP

Safety is the FAA's primary mission. Safety will remain the FAA's top priority as the aviation industry readjusts itself to a world transformed by terrorism and economic challenges. Increasingly, solutions to these challenges depend upon secure and dependable software-based systems and the ability to produce them in a timely, predictable, and cost-effective manner.

FAA work straddles the SDP and HCSS PCAs.(Please see the discussion of FAA activities in the HCSS section.)

High Confidence Software and Systems

Definition of HCSS PCA

The activities funded under the High Confidence Software and Systems (HCSS) PCA focus on the basic science and information technologies necessary to achieve affordable and predictable high levels of safety, security, reliability, and survivability in U.S. national security- and safety-critical systems. These systems play key roles in critical domains such as aviation, health care, national defense, and infrastructure. Many complex software-and information-intensive systems that have high consequences of failure must be certified as to their safety and security. Currently, however, this certification – even when possible at all – requires overwhelming cost, time, and effort, discouraging and delaying innovation of new technologies and processes. The overall HCSS goal, then, is to develop and demonstrate revolutionary capabilities for system development and assurance that balance and reduce risk, cost, and effort to achieve systems that behave in predictable and robust ways. HCSS R&D will help transform our ability to feasibly build certifiably dependable systems in the challenging environment of an increasingly interconnected and automated society.

Broad Areas of HCSS Concern

- Security and privacy
- Safety, robustness, reliability of software and systems
- Trust, risk, and accountability
- Assured development and certification of software and systems
- Survivability

Technical Goals

- Provide a sound theoretical, scientific, and technological basis for assured construction of safe, secure systems
- Develop hardware, software, and system engineering tools that incorporate ubiquitous, application-based, domain-based, and risk-based assurance
- Reduce the effort, time, and cost of assurance and quality certification processes
- Provide a technology base of public domain, advanced-prototype implementations of high confidence technologies to enable rapid adoption
- Provide measures of results

Illustrative Technical Thrusts

- Foundations of assurance and composition
- Correct-by-construction system design and software technologies
- Evidence and measurement technologies for verification and validation
- Authentication, access control, intrusion detection, trust models, and forensics
- Dependable open, distributed, and networked systems
- Secure and reliable hardware, network, operating system, and middleware technologies
- Dependable and survivable real-time, embedded, and control system technologies
- Verification and certification technologies
- Dependable technologies for transportation, medical

HCSS Agencies

NSF NASA NIST
NSA DARPA NIH

Participatimg Agencies

AFRL FAA FDA ONR

HCSS PCA Budget Crosscut

FY 2004 Estimate	FY 2005 Request
$144.4 M	$152.5 M

devices and health systems, power generation and distribution systems, financial services, and other critical infrastructures

- Experimentation and reference HCSS implementations
- Assured open source software

HCSS PCA: Coordination and Activities

HCSS Highlights

In FY 2001, the word "Software" was added to the name of the prior High Confidence Systems PCA to reflect the central role played by software in the overall reliability, security, and manageability of the Nation's most complex and critical computing and communications systems. The recommendation to make software a top priority of Federal IT R&D activities had been highlighted by the PITAC in its 1999 report on Federal IT R&D investments. The purview of the High Confidence Software and Systems PCA now includes R&D in all aspects of software development for very-high-assurance trusted systems.

Through monthly meetings, the HCSS Coordinating Group (CG) shares information on agency research programs, upcoming meetings, and workshops. The group cooperatively supports studies on HCSS topics, holds workshops in key research and programmatic areas, and invites other agencies to conferences and principal investigator (PI) meetings. FY 2004 CG activities include:

- A study on "Sufficient Evidence? Building Certifiably Dependable Systems" being conducted by the Computer Science and Telecommunications Board of the National Academies. Sponsored by NSF, NSA, and ONR; AFRL, ARO, DARPA, FAA, FDA, NASA, and NIST also participate. The study brings together a broad group of experts to assess current practices for developing and evaluating mission-critical software, with an emphasis on dependability objectives. The group is addressing system certification and examining a few application domains (e.g., medical devices and aviation systems) and their approaches to software evaluation and assurance. The goal is provide some understanding of what common ground and disparities exist. The study committee hosted a workshop on Software Certification and Dependability on April 19-20, 2004, to survey technical, business, and governmental perspectives and to promote dialogue between the research community and government and industry practitioners who develop safety-critical systems.
- The HCSS CG hosted an Open Verification Workshop on April 12, 2004.
- Several other HCSS agencies participated in NSA's High Confidence Software and System Conference, April 13-15, 2004.
- A set of aviation safety workshops is being planned to address safety issues related to the use of unmanned aerial vehicles in civilian and military airspace. AFRL, FAA, NASA, and NSF are the major planners.
- The HCSS CG is planning a workshop on medical devices software safety, with FDA, NASA, NIST, NSA, NSF, and others.

Multiagency Collaborations

In FY 2004, HCSS agencies are working together on several collaborative research projects and workshops in assurance, cybersecurity, and medical devices. For example:

- Using a new NASA testbed facility, NSF and NASA are jointly sponsoring the Highly Dependable Computing and Communications Systems Research (HDCCSR) program to promote the ability to design, test, implement, evolve, and certify highly dependable software-based systems.
- DARPA, NSF, and other agencies supported the 2003 kickoff of the Embedded Software Consortium for Hybrid and Embedded Software and Systems (ESCHER, which is included in both the HCSS and SDP PCAs). This group, which has industry support and participation, will focus on system design tools, open source system software, and reference implementations.
- NSF is supporting a cybersecurity study by the Computer Science and Telecommunications Board (CSTB) of the National Academy of Science and invites participation by other agencies.
- FDA and NSF are exploring a joint project to promote participation by computer-science students at FDA. Students will work to facilitate the transition of software methods and to expand FDA's expertise in identifying needs for software-enabled medical devices.

HCSS R&D Programs By Agency

Selected FY 2004 Activities and FY 2005 Plans

HCSS **NSF** HCSS

NSF's HCSS activities reside in the Cyber Trust and Science of Design themes in the Computer and Information Science and Engineering (CISE) Directorate, and in the NSF-wide Information Technology Research (ITR) Program as follows:

Cyber Trust – initiative across CISE divisions that envisions a society in which:

- Computing systems operate securely and reliably
- Computing systems protect sensitive information
- Systems are developed and operated by a well-trained and diverse workforce

This program supports research on foundations, network security, systems software, and information systems. It sponsors integrated education and workforce activities. Cyber Trust workshops, including PI workshops, are open to participants from other government agencies. In other current research efforts, NSF is seeking help from other agencies in identifying technology transfer opportunities and creating and distributing relevant cyber trust data sets.

Science of Design – a crosscutting initiative that emphasizes design of software-intensive computing, information, and communications systems. The goal is to improve the development, evolution, and understanding of systems of large-scale scope and complexity. These are systems for which software is the main means of conceptualization, definition, modeling, analysis, development, integration, operation, control, and management. A workshop was held November 2-4, 2003, in Northern Virginia to develop the program's foundations.

ITR Program – emphasizes national priorities including national and homeland security, which includes research related to critical infrastructure protection and SCADA systems.

Other CISE program activities – CISE's Distributed Computing, Embedded and Hybrid Systems, Networking, and Foundations of Computing Processes and Artifacts programs also include HCSS work.

The following current projects are representative of NSF support for efforts addressing aspects of trustworthy systems:

- Cryptography
  - Information Theoretic Secure Hyper-Encryption and Protocols
- Data, Security, and Privacy
  - DataMotion: Dealing With Fast-Moving Data
  - Deployment-Oriented Security and Content Protection
  - Sensitive Information in a Wired World
- High Confidence Control
  - A Unified Framework for Distributed Control with Limited and Disrupted Communication
  - Algorithmic Synthesis of Embedded Controller
  - Symbolic Approaches to Analysis and Hybrid Systems
- Prevention, Detection, and Response
  - A Semantic-Based Approach for Automated Response to Attacks
  - Architectural Solutions for Preventing Distributed Denial of Service Attacks
  - Automated and Adaptive Diversity for Improving Computer Systems Security
  - Forensix: Large-scale Tamper-resistant Computer Forensic System
  - Intrusion Detection Techniques for Mobile Ad Hoc Networks
- Systems Software for Protecting Critical Infrastructure
  - Distributed Authentication and Authorization: Models, Calculi, Methods
  - High-Assurance Common Language Runtime
  - Key Management for Secure Dynamic Group Communications
  - Language-Based Software Security
  - Practice-Oriented Provable Security for Higher-Layer Protocols: Models, Analyses and Solutions
  - Security and Privacy for Publish-Subscribe Systems
  - Survivable Trust for Critical Infrastructure
  - Trusted Peer-To-Peer Systems

In FY 2005 , NSF will continue HCSS R&D in:

- *Cyber Trust – research aimed at creating systems that are more predictable, more accountable, and less vulnerable to attack and abuse; developed, configured, operated, and evaluated by a well-trained and diverse workforce; and used by a public educated in their secure and ethical operation*
- *Disciplinary research in science and technology for the design and implementation of high-confidence networks, embedded and control systems, computer hardware design, operating systems, and distributed systems. CISE will also support research in assurance technology and methods that help to verify safety, security, timeliness, and correctness aspects of critical systems*
- *Research projects under ITR aimed at dramatically increasing our ability to build high-confidence security- and safety-critical systems*

Selected new multiyear project awards made in August 2004 include:

- *Byzantine Fault Tolerance for Large-Scale, High-Performance Distributed Storage Systems*
- *The Design and Use of Digital Identities*
- *Graph-Based Refinement Strategies for Hybrid Systems*
- *IIT-based Collaboration framework for Preparing against, Responding to and Recovering from Disasters involving Critical Physical Infrastructures*
- *Panoply: Enabling Safe Ubiquitous Computing Environments*
- *Privacy-Preserving Data Integration and Sharing*
- *Toward a Multi-Layered Architecture for Reliable and Secure Large-Scale Networks: The Case of an Electric Power Grid*

HCSS NSA HCSS

Information Assurance Research Group (IARG) – promotes HCSS research through three product-assurance capability threads:

- Trusted by design – to help software engineers achieve assured designs and reduce the cost of certifying the security of complex information systems
- Trusted by analysis – to assess the confidence in a system that has been built outside of NSA control and whose assurance is unknown
- Containment – to balance granularity of protection against ease of use and cost

The HCSS roadmap for IARG comprises three areas of research:

- **Foundations to develop the supporting theory and scientific basis for high-confidence systems** such as automatic theorem proving, design and analysis of protocols, interoperability and composition and decomposition of agents, and systems security and survivability architectures. Current work includes:
  - National Academy of Sciences Certification Study, focused on addressing system certification and approaches to software evaluation and assurance.
  - Protocol Specification and Synthesis, effort focused on foundational methods of secure communication, with the goal of providing methods and tools upon which the design, analysis, and implementation of security structures might be carried out.
- **Tools and technologies for building high-confidence systems of the future** through the development of analysis, evaluation, and vulnerability tools and techniques. Projects include:
  - Specware, an environment supporting the design, development, and automated synthesis of correct-by-construction software
  - Cryptol, a programming language focused solely on the domain of cryptography, and recently adopted by General Dynamics
  - Vulnerability discovery, focused on developing and demonstrating a support environment for the analyst who is interested in software system vulnerabilities
  - Java Program Verification Condition Generator, a tool that uses formal analysis to eliminate classes of errors during software development of Java programs
  - Formal analysis of hardware/software co-design
  - Biospark, reliability engineering in biometrics that teams HCSS, smart card, and biometrics researchers
  - Polyspace, a project focused on evaluating the fitness for use of the commercial Polyspace static verifier for detecting run-time software errors
- **Engineering and experimentation to demonstrate the effectiveness and efficiency of HCSS technologies** on diverse hardware and software

platforms. Projects include the following:

- Trusted Web server, focused on developing a cross-domain server that can be certifiable for simultaneous connection to networks spanning two or three domains
- Osker (the Oregon Separation Kernel), a prototype POSIX-compliant operating system in Haskell (see the next bullet) that provably achieves a strict separation between processes according to a specified security model. Osker is a challenge application for the Programatica project, a system for developing high-assurance software.
- Haskell on Bare Metal (HBM), an adaptation of the Haskell runtime system, to replace the usual operating system layers between application and hardware
- Java applet generation, an automatic generator that produces Java Card applets from high-level formal specifications
- AAMP7 development environment, a partition-aware development environment for Rockwell Collins's AAMP7 microprocessor that will allow rapid development of partitioned AAMP applications

NSA's FY 2005 planned activities in HCSS include:

- *Host 5th Annual HCSS Conference*
- *Continue joint sponsorship of National Academy of Sciences study on software certification*
- *Initiate joint sponsorship of Open Verification activities with HCSS CG members, resulting in sponsorship of IFIP working conference of Verified Software as well as a Safe Code workshop*
- *Sponsor and conduct research through NSA IARG within the following research themes:*
  - *Product assurance: HCSS tasks focused on trusted development and containment mechanisms*
  - *Trusted Development Thread, which attempts to achieve assured software and system designs and implementations through enhancement of assured development and analysis techniques throughout the entire software and system lifecycle*
  - *Containment Thread, which is focused on mitigating the risk posed by our inability to build systems whose components are all perfectly assured, thereby limiting the impact of improper software and system behavior. The primary challenge in designing containment mechanisms comes in balancing granularity of protection against ease of use and cost.*
  - *Transparency: HCSS task focused on supporting the development of critical architectures and components necessary to support information assurance*
  - *High Assurance Platform: HCSS tasks focused on supporting very promising industrial partnerships through virtualization and measurement capabilities*

HCSS NASA HCSS

NASA missions have several critical needs that HCSS R&D helps address:

- Mission- and safety-critical software
- High-confidence software within predictable cost and schedule
- High confidence for new types of software, such as for model-based autonomy and adaptive control
- Sustained engineering (for example, the ISS and the Space Shuttle)
- Security for ground and radio frequency networks

Several major programs span the agency's technical readiness level (TRL) scale, which runs from 1 to 9 (9 denotes a capability that has served on the space shuttle for 50 flights). High-TRL work has a strong process orientation, mid-TRL is work in transition to practice, and low-TRL work involves fundamental research. HCSS-related efforts include:

Computing, Information and Communications Technology (CICT) – (low- to mid-TRL) project aims to develop automated mathematical techniques for the software development process, yielding tools for cost-effective development of high confidence, highly reliable software systems for aerospace applications. Its goal is to develop technologies with enhanced capabilities to:

- Analytically verify the next generation of aerospace software:
  - Scalable software model checking
  - Automated program abstraction
  - State-space search algorithms
  - Formal method verification of integrated modular

avionics design

- Produce certifiable program synthesis for the following technologies:
  - Program generation through automated reasoning
  - Product-oriented certification methods
  - Automated tools that certify automatically synthesized code
- Develop adaptive, integrated software verification and monitoring technology, including:
  - Runtime monitors generated from requirements specifications
  - Automated behavioral verification
  - Machine learning to optimize exploration of potential behaviors
  - Automated generation of software fault recovery

These capabilities would then be applied to specific missions such as the ISS and the Mars Lander.

Highly Dependable Computing Platform Testbed – provides a modern software platform for real-time embedded systems. The approach (low- to mid-TRL) is to evaluate real-time Java to address in-flight software demands and use the Mission Data Systems (MDS) framework and software as a testbed. While NASA typically runs older hardware on the ISS and the Hubble telescope because that hardware is known to be hardened against radiation, it develops software on modern workstations and then ports that software to the older hardware. The real-time Java needs to have demonstrably lightweight CPU usage and provide the desired throughput and response. NASA needs to be sure that timing jitters do not surface and cause problems.

Mission Data Systems (MDS) – (mid-TRL) developing a reusable infrastructure for flight and ground software for the mission to Mars in 2009. In preparation for the launch, all needed technologies should be in place in 2005. MDS is integrating the best systems engineering and software engineering practices for autonomous control of physical systems. The program was developed for unmanned space science missions involving spacecraft, landers, rovers, and ground systems. It is broadly applicable to mobile and immobile robots that operate autonomously to achieve goals specified by humans. It is also architecturally suited for complex interactive systems where "everything affects everything."

As complexity grows, the line between specifying behavior and designing behavior is blurring. For each of the items in the following illustrative list, systems engineers need to know and want to specify the item, while software engineers want to design software that knows the item:

- How a system is put together (connections and other interactions)
- What functions each element performs (models of behavior)
- How system elements might fail (models of faulty behavior)
- What the environment is like and how it affects the system (more models)
- What the system must be able to do (scenarios and their objectives)
- What operating constraints the system must honor (flight rules, etc.)
- What resources the system must manage (power, data storage, etc.)

The MDS approach is through product line practice to exploit commonalities:

- Define a reference architecture to which missions and products conform
- Provide framework software to be used and adapted
- Define processes for systems engineering and software development

An example is state analysis for embedded systems. The Mars science lab now has some 10,000 state variables. The relationship between each pair (for example a disk drive's power and the heat it produces) is described and the software is designed to include rules on determining and controlling state. This effort helps systems engineers and software engineers use the same vocabularies.

Office of Safety and Mission Assurance Research Program – mid-TRL effort that encompasses the following:

- Software assurance practices for auto-generated code:
  - Evaluation of available artifacts from autocode processes
  - Verification of the code generator
- Software assurance practices for commercial off-the-shelf integration:
  - V&V of interface to COTS

– Validation of a COTS application for an intended purpose

- Software assurance practices for reused or heritage software:
  - Reuse or heritage factors that impact software risk
  - Appropriate level of software assurance for reused or heritage code
- Reliability of operating systems
- Tandem experiment to improve software assurance
- Independent V&V (IV&V):

IV&V is verification and validation performed by an organization that is technically, managerially, and financially independent. IV&V focuses on mission critical software, provides addition reviews and analyses, and provides in-depth evaluation of life cycle products that have the highest level of risk. Examples of IV&V activities include the following:

- Validation of design to meet system needs and requirements
- Traceability of safety-critical requirements
- Code analysis of mission-critical software components
- Design analysis of selected critical algorithms

Software Engineering Initiative (SEI) – high-TRL effort begun to respond to the growing complexity, size, and sophistication of software components (for example, the two Mars missions that landed in January 2004 involve 625,000 lines of source code). The goal of the SEI is to advance software engineering development, assurance, and management practices to meet NASA's science and technology objectives. Elements of this initiative include:

- Plans from each center to improve software process and products
- Use of the Carnegie Mellon University Software Engineering Institute's Capability Maturity Models (CMM) as benchmarks for assessments
- Infusion of the "best of the best" software engineering research and technology
- Software metrics to monitor the initiative's progress and to provide early warning of problems
- Effective guidelines, principles, and standards
- Enhanced knowledge and skills in software engineering through training, education, and information exchange
- Improved software acquisition capabilities

Software Assurance Program – (high TRL) seeks the following:

- Software risk mitigation
- Improved quality of software products while using risk mitigation techniques
- Project management insight into software development processes and products throughout the life cycle
- Early error detection, problem prevention, and risk identification and mitigation
- Improve the quality of future products and services

The level of software assurance needed is dependent on the software size, complexity, criticality, and level of risk. Software assurance covers practices for auto-generated code, COTS integration, and reused or heritage software. Software assurance work is performed in the following areas: standards; guidance; policy; contractor evaluation criteria; metrics; means to classify software across NASA; IV&V; research; benchmarking; and outreach.

Software assurance involves both software safety and software reliability, as follows:

- Software safety includes a systematic approach to identifying, analyzing, tracking, mitigating, and controlling software hazards and hazardous functions (data and commands) to ensure safer software operation within a system.
- Software reliability is the process of optimizing the software through emphasis on requiring and building in software error prevention, fault detection, isolation, recovery, tolerance, and/or transition to planned reduced functionality states. It also includes a process for measuring and analyzing defects in the software products during development activities in order to find and address possible problem areas within the software.

HCSS **DARPA** HCSS

Self-Regenerative Systems (SRS) – aims to develop a military exemplar system that shows it is possible to: provide 100 percent of critical functions at all times in spite of attacks; learn about one's own vulnerabilities to improve survivability over time; and regenerate service after attack. The result of SRS activities will be intrusion-tolerant systems that gracefully degrade and recover after an attack while maintaining some level of system performance instead of crashing. The development phase will involve self-regenerative systems that restore performance to full operational capability. SRS technical areas include the following:

- Biologically Inspired Diversity to reduce common software vulnerabilities to attack by providing different versions of software with different implementations and configurations
- Cognitive Immunity and Healing systems that incorporate biologically inspired response strategies and machine learning to identify and correct root causes of vulnerabilities
- Reasoning About Insider Threats to pre-empt insider attacks or detect system overrun by combining and correlating information across system layers, inferring user goals, and enabling effective anomaly detection
- Granular, Scalable Redundancy to survive massive attacks or extreme hostility by approach exploiting environment knowledge to scale or perform and develop probabilistic consistency protocols that will survive extremely hostile environments and provide "good enough" service

Security-Aware Systems.– goal is to minimize unavoidable cyber risk to military missions by having the system itself smoothly adapt to changing resources, building blocks, security requirements, mission goals, and threats. A security-aware system will reason about its own security attributes, capabilities, and the utility of its functions with respect to a mission context. It will dynamically adapt to provide desired levels of service while minimizing risk and providing coherent explanations of the relative safety of service level alternatives.

In FY 2005, work will continue on the following DARPA effort:

- *Self-Regenerative Systems (SRS)*

The following DARPA effort is new for FY 2005:

Security-Aware Critical Software (SACS) program – *will create a new generation of software that provides a comprehensive picture of security properties and current status, presenting this information at multiple levels of abstraction and formality. SACS will thus make security properties and status transparent to decision makers, which will increase the speed and confidence with which military systems can be securely and dynamically reconfigured, particularly under stressful conditions. SACS will enable construction of a security-aware system that can reason about its own security attributes and capabilities and the utility of its functions with respect to a mission context. The software will dynamically adapt to provide desired levels of service while minimizing risk and providing coherent explanations of the relative safety of service-level alternatives.*

HCSS **NIST** HCSS

Two divisions in the Information Technology Laboratory at NIST – the Software Diagnostics and Conformance Testing Division (SDCTD) and the Computer Security Division (CSD) – are the primary organizations involved in HCSS activities. The SDCTD mission is to develop software testing tools and methods that improve quality, conformance to standards, and correctness, and to work with industry to develop forward-looking standards. Five technical areas of SDCTD involve HCSS R&D:

Electronic Commerce – focuses on extensible markup language (XML), a universal interchange format including core technologies. More generalized than HTML and can be used for tagging data streams more precisely and extensively. World Wide Web Consortium (W3C) interoperability testing will be conducted to evaluate interoperability in both messaging and smart card services. NIST aims to develop and automate consistent, complete, and logical specifications and turn these into performance testing for eventual commercial use.

E-Health – developing Health Level Seven (HL-7) standards and conformance and a standards roadmap so that medical devices, hospital systems, and other health care service provider systems can talk to each other while protecting patient privacy. NIST is working with the Department of Veterans Affairs on access control, sign-on, and other procedures, acting as a trusted impartial "third party" among providers, researchers, manufacturers, and others to promote effective access controls.

Computer Forensics – working with the FBI and the National Institute of Justice to develop a National Software Reference Library (NSRL) and specifications and evaluations of computer forensics tools to use in efficiently analyzing seized property such as disk drives and verifying that rules of evidence are observed.

Pervasive Computing – addressing development of wireless service discovery protocols for wireless devices such as palm pilots to assure trustworthy interactions.

Test Method Research – fundamental work in object-oriented component testing and in automatically generating tests from formal specifications in a cost-effective manner.

FY 2004 new opportunities – conformance testing for medical devices and test suites for medical device communication standards; using the NSRL for data reduction, integrity management, and computer security applications; and investigating grid computing vulnerabilities to identify requirement for maintaining system robustness.

The mission of the CSD is to improve information systems' security by: raising awareness of IT risks, vulnerabilities, and protection requirements; researching, studying, and advising agencies of IT vulnerabilities and devising techniques for cost-effective security and privacy of sensitive Federal systems; developing standards, metrics, test and validation programs; and developing guidance to increase secure IT planning, implementation, management, and operation. CSD programs encompass the following:

Security Technologies – cryptographic standards, key management, public key infrastructure (PKI), identity management, protocols and e-government, and agency e-government support

Systems and Network Security – technical guidelines, checklists, smart cards, wireless/mobile, Intrusion Detection System (IDS), ICAT, IP Security Protocol (IPSec), authorization management, automated testing, and quantum cryptography

Management and Assistance Program – outreach, expert assistance, policy, and guidelines, and Information Security and Privacy Advisory Board (ISPAB)

Security Testing and Metrics – security control development, certification and accreditation, cryptographic module validation, laboratory accreditation, and the National Information Assurance Partnership (NIAP)

FY 2004 new opportunities – a Standard Reference Model (SRM) for source code security to develop a metric for automated tools to review security properties of software and a database of known security flaws from "buffer overflow" through "trap doors"; trust and confidence taxonomy (reliability, security, interoperability) or "-ilities" toolkit

NIST FY 2005 plans in HCSS-related R&D include:

- *Continue work in electronic commerce, e-health, computer forensics, test method research, security technologies, systems and network security, management and assistance, and security testing and metrics*
- *Possible new activity in high-confidence methods for voting and vote counting*

Participating Agencies

HCSS FAA HCSS

FAA's Office of the Assistant Administrator for Information Services and CIO focuses on security, processes (enterprise architecture), and education issues. The Office must co-sponsor its research, have an FAA customer, and either have a short-term focus or collaborate with others on longer-term issues.

Over the past three years, FAA has evolved a systematic approach to defending the air traffic control system against cyber attack:

- Harden individual system and network elements
- Isolate elements to avoid "viral" spread
- Replicate elements to avoid service disruption

This strategy is difficult because of the size and complexity of the air-traffic-control system, the increased use of COTS products, and the safety-critical nature of the air-traffic control system. Significant challenges such as building trustworthy systems with untrustworthy components remain. FY 2004 AIO activities are:

- Rapid quarantine capability
- Wireless information systems security
- A Common Criteria test lab
- Integrity and confidentiality lab
- Estimating security costs
- A software reliability pilot
- Biometrics for single sign-on
- Data mining for vulnerabilities

HCSS FDA HCSS

FDA, through its Center for Devices and Radiological Health, with other agencies develops medical devices that require high-confidence, assured, safe software to deliver quality medical care. The FDA Office of Science and Technology leverages funds to work with other agencies. Research interests focus on formal methods of design in three areas:

- Safety and safety modeling
- Certification issues
- Forensic analysis

Specific research projects include:

- Life Support for Trauma and Transport (LSTAT), an intelligent litter platform with Walter Reed Army Medical Center (safety and safety modeling) • Proton beam therapy device (safety and safety modeling)
- Software for an infusion pump with a control loop, which led to an initiative to develop similar control loop software for a ventilator device (certification)
- Blood bank software regulation (certification)
- Radiation-treatment planning systems that employ reverse engineering of C programs to look for inconsistencies and errors in analysis of brain tumors (forensics)

HCSS AFRL HCSS

The Air Force Research Laboratory in Dayton, Ohio is working on developing certifiability requirements for autonomous aircraft that operate in civilian and military airspace.

Social, Economic, and Workforce Implications of IT and IT Workforce Development

Definition of SEW PCA

The activities funded under the Social, Economic, and Workforce Implications of IT and IT Workforce Development (SEW) PCA focus on the nature and dynamics of IT impacts on technical and social systems as well as the interactions between people and IT devices and capabilities; the workforce development needs arising from the growing demand for workers who are highly skilled in information technology; and the role of innovative IT applications in education and training. SEW also supports efforts to transfer the results of IT R&D to the policymaking and IT user communities in government at all levels and the private sector. Amid today's rapid global transformations driven by IT, SEW research aims to provide new knowledge to help society anticipate, identify, understand, and address the diverse issues of the digital age.

Broad Areas of SEW Concern

- Economic, organizational, social, and educational transformations driven by new information technologies
- Participation in the digital society, including e-government
- Intellectual property and privacy rights
- Innovative applications of IT in education
- IT workforce development

Technical Goals

- Develop baseline empirical findings about the complex interactions between people and IT
- Support training to expand the skilled IT workforce
- Increase understanding of intellectual property and privacy issues in the digital society
- Promote linkages between the SEW research community and policymakers
- Demonstrate innovative IT applications for education

Illustrative Technical Thrusts

- Interactions and complex interdependencies of information systems and social systems
- Collaborative knowledge environments for science and engineering
- Management of knowledge-intensive dynamic systems
- Tools and technologies and tools for social-network analysis
- Application of information technology to law and regulation
- Technologies and tools to facilitate large-scale collaborative research through distributed systems
- Technologies in and theories of electronic business, supply chains, economics of IT, productivity, and related areas
- Innovation in computational modeling or simulation in research or education
- Advanced graduate training in the strategically important IT fields of bioinformatics and computational science

SEW Agencies

		Participating Agency
NSF	DOE/NNSA	
NIH	DOE/SC	
NASA		GSA

SEW PCA Budget Crosscut

FY 2004 Estimate	FY 2005 Request
$120.9 M	$130.9 M

- Efforts to eliminate barriers to IT workforce participation for women and minorities
- Experimentation with cutting-edge networked applications for engineering training and K-14 science and mathematics education

SEW PCA: Coordination and Activities

SEW Highlights

The SEW PCA has two related but distinct components: 1) education and workforce development activities and 2) activities involving the socioeconomic implications of IT. NSF is the sole NITRD agency pursuing research in the latter area, while the other SEW agencies' investments lie in the former. Because of the breadth of this portfolio, the SEW Coordinating Group (CG) has taken on a character somewhat different from those of the CGs devoted to IT research topics that engage multiple agencies. The SEW CG has developed a program of briefings on themes of interest to agencies beyond the IT R&D community. Subjects have included intellectual property issues in open source software, issues in creating a national health information infrastructure, and trends in IT workforce demographics and their implications for education and training.

Since FY 2002, the CG also has supported, through its Universal Accessibility Team, the development of a new program of workshops sponsored by GSA and the Federal CIO Council to foster collaboration among government and community implementers of IT and to demonstrate promising IT capabilities emerging from Federal research.

Each of these evolutionary directions has sought to position the SEW CG as a communications link between IT researchers and policymakers and implementers of IT applications. In FY 2004, the SEW CG's principal activity is an examination of how its role and structure have changed since the PCA's inception.

Universal Accessibility Team Activities

The monthly Collaboration Expedition Workshops, completing their third year in FY 2004, bring together a diverse group of IT researchers, developers, and implementers representing such fields as emergency preparedness and response, health care, environmental protection, and citizen services. Drawing between 60 and 100 participants each month, the workshops also assist Federal program managers in coordinating necessary steps to implement the Administration's Federal Enterprise Architecture Program in their agencies.

The workshops have developed into a crossroads for software and system developers, IT managers and implementers, and public services practitioners across all levels of government, in the private sector, and in local communities. Each monthly meeting includes demonstrations of emerging technologies and prototype applications for developing intergovernmental "citizen-centric" services, and discussions of barriers and opportunities in applying technologies to enhance citizen-government interactions. The goal is to accelerate multi-sector partnerships around IT capabilities that help government work better on behalf of all citizens. Sample meeting topics include:

- Taxonomies and the Semantic Web (XML and XML Schema, Resource Description Framework/Schema [RDF/S], DARPA Agent Markup Language [DAML]+Ontology Inference Layer [OIL] and the new Web Ontology Language OWL, derived from DAML+OIL)
- Extensible Federal Enterprise Architecture components that transcend "stove-piping" through open standards technologies
- Realistic citizen-service scenarios for benchmarking performance
- Audio e-book technology
- Multi-channel communication and information services, including dynamic knowledge repositories
- Web-based collaboration

SEW R&D Programs By Agency

Selected FY 2004 Activities and FY 2005 Plans

SEW **NSF** SEW

NSF's SEW research portfolio encompasses a broad range of efforts, from studies of the socioeconomic implications of the ongoing digital revolution, to explorations of how IT can enhance government-citizen interactions, to research on ways to expand and better prepare the Nation's IT workforce, to R&D on innovative applications of information technologies in education and training.

In FY 2004, elements of SEW research are supported under the following programs:

Information Technology Research (ITR) Program

This major foundation-wide interdisciplinary priority area in FY 2004 began its fifth and final year of grant awards with a new focus on Information Technology Research for National Priorities. SEW interests are highlighted in NSF's call for ITR proposals that integrate research and education activities or foster the development of a diverse IT workforce. In addition, the solicitation calls for SEW research related to the following national priorities:

- Advances in Science and Engineering (ASE), which could include research on technologies and tools to facilitate large-scale collaborative research through distributed systems
- Economic Prosperity and Vibrant Civil Society (ECS), which seeks projects that investigate the human and socio-technical aspects of current and future distributed information systems for economic prosperity and a vibrant civil society. Examples of topics include human and social aspects of distributed information systems for innovation, business, work, health, government, learning, and community, and their related policy implications.
- National and Homeland Security (NHS), which includes research on critical infrastructure protection and technologies and tools for understanding threats to national security

About 100 ITR grants awarded in prior fiscal years for SEW-related research are continuing in FY 2004. Under the overarching FY 2004 theme, new ITR proposals must address one or more of four specified technical focus areas. The following two of the four directly address SEW research interests:

- Interactions and complex interdependencies of information systems and social systems
- Innovation in computational modeling or simulation in research or education

Computer and Information Science and Engineering (CISE) Directorate Programs

In the FY 2004 divisional reorganization within the CISE Directorate, a substantial portion of SEW-related research is housed in the Systems in Context cluster of the Division of Information and Intelligent Systems (IIS). As part of the directorate's overarching FY 2004 emphasis on education and workforce issues, SEW-related research in educational technologies and IT workforce development is also supported under the Combined Research and Curriculum Development and Educational Innovation Program and the Information Technology Workforce Program. These two programs reside in the Education and Workforce Cluster of the Division of Computer and Network Systems (CNS).

CISE/IIS/Systems in Context

This IIS cluster includes three research thrust areas funding SEW-related work in FY 2004. They are:

Digital Government – supporting research that advances IT applications for governmental missions and/or research that enhances understanding of the impact of IT on structures, processes, and outcomes within government, both from the perspective of government agencies and from the viewpoint of the citizenry at large. Sample research topics:

- The capture of government decision-making processes
- The application of information technology to law and regulation
- Software engineering of large-scale government projects
- Online campaigning and e-voting
- New forms of IT-enabled civic engagement and interaction
- Failures and successes of governmental IT adoption

- Implications and impact of IT on democracy and forms of governance

Digital Society and Technologies – FY 2004 emphases in this thrust area are:

- Universal Participation in a Digital Society – research seeking to understand the underlying processes by which IT shapes and transforms society and society simultaneously shapes and transforms new IT, and how these transformations impact the goal of universal participation in our democratic society. Areas of study include e-commerce, digital science, the IT workforce, community networking, and digital governance.
- Collaborative Intelligence – includes theories, models, and technologies for distributed, intelligent, collective action among humans, agents, robots, and other embedded devices. Focus on:
  - The science of collaboration (design principles, mixed initiative and adjustable autonomy problems, and implicit and explicit, affective and instrumental human-machine interactions)
  - Distributed intelligence (knowledge representation, management and fusion, science of coordination, and division of labor)
  - Systems for managing trust, reputation, and other critical elements in heterogeneous, dynamic, distant relationships
- Management of Knowledge Intensive Enterprises – research to understand how structured, global collections of knowledge can be brought to bear on complex decision-making processes so that processes can rapidly reconfigure and reschedule resources while the enterprise remains stable enough to carry out routine processes and achieve high levels of performance. The focus is on:
  - Adaptive scheduling and control of product dynamics, rapid reconfiguration and rescheduling of human and machine resources
  - Learning hidden workflow rules to optimize workflow
  - Distributed decision-making and appropriate schemes for distributing decision authority throughout hierarchies, understanding how information is shared, partitioned, and flows to the right places
  - How we measure the productivity of dynamically, re-configuring business processes, local vs. global problems such as performance across levels of analysis
  - Collaborative knowledge representation, acquisition, retrieval, and inference
- Knowledge Environments for Science and Engineering – research focused on:
  - Identifying the requirements of distributed scientific practices, how scientific practices are changing (e.g., due to more complex data sets, more interdisciplinary teams) and to what consequence
  - Understanding barriers to adoption and use, building trust across geographic boundaries, and IT strategies for sharing resources
  - Understanding the governance issues related to distributed work practices, facilities, and shared resources
  - Understanding copyright restrictions, information privacy and open source software issues related to collecting and harvesting knowledge across geographic and social boundaries
- Transforming Enterprise – research investigates:
  - Technologies and theories of electronic business, supply chains, economics of IT, productivity, etc.
  - Technologies and theories of collaborative and distributed work, including the development and use of collective knowledge representations and open source software development of human and machine resources
  - Understanding the various legal, social, and cultural issues when information, software, and autonomous proxies flow across boundaries
  - Understanding how to value information and evaluate risks and reputations in transactions with distant strangers
  - Understanding and mitigating information balkanization

Universal Access – FY 2004 emphasis is on developing new knowledge about how IT can empower people with disabilities, young children, seniors, and members of other traditionally under-represented groups to participate fully in the digital society.

CISE/CNS/Education and Workforce

This CNS cluster supports projects that integrate research and education across CISE, study the causes of the current lack of diversity in the IT workforce, and lead to a broadening of participation by all under-represented groups. The cluster works closely with all CISE divisions to achieve these goals. It also coordinates CISE participation in a

portfolio of NSF-wide education and workforce programs. SEW-related work:

- Information Technology Workforce Program – projects to develop and implement promising strategies to reverse the underrepresentation of women and minorities in IT education and careers

CISE/Combined Research and Curriculum Development and Educational Innovation Program

FY 2004 efforts focus on the design, development, testing, and dissemination of innovative IT approaches for increasing the effectiveness of educational experiences, including integration of research results into courses and curricula

NSF's FY 2005 plans for SEW-related R&D include:

- *A series of workshops on open standards, grid computing and innovation diffusion, dynamic knowledge repositories and collaboration, and agile frameworks for broad economic prosperity*
- *Approximately 120 new group projects under the ITR for National Priorities program, which focuses on IT advances in science and engineering, IT for economic prosperity and vibrant civil society, and IT for national and homeland security*
- *Digital Society and Technologies Program will focus on research to understand the challenges facing enterprises in dynamic environments and the ways in which IT can allow for complex, distributed decision-making, rapid reconfiguration, and resource scheduling and reallocation while achieving high levels of performance*

SEW NIH SEW

NIH's National Library of Medicine (NLM) is the pioneer supporter of advanced training in the emerging field of bioinformatics, whose practitioners bring both high-end computer science and medical expertise to the health-care arena. The need for bioinformatics skills spans biomedical research applications, telemedicine, and large-scale health care systems. The NLM program of institutional support and individual fellowships is establishing an academic training infrastructure and expanding the ranks of bioinformatics professionals, who are still far too few in number to fill the growing nationwide demand for these practitioners.

Training efforts in FY 2004 include:

- Institutional Grants: NLM supports training grants in medical informatics at ten major universities. They are intended to train predoctoral and postdoctoral students for research in medical informatics. Some offer special track training in such informatics-intensive fields as radiation oncology and biotechnology.
- Individual Fellowships: Informatics Research Training. Post-doctoral health science workers who identify a mentor, institution, research project, and curriculum are eligible to compete for these fellowships. They complement the institutional programs (described above) by making it possible for students to enter informatics training at any institution with appropriate mentor and facilities.
- Applied Informatics: These fellowships support training for those who will design or manage large information systems, or adapt applications developed by research workers to actual use in a clinic setting, classroom, laboratory, library, or office. Although many applicants will have doctorates, nurses, librarians, and other health professionals without doctoral degrees are encouraged to apply.
- Integrated Academic Information Management Systems (IAIMS): Over the last decade, NLM has supported the development of IAIMS in selected major medical centers. The experience gained by those who, through on-the-job training, have developed and implemented complex integrated information systems will now be systematically exploited through designated training slots at IAIMS sites.
- Biotechnology: For recent doctoral graduates, the National Research Council Research Associateship Program provides an opportunity for concentrated research in association with selected members of the NCBI scientific staff.
- Medical Informatics: The Marine Biological Laboratory, Woods Hole, Massachusetts, conducts an annual NLM-sponsored one-week course in medical informatics. Thirty trainees are selected from applicants in health professions, research, and librarianship. They receive intensive hands-on experience with a variety of medical information systems, including medical informatics, expert systems, and molecular biology databases.

- HBCU Toxicology Information Outreach: NLM's Toxicology Information Program (TIP) supports projects designed to strengthen the capacity of historically black colleges and universities (HBCUs) to train medical and other health professionals in the use of toxicological, environmental, occupational, and hazardous wastes information resources developed at NLM. A number of HBCUs with schools of medicine, pharmacy, and nursing are participating in the program, which includes training and free access to NLM's databases.
- Medical Informatics Elective: The Computer Science Branch, Lister Hill National Center for Biomedical Communications (LHNCBC) at NLM, conducts an eight-week elective in Medical Informatics, as part of NIH's Clinical Electives Program. Each spring this elective combines an extensive seminar series by senior figures in the field with an independent research project under the preceptorship of an LHNCBC professional. Eight to fourteen fourth-year medical students are admitted each year.
- Medical Informatics Training Program: LHNCBC conducts a Medical Informatics Training Program to provide support for faculty members, postdoctoral scientists, graduate students, and undergraduate students for research participation at the Center through visits of a few months to several years.

SEW NASA SEW

Learning Technologies Project (LTP) – NASA's educational technology incubator. LTP funds activities and collaborates with endeavors that incorporate NASA content with revolutionary technologies or innovative use of entrenched technologies to enhance education at all levels in the areas of science, technology, engineering, and mathematics (STEM). LTP's mission is to efficiently develop world-class educational products that:

- Are poised for the widest possible market diffusion, that inspire and educate
- Use innovative methods or emerging technologies
- Address educational standards utilizing NASA data. These products account for diverse learning environments whenever applicable and wherever possible.

FY 2004 projects include:

- Animated Earth – developing IT capabilities to provide Internet-accessible animated visualizations of important Earth Science processes, events, and phenomena to students, educators and the public, using NASA remote sensing and model data; includes work to determine which standards and protocols will be adopted to convey these visualizations over the Internet.and to implement and document a public server-side infrastructure to deliver these visualizations using the chosen standards and protocols.

• Information Accessibility Lab – research aiming to provide software tools that enable development of assistive instructional software applications for sensorily impaired K-12 students; utilize a combination of graphing, sonification, and mathematical analysis software to represent mathematical and scientific information; and provide unique, NASA-technology teaching tools that enhance STEM education for sensorily impaired students.

• What's the Difference? – research to: develop a simple to use software component that uses richly visual and highly interactive comparisons to teach science and math concepts; design the component so that additional information and new information sets can be developed and easily added by curriculum developers; enable developers of educational software applications to utilize the visual comparison componentry and information in their applications; provide information sets and tool capabilities beyond those delivered for this project's phase 1 effort.

• Virtual Lab – the goal of this R&D effort is to: provide students and their educators with virtual but realistic software implementations of sophisticated scientific instruments commonly used by NASA scientists and engineers; design and implement the virtual instruments such that additional specimens can be added easily and additional instruments can be used to study the same specimens; build on the LTP Phase 1 Virtual Lab by expanding the set of specimens for the Virtual Scanning Electron Microscope; provide mechanisms to enable independent applications to invoke and contain the virtual instruments.

SEW DOE/SC & DOE/NNSA SEW

DOE Computational Science Graduate Fellowship (CSGF): Funded by the Office of Science and Office of Defense Programs, this program works to identify and provide support for some of the very best computational science graduate students in the nation. The fellowships have supported more than 120 students at approximately 50 universities. The program is administered by the Krell Institute in Ames, Iowa.

The CSGF program partnership of DOE/SC and DOE/NNSA to develop the next generation of leaders in computational science will continue in FY 2005.

SEW DOE/NNSA SEW

Academic Strategic Alliance Program (ASAP): Through partnerships with universities, this element of the Advanced Simulation and Computing (ASC) Program pursues advances in the development of computational science, computer systems, mathematical modeling, and numerical mathematics important to the ASC effort. By supporting student involvement in these research efforts, ASAP aims to strengthen education and research in areas critical to the long-term success of ASC and the Stockpile Stewardship Program (SSP).

The ASAP will continue in FY 2005.

Participating Agency

SEW GSA SEW

Collaboration Expedition Workshop Series – monthly open workshops for representatives of Federal, state, and local government, community organizations, and the private sector to explore how to create a citizen-centric governance infrastructure supported by new technologies. Sponsored by GSA's Office of Intergovernmental Solutions in conjunction with the Emerging Technology Subcommittee of the Federal CIO Council. Objectives are to:

- Accelerate mutual understanding of the Federal Enterprise Architecture (FEA) initiative and commitments toward intergovernmental collaboration practices needed to implement the E-government Act of 2002
- Accelerate maturation of candidate technologies
- Expand intergovernmental collaboration and facilitate development of communities of practice
- Provide a forum for discussion of and experimentation with new IT capabilities and "best practices" in IT applications and deployment

By bringing IT developers and researchers together with practitioners across a broad range of government and citizen service sectors, the Expedition Workshops are developing a collaborative "incubator" process to facilitate emergence of a lasting enterprise architecture by providing:

- Realistic citizen-service scenarios for benchmarking performance
- Innovation practitioners with multilateral organizing skills
- Faster maturation and transfer of validated capabilities among intergovernmental partners
- Extensible e-gov components that transcend "stove-piping" through open standards technologies (OMB Circular A-119)

Plans for FY 2005 include:

- *Continue to develop participant skills in managing an IT innovation life-cycle process that scans emerging technology and fosters collaborative prototyping*
- *Expand network of intergovernmental partners to include new state and local government actors, non-governmental organizations, and Federal leaders with e-government responsibilities*
- *Identify 1-3 emerging technologies with the greatest potential for affecting the lasting value of enterprise architecture and incorporate findings into information/service offerings of the sponsors*

Interagency Working Group on IT R&D

Co-Chairs
Peter A. Freeman, NSF
David B. Nelson, NCO

NSF
Representative
Peter A. Freeman
Alternates
Deborah L. Crawford
C. Suzanne Iacano

DARPA
Representative
Ronald J. Brachman
Alternate
Barbara L. Yoon

NIH
Representative
Michael Marron
Alternates
Michael J. Ackerman
Robert L. Martino
Judith L. Vaitukaitis
Karen Skinner

NASA
Representative
Walter F. Brooks
Alternate
Vacant

DOE Office of Science
Representative
C. Edward Oliver
Alternates
Daniel A. Hitchcock
Norman H. Kreisman

NSA
Representative
George R. Cotter
Alternate
James Rupp

NIST
Representative
Kamie M. Roberts
Alternate
Larry H. Reeker

NOAA
Representative
William T. Turnbull

AHRQ
Representative
J. Michael Fitzmaurice

ODUSD (S&T)
Representative
André M. van Tilborg
Alternate
Robert Gold

EPA
Representative
Gary L. Walter
Alternate
Robin L. Dennis

DOE/NNSA
Representative
Robert Meisner

OMB
Representative
David S. Trinkle

NSTC
Representative
Charles H. Romine

NCO
Representative
David B. Nelson
Alternate
Sally E. Howe

PCA Coordinating Groups and Team Chairs

High End Computing (HEC) Coordinating Group
Co-Chairs
Frederick C. Johnson, DOE/SC
José L. Muñoz, NSF

Human Computer Interaction and Information Management (HCI&IM) Coordinating Group
Co-Chairs
William S. Bainbridge, NSF
Jean C. Scholtz, NIST

Large Scale Networking (LSN) Coordinating Group
Co-Chairs
Daniel A. Hitchcock, DOE/SC
George O. Strawn, NSF

LSN Teams:

Middleware and Grid Infrastructure Coordination (MAGIC) Team
Co-Chairs
Mary Anne Scott, DOE/SC
Kevin L. Thompson, NSF

Joint Engineering Team (JET)
Co-Chairs
Douglas G. Gatchell, NSF
George R. Seweryniak, DOE/SC
Vice-Chair
Paul E. Love, Internet2

Networking Research Team (NRT)
Co-Chairs
Thomas Ndousse-Fetter, DOE/SC
Taieb F. Znati, NSF

Software Design and Productivity (SDP) Coordinating Group
Co-Chairs
Frank D. Anger, NSF (deceased)
Thuc T. Hoang, DOE/NNSA

High Confidence Software and Systems (HCSS) Coordinating Group
Co-Chairs
Helen D. Gill, NSF
Vacant

Social, Economic, and Workforce Implications of IT and IT Workforce Development (SEW) Coordinating Group
Co-Chairs
C. Suzanne Iacono, NSF
Susan B. Turnbull, GSA

Universal Accessibility Team
Chair
Susan B. Turnbull, GSA

Participation in Federal NITRD Activities

The following are criteria developed by the multiagency IT research program that agencies considering participation can use to assess whether their research activities fit the NITRD profile.

NITRD Goals

Assure continued U.S. leadership in computing, information, and communications technologies to meet Federal goals and to support U.S. 21st century academic, industrial, and government interests

Accelerate deployment of advanced and experimental information technologies to maintain world leadership in science, engineering, and mathematics; improve the quality of life; promote long-term economic growth; increase lifelong learning; protect the environment; harness information technology; and enhance national security

Advance U.S. productivity and industrial competitiveness through long-term scientific and engineering research in computing, information, and communications technologies

Evaluation Criteria for Participation

Relevance of Contribution

The research must significantly contribute to the overall goals of the Federal Networking and Information Technology Research and Development (NITRD) Program and to the goals of one or more of the Program's seven Program Component Areas – High End Computing Infrastructure and Applications (HEC I&A), High End Computing Research and Development (HEC R&D), Human Computer Interaction and Information Management (HCI & IM), Large Scale Networking (LSN), Software Design and Productivity (SDP), High Confidence Software and Systems (HCSS), and Social, Economic, and Workforce Implications of Information Technology and Information Technology Workforce Development (SEW) – in order to enable the solution of applications problems that address agency mission needs and that place great demands on the technologies being developed by the Program.

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 Program Component Areas.

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 to conduct the proposed work, promote prospects for coordinated or joint funding, and address long-term resource implications.

Agency Approval

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

Agency NITRD Budgets by Program Component Area

FY 2004 Budget Estimates and FY 2005 Budget Requests

(dollars in millions)

Agency	High End Computing Infrastructure and Applications (HEC I&A)	High End Computing Research and Development (HEC R&D)	Human Computer Interaction and Information Management (HCI & IM)	Large Scale Networking (LSN)	Software Design and Productivity (SDP)	High Confidence Software and Systems (HCSS)	Social, Economic, and Workforce (SEW)	Totals
NSF	*207.0*	*87.3*	*144.7*	*94.6*	*61.0*	*65.1*	*93.9*	*754*
NSF	198.8	90.8	147.6	96.2	56.4	66.7	103.5	760
NIH	*103.4*	*6.3*	*128.0*	*103.0*	*4.1*	*10.8*	*12.3*	*368*
NIH	101.6	6.3	131.5	104.4	4.3	11.0	12..3	371
DOE Office of Science	*85.9*	*81.4*	*16.2*	*30.8*			*3.5*	*218*
DOE Office of Science	99.2	80.0		43.2			3.5	226
NASA	*67.4*	*33.0*	*67.1*	*43.4*	*59.2*	*24.2*	*6.7*	*301*
NASA	53.7	1.7	39.2	29.6	62.4	17.9	6.7	211
DARPA		*77.0*	*75.4*	*10.7*	*13.3*	*11.3*		*188*
DARPA		64.3	61.2	11.0		17.0		154
NSA		*31.4*		*1.5*		*25.4*		*58*
NSA		31.9		1.5		27.7		61
AHRQ			*31.0*	*24.4*				*55*
AHRQ			31.1	24.4				56
NIST	*2.5*	*1.2*	*4.4*	*3.9*	*4.8*	*7.6*		*24*
NIST	3.5	1.2	6.4	4.2	4.8	12.2		32
NOAA	*13.3*	*1.8*	*0.5*	*2.6*	*1.5*			*20*
NOAA	13.4	1.8	0.5	2.9	1.5			20
EPA	*2.3*		*2.0*					*4*
EPA	2.3		2.0					4
Subtotals	*481.7*	*319.2*	*469.2*	*314.9*	*144.0*	*144.4*	*116.4*	*1,990*
Subtotals	472.5	277.9	419.5	317.4	129.3	152.5	126.0	1,895
DOE/NNSA	*34.5*	*35.9*		*14.9*	*35.3*		*4.5*	*125*
DOE/NNSA	33.2	39.6		15.4	36.7		4.9	130
TOTALS[a]	*516.2*	*355.1*	*469.2*	*329.8*	179.3	*144.4*	*120.9*	*2,115*
TOTALS[a]	505.7	317.5	419.5	332.8	166.0	152.5	130.9	2,025

Note:

[a] These totals include discrepancies from numbers released with the President's FY 2005 Budget due to a combination of rounding, shifts in program estimates, and inadvertent omissions in the numbers shown for DOC and HHS.

Agency Contacts

AHRQ

* J. Michael Fitzmaurice, Ph.D., FACMI
Senior Science Advisor for Information Technology, Immediate Office of the Director
Agency for Healthcare Research and Quality
540 Gaither Road, Suite 3026
Rockville, MD 20850
(301) 427-1227
FAX: (301) 427-1210

DARPA

* Ronald J. Brachman, Ph.D.
Director, Information Processing Technology Office
Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington, VA 22203-1714
(703) 696-2264
FAX: (703) 696-4534

Robert B. Graybill
Program Manager, Information Processing Technology Office
Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington, VA 22203-1714
(703) 696-2220
FAX: (703) 696-4534

Gary M. Koob, Ph.D. (retired)
Program Manager, Information Processing Technology Office
Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington, VA 22203-1714
(703) 696-7463
FAX: (703) 696-4534

Mari Maeda, Ph.D.
Program Manager, Information Processing Technology Office
Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington, VA 22203-1714
(571) 218-4215
FAX: (703) 248-1866

Dylan Schmorrow, LCDR
Program Manager, Information Processing Technology Office
Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington, VA 22203-1714
(703) 696-4466
FAX: (703) 696-4534

** Barbara L. Yoon, Ph.D.
Deputy Director, Information Processing Technology Office
Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington, VA 22203-1714
(703) 696-7441
FAX: (703) 696-4534

DOE/NNSA

Thuc T. Hoang, Ph.D.
Computer Engineer, ASC PathForward & PSE Programs, Office of Advanced Simulation and Computing
National Nuclear Security Administration
Department of Energy
NA-114, Forrestal Building
1000 Independence Avenue, S.W.
Washington, DC 20585
(202) 586-7050
FAX: (202) 586-0405

Dimitri F. Kusnezov, Ph.D.
Director, Office of Advanced Simulation and Computing
National Nuclear Security Administration
Department of Energy
NA-114, Forrestal Building
1000 Independence Avenue, S.W.
Washington, DC 20585
202-586-1800
FAX: 202-586-0405

Sander L. Lee
Computer Scientist, Office of Integrated Computing Systems, Office of Advanced Simulation and Computing
National Nuclear Security Administration
Department of Energy
NA-114, Forrestal Building
1000 Independence Avenue, S.W.
Washington, DC 20585
202-586-2698
FAX: 202-586-0405

* Robert Meisner
Director, Office of Integrated Computing Systems, Office of Advanced Simulation and Computing
National Nuclear Security Administration
Department of Energy
NA-114, Forrestal Building
1000 Independence Avenue, S.W.
Washington, DC 20585-0104
(202) 586-0908
FAX: 202-586-0405

*** IWG Representative**
**** IWG Alternate**

DOE/SC

** Daniel A. Hitchcock, Ph.D.
Senior Technical Advisor for Advanced Scientific Computing Research, Office of Advanced Scientific Computing Research (OASCR)
Department of Energy
OASCR, SC-30
Germantown Building
1000 Independence Avenue, S.W.
Washington, DC 20585-1290
(301) 903-6767
FAX: (301) 903-4846

Frederick C. Johnson, Ph.D.
Program Manager, Mathematical, Information, and Computational Sciences (MICS) Division, Office of Advanced Scientific Computing Research (OASCR)
Department of Energy
OASCR/MICS, SC-31
Germantown Building
1000 Independence Avenue, S.W.
Washington, DC 20585-1290
(301) 903-3601
FAX: (301) 903-7774

Gary M. Johnson, Ph.D.
Program Manager, Mathematical, Information, and Computational Sciences (MICS) Division, Office of Advanced Scientific Computing Research (OASCR)
Department of Energy
OASCR/MICS, SC-31
Germantown Building
1000 Independence Avenue, S.W.
Washington, DC 20585-1290
(301) 903-4361
FAX: (301) 903-7774

** Norman H. Kreisman
Advisor, International Technology
Department of Energy, SC5
Mail Stop 3H049-FORS
Forrestal Building
1000 Independence Avenue, S.W.
Washington, DC 20585
(202) 586-9746
FAX: (202) 586-7719

Thomas Ndousse-Fetter, Ph.D.
Program Manager for Networking, Mathematical, Information, and Computational Sciences (MICS) Division, Office of Advanced Scientific Computing Research (OASCR)
Department of Energy
OASCR/MICS, SC-31
Germantown Building
1000 Independence Avenue, S.W.
Washington, DC 20585-1290
(301) 903-9960
FAX: (301) 903-7774

* C. Edward Oliver, Ph.D.
Associate Director, Office of Advanced Scientific Computing Research (OASCR)
Department of Energy
OASCR, SC-30
Germantown Building
1000 Independence Avenue, S.W.
Washington, DC 20585-1290
(301) 903-7486
FAX: (301) 903-4846

Mary Anne Scott, Ph.D.
Program Manager, Mathematical, Information, and Computational Sciences (MICS) Division, Office of Advanced Scientific Computing Research (OASCR)
Department of Energy
OASCR/MICS, SC-31
Germantown Building
1000 Independence Avenue, S.W.
Washington, DC 20585-1290
(301) 903-6368
FAX: (301) 903-7774

George R. Seweryniak
Program Manager, Mathematical, Information, and Computational Sciences (MICS) Division, Office of Advanced Scientific Computing Research (OASCR)
Department of Energy
OASCR/MICS, SC-31
Germantown Building
1000 Independence Avenue, S.W.
Washington, DC 20585-1290
(301) 903-0071
FAX: (301) 903-7774

EPA

** Robin L. Dennis, Ph.D.
Senior Science Program Manager, MD E243-01
Environmental Protection Agency
Research Triangle Park, NC 27711
(919) 541-2870
FAX: (919) 541-1379

* Gary L. Walter
Computer Scientist, MD E243-01
Environmental Protection Agency
Research Triangle Park, NC 27711
(919) 541-0573
FAX: (919) 541-1379

FAA

Ernest R. Lucier
Chief Scientist for Information Technology-Acting
Federal Aviation Administration
FAA/AIO-4
800 Independence Avenue, S.W.
Washington, DC 20591
(202) 385-8157
FAX: (202) 366-3064

Hal Pierson
Research Scientist
Federal Aviation Administration
FAA/AIO-4
800 Independence Avenue, S.W.
Washington, DC 20591
(202) 385-8153

FDA

Paul L. Jones
Senior Systems/Software Engineer, MCSE, CDP, CSQE
Food and Drug Administration
12720 Twinbrook Parkway (HZ-141)
Rockville, MD 20857
(301) 443-2536 x 164
FAX: (301) 443-9101

GSA

Susan B. Turnbull
Director, Center for IT Accommodation
General Services Administration
1800 F Street, N.W. (MKC) Room 2236
Washington, DC 20405
(202) 501-6214
FAX: (202) 219-1533

NASA

Terry Allard, Ph.D.
Acting Division Director, Mission and Science Measurement Technology
National Aeronautics and Space Administration
Code RD - Room 6J11
300 E Street, N.W.
Washington, DC 20546-0001
(202) 358-1891
FAX: (202) 358-3550

* Walter F. Brooks, Ph.D.
Chief, NASA Advanced Supercomputing Facility, NASA Ames Research Center
National Aeronautics and Space Administration
Mail Stop 258-5
Moffett Field, CA 94035
(650) 604-5699

Ricky W. Butler
Senior Research Engineer, Airborne Systems Competency Division, NASA Langley Research Center
National Aeronautics and Space Administration
Mail Stop 130
Hampton, VA 23681-2199
(757) 864-6198
FAX: (757) 864-4234

James R. Fischer
Project Manager, Earth and Space Sciences Project, NASA Goddard Space Flight Center
National Aeronautics and Space Administration
Code 930 - NASA/GSFC
Greenbelt, MD 20771
(301) 286-3465
FAX: (301) 286-1634

Kenneth Freeman
HPCC/NREN Project Manager, NASA Ames Research Center
National Aeronautics and Space Administration
Mail Stop 233-21
Moffett Field, CA 94035-1000
(650) 604-1263
FAX: (650) 604-3080

Marjory Johnson
Associate Project Manager, NREN, NASA Ames Research Center
National Aeronautics and Space Administration
Mail Stop 233-21
Moffett Field, CA 94035-1000
(650) 604-6922
FAX: (650) 604-3080

Kevin L. Jones
Network Engineer, NASA Ames Research Center
National Aeronautics and Space Administration
Mail Stop 233-21
Moffett Field, CA 94035-1000
(650) 604-2006
FAX: (650) 604-3080

Patricia M. Jones, Ph.D.
NASA Ames Research Center
National Aeronautics and Space Administration
Mail Stop 262-11
Room 204
Moffett Field, CA 94035-1000
(650) 604-1345
FAX: (650) 604-3323

Nand Lal, Ph.D.
Computer Scientist, Digital Library Technologies, NASA Goddard Space Flight Center
National Aeronautics and Space Administration
Code 933 - NASA/GSFC
Greenbelt, MD 20771
(301) 286-7350
FAX: (301) 286-1775

Tsengdar Lee, Ph.D.
Information Systems Specialist
National Aeronautics and Space Administration
300 E Street, S.W.
Washington, DC 20546
(202) 358-0860
FAX: (202) 358-2770

Michael R. Lowry
Senior Research Scientist and Area Chief, Automated Software Engineering, NASA Ames Research Center
National Aeronautics and Space Administration
Mail Stop 269-2
Moffett Field, CA 94035-1000
(650) 604-3369
FAX: (650) 604-3594

Piyush Mehrotra, Ph.D.
Senior Research Scientist, NASA Ames Research Center
National Aeronautics and Space Administration
M/S T27A-1
Moffett Field, CA 94035
(650) 604-5126
FAX: (650) 604-3957

Paul S. Miner, Ph.D.
Senior Research Engineer, NASA Langley Research Center
National Aeronautics and Space Administration
Mail Stop 130
Hampton, VA 23681-2199
(757) 864-6201
FAX: (757) 864-4234

Stephen Scott Santiago
Chief Information Officer, NASA Ames Research Center
National Aeronautics and Space Administration
Mail Stop 233-7
Moffett Field, CA 94035-1000
(650) 604-5015
FAX: (650) 604-6999

*** IWG Representative**
**** IWG Alternate**

Michael G. Shafto, Ph.D.
Program Manager, Human-Centered Computing, NASA Ames Research Center
National Aeronautics and Space Administration
Mail Stop 269-4
Moffett Field, CA 94035-1000
(650) 604-6170
FAX: (650) 604-6009

Eugene Tu, Ph.D.
Program Manager, Computing, Information, and Communications Technologies (CICT) Office, NASA Ames Research Center
National Aeronautics and Space Administration
Mail Stop 258-2
Moffett Field, CA 94035-1000
(650) 604-4486
FAX: (650) 604-4377

NIH

** Michael J. Ackerman, Ph.D.
Assistant Director for High Performance Computing and Communications, National Library of Medicine
National Institutes of Health
Building 38A, Room B1N30
8600 Rockville Pike
Bethesda, MD 20894
(301) 402-4100
FAX: (301) 402-4080

Eric Jakobsson, Ph.D.
Director, National Institute of General Medical Sciences (NIGMS) Center for Bioinformatics and Computational Biology, and Chair, NIH Biomedical Information Science and Technology Iniitiative Consortium
National Institutes of Health
Natcher Building, 2As55
Bethesda, MD 20892
(301) 451-6446
FAX: (301) 480-2802

Donald A.B. Lindberg, M.D.
Director, National Library of Medicine
National Institutes of Health
Building 38, Room 2E17B
8600 Rockville Pike
Bethesda, MD 20894
(301) 496-6221
FAX: (301) 496-4450

Jacob Maizel, Ph.D.
Biomedical Supercomputer Center, National Cancer Institute, Frederick Cancer Research and Development Center
National Institutes of Health
P.O. Box B, Building 469, Room 151
Frederick, MD 21702-1201
(301) 846-5532
FAX: (301) 846-5598

* Michael Marron, Ph.D.
Director, Division for Biomedical Technology Research and Research Resources, National Center for Research Resources
National Institutes of Health
6701 Democracy Boulevard, Room 962
Bethesda, MD 20892-4874
(301) 435-0755
FAX: (301) 480-3659

** Robert L. Martino, Ph.D.
Director, Division of Computational Bioscience and Chief, Computational Bioscience and Engineering Laboratory, Center for Information Technology (CIT)
National Institutes of Health
12 South Drive, MSC 5654
Building 12A, Room 2033
Bethesda, MD 20892-5654
(301) 496-1112
FAX: (301) 402-2867

** Karen Skinner, Ph.D.
Deputy Director for Science and Technology Development, Division of Neuroscience and Behavior Research, National Institute on Drug Abuse
National Institutes of Health
6001 Executive Boulevard, Room 4255
Bethesda, MD 20892-9651
(301) 435-0886

** Judith L. Vaitukaitis, M.D.
Director, National Center for Research Resources
National Institutes of Health
31B Center Drive
Building 31, Room 3B11
Bethesda, MD 20892-2128
(301) 496-5793
FAX: (301) 402-0006

NIST

Paul E. Black, Ph.D.
Computer Scientist, Information Technology Laboratory
National Institute of Standards and Technology
100 Bureau Drive, Stop 8970
Gaithersburg, MD 20899-8970
(301) 975-4794
FAX: (301) 926-3696

Judith E. Devaney, Ph.D.
Group Leader, Scientific Applications and Visualization Group, Mathematical and Computational Sciences Division, Information Technology Laboratory
National Institute of Standards and Technology
100 Bureau Drive, Stop 8911
Gaithersburg, MD 20899-8911
(301) 975-2882
FAX: (301) 975-3218

Simon J. Frechette
SIMA Program Manager
National Institute of Standards and Technology
100 Bureau Drive, Stop 8260
Gaithersburg, MD 20899-8260
(301) 975-3335
FAX: (301) 258-9749

Cita M. Furlani
Chief Information Officer
National Institute of Standards and Technology
100 Bureau Drive, Stop 1800
Gaithersburg, VA 20899-1800
(301) 975-6500
FAX: (301) 975-4160

Martin Herman, Ph.D.
Chief, Information Access and User Interfaces Division, Information Technology Laboratory
National Institute of Standards and Technology
100 Bureau Drive, Stop 8940
Gaithersburg, MD 20899-8940
(301) 975-4495
FAX: (301) 670-0939

Douglas Montgomery
Manager, Internetworking Technologies Group, Advanced Network Technologies Division, Information Technology Laboratory
National Institute of Standards and Technology
100 Bureau Drive, Stop 8920
Gaithersburg, MD 20899-8920
(301) 975-3630
FAX: (301) 590-0932

Steven R. Ray, Ph.D.
Chief, Manufacturing Systems Integration Division, Manufacturing Engineering Laboratory
National Institute of Standards and Technology
100 Bureau Drive, Stop 8260
Gaithersburg, MD 20899-8260
(301) 975-3524
FAX: (301) 258-9749

** Larry H. Reeker, Ph.D.
Senior Computer Scientist, Information Technology Laboratory
National Institute of Standards and Technology
100 Bureau Drive, Stop 8901
Gaithersburg, MD 20899-8901
(301) 975-5147
FAX: (301) 948-1784

* Kamie M. Roberts
Acting Deputy Director, Associate Director for Federal and Industrial Relations, and Chief of Staff, Information Technology Laboratory
National Institute of Standards and Technology
100 Bureau Drive, Stop 8901
Gaithersburg, MD 20899-8901
(301) 975-2982
FAX: (301) 948-1784

Ronald S. Ross, Ph.D.
Computer Science Division, Information Technology Laboratory
National Institute of Standards and Technology
100 Bureau Drive, Stop 8930
Gaithersburg, MD 20899-8930
(301) 975-5390

Jean C. Scholtz, Ph.D.
Information Access Division, Information Technology Laboratory
National Institute of Standards and Technology
100 Bureau Drive, Stop 8940
Gaithersburg, MD 20899-8940
(301) 975-2520
FAX: (301) 975-5287

David Su, Ph.D.
Chief, Advanced Network Technologies Division, Information Technology Laboratory
National Institute of Standards and Technology
100 Bureau Drive, Stop 8920
Gaithersburg, MD 20899-8920
(301) 975-6194
FAX: (301) 590-0932

NOAA

Brian Gross
Geophysical Fluid Dynamics Laboratory
National Oceanic and Atmospheric Administration
Forrestal Campus, U.S. Route 1
P.O. Box 308
Princeton, NJ 08542
(609) 452-6597
FAX: (609) 987-5063

Mike Kane
National Oceanic and Atmospheric Administration
Room 9855
1315 East-West Highway
Silver Spring, MD 20910
(301) 713-3575 x200
FAX: (301) 713-4040

Ants Leetmaa, Ph.D.
Director, Geophysical Fluid Dynamics Laboratory
National Oceanic and Atmospheric Administration
Forrestal Campus, U.S. Route 1
P.O. Box 308
Princeton, NJ 08542-0308
(609) 452-6502
FAX: (609) 987-5070

Alexander E. MacDonald, Ph.D.
Director, Forecast Systems Laboratory
National Oceanic and Atmospheric Administration
325 Broadway
Boulder, CO 80303
(303) 497-6378
FAX: (303) 497-6821

* William T. Turnbull
Deputy Chief Information Officer and Director, HPCC Office
National Oceanic and Atmospheric Administration
Room 9636
1315 East-West Highway
Silver Spring, MD 20910
(301) 713-9600 x 133
FAX: (301) 713-4040

Louis Uccellini
Director, National Centers for Environmental Prediction
National Oceanic and Atmospheric Administration
5200 Auth Road, Room 101
Camp Springs, MD 20746
(301) 763-8000
FAX: (301) 763-8434

*** IWG Representative**
**** IWG Alternate**

Gary M. Wohl
New Technology Coordinator, National Weather Service
National Oceanic and Atmospheric Administration
5200 Auth Road, Room 307
Camp Springs, MD 20746-4304
(301) 763-8000 x7157
FAX: (301) 763-8381

NSA

* George R. Cotter
Office of Corporate Assessments
National Security Agency
9800 Savage Road, Suite 6217
Fort George G. Meade, MD 20755-6217
(301) 688-6434
FAX: (301) 688-4980

Candace S. Culhane
Computer Systems Researcher
National Security Agency
9800 Savage Road, Suite 6107
Fort George G. Meade, MD 20755-6107
(443) 479-0569
FAX: (301) 688-3633

Craig Holcomb
Senior Computer Scientist/Technical Director, NSA/CSS CIO, Information and Technology Governance and Policy Office
National Security Agency
9800 Savage Road, Suite 6459
Fort George G. Meade, MD 20755-6459
(301) 688-8762
FAX: (301) 688-4995

William Bradley Martin
Senior Computer Scientist
National Security Agency
9800 Savage Road, Suite 6511
Fort George G. Meade, MD 20750-6511
(301) 688-1057
FAX: (301) 688-6766

** James Rupp
National Security Agency
9800 Savage Road, Suite 6468
Fort George G. Meade, MD 20755-6468
(443) 479-8944

William J. Semancik, Ph.D.
Director, Laboratory for Telecommunications Sciences
National Security Agency
c/o U. S. Army Research Laboratory
Adelphi Laboratory Center
2800 Powder Mill Road, Building 601, Room 131
Adelphi, MD 20783-1197
(301) 688-1709
FAX: (301) 291-2591

James Widmaier
Senior Infosec Researcher
National Security Agency
9800 Savage Road, Suite 6511
Fort George G. Meade, MD 20755-6511
(301) 688-1043
FAX: (301) 688-0255

NSF

S. Kamal Abdali, Ph.D.
Division Director, Computing and Communication Foundations Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1115
Arlington, VA 22230
(703) 292-8910
FAX: (703) 292-9059

Gregory R. Andrews, Ph.D.
Division Director, Computer and Network Systems Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1175
Arlington, VA 22230
(703) 292-8950
FAX: (703) 292-9010

William S. Bainbridge, Ph.D.
Deputy Division Director and Program Director, Information and Intelligent Systems Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1125
Arlington, VA 22230
(703) 292-8930
FAX: (703) 292-9073

Lawrence E. Brandt
Program Manager, Information and Intelligent Systems Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1125
Arlington, VA 22230
(703) 292-8911
FAX: (703) 292-9073

John C. Cherniavsky, Ph.D.
Senior EHR Advisor for Research, Research, Evaluation and Communication Division, Directorate for Education and Human Resources
National Science Foundation
4201 Wilson Boulevard, Suite 855
Arlington, VA 22230
(703) 292-5136
FAX: (703) 292-9046

** Deborah L. Crawford
Deputy Assistant Director, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1105
Arlington, VA 22230
(703) 292-8900
FAX: (703) 292-9074

Frederica Darema, Ph.D.
Senior Science and Technology Advisor, Computer and Network Systems Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1175
Arlington, VA 22230
(703) 292-8950
FAX: (703) 292-9010

* Peter A. Freeman, Ph.D.
Assistant Director, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1105
Arlington, VA 22230
(703) 292-8900
FAX: (703) 292-9074

James C. French, Ph.D.
Program Director, Information and Intelligent Systems Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1125
Arlington, VA 22230
(703) 292-8936
FAX: (703) 292-9073

Douglas G. Gatchell
Program Director, Shared Cyberinfrastructure Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1145
Arlington, VA 22230
(703) 292-8962
FAX: (703) 292-9060

Helen D. Gill, Ph.D.
Program Director, Computer and Network Systems Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1175
Arlington, VA 22230
(703) 292-8950
FAX: (703) 292-9010

Sol J. Greenspan, Ph.D.
Program Director, Computing and Communication Foundations Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1115
Arlington, VA 22230
(703) 292-8910
FAX: (703) 292-9059

** C. Suzanne Iacono, Ph.D.
Program Director, Information and Intelligent Systems Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1125
Arlington, VA 22230
(703) 292-8930
FAX: (703) 292-9073

Sangtae Kim, Ph.D.
Division Director, Computer and Network Systems Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1145
Arlington, VA 22230
(703) 292-8970
FAX: (703) 292-9060

Carl E. Landwehr, Ph.D.
Program Director, Computer and Network Systems Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1175
Arlington, VA 22230
(703) 292-8950
FAX: (703) 292-9010

Stephen R. Mahaney, Ph.D.
Senior Advisor for Budget Management and Planning and Policy, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1105
Arlington, VA 22230
(703) 292-8900
FAX: (703) 292-9074

José L. Muñoz, Ph.D.
Deputy Division Director, Shared Cyberinfrastructure Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1145
Arlington, VA 22230
(703) 292-8970
FAX: (703) 292-9060

Michael J. Pazzani, Ph.D.
Division Director, Information and Intelligent Systems Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1125
Arlington, VA 22230
(703) 292-8930
FAX: (703) 292-9073

Edwina L. Rissland, Ph.D.
Program Director, Information and Intelligent Systems Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1125
Arlington, VA 22230
(703) 292-8918
FAX: (703) 292-9073

George O. Strawn, Ph.D.
Chief Information Officer, Office of Information and Resource Management
National Science Foundation
4201 Wilson Boulevard, Suite 305
Arlington, VA 22230
(703) 292-8102
FAX: (703) 292-9084

Kevin L. Thompson
Program Director, Shared Cyberinfrastructure Division, Directorate for Computer and Information Science and Engineering
National Science Foundation
4201 Wilson Boulevard, Suite 1145
Arlington, VA 22230
(703) 292-8962
FAX: (703) 292-9060

ODUSD (S&T)

Larry P. Davis, Ph.D.
Deputy Director, High Performance Computing Modernization Office, Office of the Deputy Under Secretary of Defense (Science and Technology)
Department of Defense
1010 North Glebe Road, Suite 510
Arlington, VA 22201
(703) 812-8205
FAX: (703) 812-9701

** Robert Gold
Information Systems Directorate, Office of the Deputy Under Secreatary of Defense (Science and Technology)
Department of Defense
1777 North Kent Street, Suite 9030
Rosslyn, VA 22209
(703) 588-7411
FAX: (703) 588-7560

John Grosh
General Engineer, Office of the Deputy Under Secretary of Defense (Science and Technology)
Department of Defense
1777 North Kent Street, Suite 9030
Rosslyn, VA 22209
(703) 588-7413
FAX: (703) 588-7560

Cray J. Henry
Director, High Performance Computing Modernization Office, Office of the Deputy Under Secretary of Defense (Science and Technology)
Department of Defense
1010 North Glebe Road, Suite 510
Arlington, VA 22201
(703) 812-8205
FAX: (703) 812-9701

Rodger Johnson
Program Manager, Defense Research and Engineering Network, High Performance Computing Modernization Office, Office of the Deputy Under Secretary of Defense (Science and Technology)
1010 North Glebe Road, Suite 510
Arlington, VA 22201
(703) 812-8205
FAX: (703) 812-9701

* André M. van Tilborg, Ph.D.
Director, Information Systems
Office of the Deputy Under Secretary of Defense (Science and Technology)
Department of Defense
1777 North Kent Street, Suite 9030
Rossyln, VA 22209
(703) 588-7443
FAX: (703) 588-7560

ONR

Helen M. Gigley, Ph.D.
Program Officer
Office of Naval Research
800 North Quincy Street
Code 342
Arlington, VA 22217-5660
(703) 696-2160
FAX: (703) 292-9084

Ralph F. Wachter, Ph.D.
Program Manager
Office of Naval Research
800 North Quincy Street
Code 311/361
Arlington, VA 22217-5660
(703) 696-4304
FAX: (703) 696-2611

USAF

Raymond A. Bortner
Senior Electronic Engineer, AFRL/VACC
Air Force Research Laboratory
2130 Eighth Street, Room 270
Wright-Patterson AFB, OH 45433
(937) 255-8292
FAX: (937) 255-8297

*** IWG Representative**
**** IWG Alternate**

Executive Office of the President

OMB

David S. Trinkle
Program Examiner
Office of Management and Budget
Room 8225
New Executive Office Building
725 17th Street, N.W.
Washington, DC 20503
(202) 395-4706
FAX: (202) 395-4652

NSTC

Charles H. Romine
Office of Science and Technology Policy
Executive Office of the President
New Executive Office Building
725 17th Street, N.W.
Washington, DC 20502
(202) 456-6054
FAX: (202) 456-6021

Glossary

3D
Three-dimensional

3MRA
EPA's Multimedia, Multipathway, Multireceptor Risk Assessment modeling system

6TAP
IPv6 Transit Access Point

A

AAMP7
Latest member of the Advanced Architecture Microprocessor family that supports high-assurance application development exploiting intrinsic partitioning

ACE
NIST's Automatic Content Extraction program

ACRT
DOE/SC's Advanced Computing Research Testbeds activity

ADAPT
NASA's Automatic Data Alignment and Placement Tool

AeroDB
Aeronautical Database

AEX
NIST's Automating Equipment Information Exchange project

AFRL
Air Force Research Laboratory

AFWA
Air Force Weather Agency

AHRQ
Agency for Healthcare Research and Quality

AI&CS
NSF/CISE's Artificial Intelligence and Cognitive Science program

AIO
FAA's Administrative Information Office

Altix
A line of computers produced by SGI, Inc.

ALU
Arithmetic logic unit

ANL
DOE's Argonne National Laboratory

ANTD
NIST's Advanced Network Technologies Division

APDEC
DOE/SC's Algorithmic and Software Framework for Applied Partial Differential Equations ISIC

API
Application program interface

AQMA
EPA's Air Quality Modeling Applications program

AQUAINT
NSA/NIST's Advanced Question and Answering for Intelligence program

ARDA
Intelligence community's Advanced Research and Development Activity

ARL
Army Research Laboratory

ARO
Army Research Office

ASAP
DOE/NNSA's Academic Strategic Alliance Program

ASC
DOE/NNSA's Advanced Simulation and Computing program (formerly ASCI, for Accelerated Strategic Computing Initiative)

ASC
Army Signal Command, a DoD/HPCMPO Major Shared Resource Center

ASCR or OASCR
DOE/SC's Office of Advanced Scientific Computing Research

ASC Blue Mountain
DOE/NNSA program's massively parallel 3-teraflops SGI Origin supercomputing platform at Los Alamos National Laboratory

ASC Purple
DOE/NNSA program's 100-teraflops IBM SMP supercomputing platform under development, in tandem with Blue Gene/L, at Lawrence Livermore National Laboratory

ASC Q
DOE/NNSA program's 20-teraflops system, based on HP/Compaq technology, at Los Alamos National Laboratory

ASC Red
ASC program's first teraflops-level platform, by Intel, installed at Sandia National Laboratory in 1996

ASC Red Storm
DOE/NNSA program's 40-teraflops Cray system under development at Sandia National Laboratory

ASC White
DOE/NNSA program's massively parallel, 12-teraflops IBM supercomputing platform at Lawrence Livermore National Laboratory

ATDnet
DoD's Advanced Technology Demonstration network

ATM
Asynchronous transfer mode

B

BECON
NIH's Bioengineering Consortium

BGP
Border gateway protocol

BIO
NSF's Biological Sciences Directorate

BioAPI
Biometric application programming interface

BISTI
NIH's Biomedical Information Science and Technology Initiative

BISTIC
NIH's Biomedical Information Science and Technology Initiative Consortium

BLAS
Basic linear algebra subroutines

Blue Gene/L
Scalable experimental new supercomputing system being developed by IBM in partnership with DOE/SC and DOE/NNSA; expected to achieve 300-teraflops+ processing speeds

BOSSnet
BoSton-South network, MIT's Boston to Washington, D.C., fiber-optic network

C

CAD
Computer-aided design

CAF
Co-Array Fortran

CAIB
NASA Columbia Accident Investigation Board

CalREN2
California Research and Education Network 2

CalTech
California Institute of Technology

Canarie2
Canadian Network for the Advancement of Research, Industry, and Education 2

CAPO
NASA's Computer-Aided Parallelizer and Optimizer

CAST
DARPA's Compact Aids for Speech Translation program

CBEFF
Common Biometric Exchange File Format

CCAMP
Common control and measurement plane

CCF
NSF/CISE's Computing and Communications Foundations Division

CCM
DoD/HPCMPO's Computational Chemistry and Materials Science CTA

CDROM
Compact disc read-only memory

CEA
DoD/HPCMPO's Computational Electromagnetics and Acoustics CTA

CEN
DoD/HPCMPO's Computational Electronics and Nanoelectronics CTA

CERN
European Laboratory for Particle Physics

CEV
NASA's crew exploration vehicle

CFD
Computational fluid dynamics

CFD
DoD/HPCMPO's Computational Fluid Dynamics CTA

CG
Coordinating Group

CHEETAH
NSF's Circuit-switched High-speed End-to-End ArcHitecture network

CHSSI
DoD/HPCMPO's Common High Performance Computing Software Initiative

CI-TEAM
NSF/CISE's Cyber-Infrastructure Training, Education, Advancement, and Mentoring activity

CI
Cyberinfrastructure

CIAPP
NASA'S Computational Intelligence for Aerospace Power and Propulsion systems

CICT
NASA's Computing, Information, and Communications Technology program

CIF
Common industry format

CIO
Chief information officer

CIP
Critical infrastructure protection

CISE
NSF's Computer and Information Science and Engineering Directorate

CMAQ
EPA's Community Multi-Scale Air Quality modeling system

CMM
Carnegie Mellon University's Software Engineering Institute's Capability Maturity Models

CNS
NSF/CISE's Computer and Networking Systems Division

CoE
EPA's Center of Excellence

CONUS
Continental United States

CORBA
Common object request broker architecture

COTS
Commercial-off-the-shelf

CPU
Central processing unit

CSD
NIST/ITL's Computer Security Division

CSGF
Computational Science Graduate Fellowship program supported by DOE/SC and DOE/NNSA

CSM
DoD/HPCMPO's Computational Structural Mechanics CTA

CSTB
Computer Science and Telecommunications Board of the National Research Council

CSTL
NIST's Chemical Science and Technology Lab

CSU
Colorado State University

CTA
DoD/HPCMPO's Computational Technology Area

D

DAML
DARPA Agent Markup Language

DARPA
Defense Advanced Research Projects Agency

DCE
Distributed computing environment

DETER
NSF's cyber DEfense Technology Experimental Research network

DHHS
Department of Health and Human Services

DHS
Department of Homeland Security

DiffServ
Differentiated services

DisCom
DOE/NNSA's Distance Computing program

DiVAs
DOE/SC's Distributed Visualization Architecture workshops

DIVERSE
NIST's Device-Independent Virtual Environments – Reconfigurable, Scalable, Extensible

DLMF
NIST/NSF's Digital Library of Mathematical Functions

DMF
SGI's Data Management Facility, a mass storage system

DMTF
Distributed Management Task Force

DNA
Deoxyribonucleic acid

DNS
Domain Name System

DoD
Department of Defense

DoD/HPCMPO
DoD/High Performance Computing Modernization Program Office

DoD/OSD
DoD/Office of the Secretary of Defense

DOE
Department of Energy

DOE/NNSA
DOE/National Nuclear Security Administration

DOE/SC
DOE/Office of Science

DOJ
Department of Justice

DRAGON
NSF's Dynamic Resource Allocation (via GMPLS) Optical Network

DREN
Defense Research and Engineering Network

DS
Data Send, a network transmission speed standard

DSM
Distributed shared memory

DTC
NOAA's Development Test Center

DUC
NIST's Document Understanding Conference

E

EARS
DARPA/NIST Effective, Affordable, Reusable Speech-To-Text program

ebXML
electronic business Extensible Markup Language

ECCO
NASA's Estimating the Circulation and Climate of the Ocean project

ECEI
NIST's Electronic Commerce for the Electronics Industry

ECS
NASA's Engineering for Complex Systems program

EDRAM
Erasable dynamic random-access memory

EHR
NSF's Education and Human Resources Directorate

EHR
Electronic health record

EHRS
Electronic health record system

EINs
NSF's Experimental Infrastructure Networks

ENG
NSF's Engineering Directorate

EPA
Environmental Protection Agency

EQM
DoD/HPCMPO's Environmental Quality Modeling and Simulation CTA

ERDC
U.S. Army's Engineering Research and Development Center, a DoD/HPCMPO Major Shared Resource Center

ESCHER
Embedded Systems Consortium for Hybrid and Embedded Research, a joint effort of DARPA and NSF

ESMF
Earth System Modeling Framework

ESnet
DOE/SC's Energy Sciences network

ESTO
NASA's Earth Sciences Technology Office

ETF
NSF's Extensible Terascale Facility

F

FAA
Federal Aviation Administration

FBI
Federal Bureau of Investigation

FDA
Food and Drug Administration

FEA
Federal Enterprise Architecture

FMS
NOAA's Flexible Modeling System

FMS
DoD/HPCMPO's Forces Modeling and Simulation CTA

FNMOC
U.S. Navy's Fleet Numerical Meteorology and Oceanography Center

FPGA
Field programmable gate array

FpVTE
Fingerprint vendor technology evaluation

FSL
NOAA's Forecast Systems Laboratory

FTP
File transfer protocol

FY
Fiscal year

G

Gbps
Gigabits per second

Gbytes, or GB
Gigabytes per second

GEO
NSF's Geosciences Directorate

GFDL
NOAA's Geophysical Fluid Dynamics Laboratory

Gflops
Gigaflops, billions of floating point operations per second

GGF
Global Grid Forum

GGMAO
NASA's Goddard Global Modeling and Assimilation Office

GHz
Gigahertz

GigE
Gigabit Ethernet

GIPS
Giga-integers per second

GIS
Geographical information system

GISS
NASA's Goddard Institute for Space Studies

GLASS
NIST's GMPLS/Optical Simulation tool project

GLIS
NASA's Goddard Land Information Systems project

GMPLS
Generalized multi-protocol label switching

GOPS
Giga-operations per second

GSA
General Services Administration

GSFC
NASA's Goddard Space Flight Center

GUI
Graphical user interface

GUPs
Giga-updates per second

H

H-ANIM
Human animation

HARD
NIST's High Accuracy Retrieval from Documents program

HBCUs
Historically black colleges and universities

HBM
Haskell on Bare Metal

HCI&IM
Human Computer Interaction and Information Management, one of NITRD's seven Program Component Areas

HCSS
High Confidence Software and Systems, one of NITRD's seven Program Component Areas

HDCCSR
Highly Dependable Computing and Communications Systems Research program, a joint effort of NASA and NSF

HEC
High-end computing

HEC I&A
HEC Infrastructure and Applications, one of NITRD's seven Program Component Areas

HEC R&D
HEC Research and Development, one of NITRD's seven Program Component Areas

HECRTF
High-End Computing Revitalization Task Force

HEC-URA
HEC University Research Activity, funded by DARPA, DOE/SC, and NSF

HIT
AHRQ's Patient Safety Health Information Technology Initiative

HL-7
Health Level Seven, an electronic medical-records standard

HP
Hewlett Packard

HPC
High-performance computing

HPCC
High-performance computing and communications

HPCMPO
DoD's High Performance Computing Modernization Program Office

HPCS
DARPA's High Productivity Computing Systems program

HRI
NIST's Human Robot Interaction activity

HSAIs
DoD/HPCMPO's High Performance Computing Software Applications Institutes

HTML
HyperText Markup Language

HVAC
Heating, ventilation, air conditioning

HVAC/R
Heating, ventilation, air conditioning, and refrigeration

I

IAD
NIST's Information Access Division

IAFIS
FBI's Integrated Automated Fingerprint Identification System

IAIMS
NIH/NLM's Integrated Academic Information Management Systems

IARG
NSA's Information Assurance Research Group

IBM
International Business Machines

ICAT
NIST/ITL/CSD's searchable index of information on computer vulnerabilities

ICR
NITRD Program's Interagency Coordination Report

IDC
A global market intelligence and advisory firm

IDE
Integrated development environment

IDS
NIST's Intrusion Detection System

IEEE
Institute of Electrical and Electronics Engineers

IETF
Internet Engineering Task Force

IIS
NSF/CISE's Information and Intelligent Systems Division

iLab
NASA's Information Power Grid Virtual Laboratory, a software package for creating and monitoring parameter studies of grid resources

IMPI
NIST's interoperable message-passing interface

IMT
DoD/HPCMPO's Integrated Modeling and Test Environments CTA

INCITE
DOE/SC's Innovative and Novel Computational Impact on Theory and Experiment program

INCITS
International Committee for IT Standards

I/O
Input/output

IP
Internet protocol

IPG
NASA's Information Power Grid

IPSec
IP security

IPTO
DARPA's Information Processing Technology Office

IPv6
Internet protocol, version 6

ISD
NIST/MEL's Intelligent Systems Division

ISICs
DOE/SC's Integrated Software Infrastructure Centers

ISO
International Organization for Standardization

ISPAB
NIST's Information Security and Privacy Advisory Board

ISS
International Space Station

IT
Information technology

ITL
NIST's Information Technology Laboratory

ITR
NSF's Information Technology Research program

IT R&D
Information technology research and development

IU
Indiana University

IUSR
NIST's Industry USability Reporting activity

IV&V
Independent verification and validation

iVDGL
NSF/CISE's International Virtual Data Grid Laboratory

IWG
Interagency Working Group

IXO
DARPA's Information Exploitation Office

J

Java
An operating system-independent programming language

Javaspaces
Java-based component of the Jini open architecture for networking

JET
LSN's Joint Engineering Team

JETnets
Federal research networks supporting networking researchers and advanced applications development

Jini
Java-based open architecture for development of network services

K

Ka-band
Kurtz above-band, a portion of the K band of the electromagnetic spectrum ranging from about 18 to 40 GHz; often used in communications satellites

L

LANL
DOE's Los Alamos National Laboratory

LBNL or LBL
DOE's Lawrence-Berkeley National Laboratory

LEAD
NSF's Linked Environments for Atmospheric Discovery

LHNCBC
NIH/NLM's Lister Hill National Center for Biomedical Communications

LSN
Large Scale Networking, one of NITRD's seven Program Component Areas

LSTAT
FDA's Life Support for Trauma and Transport device

LTP
NASA's Learning Technologies Project

LTS
NSA's Laboratory for Telecommunications Sciences

M

MAA
EPA's Multimedia Assessments and Applications framework

MAGIC
LSN's Middleware and Grid Infrastructure Coordination team

MAX
Mid-Atlantic Exchange

Mbytes, or MB
Megabytes

MCHIP
DARPA's Multipoint Congramoriented High-performance Internet Protocol

MDS
NASA's Mission Data Systems program

MEL
NIST's Manufacturing Engineering Laboratory

MHz, or MH
Megahertz

MIT
Massachusetts Institute of Technology

MLP
Multi-level parallelism

MoBIES
DARPA's Model-Based Integration of Embedded Systems program

MOU
Memorandum of understanding

MPEG-7
Moving Picture Experts Group's multimedia content description interface, release 7

MPI
Message-passing interface

MPICH
Freely available, high-performance portable implementation of MPI

MPLS
Multi-protocol label switching

MPS
NSF's Mathematical and Physical Sciences Directorate

MSID
NIST's Manufacturing Systems Integration Division

MSRCs
DoD/HPCMPO's Major Shared Resource Centers

MT
NIST's Machine Translation program

MTBF
Mean time between failures

N

NAP
Network access point

NASA
National Aeronautics and Space Administration

NAVO
U.S. Navy's Naval Oceanographic Office

NCAR
NSF-supported National Center for Atmospheric Research

NCBI
NIH/NLM's National Center for Biotechnology Information

NCEP
NOAA's National Center for Environmental Prediction

NCO
National Coordination Office for Information Technology Research and Development

NCSA
NSF-supported National Center for Supercomputing Applications

NERSC
DOE/SC's National Energy Research Scientific Computing Center

NESC2
National Environmental Scientific Computing Center

NEST
DARPA's Networked Embedded Systems Technology program

NEXRAD
NOAA's Next Generation Weather Radar System

NGIX
Next Generation Internet Exchange point

NHII
National Health Information Infrastructure

NIAP
NIST/NSA National Information Assurance Partnership

NIGMS
NIH's National Institute of General Medical Sciences

NIH
National Institutes of Health

NIMD
NIST's Novel Intelligence for Massive Data activity

NISN
NASA Integrated Services Network

NIST
National Institute of Standards and Technology

NITRD
Networking and Information Technology Research and Development

NLM
NIH's National Library of Medicine

NLR
National LambdaRail

NMI
NSF Middleware Initiative

NOAA
National Oceanographic and Atmospheric Administration

NO_x
Nitrous oxide

NPACI
NSF-supported National Partnership for Advanced Computational Infrastructure

NREN
NASA Research and Education Network

NRL
Naval Research Laboratory

NRNRT
National radio network research testbed

NRT
LSN's Networking Research Team

NRT
NSF/CISE's Networking Research Testbed program

NS-500
A GigE security device that uses Virtual Private Network

NSA
National Security Agency

NSF
National Science Foundation

NSRL
National Software Reference Library, joint effort of DOJ, FBI, and NIST

NSTC
White House National Science and Technology Council

NWChem
Chemistry modeling software developed by DOE's Pacific Northwest National Laboratory

NWS
NOAA's National Weather Service

O

O3K
An SGI Origin computer

OASCR or ASCR
DOE/SC's Office of Advanced Scientific Computing Research

OAV
Organic air vehicle

OC-x
Optical carrier rate, an optical network transmission standard. The base rate (OC-1) is 51.84 Mbps; higher transmission speeds such as OC-3 or OC-48 are multiples of OC-1.

OCONUS
Outside the continental United States

ODDR&E
DoD's Office of the Director, Defense Research and Engineering

OIL
Ontology inference layer

OMB
White House Office of Management and Budget

ONR
Office of Naval Research

OOF
NIST's Object-Oriented Finite Element Modeling of Material Microstructures software

OOMMF
NIST's Object-Oriented MicroMagnetic Computing Framework

OpenMP
An open message-passing standard

ORNL
DOE's Oak Ridge National Laboratory

OS
Operating system

Osker
Oregon separation kernel

OSPF
Open shortest path first, an interior gateway protocol

OSS
Open source software

OSTP
White House Office of Science and Technology Policy

OWL
Web Ontology Language derived from DAML + OIL

P

PC
Personal computer

PCA
DARPA's Polymorphous Computing Architectures program

PCA
Program Component Area

PCMon
PC-based network traffic monitoring, a monitoring and measurement tool that enables detailed analysis of individual traffic flows

PDE
Partial differential equation

PE
Processing element

PerMIS
NIST's Performance Metrics for Intelligent Systems workshop series

PET
DoD/HCPMPO's Programming Environment and Training program

Petaflops
Quadrillions of floating point operations per second

PHAML
NIST's Parallel Hierarchical Adaptive MultiLevel project

PI
Principal investigator

PIM
Processor-in-memory

PITAC
President's Information Technology Advisory Committee

PKI
Public key infrastructure

PlanetLab
NSF' overlay testbed for disruptive network services

PMEL
NOAA's Pacific Marine Environmental Laboratory

POS
Packet over SONET

PPPL
Princeton Plasma Physics Laboratory

PSC
NSF-supported Pittsburgh Supercomputing Center

PSE
Patient safety event

PSE
DOE/NNSA's Problem Solving Environment program

PU
Purdue University

PVM
Parallel virtual machine

Q

QCD
Quantum chromodynamics

QFS
A Sun Microsystems computer

QoS
Quality of service

R

R&D
Research and development

R&E
Research and engineering

RAiN
NIST's Resilient Agile Networking program

RAMAS
NASA's Resource Allocation, Measurement and Adjustment System

RAS
Reliability, availability, serviceability

RAW
DARPA's Reconfigurable Architecture Workstation activity

RDF/S
Resource description framework/schema

RFI
Request for information

RFP
Request for proposals

RLV
NASA's reusable launch vehicle

RMS
DARPA's Rapid Multilingual Support one-way translation software

RONs
Regional operational networks

RSS
NOAA's Really simple syndication

S

S&T
Science and technology

SACS
DARPA's Security-Aware Critical Software program

SAPP
DOE/SC's Scientific Applications Pilot Programs

SAS
DoD/HPCMPO's Software Application Support program

SBE
NSF's Social, Behavioral, and Economic Sciences Directorate

SCADA
Supervisory control and data acquisition

SCI
NSF/CISE's Shared Cyberinfrastructure Division

SciDAC
DOE/SC's Scientific Discovery through Advanced Computing program

SCS
Simulation and Computer Science, part of DOE/NNSA's ASC program

SDCTD
NIST/ITL's Software Diagnostics and Conformance Testing Division

SDSC
NSF-supported San Diego Supercomputing Center

SDP
Software Design and Productivity, one of NITRD's seven Program Component Areas

SEC
DARPA's Software-Enabled Control program

SEI
NASA's Software Engineering Initiative

SEW
Social, Economic, and Workforce Implications of IT and IT Workforce Development, one of NITRD's seven Program Component Areas

SGI
Silicon Graphics, Inc.

SI
System Integrator, type of award made under NSF Middleware Initiative

SIMA
NIST/MEL's Systems Integration for Manufacturing Applications program

SIP
Session-initiated protocol

SIP
DoD/HPCMPO's Computational Signal/Image Processing CTA

SMP
Symmetric multiprocessing

SMT/CMP
Simultaneous multithreading/chip multiprocessing, a kind of high-throughput processor

SoD
NSF/CISE's Science of Design program

SODA
Service-oriented data access

SONET
Synchronous optical network

SP
IBM line of scalable power parallel systems

SPC
DoD's Software Protection Center

SRM
NIST's standard reference model

SRS
Self-regenerative systems

SSI
Single system image, an operating system architecture

SSP
DOE/NNSA's Stockpile Stewardship Program

StarLight
NSF-supported international optical network peering point in Chicago

STEM
Science, technology, engineering, and mathematics

STI
NSF's Strategic Technologies for the Internet program

STM1
Synchronous transport mode 1 (155 Mbps)

STT
Speech-to-text

SUNMOS
Sandia University of New Mexico Operating System

SuperMUSE
EPA's Supercomputer for Model Uncertainty and Sensitivity Evaluation; 400-node Intel-based supercomputing cluster

SuperNet
DARPA's advanced research network

SV2
A Cray scalar-vector supercomputer

T

T&E
Test and evaluation

TACC
Texas Advanced Computing Center

TAF-J
NASA's Tool Agent Framework-Java

Tbyte, or TB
Terabyte, a trillion bytes

TDT
NIST's Topic Detection and Traction program

TCO
Total cost of ownership

TCP
Transmission control protocol

TCS
NSF-supported University of Pittsburgh terascale computing system

TES
Transportable Earth Station, NREN mobile ground control center for distributed and grid-based scientific experiments

Tflop, or TF
Teraflop, a trillion floating point operations

TIDES
DARPA/NIST's Translingual Information Detection, Extraction and Summarization program

TIP
NLM's Toxicology Information Program

TNT
NIST's Template Numerical Toolkit

TOPS
DOE/SC's Terascale Optimal PDE Solvers ISIC

TREC
The DARPA/NIST Text Retrieval Conference

TRIPS
DARPA PCA program's Tera-op Reliable Intelligently Adaptive Processing System-x

TRL
NASA's Technology Readiness Level scale

TSTT
DOE/SC's Terascale Simulation Tools and Technologies ISIC

U

UA
NIST's Uncertainty Analysis framework

UA
Universal access

UAV
Unmanned air vehicle

UCAID
University Consortium for Advanced Internet Development

UCAR
University Corporation for Atmospheric Research

UltraScience Net
DOE/SC's experimental research network

UnitsML
NIST/DOE/SC's XML schema for encoding units of measurement

UPC
Unified parallel C, a programming language

US VISIT
DHS's United States Visitor and Immigrant Status Indicator Technology

UWB
Ultrawide band

V

V&V
Verification and validation

VACE
ARDA's Video Analysis and Content Extraction program

vBNS+
NSF's very high-performance Backbone Network Services plus

VIEWS
DOE/NNSA's Visual Interactive Environment for Weapons Simulation

VPN
Virtual private network

VRML
Virtual Reality Modeling Language

W

W3C
World Wide Web Consortium

WAN
Wide-area network

WAN-in-Lab
NSF's wide-area-network laboratory project, part of the EIN program

Web3D
Communicating with real-time 3D across applications, networks, and XML Web services

WPAN
Wireless personal area network

WRF
NOAA's Weather Research and Forecast program

X

X1
Cray, Inc. scalable, hybrid scalar-vector high-end computing system developed with support from NSA and ODDR&E

X1e
NSA-supported next generation of Cray X1 known as Black Widow; under development now, is expected to be available in 2006

X3D
An extensible open file format standard for 3D visual effects, behavioral modeling, and interaction

XML
Extensible markup language that is more generalized than HTML and can be used for tagging data streams more precisely and extensively

National Coordination Office (NCO) for Information Technology Research and Development (IT R&D)

David B. Nelson
Director

Sally E. Howe, Ph.D.
Associate Director
FY 2005 Blue Book Executive Editor

Martha K. Matzke
FY 2005 Blue Book Editor

Suite II-405
4201 Wilson Boulevard
Arlington, VA 22230
(703) 292-4873
FAX:(703) 292-9097
nco@nitrd.gov

Internet Web Site
http://www.nitrd.gov/

Acknowledgments

The Supplement to the President s FY 2005 Budget reflects work and contributions by a great many individuals in Federal agencies, national laboratories, universities, and the IT industry. The editors thank all the NITRD participants who helped compile and report the FY 2004 activities and FY 2005 plans that make up this detailed guide to the NITRD Program, including the work of the NITRD IWG and Coordinating Groups, the rich assortment of multiagency collaborations across NITRD s seven Program Component Areas, and the summaries of current agency activities by PCA.

Special thanks to Diane Theiss for contributing organizational skills, attention to detail, and technical support that substantially eased the editorial production process, to Alan Tellington for IT expertise and problem solving, and to Frankie King for preparation of much of the programmatic information contained in this report.

Contributors

Michael J. Ackerman, NIH
Frank D. Anger, NSF (deceased)
William F. Bainbridge, NSF
Nekeia Bell, NCO
Walter F. Brooks, NASA
Stephen Clapham, NCO
George R. Cotter, NSA
Candace S. Culhane, NSA
Steven Fine, NOAA (formerly EPA)
J. Michael Fitzmaurice, ARHQ
Peter A. Freeman, NSF
Robert Graybill, DARPA
Pamela Green, NSF
Helen Gill, NSF
Cray J. Henry, HPCMPO
Daniel A. Hitchcock, DOE/SC
Thuc T. Hoang, DOE/NNSA
C. Suzanne Iacono, NSF
Frederick C. Johnson, DOE/SC
Betty J. Kirk (formerly NCO)
Frankie King, NCO
Gary M. Koob (formerly DARPA)
Norman H. Kreisman, DOE
Ernest R. Lucier, FAA
Stephen R. Mahaney, NSF
William Bradley Martin, NSA
Michael Marron, NIH
José Muñoz, NSF
William Bradley Martin, NSA
Piyush Mehrotra, NASA
Grant Miller, NCO
Virginia Moore, NCO
José L. Muñoz, NSF
Hal Pierson, FAA
Steven R. Ray, NIST
Larry H. Reeker, NIST
James Rupp, NSA
Jean C. Scholtz, NIST
William J. Semancik, NSA
Karen Skeete, NCO
Karen Skinner, NIH
Mary Anne Scott, DOE/SC
George O. Strawn, NSF
Alan Tellington, NCO
Diane R. Theiss, NCO
Susan B. Turnbull, GSA
William T. Turnbull, NOAA
Andre M. von Tilborg, ODDR&E
Gary Walter, EPA

Abstract

The Federal agencies that participate in the Networking and Information Technology Research and Development (NITRD) Program coordinate their IT R&D activities both to support critical agency missions and to maintain U.S. leadership in high-end computing, advanced networking, high-quality and high-confidence software, information management, and related information technologies. NITRD also supports R&D on the socieconomic implications of IT and on development of a highly skilled IT workforce. The NITRD Program's collaborative approach enables agencies to leverage strengths, avoid duplication, and increase the interoperability of research accomplishments to maximize the utility of Federal R&D investments.

The Supplement to the President's FY 2005 Budget summarizes FY 2004 NITRD accomplishments and FY 2005 plans, as required by the High-Performance Computing Act of 1991. The document serves as a directory of today's NITRD Program, providing detailed descriptions of its coordination structures and activities as well as extended summaries of agencies' current R&D programs in each of the seven NITRD Program Component Areas. The FY 2005 supplement highlights in particular the myriad formal and informal multiagency collaborative activities that make the NITRD Program a uniquely successful effort of its kind within the Federal government.

The NITRD Program is a top-priority R&D focus of the President's FY 2005 Budget.

Cover Design and Imagery

The cover's graphic design is the work of James J. Caras of NSF's Design and Publishing Division.

Front-cover images, clockwise from top left: 1) Visualization of brain fiber tracts and fractional anisotropy isosurface in a DT-MRI scan; developed by computational scientists with the NIH/BISTI Program of Excellence for Computational Bioimaging and Visualization at the University of Utah's Scientific Computing and Imaging Institute. *Credit: Gordon Kindlmann, Dr. David Weinstein. http://www.sci.utah.edu/stories/2004/sum_tensorfields.html*
2) In a project that has won awards internationally, Tufts University mathematician Bruce Boghosian and University of London chemist Peter Coveney used the NSF-supported TeraGrid – including 17 terabytes of computational, storage, and visualization resources at ANL, NCSA, PSC, and SDSC – and linked sites in the U.K. to simulate the behaviors of amphiphilic fluids, or surfactants. The work produced the discovery of a fleeting fluid state that is structurally a gyroid (pictured), a shape associated with solids. *http://www.psc.edu/science/physics.html*
3) New high-speed chipset designed by engineers at NOAA, NASA, and DoD and produced by Northrup Grumman and others as a first of its kind, hardened for the extreme conditions of space applications; the set is able to transmit data at 100 Mbps between satellite sensors and ground stations; it will be applied in the joint NOAA-DoD National Polar-Orbiting Environmental Satellite System (NPOESS) to become operational in 2009.

Back-cover images, clockwise from top: 1) Step from a simulation of a proton beamline interacting with surrounding electrons. The accelerator modeling project in DOE/SC's SciDAC initiative develops such simulations to better understand beamline behavior and to optimize performance. The representation for particles in these simulations is a 6D dataset composed of 3 spatial and 3 phase dimensions. Image shows trajectories of electrons selected interactively with a box widget in the projection of the last simulation step along the z direction. The proton beam is rendered as volume density data. The trajectories are rendered as splines colored by the magnitude of the velocities. *Credit: Andreas Adelmann, LBNL* **2)** Test simulation image of the interaction of strong shocks with material surfaces in fluid phases, a key focus of the Virtual shock physics Test Facility (VTF) at the CalTech Center for the Simulation of Dynamic Response in Materials supported by DOE/NNSA's ASC program. The center's goals are to: facilitate simulation of a variety of phenomena in which strong shock and detonation waves impinge on targets consisting of various combinations of materials, compute the subsequent dynamic response of the target materials, and validate these computations against experimental data. *http://www.cacr.caltech.edu/ASAP/*
3) On June 30, 2004, after seven years and a billion miles, NASA's Cassini-Huygens spacecraft reached Saturn and began sending close-up image data. This view of the planet's rings from the inside out shows primarily the outer turquoise A ring, with an icy composition. The reddish bands are inner rings composed of smaller, "dirtier" particles than the A ring; the single red band in the A ring is known as the Encke gap. The image was taken with the Ultraviolet Imaging Spectrograph instrument, which is capable of resolving the rings to show features up to 97 kilometers (60 miles) across, roughly 100 times the resolution of ultraviolet data obtained by the Voyager 2 spacecraft. *Credit: NASA/JPL/University of Colorado*

Human-Computer Interaction and Information Managem

Large-Scale Networking

HEC I&A and R&D Programs by Agency:
FY 2004 Activities and FY 2005 Plans

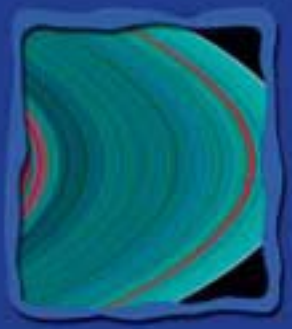
Software Design and Productivity

Agency NI

Supplement to the President's FY 2005 Budget

Interagency Wo
on IT W

Social, Economic, and Workforce Implications of IT
and IT Workforce Development



The NITRD Agencies

The NITRD Program To

National Science and Technology Council
Interagency Working Group on
Information Technology Research and Development
National Coordination Office for
Information Technology Research and Development
Suite II - 405, 4201 Wilson Blvd. Arlington, VA 22230
(703) 292-4873 http://www.nitrd.gov