



**THE NITRD PROGRAM:  
FY 2004  
INTERAGENCY COORDINATION  
REPORT**

Interagency Working Group on  
Information Technology Research and Development  
(IWG/ITR&D)

Second Printing—OCTOBER 2004

First Printing—SEPTEMBER 2004

# TABLE OF CONTENTS

<b>EXECUTIVE SUMMARY .....</b>	<b>v</b>
<b>1. ORGANIZATIONAL STRUCTURE AND OVERARCHING ACTIVITIES OF THE NITRD PROGRAM .....</b>	<b>1</b>
1.1 GOALS OF THE NITRD PROGRAM .....	1
1.2 THE NITRD AGENCIES .....	1
1.2.1 Principal NITRD Agencies .....	1
1.2.2 Participation of Other Federal Agencies in NITRD Program Activities.....	2
1.3 THE NITRD PROGRAM AUTHORIZING LEGISLATION .....	2
1.4 OFFICE OF SCIENCE AND TECHNOLOGY POLICY (OSTP) AND NATIONAL SCIENCE AND TECHNOLOGY COUNCIL (NSTC) OVERSIGHT OF THE NITRD PROGRAM.....	2
1.4.1 Special NSTC Project: High-End Computing Revitalization Task Force (HECRTF).....	3
1.5 INTERAGENCY WORKING GROUP ON INFORMATION TECHNOLOGY RESEARCH AND DEVELOPMENT (IWG/ITR&D): COORDINATION OF THE NITRD PROGRAM .....	3
1.5.1 Program Component Areas (PCAs) .....	3
1.5.2 Overarching Activities Under the IWG.....	5
1.6 ABOUT THE NCO .....	9
<b>2. HIGH-END COMPUTING (HEC) .....</b>	<b>11</b>
2.1 INTRODUCTION.....	11
2.2 DEFINITION OF THE HEC I&A PCA.....	11
2.3 DEFINITION OF THE HEC R&D PCA .....	12
2.4 HEC CG ACTIVITIES.....	14
2.5 HEC MULTI-AGENCY ACTIVITIES.....	15
2.5.1 Earth System Modeling Framework.....	16
2.6 HIGH-END COMPUTING REVITALIZATION TASK FORCE (HECRTF).....	17
2.6.1 HECRTF Report Released at House Science Committee Hearing .....	17
2.6.2 Charge to the HECRTF .....	17
2.6.3 HECRTF Report: <i>Federal Plan for High-End Computing</i> .....	17
2.6.4 Workshop and Report on the Roadmap for the Revitalization of High-End Computing.....	19
2.6.5 HEC–University Research Activity (HEC-URA) .....	20
2.6.6 Implementation of the HECRTF Plan .....	21
2.6.7 Study on High-End Computing in Japan .....	21
2.7 HEC AGENCY ACTIVITIES .....	23
2.7.1 National Science Foundation (NSF).....	23
2.7.2 Department Of Energy/Office of Science (DOE/SC).....	26
2.7.3 Defense Advanced Research Projects Agency (DARPA) .....	31
2.7.4 National Aeronautics and Space Administration (NASA) .....	34
2.7.5 DOE/National Nuclear Security Administration (DOE/NNSA).....	37
2.7.6 National Security Agency (NSA).....	38
2.7.7 National Oceanic and Atmospheric Administration (NOAA) .....	40
2.7.8 National Institute of Standards and Technology (NIST) .....	41
2.7.9 Environmental Protection Agency (EPA) .....	43
2.7.10 Department of Defense/High Performance Computing Modernization Program Office (DoD/HPCMPO) .....	45
<b>3. HUMAN COMPUTER INTERACTION AND INFORMATION MANAGEMENT (HCI&amp;IM) .....</b>	<b>49</b>
3.1 INTRODUCTION.....	49
3.2 DEFINITION OF THE HCI&IM PCA .....	49
3.3 HCI&IM CG ACTIVITIES.....	51
3.3.1 Release of the “Human–Computer Interaction and Information Management Research Needs” Report. 51	51
3.3.2 HCI&IM R&D Planning and Coordination.....	51
3.3.3 HCI&IM CG Special Meeting in December 2003 .....	52

3.4	HCI&IM MULTI-AGENCY ACTIVITIES.....	52
3.4.1	DARPA.....	52
3.4.2	NASA.....	52
3.4.3	AHRQ.....	52
3.4.4	NIST.....	53
3.5	HCI&IM AGENCY ACTIVITIES.....	53
3.5.1	NSF.....	53
3.5.2	NIH.....	55
3.5.3	DARPA.....	56
3.5.4	NASA.....	59
3.5.5	AHRQ.....	60
3.5.6	DOE/SC.....	61
3.5.7	NIST.....	62
3.5.8	NOAA.....	66
3.5.9	EPA.....	66
4.	LARGE SCALE NETWORKING (LSN).....	69
4.1	INTRODUCTION.....	69
4.2	DEFINITION OF THE LSN PCA.....	69
4.3	LSN CG ACTIVITIES.....	70
4.3.1	Network Testbed Infrastructure.....	71
4.3.2	Network Performance Measurement.....	71
4.3.3	Grid Outreach.....	72
4.3.4	Additional Topics for LSN and IWG Coordination.....	72
4.4	LSN MULTI-AGENCY ACTIVITIES.....	72
4.4.1	Common Interests.....	72
4.4.2	Coordination of Network Performance Measurement.....	73
4.4.3	Optical Networking Workshop.....	73
4.4.4	Coordination of Meetings.....	73
4.5	LSN TEAM ACTIVITIES.....	74
4.5.1	Joint Engineering Team (JET).....	74
4.5.2	Middleware and Grid Infrastructure Coordination (MAGIC) Team.....	76
4.5.3	Networking Research Team (NRT).....	77
4.6	LSN AGENCY ACTIVITIES.....	78
4.6.1	NIH/NLM.....	78
4.6.2	NSF.....	79
4.6.3	DOE/SC.....	82
4.6.4	NASA.....	84
4.6.5	DARPA.....	86
4.6.6	DOE/NNSA.....	87
4.6.7	NIST.....	88
4.6.8	NOAA.....	90
4.6.9	NSA.....	91
4.6.10	DoD/HPCMPO.....	91
5.	SOFTWARE DESIGN AND PRODUCTIVITY (SDP).....	93
5.1	INTRODUCTION.....	93
5.2	DEFINITION OF THE SDP PCA.....	93
5.3	SDP CG ACTIVITIES.....	94
5.4	SDP MULTI-AGENCY ACTIVITIES.....	96
5.4.1	NASA.....	96
5.4.2	NSF.....	97
5.4.3	DOE/NNSA.....	97
5.4.4	DARPA.....	97
5.4.5	NIST.....	97
5.5	SDP AGENCY ACTIVITIES.....	97
5.5.1	NASA.....	97

5.5.2	NSF .....	102
5.5.3	DOE/NNSA.....	103
5.5.4	DARPA .....	103
5.5.5	NIH .....	104
5.5.6	NIST.....	105
5.5.7	NOAA .....	106
5.5.8	FAA.....	107
6.	HIGH CONFIDENCE SOFTWARE AND SYSTEMS (HCSS) .....	109
6.1	INTRODUCTION.....	109
6.2	DEFINITION OF THE HCSS PCA.....	109
6.3	HCSS CG ACTIVITIES.....	110
6.4	HCSS MULTI-AGENCY ACTIVITIES.....	111
6.5	HCSS AGENCY ACTIVITIES.....	112
6.5.1	NSF .....	112
6.5.2	NSA.....	114
6.5.3	NASA.....	116
6.5.4	DARPA .....	120
6.5.5	NIST.....	121
6.5.6	AFRL .....	123
6.5.7	FAA.....	123
6.5.8	FDA.....	124
7.	SOCIAL, ECONOMIC, AND WORKFORCE IMPLICATIONS OF IT AND IT WORKFORCE DEVELOPMENT (SEW) .....	125
7.1	INTRODUCTION.....	125
7.2	DEFINITION OF THE SEW PCA .....	125
7.3	SEW CG ACTIVITIES .....	126
7.4	SEW UNIVERSAL ACCESSIBILITY TEAM ACTIVITIES.....	126
7.5	SEW AGENCY ACTIVITIES .....	127
7.5.1	NSF .....	127
7.5.2	NIH .....	131
7.5.3	NASA.....	132
7.5.4	DOE/NNSA.....	133
7.5.5	DOE/NNSA and DOE/SC.....	134
7.5.6	GSA.....	134
	Appendix A: Overview of NITRD Agency Missions .....	135
	Appendix B: Interagency Working Group on ITR&D.....	147
	Appendix C: PCA Coordinating Groups and Team Chairs.....	147
	Appendix D: Report Contributors and NITRD Contacts .....	149
	Appendix E: Glossary and List of Acronyms .....	157



## EXECUTIVE SUMMARY

This Interagency Coordination Report (ICR) provides a comprehensive description of the FY 2004 activities of the multi-agency \$2 billion Federal *Networking and Information Technology Research and Development (NITRD) Program* and the organizations that implement them. The NITRD Program is one of several formal interagency programs in the Federal government. It provides agencies that perform ITR&D with the ability to plan, budget, coordinate, implement, and assess their ITR&D programs. The agencies that formally participate in the NITRD Program in FY 2004 are:

- Agency for Healthcare Research and Quality (AHRQ)
- Defense Advanced Research Projects Agency (DARPA)
- Defense Information Systems Agency (DISA)
- Department of Energy/National Nuclear Security Administration (DOE/NNSA)
- Department of Energy/Office of Science (DOE/SC)
- Environmental Protection Agency (EPA)
- National Aeronautics and Space Administration (NASA)
- National Institute of Standards and Technology (NIST)
- National Institutes of Health (NIH)
- National Oceanographic and Atmospheric Administration (NOAA)
- National Science Foundation (NSF)
- National Security Agency (NSA)

Additionally, six agencies that are not formal NITRD members participate in the FY 2004 NITRD Program activities. They are:

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

The NITRD Program is coordinated through the Interagency Working Group on Information Technology Research and Development (IWG/ITR&D) of the National Science and Technology Council (NSTC). (The NITRD Program, the NSTC, and the IWG are described in chapter 1 of this report.) This FY 2004 ICR is the successor to past Implementation Plans (IPs).

The ICR focuses on the building blocks of the NITRD Program – seven Program Component Areas (PCAs) and six Coordinating Groups (CGs). The PCAs are the subject areas where agency investments are made in ITR&D and the CGs are the structured groups of NITRD agency program managers that guide the PCA investments through planning, coordinating, implementing, and assessing NITRD Program activities. The ICR's focus on the PCAs and the CGs enables an in-depth view of the coordinated NITRD activities.

The seven PCAs are:

- High-End Computing Infrastructure and Applications (HEC I&A)
- HEC 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 Implications of IT and IT Workforce Development (SEW)

CGs map one-to-one and onto the PCAs, except in the case of the HEC CG, which coordinates the activities in both the HEC I&A and HEC R&D PCAs.

The coordinated FY 2004 NITRD Program activities described in this report are wide and varied and include:

- The release of NITRD agency reports including:
  - Budget-related reports such as the FY 2004 *Supplement to the President's Budget* (Blue Book), formally entitled, *NITRD: Advanced Foundations for American Innovation*
  - A special OSTP/NSTC-sanctioned interagency report entitled the *Federal Plan for High-End Computing*
  - A special IWG-sanctioned report called *Grand Challenges: Science, Engineering, and Societal Advances Requiring Networking and Information Technology Research and Development*
  - Individual CG-sanctioned research needs reports such as the HCI&IM CG's *Human-Computer Interaction and Information Management Research Needs*
- The development of ITR&D research scopes such as SDP's taxonomy
- The assembly of NITRD agency representatives in quarterly IWG meetings, monthly CG meetings, and Task Forces such as the Grand Challenges Task Force
- The assembly of Federal, university, and industry experts in science, technology, and engineering in multi-agency activities such as the NASA-hosted workshop on Optical Network Testbeds (ONT)
- Single agency activities such as DOE/SC's Scientific Discovery through Advanced Computing (SciDAC) program

## ABOUT THIS REPORT

The material in this report comes from detailed summaries of activities and plans presented by agency representatives at Special Meetings of the CGs held between November 2003 and March 2004. This ICR is a snapshot of FY 2004 activities and plans at the time these summaries were made. Many of these activities are multi-year. Occasionally, FY 2003 or earlier activities are reported. An example is workshops to plan R&D activities that are funded in FY 2004. Some agency FY 2004 activities and plans changed after this snapshot was taken.

The report was prepared by the National Coordination Office for Information Technology Research and Development (NCO/ITR&D). It is organized into seven chapters. Chapter 1 describes the organizational structure of the NITRD Program, its authorizing legislation, the short versions of the recently updated PCA definitions, and the IWG-level NITRD activities. Each of the subsequent six chapters describes activities of one CG. They are in descending order of their NITRD budgets. Each chapter includes:

- A detailed PCA definition(s)
- *FY 2004 CG activities* that are sponsored by a CG and involve more than one CG member agency
- *FY 2004 multi-agency activities* that are cooperatively implemented by multiple CG member agencies
- *FY 2004 agency activities* that address an agency's mission goals

The CG chapters are followed by five appendices and acknowledgements. Appendix A provides an overview of the missions of the NITRD agencies, Appendix B lists the members of the IWG, Appendix C lists the PCA Coordinating Groups and team chairs, Appendix D lists the report contributors and NITRD contacts, and Appendix E provides the glossary and spells out the acronyms found throughout this report.



# **1. ORGANIZATIONAL STRUCTURE AND OVERARCHING ACTIVITIES OF THE NITRD PROGRAM**

The NITRD Program is one of a small number of formal interagency programs in the Federal government. It provides agencies that perform ITR&D with the ability to plan, budget, coordinate, implement, and assess their ITR&D activities.

## **1.1 GOALS OF THE NITRD PROGRAM**

The NITRD Program goals are to:

- Assure continued U.S. leadership in computing, networking, and information technologies to meet Federal goals and to support U.S. 21<sup>st</sup> 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, networking, and information technologies

## **1.2 THE NITRD AGENCIES**

### **1.2.1 Principal NITRD Agencies**

Agencies that formally participate in the NITRD Program are:

- Agency for Healthcare Research and Quality (AHRQ)
- Defense Advanced Research Projects Agency (DARPA)
- Defense Information Systems Agency (DISA)
- Department of Energy/National Nuclear Security Administration (DOE/NNSA)
- Department of Energy/Office of Science (DOE/SC)
- Environmental Protection Agency (EPA)
- National Aeronautics and Space Administration (NASA)
- National Institute of Standards and Technology (NIST)
- National Institutes of Health (NIH)
- National Oceanic and Atmospheric Administration (NOAA)
- National Science Foundation (NSF)
- National Security Agency (NSA)

## **1.2.2 Participation of Other Federal Agencies in NITRD Program Activities**

Participation in NITRD activities is open to other Federal agencies. In FY 2004, the following agencies participated in NITRD activities:

- 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 agencies listed in sections 1.2.1 and 1.2.2 participated in the production of this report.

## **1.3 THE NITRD PROGRAM AUTHORIZING LEGISLATION**

The NITRD Program stems from the High-Performance Computing (HPC) Act of 1991 (P.L. 102-194) as amended by the Next Generation Internet (NGI) Act of 1998 (P.L. 105-305).<sup>1</sup> These Acts authorize Federal agencies to set goals and prioritize their investments in networking and information technology research and development (including high performance computing) and authorize the establishment of what is now known as the President's Information Technology Advisory Committee (PITAC).

## **1.4 OFFICE OF SCIENCE AND TECHNOLOGY POLICY (OSTP) AND NATIONAL SCIENCE AND TECHNOLOGY COUNCIL (NSTC) OVERSIGHT OF THE NITRD PROGRAM**

The Office of Science and Technology Policy (OSTP)<sup>2</sup> was established by the National Science and Technology Policy, Organization, and Priorities Act of 1976.<sup>3</sup> OSTP advises and oversees the formulation of policies and budget developments on all questions in which science and technology (S&T) are important elements; it articulates and directs the President's S&T policies and programs; and it creates and sustains scientific partnerships among Federal, state, and local government, industry, and academic organizations.

The Cabinet-level NSTC,<sup>4</sup> established by Executive Order on November 23, 1993, is the principal means for the President to coordinate science, space, and technology policies across the Federal government. The President chairs the NSTC and the membership consists of the Vice President, the Assistant to the President for Science and Technology, Cabinet Secretaries, agency heads with significant S&T responsibilities, and other White House officials. The Assistant to

---

<sup>1</sup> Both Acts can be found at <http://www.nitrd.gov/congressional/laws/index.html>

<sup>2</sup> <http://www.ostp.gov>

<sup>3</sup> The Act (P.L. 94-282) can be found at <http://www.thomas.loc.gov/>

<sup>4</sup> <http://www.nstc.gov>

the President for Science and Technology directs both the OSTP and the NSTC on behalf of the President.

The Interagency Working Group for ITR&D is the organizational framework established by the NSTC to coordinate the work of the NITRD Program. The IWG reports to the Committee on Technology of the NSTC.

#### **1.4.1 Special NSTC Project: High-End Computing Revitalization Task Force (HECRTF)**

As part of the President's FY 2004 Budget, OSTP sanctioned the creation of the HECRTF whose work was chartered under the auspices of the NSTC. The HECRTF was charged with developing a plan for undertaking and sustaining a robust Federal high-end computing program to maintain U.S. leadership in S&T. Following approval by the members of the NSTC Committee on Technology, the HECRTF released its *Federal Plan for High-End Computing*<sup>5</sup> in May 2004. The HECRTF and its activities are described in section 2.6.

### **1.5 INTERAGENCY WORKING GROUP ON INFORMATION TECHNOLOGY RESEARCH AND DEVELOPMENT (IWG/ITR&D): COORDINATION OF THE NITRD PROGRAM**

The IWG/ITR&D<sup>6</sup> (henceforth called the "IWG") provides hands-on coordination of the NITRD Program. It comprises senior ITR&D managers from the participating agencies and representatives from the Office of Management and Budget (OMB), OSTP, and the National Coordination Office for Information Technology Research and Development (NCO/ITR&D).<sup>7</sup> The IWG meets quarterly. The NCO supports the IWG.

The building blocks of the NITRD Program are called Program Component Areas (PCAs) and Coordinating Groups (CGs). PCAs – the conceptual building blocks – are the subject areas across which R&D investments are made by the NITRD agencies. CGs – the operational building blocks – are the groups of agency representatives that coordinate the activities funded under the PCAs.

#### **1.5.1 Program Component Areas (PCAs)**

There are seven PCAs in the NITRD Program. The two HEC PCAs – HEC Infrastructure and Applications (I&A) and HEC Research and Development (R&D) – are coordinated by the HEC CG. The PCAs are briefly defined below (detailed definitions appear at the beginning of each CG chapter):

**High-End Computing Infrastructure and Applications (HEC I&A)** – The activities and facilities funded under the NITRD Program's HEC I&A PCA include R&D infrastructure and

---

<sup>5</sup> [http://www.nitrd.gov/pubs/2004\\_hecertf/20040510\\_hecertf.pdf](http://www.nitrd.gov/pubs/2004_hecertf/20040510_hecertf.pdf)

<sup>6</sup> <http://www.nitrd.gov/iwg/program.html>

<sup>7</sup> <http://www.nitrd.gov/about/index.html>

R&D to extend the state of the art in computing systems, science and engineering applications, and data management necessary to keep the United States at the forefront of 21st century science and engineering discoveries.

**High-End Computing Research and Development (HEC R&D)** – The activities funded under the HEC R&D PCA include R&D to optimize the performance of today’s high-end computing systems and to develop future generations of high-end computing systems to address Federal agency mission needs and in turn many of society’s most challenging large-scale computational problems in order to strengthen the Nation’s global leadership in the sciences, engineering, and technology.

**Human Computer Interaction and Information Management (HCI&IM)** – The activities funded under the HCI&IM PCA have the goal of increasing the benefit of computer technology to humans through the development of future user interaction technologies, cognitive systems, information systems, and robotics.

**Large Scale Networking (LSN)** – The activities funded under the 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.

**Software Design and Productivity (SDP)** – The activities funded under the SDP PCA focus on achieving fundamental advances in concepts, methods, techniques, and tools for software design, development, and maintenance that address the widening gap between society’s need for usable and dependable software-based systems and our ability to produce them in a timely, predictable, and cost-effective manner.

**High Confidence Software and Systems (HCSS)** – The activities funded under the 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 in critical domains such as aviation, healthcare, national defense, and infrastructure.

**Social, Economic, and Workforce Implications of IT and IT Workforce Development (SEW)** – The activities funded under the 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 ITR&D to the policymaking and IT user communities in government at all levels and the private sector.

**1.5.1.1 FY 2004 NITRD AGENCY BUDGET REQUESTS BY PCA (dollars in millions)<sup>8</sup>**

Agency	HEC I&A	HEC R&D	HCI&IM	LSN	SDP	HCSS	SEW	TOTAL
NSF	218.1	97.9	125.3	103.4	55.0	59.9	74.0	<b>734</b>
NIH	87.6	41.7	99.0	132.2	9.2	3.7	12.2	<b>386</b>
NASA	45.9	34.6	67.1	28.9	59.2	24.2	6.7	<b>267</b>
DARPA		108.5	78.4	18.2	13.3	4.0		<b>222</b>
DOE/SC	88.9	51.3	16.4	30.0			3.5	<b>190</b>
AHRQ			32.0	25.0				<b>57</b>
NSA		21.3		1.9		28.1		<b>51</b>
NIST	3.5		6.2	3.2	7.5	2.0		<b>22</b>
NOAA	13.5	1.8	0.5	2.8	1.5			<b>20</b>
EPA	1.6		0.2					<b>2</b>
DOE/NNSA	41.5	37.3		14.4	32.8		4.4	<b>130</b>
DISA			46.2	13.2		6.1		<b>66</b>
<b>TOTAL</b>	<b>500.6</b>	<b>394.4</b>	<b>471.3</b>	<b>373.3</b>	<b>178.5</b>	<b>128.1</b>	<b>100.8</b>	<b>2,147</b>

**1.5.1.2 Coordinating Groups**

The NITRD Program’s six CGs comprise program managers who plan, coordinate, implement, and assess the associated PCA investments. The CGs meet approximately monthly and report to the IWG.

**1.5.2 Overarching Activities Under the IWG**

In FY 2004, the IWG sanctioned the following activities:

- The preparation of the FY 2005 Blue Book, which is required by law
- The preparation of the Grand Challenges Task Force report
- The CG Special Meetings
- The preparation of the ICR
- The PCA/CG Task Group that, among other things, shepherded the updates to the PCA definitions

**1.5.2.1 FY 2004 Supplement to the President’s Budget (“The Blue Book”)**

Officially called Supplements to the President’s Budget, Blue Books provide descriptive accounts of the extensive and diverse portfolio of activities funded under the NITRD Program and a breakdown of the corresponding budget details of these activities. Required by the HPC Act of 1991, the NCO/ITR&D prepares Blue Books on an annual basis with input from the NITRD agencies.

---

<sup>8</sup> FY 2004 Blue Book, Agency NITRD Budgets by Program Component Area, p. 39.

Released in September 2003, the FY 2004 Blue Book, entitled *Networking and Information Technology Research and Development: Advanced Foundations for American Innovation*,<sup>9</sup> highlighted outcomes of Federal NITRD activities and specifically concentrated on the critical role of fundamental IT research executed by the Federal agencies. This IT research resulted in many of the advanced foundations for innovation in key dimensions of national interests including foundations for national security, scientific leadership, research and learning, and 21<sup>st</sup> Century society.

### **1.5.2.2 Relationship Between the ICR and the FY 2005 Blue Book**

The ICR documents the NITRD Program activities within a single fiscal year, with an emphasis on coordination. These activities are described from both multi-agency and agency-specific coordination perspectives. The FY 2005 Blue Book is an abbreviated adaptation of the ICR that is accompanied by the agency budgets by PCA.

### **1.5.2.3 FY 2004 Grand Challenges Task Force and Booklet**

In November 2002, 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 ITR&D. The goal was to update the list presented in the FY 1994 Blue Book that was called for in the HPC Act of 1991. The Task Force comprised volunteers from ten NITRD agencies and FAA, OSTP, and the NCO.

The work of the Task Force culminated in the first NITRD report of its kind: *Grand Challenges: Science, Engineering, and Societal Advances Requiring Networking and Information Technology Research and Development*<sup>10</sup> which was released in November 2003.

Recognizing that IT advances will enhance existing applications and enable new ones that can have even greater impact, the Task Force updated the “grand challenge” definition. The new definition of a NITRD grand challenge is:

*A long-term science, engineering, or societal advance, whose realization requires innovative breakthroughs in information technology research and development (ITR&D) and which will help address our country’s priorities.*

The Task Force identified 16 new NITRD Illustrative Grand Challenges. Its goal in doing so was to identify a set of grand challenges that are expected to yield significant breakthroughs of practical importance to mankind. Criteria to guide the development of the grand challenges are reflected in the template used to describe them:

- Description of the multi-decade grand challenge
- Focus of the grand challenge in the next ten years

---

<sup>9</sup> <http://www.nitrd.gov/pubs/blue04>. 2,500 printed copies have been distributed.

<sup>10</sup> [http://www.nitrd.gov/pubs/200311\\_grand\\_challenges.pdf](http://www.nitrd.gov/pubs/200311_grand_challenges.pdf). 1,500 printed copies have been distributed.

- Benefits of the grand challenge to the social, economic, political, scientific, and technological well-being of mankind
- Relationship of the grand challenge to our country's national priorities. In consultation with OSTP the Task Force identified six national priority areas that reflect the country's broad scientific, military, social, economic, and political values and goals. They are:
  - Leadership in S&T
  - Homeland and National Security
  - Health and Environment
  - Economic Prosperity
  - A Well-Educated Populace
  - A Vibrant Civil Society
- Relationship to Information Technology Hard Problem Areas (ITHPAs) (described below)
- Indications of progress

The 16 new NITRD Illustrative Grand Challenges are:

- 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

The Task Force also developed a list of 14 IT Hard Problem Areas (ITHPAs) – broad categories of topics of interest to the NITRD Program and the ITR&D community and for which solutions or advances are required to achieve progress towards the grand challenges. The 14 ITHPAs are:

- 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

The Grand Challenges booklet provides further details about the grand challenges and the ITHPAs.

#### ***1.5.2.4 FY 2004 PCA/CG Task Group And Update To The PCA Definitions***

At its September 2003 meeting, the IWG chartered the NITRD PCA/CG Task Group to assess the impact of their R&D advances on agency NITRD investments and to update the existing PCA definitions.

Composed of the CG Co-Chairs and supported by the NCO, the Task Group was given the following charge:

- Examine the current foci of PCAs and CGs
- Recommend any changes to the PCAs and CGs to better serve the needs of the NITRD Program and participating agencies

The Task Force identified broad issues affecting both the PCAs and the CGs and reported the following actions:

- The Task Force established Temporary Linkages Groups (TLGs) in open source software and security. These two topics are crosscutting issues that are of interest to several CGs. The TLGs will operate in the same manner as the CGs.
- The CGs and the Task Force revised the PCA definitions. The abbreviated versions of the definitions will be used in the Executive Summary of the FY 2005 Blue Book and section 1.5.1 of this ICR. The detailed versions appear in the main chapters of the Blue Book and open each CG chapter of this report.

#### ***1.5.2.5 FY 2004 Special Meetings***

Between November 2003 and March 2004, each CG held a Special Meeting at which the CG members summarized current and future agency and, in some cases, Team activities over the course of FY 2004. The information gathered at the Special Meetings provided information useful to the CGs in fulfilling their responsibilities and served as input to the work of the IWG's PCA/CG Task Group and to the preparation of this ICR.

At each Special Meeting, agencies were asked to present the following information:

- Major agency plans and agency workshops, conferences, principal investigator (PI) meetings, planning meetings, and reports
- Multi-agency plans and multi-agency workshops, conferences, and reports that are sponsored by a CG member agency, but not necessarily by a CG, and that involve more than one CG member agency

## **1.6 ABOUT THE NCO**

The NCO/ITR&D provides the technical and administrative support for the IWG, the PCA Coordinating Groups, and the President's Information Technology Advisory Committee (PITAC). The HPC Act of 1991 authorizes the functions performed by the NCO to support the IWG and the CGs. Executive Order 13035 authorizes NCO support for the PITAC. The National Science Foundation (NSF) serves as the host agency for the NCO.



## **2. HIGH-END COMPUTING (HEC)**

### **2.1 INTRODUCTION**

High-End Computing (HEC) programs of the Federal agencies in the NITRD Program are reported under two Program Component Areas (PCAs): HEC Infrastructure and Applications (HEC I&A) and HEC Research and Development (HEC R&D). The first two sections of this chapter give their detailed definitions.

### **2.2 DEFINITION OF THE HEC I&A PCA**

The activities and facilities funded under the NITRD Program's HEC I&A PCA include R&D infrastructure and R&D to extend the state of the art in computing systems, science and engineering applications, and data management necessary to achieve the breakthroughs necessary to keep the United States at the forefront of 21<sup>st</sup> 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 S&T, 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 subatomic to supernova) and time scales (such as nano-second to multi-century climate model) 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 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.

In order 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, tools to facilitate high-end computation, storage, and visualization of large data sets encountered in the biomedical sciences, climate modeling and weather forecasting, computational aerosciences, crisis management, Earth and space sciences, and a vast range of human activities.

Broad areas of HEC I&A concerns include:

- Algorithm R&D
- Data management and understanding
- Programming environments
- Scientific applications
- Computational facilities

To address these concerns, HEC I&A R&D pursues the following technical goals:

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

Illustrative technical thrusts of HEC I&A programs include:

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

## **2.3 DEFINITION OF THE HEC R&D PCA**

The activities funded under the HEC R&D PCA include R&D to optimize the performance of today's high-end computing systems and to develop future generations of high-end computing systems to address Federal agency mission needs and, in turn, many of society's most challenging large-scale computational problems in order to strengthen the Nation's global leadership in the sciences, engineering, and technology.

Current high-end systems typically 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) these 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 applications. HEC R&D focuses on teraflop<sup>11</sup> through petaflop<sup>12</sup> 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 include:

- 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

To address these concerns, HEC R&D pursues the following technical goals:

- Parallel architecture designed for scientific application requirements
- Parallel I/O and file systems for sustained high data throughput

---

<sup>11</sup> 10<sup>12</sup> floating point operations per second

<sup>12</sup> 10<sup>15</sup> floating point operations per second

- Systems scalable to hundreds of thousands of processors
- Reliable and fault tolerant systems
- Improved programming model expressibility and parallel compiler support
- Effective performance measurement and optimization tools
- Improved ease of use and time to solution

Illustrative technical thrusts of HEC R&D programs include:

- 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

## **2.4 HEC CG ACTIVITIES**

The HEC Coordinating Group (HEC CG) coordinates activities funded under the HEC I&A and HEC R&D PCAs by providing a forum for Federal agency program managers to coordinate and collaborate on HEC research programs and on implementing Federal high-end computing activities. It is charged with:

- Encouraging and facilitating interagency coordination and collaborations in HEC I&A and R&D
- Addressing requirements for HEC 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

At HEC CG meetings, HEC agencies inform each other of upcoming agency meetings and workshops so that program managers from all HEC agencies can participate and benefit. Between March 2003 and March 2004, HEC agencies participated in the HECRTF and the HEC CG did not meet. (The HECRTF is discussed in section 2.5.)

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

## 2.5 HEC MULTI-AGENCY ACTIVITIES

The HEC agencies have common interests 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/NNSA, 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
- 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
- HPCS Phase II: DARPA, DOE/NNSA, DOE/SC, NASA, NSA, NSF
- HPCS productivity metrics:<sup>13</sup> 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/NNSA, DOE/SC (report expected in 2004)<sup>14</sup>
- Optical switches and interconnects: DARPA, DOE/NNSA, NSA
- Quantum information science: DARPA, NIST, NSA
- Quarterly reviews of Cray X1e/Black Widow R&D programs: DoD/HPCMPO, DOE/NNSA, DOE/SC, NASA, NRO, NSA
- Reviews of ASC White and ASC Q software environments: DOE/NNSA, DOE/SC
- Reviews of SV2: DOE/NNSA, DOE/SC, NSA
- 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: DARPA, DoD/HPCMPO, NIST, NOAA, NSA
- Unified Parallel Compiler (UPC): DOE/SC, NSA
- Weather Research and Forecast (WRF): NOAA, NSF/NCAR

---

<sup>13</sup> <http://www.highproductivity.org/>

<sup>14</sup> <http://www.nap.edu/> [NAP is the National Academies Press.]

The HEC agencies cooperate on many multi-agency workshops such as:

- ORNL Cray X1 Review (February 2004)<sup>15</sup>
- Interagency HPC Metrics Workshop (June 28, 2004)

Additional multi-agency workshops are noted below in agency activity descriptions. Other agencies are usually invited to attend single-agency workshops.

### **2.5.1 Earth System Modeling Framework**

Another broad-based 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). 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 in the ESMF effort:

- 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)
  - Goddard Global Modeling and Assimilation Office (GGMAO)
  - Goddard Institute for Space Studies (GISS)
  - Goddard Land Information Systems project (GLIS)

---

<sup>15</sup> <http://csm.ornl.gov/DOE/Feb2004/DOE-04index.html>

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) standard services such as time management and logging messages.

## **2.6 HIGH-END COMPUTING REVITALIZATION TASK FORCE (HECRTF)**

A primary focus of the HEC CG during FY 2004 has been interagency participation in and cooperation on the High-End Computing Revitalization Task Force (HECRTF) commissioned by OSTP in coordination with the NSTC. Participating Federal organizations were DoD/HPCMPO, DoD/OSD, DOE/NNSA, DOE/SC, EPA, NASA, NIST, NOAA, NSA, NSF, OMB, OSTP, and the NCO.

### **2.6.1 HECRTF Report Released at House Science Committee Hearing**

On May 13, 2004, following approval by the members of the NSTC Committee on Technology, the HECRTF released its final report entitled *Federal Plan for High-End Computing* (henceforth called the *Federal Plan*) at a House Science Committee hearing<sup>16</sup> on H.R. 4218, the High Performance Computing Act of 2004.<sup>17</sup>

### **2.6.2 Charge to the HECRTF**

The report fulfilled the charge given to the HECRTF to develop a five-year plan to guide future Federal investments in high-end computing.<sup>18</sup> The Task Force was given subtasks to develop:

- A roadmap for core technology development including key technologies and strategies for their development
- A roadmap for HEC computing capability, capacity, and accessibility including plans to address capability or capacity gaps and performance targets for proposed HEC systems linked to applications and user needs
- Findings and recommendations about Federal procurement of HEC systems including strategies for performance measures, methods for deriving system performance targets, and discussion of total cost of ownership
- A roadmap for the Federal role in HEC R&D, utilization, and procurement that integrates the three components above

### **2.6.3 HECRTF Report: *Federal Plan for High-End Computing***

The HECRTF identified key areas for HEC R&D:

- Hardware

---

<sup>16</sup> <http://www.house.gov/science/hearings/full04>

<sup>17</sup> <http://thomas.loc.gov/cgi-bin/query>

<sup>18</sup> <http://www.nitrd.gov/hecrtf-outreach/index.html>, which includes the full OSTP charge to the HECRTF

- Microarchitecture
- Memory
- Interconnect
- Power, cooling, and packaging
- I/O and storage
- Software
  - Operating systems
  - Languages, compilers, and libraries
  - Software tools and development environments
  - Algorithms
- Systems
  - System architecture
  - System modeling and performance analysis
  - Programming models
  - Reliability, availability, serviceability, and security

To address the requirements for HEC resources from a broad range of scientific disciplines across the Federal government, two classes of issues were identified:

- Architectural availability – the Government could help close the gap between scientific and defense needs and commercial IT by investing in research and prototype development to ensure a mix of architectures to meet Federal computing needs.
- HEC capacity to address agency mission needs – Federal S&T agencies identified high-end computing requirements that are about three times the current capacity and are growing about 80 percent per year.

The HECRTF made recommendations about access, availability, and leadership-class systems:

- *Accessibility*
  - Agencies whose researchers currently obtain HEC resources from other agencies should examine options for formalizing the provision of these resources to their research communities through cooperative agreements. For example, agencies with longstanding HEC programs, including DoD, DOE, and NSF, could serve as models for doing this.
  - Each agency should assess and make arrangements to provide for its own resource needs based on mission priorities.
- *Availability*
  - Agencies should examine the value of reallocating resources to respond to the growing demand for computational resources and the increasing stress that will result from an already overburdened system.

- Agencies might also assess and adjust the relative balance among research and engineering (R&E) modes (theory, experiment, and computation) to ensure optimal allocation of resources.
- *Leadership Systems*
  - High cost limits the number of Leadership Systems to a few.
  - Agencies should manage Leadership Systems as national resources for all participating agencies.
  - Agencies should operate Leadership Systems as an open user facility with cooperative stewardship practices similar to those used by large national resources such as the Hubble Space Telescope and the Spallation Neutron Source.
  - Access to these systems should be governed by a peer review process managed by a council of representatives from the agencies involved.

The report also proposed three multi-agency HEC pilot projects to improve the efficiency of Federal procurement practices:

- *Benchmarking* — to develop a single suite of benchmarks based on applications and use the suite in a pilot collective acquisition cycle
- *Enhanced Total Cost of Ownership (TCO)* — to develop best practices for determining TCO. Elements include acquisition and maintenance, personnel, extra-center communications, user productivity (including applications software development cost), as well as Grid and distributed computing, large remote data sets, long-haul communications, and cost of systems software and middleware.
- *Procurement* — to develop a common solicitation and use a single suite of benchmarks. Each participating agency would weight the benchmarks and apply relevant parts of the TCO model according to its needs.

#### **2.6.4 Workshop and Report on the Roadmap for the Revitalization of High-End Computing**

To solicit inputs from the national community including researchers, applications developers, commercial providers, HEC managers, and other interested parties, the HECRTF participated in an independent workshop convened by the Computing Research Association on *The Roadmap for the Revitalization of High-End Computing* on June 16-18, 2003 in Washington, D.C. Participation in the workshop was limited to approximately 200 individuals who were selected from a pool of candidates who submitted white papers on key aspects of supercomputer requirements, technologies, software, applications, metrics, and management plus key participants from Federal agencies. Based on workshop deliberations, CRA prepared a report entitled *Workshop on the Roadmap for the Revitalization of High-End Computing*<sup>19</sup> that includes

---

<sup>19</sup> <http://www.cra.org/Activities/workshops/nitrd/>

findings and recommendations to advance the state of high-end computing in the United States. These include:

- *Develop enabling technologies:* A variety of new device technologies and three-dimensional integration and packaging concepts show promise in ameliorating the interconnect bandwidth and heat dissipation problems. Software for large-scale systems requires scaling demonstration.
- *COTS-based architectures:* Develop COTS enabling technology including memory-class ports, higher-speed signaling, higher-radix routers, and FPGAs.
- *Custom architectures need proof of concept demonstration:* Spatially direct-mapped architecture, vectors, streaming architecture, processor-in-memory, and special purpose devices.
- *Runtime and operating systems:* Beyond UNIX, alternate resource management models are needed for dynamic adaptation.
- *Programming environment and tools:* Multidisciplinary research will require improved quality, availability, and usability of the software tools used throughout an application's life cycle.
- *Performance analysis:* Develop modeling to predict future system performance based on time-to-solution for applications
- *Applications driven system requirements:* A wide range of applications need system speedups in sustained performance of 50 to 100 fold.
- *Procurement, accessibility, and cost-of-ownership:* Functional specifications should be used for science requirements and total cost of ownership, together with technical and risk assessments, should be the primary evaluation criterion.

These results provided input to the preparation of the HECRTF report.

### **2.6.5 HEC–University Research Activity (HEC-URA)**

DARPA, DOE/SC, NSA, and NSF collaboratively developed the HEC-URA<sup>20</sup> that began in FY 2004. It is an outgrowth of the HECRTF activity. It funds universities to conduct research in software specifically for high-end computing. Their 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

---

<sup>20</sup> <http://www.nitrd.gov/hecrtf-outreach/hec-ura/>

- Integration at the system level
- Development of new technologies (serial number one)
- Test and evaluation to reduce the risk for development, engineering, and government procurement

They 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 (FY 2004), 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. To achieve these ends, the HEC-URA agencies coordinated their planning, with DOE/SC and DARPA focusing on operating systems for extreme-scale scientific computations, NSF and DARPA focusing on languages, compilers, and libraries, and NSF focusing on software tools and development environments.<sup>21</sup>

## 2.6.6 Implementation of the HECRTF Plan

In June 2004, the HECRTF Implementation Committee began to implement the recommendations called for in the *Federal Plan*, starting with a series of activities that include:

- A meeting to coordinate FY 2006 agency HEC budget requests
- A meeting to plan cooperative development and use of HEC benchmarks and metrics. The goal is to make agency procurement solicitations and vendor responses easier and less expensive to prepare.
- Planning for workshops on testbeds for computer science research and on file systems

## 2.6.7 Study on High-End Computing in Japan

In January 2004, DOE/SC, NASA, and NSF commissioned, and the NCO/ITRD coordinated, an almost year-long study on high-end computing in Japan. The study was implemented by WTEC, Inc., an organization that conducts S&T/R&D assessments. The purpose and scope of the study were to:

- Gather information on current status and future trends in Japanese high-end computing from government agencies, research organizations, and the private sector
- Compare Japanese HEC R&D and applications activities with those in the United States
- Assess long-term HEC research in Japan, including follow-on machines to Japan's Earth Simulator (JES) and other architectures
- Review JES development process and operational experience including user experience and its impact on computer science and computational science communities
- Determine HEC areas amenable for Japan-U.S. cooperation to accelerate future advances

---

<sup>21</sup> The DOE/SC and NSF solicitations can be found at: <http://www.nitrd.gov/hecrtf-outreach/hec-ura/index.html>

The study panel includes expert computer scientists from the U.S. Government and academia. A key component of the panel's information gathering process was a March 28-April 3, 2004 fact finding mission to 21 industrial, academic, and government sites in Japan.

In May 2004, the panel completed a preliminary report of its findings and conclusions about Japan's HEC environment that addressed three topic areas: Japanese Earth Simulator, Grids and Clusters, and Business and Government. The panel reported these conclusions at a May 24-25, 2004 workshop in Washington, D.C.

- Japanese Earth Simulator
  - The JES is a superb engineering achievement and will continue to lead the world for a while. Considering the whole spectrum of HEC, the United States is ahead of Japan.
  - Three reasons the Earth Simulator came about:
    - Government funding was robust in Japan
    - Problem area was important to Japan
    - Visionary and driving force in the late Dr. Hajime Miyoshi, who had been the Director of the Earth Simulator Research and Development Center (ESRDC) and who had championed the development of the JES
  - There are no plans to expand the JES itself. However, a new heterogeneous generation machine has been proposed that is 10,000 times more powerful and that would be internationally funded.
  - Japan has had a broad-based strategic effort in high performance computing over the past decade.
  - For the JES, the investment in software was not in proportion to the investment in hardware.
  - JES is extending its applications to other fields beyond Earth sciences to include biosciences and nanotechnology.
  - There is resentment to the JES by some research groups in Japan that claim that it is too expensive and drains resources.
- Grids and Clusters
  - There is more emphasis on Grids than the study panel had expected.
  - Computational, data, and business Grids are included
  - Japan is actively concentrating on Grid computing software with the SuperSINet that will tie all of Japan together.
- Business and Government Environment
  - The CSTP is the highest policy-making body for S&T R&D in the Japanese government, with the Prime Minister presiding at its monthly meetings.
  - METI is not giving emphasis to funding high performance scientific computing.
  - Japan relies primarily on foreign software and standards.

- Over the years, Japanese government projects have encouraged new architectures that have since been commercialized.
- The new policy toward Independent Administrative Institution (IAI) at the universities seems to be a major change in government funding.
- The commercial viability of “traditional” supercomputing architectures with vector processors and high-bandwidth memory subsystems is problematic. NEC continues to be committed to traditional parallel-vector architectures targeted for high-end scientific research.
- Commodity clusters are replacing traditional high-bandwidth systems and shrinking their market. NEC, Fujitsu, and Hitachi have products in the commodity cluster space.

A full study that will include recommendations to the U.S. Government about its own levels of HEC software and hardware investments is projected for publication in FY 2005.

## **2.7 HEC AGENCY ACTIVITIES**

Each HEC agency focuses on programs and activities that address the mission needs of the agency while cooperating with other HEC agencies to meet these needs and to expand their awareness and involvement in other HEC areas of interest. This section reviews the programs and activities of the HEC agencies. The activities described below cover both HEC I&A and HEC R&D.

### **2.7.1 National Science Foundation (NSF)**

High-end computing is important to work supported by all NSF research directorates and offices:

- Biological Sciences (BIO)
- Computer and Information Science and Engineering (CISE)
- Education and Human Resources (EHR)
- Engineering (ENG)
- Geosciences (GEO)
- Mathematical and Physical Sciences (MPS)
- Office of Polar Programs (OPP)
- Social, Behavioral, and Economic Sciences (SBE)

The directorates have complementary HEC investments. For example:

- Computer architecture, networking, software, and cyberinfrastructure are funded by CISE
- HEC devices are funded by MPS and ENG
- Mathematical algorithms are funded by MPS and CISE
- Computational algorithms and libraries are funded by CISE with some funding from MPS
- Science and engineering applications are primarily developed with funding from MPS, ENG, GEO, and BIO

The main HEC programs at NSF are:

- HEC R&D
- Shared Cyberinfrastructure
- HEC-URA (discussed in section 2.6.5)

### **2.7.1.1 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:

- Computing and Communication Foundations (CCF)<sup>22</sup> 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 – including 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.
- The Information and Intelligent Systems (IIS) Division supports programs in data-driven science including bioinformatics, geoinformatics, cognitive neuroscience, and other areas.
- The Shared Cyberinfrastructure (SCI)<sup>23</sup> Division supports programs in:
  - Infrastructure development – creating, testing, and hardening next-generation deployed systems
  - Infrastructure deployment – planning, construction, commissioning, and operations

The following examples 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

---

<sup>22</sup> <http://www.cise.nsf.gov/div/index.cfm?div=ccf>

<sup>23</sup> [http://www.cise.nsf.gov/funding/pgm\\_display.cfm?pub\\_id=5584&div=sci](http://www.cise.nsf.gov/funding/pgm_display.cfm?pub_id=5584&div=sci)

- Dynamic and Static (compiler) Optimizations, for example:
  - Memory systems
    - Simultaneous multi-threading/chip multi-processing (SMT/CMP) to support latency reduction
    - Prefetching
    - Speculative execution (load/value prediction)
  - Verification and runtime fault tolerance
- Networking and Storage

NSF also funds speculative research that could provide breakthrough technologies for HEC including:

- Nanotechnology that promises unprecedented scale-up in computing and communication through research to apply them in architecture and design methods
- Post-silicon (quantum, DNA, chemical, etc.) computing that may provide radically new models of computation, algorithms, and “programming techniques”

#### **2.7.1.2 NSF Cyberinfrastructure Framework**

The NSF integrated cyberinfrastructure framework will support the communities performing research and using computing resources. The components of this framework are hardware (distributed resources for computation, communication, storage, etc.), shared software cybertools (Grid services and middleware, development tools and libraries), domain specific cybertools, and applications. This infrastructure enables discovery and innovation and education and training. The cyberinfrastructure framework will consist of:

- Computation engines (supercomputers, clusters, workstations) – both capability and capacity hardware
- Mass storage (disk drives, tapes, etc.) with persistence
- Networking (including wireless, distributed, ubiquitous)
- Digital libraries and databases
- 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 is needed to enable HEC.

#### **2.7.1.3 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 an 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)

Support for these activities will continue in FY 2005 as part of new activities in NSF/CISE/SCI.

#### **2.7.1.3.1 Extensible Terascale Facility (ETF)**

ETF, also known as TeraGrid, is a multi-year effort to build and deploy the world's largest, most comprehensive, distributed infrastructure 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 over 20 teraflops of computing power, be able to store and manage nearly 1 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.

#### **2.7.1.3.2 NSF HEC Users**

In FY 2003 NSF had 4,450 total users including 1,800 NPACI users, 1,200 NCSA users, and 1,450 PSC users. It is estimated that there were 3,000 unique users, since many users use more than one center.

### **2.7.2 Department Of Energy/Office of Science (DOE/SC)**

DOE/SC supports basic research that underpins DOE missions and constructs and operates large scientific facilities such as accelerators, synchrotron light sources, and neutron sources for the U.S. scientific community. Six offices oversee DOE/SC research:

- Basic Energy Sciences: Materials, chemistry, and nanoscale science
- Biological and Environmental Research: Global climate and genomics research

- Fusion Energy Sciences: Magnetic fusion, inertial fusion, plasma physics
- High Energy Physics: Accelerator design and petascale experimental data analysis
- Nuclear Physics: Quantum ChromoDynamics (QCD) and astrophysics
- Advanced Scientific Computing Research

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 up to 50 orders of magnitude.
- Several scientific disciplines are involved (for example, 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.

### **2.7.2.1 DOE's Scientific Discovery Through Advanced Computing (SciDAC) Program**

The DOE/SC SciDAC Program<sup>24</sup> provides terascale computing and associated information technologies to many scientific areas to foster breakthroughs through 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 capabilities by the broader scientific community and are bringing simulation to a level of parity with theory and observation in the scientific enterprise.

### **2.7.2.2 DOE/SC HEC R&D Programs**

Major recent DOE/SC accomplishments in HEC R&D that help address DOE/SC's computational science needs include:

- PVM — 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 — 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.

---

<sup>24</sup> <http://www.osti.gov/scidac/>

- 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

### **2.7.2.2.1 SciDAC Integrated Software Infrastructure Centers (ISICs)**

Three Applied Mathematics ISICs will develop new high-performance scalable numerical algorithms for core numerical components of scientific simulation, and distribute those algorithms through portable scalable high-performance numerical libraries. These three ISICs are:

- Algorithmic and Software Framework for Applied Partial Differential Equations (PDE) – 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<sup>25</sup> will work 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 four 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

---

<sup>25</sup> [http://www.osti.gov/scidac/computing/research\\_area.html](http://www.osti.gov/scidac/computing/research_area.html)

### **2.7.2.2 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.

### **2.7.2.3 DOE/SC HEC Workshops**

Workshops are a standard DOE/SC planning tool. Such workshops are either general or focused. DOE/SC has held the following general workshops:

- DOE/SC Science Computing Conference (June 19-20, 2003, Arlington, Virginia)<sup>26</sup>
- Science Case for Large Scale Simulation (SCALES) (June 24-25, 2003, Arlington, Virginia)<sup>27</sup>
- SciDAC (March 2004, Charleston, South Carolina)<sup>28</sup>

Focused workshops, which address specific application area needs, include:

- Fast Operating Systems<sup>29</sup>
- Multiscale Mathematics Workshop (May 3-5, 2004) – to bridge the wide scale between modeling on continuum or atomistic basis – expected output includes a roadmap
- High Performance File Systems (Summer 2004) – such file systems should support parallelism, be efficient, and be reasonably easy to use
- Distributed Visualization Architecture Workshops (DiVA)<sup>30</sup>
- Data Management Workshop series to address where file system ends and data management system begins<sup>31</sup>
- Raising the level of parallel programming abstraction (TBD)
- Parallel tool infrastructure (TBD)

### **2.7.2.4 DOE/SC HEC I&A Programs**

#### **2.7.2.4.1 National Energy Research Supercomputing Center (NERSC)**

NERSC<sup>32</sup> delivers high-end capability computing services and support to the entire DOE/SC research community. NERSC provides these services to the DOE community, to the other DOE laboratories, and to major universities performing work relevant to DOE missions. NERSC provides the majority of resources and services that are used to support the SciDAC program. The Center serves 2,000 users working on about 700 projects; 35 percent of users are university based, 61 percent are in national laboratories, three percent are in industry, and one percent in

---

<sup>26</sup> <http://www.doe-sci-comp.info>

<sup>27</sup> <http://www.pnl.gov/scales/>

<sup>28</sup> <http://www.osti.gov/scidac/>

<sup>29</sup> <http://www.cs.unm.edu/~fastos/>

<sup>30</sup> <http://vis.lbl.gov/Research/DiVA/Workshops/index.html#Findings>

<sup>31</sup> <http://www-conf.slac.Stanford.edu/dmw2004/>

<sup>32</sup> <http://www.nersc.gov/>

other government laboratories. The major computational resource at NERSC is an IBM SP computer whose capacity was doubled to 10 teraflops in FY 2004. These computational resources are integrated by a common high performance file storage system that facilitates interdisciplinary collaborations.

#### **2.7.2.4.2 Advanced Computing Research Testbeds (ACRT)**

This activity supports the acquisition, testing, and evaluation of advanced computer hardware to assess the prospects for meeting future DOE/SC computational needs, such as SciDAC and special purpose applications. The ACRT activity provides 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 government-wide.

#### **2.7.2.4.3 Scientific Application Pilot Programs (SAPP)**

The 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. It provides expertise in adapting high-performance algorithms to terascale computers and in adapting numerical modules to include application-specific knowledge.

#### **2.7.2.5 Innovative and Novel Computational Impact on Theory and Experiment (INCITE) Projects**

INCITE is a 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 our 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

## 2.7.3 Defense Advanced Research Projects Agency (DARPA)

DARPA has four HEC R&D programs:

- The High Productivity Computing Systems (HPCS) program that addresses all aspects of HEC systems including packaging, processor/memory interfaces, networks, operating systems, compilers, languages, and runtime systems
- Polymorphous Computing Architectures (PCA) so computers can adapt to applications
- Networked Embedded Systems Technology (NEST) that looks at hundreds of thousands of tiny simple computing nodes interacting with each other
- HEC – URA

The first three are described below. HEC-URA is described in section 2.6.5.

### 2.7.3.1 High Productivity Computing Systems (HPCS)

The HPCS program<sup>33</sup> has a goal of providing 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 speedup in performance of 10 to 40 times 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 it is running on
- *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 provide for the continuous fielding of new high performance computer technology through a phased concept, research, development, and demonstration approach. The program is in its second phase. The productivity goals of this phase are to provide:

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

---

<sup>33</sup> <http://www.darpa.mil/iptoPrograms/hpcs/index.htm>

In phase two, the productivity framework has been baselined to today's systems, used to evaluate vendors' emerging productivity techniques, and provides a reference for use in phase three to evaluate vendors' proposed designs.

Phase two subsystem performance indicators are a 3.2 petabytes/second bisection bandwidth, 64,000 giga updates per second (GUPS), 6.5 petabytes/second data streams bandwidth, and 2+ petaflops Linpack Fortran benchmark (a HPC Challenge) that augments the TOP500<sup>34</sup> benchmarks.

In FY 2004 DARPA supported two HPCS-related workshops. They included:

- The HPCS Productivity Team/Task Group that met June 29-30 and July 1, 2004, in the Washington, D.C. area.<sup>35</sup> The meeting targeted an expanded technical audience.
- The ninth international workshop on High-Level Parallel Programming Models and Supportive Environments (HIPS 2004) on April 26, 2004, in Santa Fe, New Mexico.<sup>36</sup>

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 two industry partners. Customers include DoD (NSA and NRO), DOE/NNSA, DOE/SC, NASA, NIH, NSF, DHS, and the commercial sector.

A Council on Competitiveness HPC Initiative 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. It includes:

- An HPC Advisory Committee with representatives from the private sector
- An annual private-sector user survey
- Annual user meetings
  - The first conference, entitled "HPC: Supercharging U.S. Innovation and Competitiveness," was held July 13, 2004 in Washington, DC.<sup>37</sup> Senior application users, HPC center directors, industry executives, and policy and funding decision makers participated.

### **2.7.3.2 Grand Challenge Case Study: Field Test of Autonomous Ground Vehicles**

DARPA, DOE/SC, and the private sector hold workshops and conduct industry Grand Challenge case studies. DARPA conducted its first Grand Challenge, a field test of autonomous ground vehicles, March 8-13, 2004.<sup>38</sup> Its purpose was to encourage the accelerated development of autonomous vehicle technologies that could be applied to military requirements.

---

<sup>34</sup> <http://www.netlib.org/benchmark/top500.html>

<sup>35</sup> <http://www.HighProductivity.org>

<sup>36</sup> <http://csdl.computer.org/comp/proceedings/hips/2004/2151/00/2151toc.htm>

<sup>37</sup> <http://www.hpcusersconference.com/agenda.html?print=1>

<sup>38</sup> <http://www.darpa.mil/grandchallenge/>

### **2.7.3.3 Polymorphous Computing Architectures (PCA) Program**

The Polymorphous Computing Architectures (PCA) Program has a goal to develop the computing foundation for agile systems by establishing computing systems (chips, networks, 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 milliseconds, 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

PCA hardware being developed includes:

- *Monarch/MCHIP*. A prototype in 2005 is projected to demonstrate 333Mhz, 64 Gflops, 64 GOPS, and 12 Mbytes erasable dynamic random access memory (EDRAM).
- *Smart Memories*. A prototype in 2005 is projected to demonstrate streaming applications based on modified commercial Tensilica CPU cores.
- *TRIPS*. A prototype in 2005 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. This technology provides early morphware, software, and kernel performance evaluation.

The morphware<sup>39</sup> 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, RAW, and others.

### **2.7.3.4 Networked Embedded Systems Technology (NEST)**

The NEST program will enable fine-grain fusion of physical and information processes. The quantitative target is to build dependable, real-time, distributed, embedded applications

---

<sup>39</sup> <http://www.morphware.org/PCA101/right1.html>

comprising 100 to 100,000 simple compute nodes. The nodes include physical and information system components coupled by sensors and actuators.

## **2.7.4 National Aeronautics and Space Administration (NASA)**

NASA's mission is to:

- Understand and protect our home planet
- Explore the universe and search for life
- Inspire the next generation of explorers

Until recently, NASA accomplished its mission through seven strategic enterprises:<sup>40</sup>

- Earth Science
- Space Science
- Biological and Physical Research
- Aeronautics
- Exploration Systems
- Space Flight
- Education

### **2.7.4.1 NASA HEC Accomplishments**

FY 2004 NASA HEC accomplishments include the following:

- Testbeds
  - Built the first single system image (SSI) Linux-based supercomputer, which was the world's fastest machine on the triad benchmark – an industry standard measurement of memory bandwidth – and which will be used on the Space Shuttle
- Advanced Architectures
  - Invested in new streams and PIM architectures and worked on languages, compilers, and key application kernels for stream technology (this is tied into DARPA's PCA and morphware efforts)
- Grid Technology
  - Continued work on its Information Power Grid (IPG) demonstrations in cooperation with Earth science and Shuttle payload management
- Applications
  - Developed advanced applications for aerospace, nanotechnology, and radiation biology

---

<sup>40</sup> Future documentation will be consistent with NASA's new organizational structure.

- Major advance in ocean modeling in the “Estimating the Circulation and Climate of the Ocean” (ECCO) code. (ECCO is further described in section 2.7.4.2.)
- Major advance in aerospace vehicle modeling
- Programming Environments
  - Improved scalability multi-level parallelism (MLP)
- Tools
  - Workflow and data base management programming environments
  - Aero data base/IPG Virtual Laboratory for workflow management and ensemble calculation data base management prototypes
  - Hyperwall (for viewing large datasets) used in nanotechnology research, Sloan Digital Sky Survey, and the February 1, 2003 Columbia Space Shuttle Mission accident investigation. This technology is now being transferred for general use within NASA.

#### **2.7.4.2 Estimating the Circulation and Climate of the Ocean (ECCO)**

The goal of the ECCO project 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 and optimized code
- Assisted in revamping Grid technique to improve prediction at poles
- Developed Kalman filter techniques at high resolution for data assimilation
- Developed an archive for community access to data
- Improved resolution from 1 degree to 1/4 degree
- Developed unique visualization approaches

As a result:

- Time to solution dropped from approximately 1 year to 5 days (a 50-fold speedup)
- Resolution improved by a factor of 4 — from 1 degree to 1/4 degree
- Code accuracy and physics improved
- Data assimilation at high resolution was enabled

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

#### **2.7.4.3 Supercomputing Support for the Columbia Accident Investigation Board (CAIB)**

NASA Ames supercomputing-related R&D assets and tools were used to produce time-critical analyses for the CAIB. In particular, 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 consisted of:

- 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

#### **2.7.4.4 NASA HEC Resources**

NASA HEC resources are located at Goddard Space Flight Center and Ames Research Center, which closely collaborate to provide resources to users. Current demand for computational resources is two times capacity and stated requirements are four times capacity. A 300 Terabyte archive was moved from Ames to Goddard with less than ten percent of the transfer supported by networks.

High-end computing resources include:

- Goddard, 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)

Recent increases in computational resources will significantly improve trajectory and mission simulations for digital flight crew exploration vehicle (CEV) development and launch-to-orbit simulations to perform risk assessments.

## 2.7.5 DOE/National Nuclear Security Administration (DOE/NNSA)

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)

### 2.7.5.1 ASC Focus Areas

ASC focuses on:

- *Software quality engineering* that is oriented to processes to provide for longevity of codes
- *Verification* to ensure numerical solutions are accurate
- *Validation* to ensure a problem is understood correctly
- *Certification methodology* to make improvements through scientific methods
- *Capability computing* to meet the most demanding computational needs
- *Capacity computing* to meet current stockpile stewardship 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*

### 2.7.5.2 ASC Activities

Pathforward activities include:

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

Topics of multidisciplinary academic collaborations include fires, rocket motors, astrophysics, jet engines, and shock. ASC works with industry to develop new technologies in optical switches, scalable visualization, memory correctness, spray cooling, and the Lustre File System.

### 2.7.5.3 ASC HEC I&A Platforms

ASC uses HEC systems and software from Cray, HP, IBM, Intel, SGI, and Linux Networx. It procures machines of differing technologies to support diverse requirements.

These types include:

- *Capability systems:* Vendor-integrated symmetric multiprocessing (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 high performance computing (HPC) clusters that support the largest workloads in terms of numbers of jobs supporting current weapons activities and as development vehicles for the 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 petaflops scales.

## **2.7.6 National Security Agency (NSA)**

NSA has unclassified HEC R&D programs in:

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

### **2.7.6.1 Cray X1e/Black Widow**

Under a multi-year 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 Black Widow system to be introduced in CY 2006 will provide outstanding global memory bandwidth across all processor configurations and will be scalable to hundreds of teraflops. It should be the most powerful commercially available system in the world at that time. Many agencies participated in the quarterly development reviews.

### **2.7.6.2 HPC Applications Benchmarking Study**

A select set of HPC users, the HPC User Forum<sup>41</sup>, and IDC are sponsoring a study of 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.<sup>42</sup> This

---

<sup>41</sup> <http://www.hpcuserforum.com/>

<sup>42</sup> <http://www.idc.com/hpc>

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 these computers. The HPC User Forum effort is funded by DoD/HPCMPO, NSA, and other agencies.

#### **2.7.6.3 Reconfigurable Computing R&D**

The reconfigurable computing program is 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. NSA collaborates with other agencies in this effort.

Major achievements in FY 2004:

- Demonstration on an NSA application of the C compiler on a particular reconfigurable computer achieved a 65-fold 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

#### **2.7.6.4 Fine-Grain Multithreaded Architectures and Execution Models**

Performance analysis of many real-world applications on today's large-scale massively parallel processor systems indicates that often a small fraction of the potential performance of these machines is achieved in practice. Applications that exhibit these problems are generally characterized by memory access patterns that are dynamic and irregular, and hence difficult to partition efficiently, by either programmers or compilers. This situation frequently results in processors being "starved" for data due to 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. NSA is pursuing research to develop architectures, execution models, and compilation techniques that can exploit multithreading to manage memory latency and provide load balancing. 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.

#### **2.7.6.5 Optical Technologies**

NSA is investigating high-end computer optical interconnect systems. A comparison of electrical and optical technologies showed that specifically-designed all-optical switching can significantly outperform current and future electronic switches. The effort demonstrated that optical

technologies can significantly improve bandwidth if implementation difficulties of all-optical switching can be overcome.

### **2.7.6.6 Quantum Computing**

Through its ARDA affiliate, NSA conducts research in quantum computing. Efforts focus on quantum information sciences including investigation of selective materials, related quantum properties, and managing decoherence and entanglement issues. This research is part of a multi-agency collaboration with DARPA and others.

## **2.7.7 National Oceanic and Atmospheric Administration (NOAA)**

NOAA uses HEC resources and advanced applications to provide users with advanced products and services.<sup>43</sup> NOAA's HEC strategy is to:

- Develop skills, algorithms, and techniques to fully use 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 changing view of HEC, which is more an enterprise resource.

### **2.7.7.1 NOAA HEC Resources, Activities, and Plans**

NOAA HEC resources are located at:

- Geophysical Fluid Dynamics Laboratory (GFDL)<sup>44</sup> in Princeton, New Jersey
- Forecast Systems Laboratory (FSL)<sup>45</sup> in Boulder, Colorado
- National Centers for Environmental Prediction (NCEP)<sup>46</sup> headquartered in Camp Springs, Maryland

In FY 2003, GFDL acquired a 1,408 processor SGI system.

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 post-doctoral students and contractors

NOAA plans are to:

---

<sup>43</sup> HPCC office: <http://www.cio.noaa/hpcc/index.html>.

<sup>44</sup> <http://www.gfdl.noaa.gov>

<sup>45</sup> <http://www.fsl.noaa.gov>

<sup>46</sup> <http://www.ncep.noaa.gov>

- Develop the Earth System Modeling Framework (ESMF) for climate model development and extend it to all domains (ESMF is described in section 2.5.1.)
- Explore the use of specific types of Grid technology
- Develop a common weather research computing environment
- Apply the Weather Research and Forecast (WRF) (described below) modeling system standards and framework to NCEP's Mesoscale Modeling Systems, including the continental domain Meso Eta and the nested Nonhydrostatic Meso Threats runs

#### **2.7.7.1.1 Weather Research and Forecasting (WRF) Development Test Center (DTC)**

NOAA/FSL is establishing a multi-agency DTC which will initially focus on the development of the WRF model that is designed as both an operational model and a research vehicle for the larger modeling community. WRF is a regional atmospheric model for operational numerical weather prediction and atmospheric research focusing on one to ten kilometer resolution. At the DTC, NOAA personnel will work with university researchers to develop and add specific model improvements to the modular WRF framework. 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. The WRF model will greatly increase the accuracy and specificity of weather forecasts. This work is joint with NSF and the National Center for Atmospheric Research (NCAR).

#### **2.7.7.2 NOAA Workshops and Meetings**

The semi-annual HPCC PI status meeting was held in the Spring of 2004 and the Tech Biennial Meeting to review progress will be held in October 2004.

### **2.7.8 National Institute of Standards and Technology (NIST)**

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. It also conducts fundamental research that facilitates measurement, testing, and the adoption of industry standards.

#### **2.7.8.1 NIST HEC I&A**

NIST developed the Interoperable Message Passing Interface (IMPI)<sup>47</sup> to support parallel computing research. The standard enables interoperability of different implementations of the Message Passing Interface (MPI) with no change to user source code.

In distributed computing, NIST has developed a Java/Jini/Javaspaces-based screen-saver science system.<sup>48</sup> The environment leverages Java's portability to use otherwise unused cycles on any set

---

<sup>47</sup> A tester is at <http://impi.nist.gov>.

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. The environment is being used to compute property data (virtual measurements via computer codes), augmenting property values obtained from laboratory measurement.

NIST has created infrastructure to use immersive visualization to explore scientific data sets drawn from both computer and laboratory experiments. Visualization projects include a three-wall Immersive Virtual Reality System that uses the Device Independent Virtual Environments – Reconfigurable, Scalable, Extensible (DIVERSE)<sup>49</sup> open source software and locally-built visualization software to see results at scale.<sup>50</sup> The local software includes the glyph toolbox, techniques for immersive volume visualization, surface visualization, and techniques for interaction. All the software runs unchanged on Linux machines.

NIST is developing fundamental mathematical software tools to enable high-end computing applications. Examples include:

- The Parallel Hierarchical Adaptive Multilevel (PHAML)<sup>51</sup> code, a parallel adaptive Grid refinement multigrid code for PDEs
- The Sparse Basic Linear Algebra Subroutines (Sparse BLAS)<sup>52</sup>
- The Template Numerical Toolkit (TNT)<sup>53</sup> for object-oriented linear algebra

NIST HEC applications often combine parallel computing with immersive visualization for basic sciences, such as physics, and applications, such as building structure and material strength. Some typical applications to which NIST applies these resources are:

- Nanostructure modeling and visualization (especially in nano-optics)<sup>54</sup>
- Modeling cement and concrete, including modeling the flow of suspensions and fluid flow in complex geometries<sup>55</sup>
- Computation of atomic properties
- Visualization of “smart gels”<sup>56</sup> (which respond to specific physical properties such as temperature or pressure) to gain insight into the gelling mechanism
- Visualization of tissue engineering<sup>57</sup> to study and optimize the growth of cells on scaffolding

In addition, NIST is developing applications-oriented problem-solving environments, such as Object-Oriented Finite Element Modeling of Material Microstructures (OOF)<sup>58</sup> for modeling

---

<sup>48</sup> <http://math.nist.gov/mcsd/savg/parallel/screen/>

<sup>49</sup> <http://diverse.sourceforge.net>

<sup>50</sup> <http://math.nist.gov/mcsd/savg/software/>

<sup>51</sup> <http://math.nist.gov/phaml/>

<sup>52</sup> <http://math.nist.gov/spblas/>

<sup>53</sup> <http://math.nist.gov/tnt/>

<sup>54</sup> <http://math.nist.gov/mcsd/savg/parallel/nano/>

<sup>55</sup> <http://math.nist.gov/mcsd/savg/parallel/dpd/>, <http://math.nist.gov/mcsd/savg/parallel/lb/>, <http://math.nist.gov/mcsd/savg/vis/concrete/>, <http://math.nist.gov/mcsd/savg/vis/fluid/>

<sup>56</sup> <http://math.nist.gov/mcsd/savg/vis/gel/>

<sup>57</sup> <http://math.nist.gov/mcsd/savg/vis/tissue/>

<sup>58</sup> <http://www.ctcms.nist.gov/oof/>

materials with complex microstructure and Object-Oriented MicroMagnetic Computing Framework (OOMMF)<sup>59</sup> for micromagnetics modeling.

### **2.7.8.2 NIST HEC R&D**

NIST is performing research in quantum computing and secure quantum communications, in measurement science aspects of nanotechnology, photonics, optoelectronics, and new chip designs, and in fabrication methods.

NIST activities in quantum information science<sup>60</sup> have both theoretical and experimental components. It conducts research in algorithms and architectures for scalable quantum computing and demonstrates quantum computing within specific physical systems such as ion traps, neutral atoms, and Josephson junctions. NIST has developed a testbed for high-speed quantum key distribution and conducts associated research on single photon sources and detectors and on quantum protocols.

### **2.7.9 Environmental Protection Agency (EPA)**

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 are wide ranging and diverse, extending from groundwater to the stratosphere.

EPA HEC 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 facilitating the proper characterization of scientific findings and the appropriate use of science 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.

In FY 2004, EPA's HEC programs<sup>61</sup> include:

- Multimedia Assessments and Applications Framework
- Uncertainty Analysis Framework Development
- Air Quality Modeling Applications
- HEC visualization

---

<sup>59</sup> <http://math.nist.gov/oommf/>

<sup>60</sup> <http://qubit.nist.gov/>

<sup>61</sup> HPCC Office: <http://www.epa.gov/hpcc>

- HEC infrastructure, Grid, and science subnet

### **2.7.9.1 Multimedia Assessments and Applications (MAA) Framework**

The MAA framework provides a foundation for research on how to structure compartmental models and improve model integration and interchangeability. The framework's objective is to provide software that supports composing, configuring, applying, linking, and evaluating complex systems of models. It will improve EPA's ability to simulate the interaction between individual environmental media components (for example, 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.

### **2.7.9.2 Uncertainty Analysis (UA) Framework Development**

The UA framework program is developing tools to support the analysis of model sensitivities and the effects of input uncertainties on model predictions. Specific tasks are to:

- Construct a 400-node Intel-based supercomputing cluster called Supercomputer for Model Uncertainty and Sensitivity Evaluation (SuperMUSE)
- 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

### **2.7.9.3 Air Quality Modeling Applications (AQMA)**

The AQMA program has the objective of advancing the state-of-the-art Community Multi-scale Air Quality (CMAQ) Chemistry-Transport Modeling System computational performance 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. The AQMA program is using a phased deployment that enables the States, which are the key stakeholders, to participate in the development.

### **2.7.9.4 Grid Deployment**

EPA has a goal 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 (NESCC). With the implementation of Grid middleware, agency 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 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.

#### **2.7.9.4.1 Grid Demonstration Project**

EPA is working 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 toward achieving air quality and public health goals. Air quality has been selected as a Grid demonstration project. Technical collaborators include DOE/Sandia and NOAA and pilot partners include the State of New York and regional planning organizations. Phase I of the Grid demonstration project includes delivering an optimized CMAQ model and datasets to the client community in the summer of 2004 and then eventually to run CMAQ over the Grid and at client sites.

### **2.7.10 Department of Defense/High Performance Computing Modernization Program Office (DoD/HPCMPO)**

DoD/HPCMPO's mission is to deliver world-class commercial, high-end, high performance computational capability to DoD's S&T and test and evaluation (T&E) communities, thereby facilitating the rapid application of advanced technology into superior warfighting capabilities. DoD/HPCMPO's vision is to create a pervasive culture among DoD's scientists and engineers in which they routinely use advanced computational environments to solve their most demanding problems.

HPC modernization is essential to future technology development in a wide range of areas critical to the DoD missions. 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 Vehicles
- Blast protection

#### **2.7.10.1 User Base and Requirements**

HPCMPO's S&T and T&E communities involves some 4,482 users working on 667 projects at approximately 175 sites at Air Force, Army, and Navy high performance computing centers and defense agencies such as DARPA. FY 2004 non-real-time user requirements are estimated at 134 Teraflop-years. HPCMPO requirements are categorized in ten key Computational Technology

Areas (CTAs). They are:

	<u>DoD Users</u>	
	<u>Primary</u>	<u>Total</u>
• Computational Structural Mechanics (CSM)	528	1,019
• Computational Fluid Dynamics (CFD)	1,154	1,612
• Computational Chemistry and Materials Science (CCM)	243	593
• Computational Electromagnetics and Acoustics (CEA)	295	575
• Climate/Weather/Ocean Modeling and Simulation (CWO)	284	521
• Signal/Image Processing (SIP)	239	815
• Forces Modeling and Simulation (FMS)/C4I	527	913
• Environmental Quality Modeling and Simulation (EQM)	120	295
• Computational Electronics and Nanoelectronics (CEN)	35	135
• Integrated Modeling and Test Environments (IMT)	994	1,201
• Others	63	219

HPCMPO capabilities include approximately 30,000 Gigaflop-years in FY 2004. The demand for capability is 130,000 in FY 2004.

### **2.7.10.2 High Performance Computing Centers**

Among HPCMPO's high performance computing centers, four are major shared resource centers (MSRCs) and 16 are distributed centers.

MSRCs provide complete networked HPC environments, world-class HPC compute engines, high-end data analysis and scientific visualization support, massive hierarchical storage, proactive and in-depth user support, and computational technology area expertise to nation-wide user communities. The MSRCs are ARL, ASC, ERDC, and NAVO.

There are two categories of distributed centers:

- Allocated Distributed Centers (ADCs) that support the annual service or agency project allocation and/or challenge project allocation process by providing reserved compute resources. These include large systems that support a few high priority projects. ADCs include Army High Performance Computing Research Center (AHPCRC), Arctic Region Supercomputing Center (ARSC), Maui High Performance Computing Center (MHPCC), and Space Missile Defense Command (SMDC).
- Dedicated Distributed Centers (DDCs) that support a dedicated project or projects for which each independent DDC was awarded an HPC system. These include moderate systems to support dedicated local needs. DDCs include the Arnold Engineering Development Center (AEDC), Air Force Research Laboratory/Information Directorate (AFRL/IF), Air Force Weather Agency (AFWA), Fleet Numerical Meteorology and Oceanography Center (FNMOC), Joint Forces Command (JFC/J9), Naval Air Warfare Center Aircraft Division (NAWCAD), Naval Research Laboratory – Washington, D.C. (NRL-DC), Redstone Technical Test Center (RTTC), Simulation and Analysis Facility

(SIMAF), SPAWAR Systems Center – San Diego (SSC-SD), and White Sands Missile Range (WSMR).

### **2.7.10.3 Software Applications Support**

Software Applications Support (SAS) has four components:

- Common HPC Software Support Initiative (CHSSI)
- HPC Software Applications Institutes (HSAIs)
- Programming Environment and Training (PET) program to enhance user productivity
- Software Protection Center (SPC) to develop the best techniques for the protection of HPC applications

Each is now described.

#### **2.7.10.3.1 Common HPC Software Support Initiative (CHSSI)**

The HPCMPO CHSSI supports the development of common, efficient HPC applications. CHSSI has the following features:

- It provides core software applications, models, and simulations that exploit HPC capabilities in areas of most concern to DoD.
- Multi-purpose codes perform on a range of HPC platforms, with focus on accuracy, efficiency, reusability, and scalability.
- Inter-service, multi-disciplinary teams include algorithm and code developers, applications specialists, computer scientists, and end users who are responsible for development, dissemination, technology transfer, and follow-on support.
- Collaboration with PET academics and MSRCs on innovative software tools, training, and other elements common to HPC programming environments

#### **2.7.10.3.2 HPC Software Applications Institutes (HSAIs)**

There are between five and eight HPC HSAIs in a given year. Each involves between three and twelve computational and computer scientists for a period of three to six years to develop new code or support existing code, and adjust local business practices to use science-based models and simulation.

#### **2.7.10.3.3 Programming Environment and Training (PET)**

The HSAIs are integrated with the PET program whose mission is to enhance the productivity of the DoD HPC user community. Its goals are to:

- Transfer leading-edge HPC (computational and computing) technology into DoD from other government, industrial, and academic HPC communities
- Develop, train, and support new and existing DoD HPC users with the education, knowledge, access, and tools to maximize productivity

#### **2.7.10.3.4 Software Protection Center (SPC)**

DoD implements protections on national security applications software. The SPC is responsible for developing software protection technologies, supporting the insertion of software protections into applications codes, working with the services and agencies to improve the effectiveness of export control regulations for software, and serving as an information center for best practices and guidelines. A Red Team tests and validates protection technologies and implementations.

## **3. HUMAN COMPUTER INTERACTION AND INFORMATION MANAGEMENT (HCI&IM)**

### **3.1 INTRODUCTION**

The HCI&IM CG plans, coordinates, and evaluates Federal R&D of technologies for mapping human knowledge into computing systems, communications networks, and information systems and back to human beings, all for the benefit of human understanding, analysis, and use.<sup>62</sup> The HCI&IM CG comprises expert volunteers from nine NITRD agencies – NSF, NIH, DARPA, NASA, AHRQ, DOE/SC, NIST, NOAA, and EPA – and one non-NITRD agency, FAA – who meet monthly to exchange information and coordinate their HCI&IM R&D programs and activities. The HCI&IM CG succeeds the Human Centered Systems (HuCS) CG of the late 1990s. The IM component was formally added in 2000 in response to the recommendations made in the 1999 PITAC report.<sup>63</sup>

### **3.2 DEFINITION OF THE HCI&IM PCA**

The activities funded under the NITRD Program’s HCI&IM R&D PCA have the goal of increasing 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 R&D focuses on developing systems that understand the needs of the users and adapt accordingly.

Broad areas of HCI&IM concern include:

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

To address these concerns, HCI&IM R&D pursues the following technical goals:

---

<sup>62</sup> [http://www.nitrd.gov/pubs/hci-im\\_research\\_needs\\_final.pdf](http://www.nitrd.gov/pubs/hci-im_research_needs_final.pdf)

<sup>63</sup> “Information Technology Research: Investing in Our Future,” Report of the President’s Information Technology Advisory Committee (PITAC), February 24, 1999, <http://www.nitrd.gov/pitac>.

- By their very nature, most of the problems in this multidisciplinary field require contributions from different branches of science and engineering, and the needed technological convergence will require exceptional communication and coordination efforts.
- Scientific understanding of both human and machine cognition is central to the area, linking the various parts of human-computer interaction and information management together.
- We must learn how to build machines and systems that employ the human senses of perception to their full potential, maximizing the information-flow bandwidth between people and their tools.
- We must discover new and better ways to achieve information integration, across a wide range of different modalities, media, and distributed resources.
- Tying together information management and human needs as data burgeons and becomes useful in all aspects of science, manufacturing, and commerce will require tools that are increasingly able to take on planning, aid in decision making, learn from experience, and develop knowledge from raw data.
- Flexible systems of control must be developed that manage the degree of autonomy exhibited by machines according to the constantly changing needs of humans, especially in that broad and extremely valuable area where HCI and IM unite for the benefit of people.

Illustrative technical thrusts of HCI&IM programs include:

- 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 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 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

## 3.3 HCI&IM CG ACTIVITIES

### 3.3.1 Release of the “Human-Computer Interaction and Information Management Research Needs” Report

In October 2003, the HCI&IM CG released its *Human-Computer Interaction and Information Management Research Needs* report (henceforth referred to as the *Research Needs Report*). The report identifies and illustrates the problems that underlie 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

It places agency FY 2002 HCI&IM R&D investments and R&D needs within a conceptual framework and elucidates what is meant by the words “human-computer interaction and information management.” The framework is used to depict why humans use computing systems, what humans use these systems for, and how humans and computing systems use data and information both separately and together. HCI&IM R&D needs are organized in four areas:

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

### 3.3.2 HCI&IM R&D Planning and Coordination

In FY 2004 the HCI&IM CG is using its *Research Needs Report* to identify and assess current agency investments, research gaps, and plans. In formulating these plans and directions, the focus is on:

- Refining plans in new R&D areas such as cognition, robotics, and devices
- Identifying and coordinating activities in areas of shared interests with other CGs.  
Potential areas are:
  - With LSN in moving the bits in distributed data and in middleware
  - With SDP in automatic software design and integration
  - With SEW in universal accessibility in the use of multi-agent systems, and in the impact of distributed data on organizations and groups sharing data

### **3.3.3 HCI&IM CG Special Meeting in December 2003**

Complementary to the release of the *Research Needs Report*, the HCI&IM CG held a Special Meeting in December 2003, among other things, to assess and refine its R&D agenda. The following sections reflect the presentations made at the meeting.

## **3.4 HCI&IM MULTI-AGENCY ACTIVITIES**

The following are activities in which more than one HCI&IM agency participates. The activities often also include agencies outside the NITRD Program, which are not identified here.

### **3.4.1 DARPA**

DARPA coordinates with other HCI&IM agencies on the following programs:

- With NSF in its Improving Warfighter Information Intake Under Stress Program
- With NIST and NSA in Translingual Information Detection, Extraction, and Summarization (TIDES)
- With NIST and NSA in Effective, Affordable, Reusable Speech-to-Text (EARS)

These programs are described in section 3.5.3.

### **3.4.2 NASA**

NASA collaborates with the following partners on activities that have HCI&IM R&D components:

- DOE/SC on high-end modeling and simulation
- The weather community, including NOAA, on global climate modeling
- FAA on automated vehicle cockpits and air traffic management and risk

### **3.4.3 AHRQ**

- NIH/NLM has joined with AHRQ in considering proposals submitted to AHRQ's Health Information Technology (HIT) program for funding.
- NIH/NLM will work with AHRQ to develop standards critical to improving patient safety, as part of the Patient Safety Task Force.
- NIST will work with AHRQ to develop a U.S. health data standards landscape model and application.
- The Food and Drug Administration (FDA) in HHS will work with AHRQ to develop data standards critical to improving patient safety in the use of prescription drugs.
- AHRQ invites other agency experts to participate in peer reviews of its grant proposals.

### 3.4.4 NIST

NIST coordinates with DARPA and NSA on the Human Language Technology Program<sup>64</sup> that involves the measurement and evaluation of speech recognition; speaker recognition; spoken language understanding; information search, retrieval, and filtering; document understanding; summarization; interactive speech interfaces; and on the Interactive Systems Technology Program<sup>65</sup> that involves measurement, standards, and evaluation for usability of interactive systems; Web usability testing; usability reporting formats; and accessibility testing.

## 3.5 HCI&IM AGENCY ACTIVITIES

### 3.5.1 NSF

The bulk of NSF's HCI&IM R&D investments are in its Directorate for Computer and Information Science and Engineering (CISE). In FY 2004, CISE was reorganized into four Divisions, including the Division of Information and Intelligent Systems (IIS), within which a large share of NSF's HCI&IM investments fall. Other HCI&IM activities lie in the NSF-wide Information Technology Research Program (ITR).

#### 3.5.1.1 CISE/Information and Intelligent Systems (IIS) Division

The goals of CISE/IIS<sup>66</sup> 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 leads to discovery in the sciences and engineering

CISE/IIS has three clusters with investments in HCI&IM R&D:

- The *Systems in Context* cluster supports 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 (HCI)
  - Universal Access (UA)
  - Digital Society and Technologies (DST)
  - Robotics
  - Digital Government (DG)

---

<sup>64</sup> <http://www.itl.nist.gov/iad/programs.html>

<sup>65</sup> <http://www.itl.nist.gov/iad/programs.html#USERS>

<sup>66</sup> [http://www.cise.nsf.gov/div/cluster.cfm?div=iis&cluster\\_id=3947](http://www.cise.nsf.gov/div/cluster.cfm?div=iis&cluster_id=3947)

- The *Data, Inference, and Understanding* cluster supports basic research with the goal of creating general-purpose systems for representing, sorting, accessing, and drawing inferences from data, information, and knowledge. This cluster integrates the following programs:
  - The Artificial Intelligence and Cognitive Science (AI&CS) program focuses on advancing the state of the art in AI&CS. It includes research fundamental to the development of computer systems capable of performing a broad variety of intelligent tasks and to the development of computational models of intelligent behavior across the spectrum of human intelligence.
  - Computer Vision
  - Human Language and Communication
  - Information and Data Management (IDM)
  - Digital Libraries
  
- The *Science and Engineering Informatics* cluster includes:
  - Collaborative Research in Computational Neuroscience that will continue as a collaborative program
  - Science and Engineering Information Integration and Informatics, which focuses on the development of information technology to solve a particular science or engineering problem and generalizes the solution to other related problems.
  - Information Integration to provide a uniform view to a multitude of heterogeneous, independently developed data sources through reconciling heterogeneous formats, Web semantics, decentralized data-sharing, data-sharing on advanced cyberinfrastructure, on-the-fly integration, and information integration resources.

### **3.5.1.2 NSF's Information Technology Research (ITR) Program**

Other NSF HCI&IM investments are in the NSF-wide ITR Program that involves all NSF Directorates and Offices including BIO, ENG, GEO, MPS, SBE, OPP, and EHR.

### **3.5.1.3 NSF Workshops**

The following FY 2003 workshops are indicative of the nature of NSF-sponsored HCI&IM-related workshops that are being held in FY 2004:

- HCI program workshops:
  - Association for Computing Machinery Computer Human Interaction (ACM CHI) '03 Doctoral Consortium held in Ft. Lauderdale, Florida, April 7-10, 2003<sup>67</sup>
  - Workshop on Dialogue Processing held in Edmonton, Canada, May 27 to June 1, 2003, provided a forum to assess the current state of the art of dialogue

---

<sup>67</sup> [http://cise.nsf.gov/events/wksp/display\\_workshops.cfm?evnt\\_id=40069&div=cise](http://cise.nsf.gov/events/wksp/display_workshops.cfm?evnt_id=40069&div=cise)

processing and identify key themes and directions that are driving research in the field.<sup>68</sup>

- An annual six-week workshop that trains graduate students and junior faculty in research techniques for computational linguistics. More than 30 graduates and undergraduates participated in the workshop.<sup>69</sup>
- IDM program sponsored workshops:
  - Joint Conference on Digital Libraries,<sup>70</sup> June 7-11, 2004, in Tucson, Arizona
  - International Workshop on Data Engineering for Wireless and Mobile Access (MobiDE '03)<sup>71</sup> held September 19, 2003 in San Diego, California. MobiDE is a bridge between the data and information management community and the mobile computing community.
  - IDM-2003 PIs Workshop held September 14-16, 2003<sup>72</sup> in Seattle, Washington
  - Workshop on Research Directions for Digital Libraries, June 15-17, 2003, in Chatham, Massachusetts. The workshop report is entitled “Knowledge Lost in Information.”<sup>73</sup>
  - Workshop on Bioinformatics held March 5-8, 2003 in Bangalor, India, provided an international forum for researchers, professionals, and industrial practitioners to share their knowledge of how to manage biological information. The workshop was held in conjunction with the 19th International Conference on Data Engineering (ICDE 2003)<sup>74</sup> to encourage database researchers to become involved in bioinformatics research.

### 3.5.2 NIH

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<sup>75</sup> 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 healthcare information technologies that may be more fully deployed in the workflow to improve efficiency and outcomes.

---

<sup>68</sup> [http://www.cise.nsf.gov/events/wksp/display\\_workshops.cfm/evnt\\_id=5054&di=cise](http://www.cise.nsf.gov/events/wksp/display_workshops.cfm/evnt_id=5054&di=cise)

<sup>69</sup> [http://www.cise.nsf.gov/events/wksp/workshops.cfm?event\\_id=40067&div=cise](http://www.cise.nsf.gov/events/wksp/workshops.cfm?event_id=40067&div=cise)

<sup>70</sup> <http://www.jcdl.org>

<sup>71</sup> <http://db.cs.pitt.edu/mobide03/>

<sup>72</sup> <http://www2.cs.washington.edu/nsf2003/home.html>

<sup>73</sup> <http://www.sis.pitt.edu/~dlwkshop/> contains information about the workshop and contains the report.

<sup>74</sup> [http://www.aztecsoft.com/\\_icde2003/](http://www.aztecsoft.com/_icde2003/)

<sup>75</sup> <http://www.becon.nih.gov/symposium2004.htm>

### 3.5.3 DARPA

DARPA's HCI&IM investments include 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.<sup>76</sup> The programs are:

- Compact Aids for Speech Translation (CAST)
- Improving Warfighter Information Intake Under Stress (formerly Augmented Cognition)<sup>77</sup>
- Translingual Information Detection, Extraction, and Summarization (TIDES)<sup>78</sup>
- Effective, Affordable, Reusable, Speech-to-Text (EARS)<sup>79</sup>

#### 3.5.3.1 *Compact Aids for Speech Translation (CAST)*

The goal of CAST is to develop rapid, two-way, natural language speech translation interfaces and platforms for the warfighter for use in field environments for force protection, refugee processing, and medical triage. The program will focus on overcoming the many technical and engineering challenges limiting current multilingual translation technology to enable future full-domain, unconstrained dialog translation in multiple environments. This technology will replace the DARPA RMS (One-Way) translator in two phases – the DARPA 1+1 Way and the DARPA Two-Way. The four major thrusts in this program are:

- Core Research
- DARPA 1+1 Way
- DARPA Two-Way
- Evaluation and Data Collection

Potential users are the U.S. Armed Forces and Intelligence Community in foreign countries. It involves the voluntary collection of foreign speech and language data and domain specific data from subject matter experts. It will improve the ability of operators to converse with, extract information from, and give instructions to a foreign language speaker encountered in the field.

#### 3.5.3.2 *Improving Warfighter Information Intake Under Stress*

The mission of this program is 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. Specifically, this program will empower one human's ability to successfully accomplish the functions currently carried out by three or more individuals. This will be done by enabling computational systems to dynamically adapt to users by developing the following:

---

<sup>76</sup> <http://www.darpa.mil/ipto/Programs.htm>

<sup>77</sup> <http://www.darpa.mil/ipto/Programs/augcog/index.htm>

<sup>78</sup> <http://www.darpa.mil/ipto/Programs/tides/index.htm>

<sup>79</sup> <http://www.darpa.mil/ipto/Programs/ears/index.htm>

- 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 to increase 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 (that is, audio, visual) to increase information throughput

The program has two technical challenges:

- Demonstrating enhancement of warfighter performance
  - Assess warfighter status in less than 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

A PI meeting was held January 6-8, 2004 in Orlando, Florida.<sup>80</sup>

### **3.5.3.3 *Translingual Information Detection, Extraction, and Summarization (TIDES)***

The TIDES program seeks to:

- Develop advanced language processing technology to enable English speakers to find and interpret critical information in multiple languages without requiring knowledge of those languages
- Develop sophisticated statistical modeling techniques for human language
- Take advantage of substantial increases in electronic data and computational power

---

<sup>80</sup> <http://www.disc-web.net/meetings/u8baerz/index.html>

- Develop and test 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 policy makers to know what is being said in a region by and to the local population
- Lessening dependence on scarce linguists
- Developing potential customers throughout the military and intelligence communities

A PI meeting was held July 13-15, 2004 in Philadelphia, Pennsylvania.

### **3.5.3.4 Effective Affordable Reusable Speech-To-Text (EARS)**

This program is developing speech-to-text (automatic transcription) technology whose output is substantially richer and more accurate than currently possible, so that machines can better detect, extract, summarize, and translate information. It will also enable humans to understand what was said by reading transcripts instead of listening to audio signals.

The EARS 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)
- Developing many potential customers throughout the military and intelligence communities

DARPA sponsored the Human Language Technology Conference and the North American Association for Computational Linguistics Annual Meeting that both were held May 2-7, 2004 in Boston, Massachusetts.<sup>81</sup> DARPA, NIST, and NSF were represented on the conference oversight committee.

---

<sup>81</sup> <http://www1.cs.columbia.edu/~pablo/hit-naac104>

### 3.5.4 NASA

NASA's HCI&IM investments include fundamental research in:

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

Within the aviation domain, HCI&IM investments include:

- Human-computer interaction for highly automated vehicle cockpits
- Aviation/shuttle flight control systems
- System health management systems
- International Space Station interfaces
  
- Human-computer interaction for air traffic management application
  - 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 include:

- Shuttle and International Space Station mission control
- Remote interactive automation for near-Earth and planetary missions

In the information management domain, NASA missions require investment 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

### 3.5.5 AHRQ

AHRQ's 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 healthcare quality (for example, promoting and accelerating the development, evaluation, adoption, and diffusion of IT in healthcare) and establishing healthcare data standards (for example, development, evaluation, and adoption of information standards and technology to support patient safety). A major goal is to support the development of a National Health Information Infrastructure (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 healthcare delivery processes.

In the area of transforming healthcare quality, AHRQ provides:

- Planning grants to rural and small communities for healthcare 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.

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” was held September 26-28, 2004 in Arlington, Virginia.<sup>82</sup> 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

---

<sup>82</sup> <http://www.ahrq.gov/qual/ptsafconf.htm>

effects analysis, and probabilistic risk assessment) a large body of medical error data and by disseminating scientific findings that are translated into 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

### **3.5.6 DOE/SC**

#### **3.5.6.1 National Collaboratories Program**

DOE/SC uses planning workshops to develop its programs, including those reported in the HCI&IM PCA. Its 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 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.

- They are focused on a problem of national scientific or engineering significance clearly related to DOE's mission and have high visibility.
- They involve geographically separated groups of personnel and/or facilities that are inherently required to collaborate or be used remotely for success of the project.
- They 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 the 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 include:

- 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
  - 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 process of multidisciplinary research. 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.

### **3.5.7 NIST**

NIST's HCI&IM programs are found in the Information Technology Laboratory's (ITL)<sup>83</sup> Information Access Division (IAD)<sup>84</sup> and in the Manufacturing Engineering Laboratory's (MEL)<sup>85</sup> Manufacturing Systems Integration Division (MSID)<sup>86</sup> and Intelligent Systems Division (ISD)<sup>87</sup>.

#### **3.5.7.1 Information Technology Laboratory (ITL)**

ITL/IAD houses several programs of relevance to HCI&IM. IAD's mission is to accelerate the development of technologies that allow intuitive and efficient access, manipulation, and exchange of complex information by facilitating the creation of measurement methods and standards. IAD's work objectives are to:

- Contribute to the advancement of these technologies
- Enable faster transition into the commercial marketplace
- Enable faster transition into sponsors' applications

IAD accomplishes these objectives through:

- Performance metrics, evaluation methods, test suites, and test data
- Prototypes and testbeds
- Workshops
- Standards and guidelines

IAD works in collaboration with industry, academia, and government.

---

<sup>83</sup> <http://www.itl.nist.gov>

<sup>84</sup> <http://www.itl.nist.gov/iaui/>

<sup>85</sup> <http://www.mel.nist.gov/what.htm>

<sup>86</sup> <http://www.mel.nist.gov/msid/>

<sup>87</sup> <http://www.isd.mel.nist.gov>

The five IAD HCI&IM programs are:

- The *Human Language Technology Program* that includes TREC, AQUAINT, TIDES/Document Understanding Conference (DUC), TIDES/Topic Detection and Tractation (TDT), EARS, Automatic Content Extraction (ACE), TIDES/Machine Translation (MT), Meeting Transcription, Speaker Recognition, and Language Recognition. For FY 2004, NIST plans to continue and expand evaluations for these programs in the following manner:
  - TREC – Continue novelty, Web, and Q&A, add robust retrieval, High Accuracy Retrieval from Documents (HARD), and Genomics (retrieval within that specific domain)
  - AQUAINT – Continue with dialogue testing and “what is” and “who is” questions
  - TIDES/DUC – Continue evaluations and add cross-language Arabic summarization
  - TIDES/TDT – Continue evaluations
  - EARS – Continue speech-to-text (STT) and metadata evaluations
  - ACE – Provide 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
  
- The *Biometrics Technology Program* that includes fingerprint and face testing and standards. For FY 2004, NIST plans to:
  - Continue testing the FBI’s Integrated Automated Fingerprint Identification System (IAFIS) on DHS and other data
  - Continue 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.)
  - Help integrate face recognition and multi-vendor verification into US VISIT
  - Carry out fingerprint vendor technology evaluation (FpVTE) testing and analysis
  - Continue data and performance standards development
  - Develop Multi-modal Biometric Accuracy Research Kiosk
  - ANSI and ISO standardization of the Common Biometric Exchange File Format (CBEFF) and the Biometric Application Programming Interface (BioAPI)
  - Continue chairing the International Committee for IT Standards (INCITS) Committee on Biometrics (M1) and the ISO Committee on Biometrics (SC37)
  
- The *Multimedia Technology Program* includes multi-media standards, video retrieval, and visualization and virtual reality for manufacturing. FY 2004 plans include:

- Continuation of work with industry to use the MPEG-7 (the metadata standard for accessing video information) Interoperability Test Bed
  - Development of “light-and-simple” profiles for existing industry-wide formats
  - For video retrieval, continue evaluations with a new story segmentation task and more high-level features to extract; process and ship 120 hours of new data
  - Continue X3D and Human Animation (H-ANIM) standardization
  - Continue seeking medical applications within Web3D and for human modeling
- The *Interactive Systems Technology Program* includes WebMetrics, Industry Usability Reporting (IUSR — a standard by which software development companies can test results and provide results to potential customers), Voting, Novel Intelligence for Massive Data (NIMD), Human Robot Interaction (HRI), Digital Library of Mathematical Functions (DLMF), and Accessibility Standards. FY 2004 plans for this program include:
    - Webmetrics – seek additional opportunities with W3C<sup>88</sup>
    - IUSR – extend Common Industry Format (CIF) to hardware and finalize CIF-requirements
    - Voting – R&D for developing usability and accessibility tests for certifying voting systems
    - NIMD – experiment with glass box data and develop NIMD roadmap
    - HRI – develop a user interface for robotic mobility platform
    - DLMF – usability study for DLMF Web site and animation of Virtual Reality Modeling Language (VRML) worlds
    - Accessibility standards – develop a prototype incorporating the new architecture from the INCITS Committee on Accessibility (V2)
  - The *Smart Space Technology Program* includes a smart space testbed and single molecule manipulation and measurement. The testbed is for human interaction in an environment that includes microphones and video cameras. FY 2004 plans include:
    - A 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
    - Continue upgrading Meeting Room with new capabilities. Originally done for speech, Meeting Room is now being used to analyze videos.
    - Continue INCITS V2 collaboration
    - Provide data acquisition services and analysis methods to NIST’s Chemical Science and Technology Lab and Physics Lab and upgrade real-time processing systems

---

<sup>88</sup> <http://www.w3.org/>

### **3.5.7.2 Manufacturing Engineering Laboratory (MEL)**

MEL's mission is to satisfy the measurements and standards needs of U.S. manufacturers in mechanical and dimensional metrology and in advanced manufacturing technology by conducting R&D, providing services, and participating in standards activities.

Most of MEL's HCI&IM activities fall within its Manufacturing Systems Integration Division (MSID). MSID promotes the development of technologies and standards that lead to the 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. MSID has an HCI&IM effort in system interoperability that is 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 integrating 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 and the distribution of higher level information between factories and front offices. MSID hosted an NIST/NSF International Workshop on Semantic Distance, November 10-12, 2003 in Gaithersburg, Maryland.

MEL's Intelligent Systems Division (ISD) develops test arenas for competitions, particularly for urban search and rescue (USAR) applications. The ISD developed the Reference Test Arenas for Autonomous Mobile Robots to focus research efforts, provide direction, and accelerate the advancement of mobile robot capabilities. These arenas, modeled from buildings in various stages of collapse, allow objective performance evaluation of robots as they perform a variety of USAR tasks in both physical and virtual domains. The USAR test arenas have inspired not only replicas to be fabricated, but have provided the foundation for the annual Rescue Robot Competitions that are currently part of the RoboCup – USOpen, RoboCup – Japan Open, and the International Joint Conference on Artificial Intelligence (IJCAI)/American Association for Artificial Intelligence (AAAI) conferences. The first of these competitions was done in cooperation with the AAAI at its 2000 meeting in Austin. Competitions for 2004 are scheduled for New Orleans, Osaka, Lisbon, and San Jose. Testing is available on an individual basis at NIST in Gaithersburg, and year-round arenas are also now available in Tokyo, Pittsburgh, and Rome.<sup>89</sup>

Since FY 2000, ISD also holds annual workshops on Performance Metrics for Intelligent Systems (PerMIS),<sup>90</sup> with the goal of defining measures and methodologies of evaluating performance, inter-comparing research results, and defining measurable program objectives of intelligent systems and sub-components of systems. NIST held a PerMIS workshop<sup>91</sup> in cooperation with DARPA on August 24-26, 2004 that included a session on merging of ontologies that is likely to develop into a larger conference with the cooperation of MSID and ITL.

---

<sup>89</sup> <http://www.isd.mel.nist.gov/projects/USAR/>

<sup>90</sup> [http://www.isd.mel.nist.gov/research\\_areas/research\\_engineering/Performance\\_Metrics/](http://www.isd.mel.nist.gov/research_areas/research_engineering/Performance_Metrics/)

<sup>91</sup> [http://www.isd.mel.nist.gov/PerMIS\\_2004/](http://www.isd.mel.nist.gov/PerMIS_2004/)

### **3.5.7.3 NIST Workshops**

NIST has organized or is organizing the following HCI&IM-related workshops (some are conducted with other agencies):

- PerMIS Workshop in cooperation with DARPA – August 24-26, 2004
- Biometric Consortium Conference – September 2004
- Language Identification Workshop – Spring 2004
- Speaker Identification Workshop – Spring 2004
- Meeting Recognition Technologies – May 2004
- International Biometrics Technical Meeting – March 2004
- Cognitive Task Analysis for Intelligence Analysts – January 21-23, 2004
- Voting Systems Symposium – December 2003
- International Workshop on Semantic Distance – November 10-12, 2003
- Video Retrieval Conference – November 2003
- Text Retrieval Conference – November 2003
- Automatic Content Extraction Workshop – October 2003

### **3.5.8 NOAA**

In digital libraries, NOAA<sup>92</sup> is developing methods of cataloging, searching, viewing, and retrieving NOAA data distributed across the Web. It is exploiting Geographical Information Systems (GIS) and XML technologies to display and describe data and is 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
- Plans include developing access to current National Weather Service (NWS)<sup>93</sup> model data subsets.

NOAA held its semi-annual PI status reports meeting in the spring of 2004.

### **3.5.9 EPA**

EPA 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

---

<sup>92</sup> <http://www.cio.noaa.gov/hpcc/>

<sup>93</sup> <http://www.nws.noaa.gov/>

decision process and convey important information in such a way that researchers and policy makers can better understand complex scientific data. These programs are performed in-house and as problem-driven research. Current 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

EPA 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 is underway. Metabonomics data will be harmonized with both proteomics and genomics data in order to better understand and predict chemical toxicity and support risk assessment.
- Air Quality Model Pilot – Tools and approaches to explore potential linkages between air quality and human health are being investigated. Assessments include impacts of regulatory and policy decisions. Exploration of relationships between regulations, emission sources, pollutants, and ecosystems is being completed. 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 the States' abilities to assess the impact of regulatory and policy decisions.



## 4. LARGE SCALE NETWORKING (LSN)

### 4.1 INTRODUCTION

The agencies that participate in the LSN CG plan, coordinate, and evaluate Federal R&D in advanced networking technologies. The LSN CG comprises expert volunteers from nine Federal agencies who meet monthly to exchange information and coordinate their LSN R&D programs and activities. Eight of these agencies are in the NITRD Program: DOE/SC, DARPA, NASA, NIST, NIH/NLM, NSA, NSF, and NOAA; the ninth is DoD/HPCMPO.

Three Teams report to the LSN CG. In addition to interagency coordination in specific areas, the Teams implement tasks assigned by the LSN CG. These Teams are:

- *Joint Engineering Team (JET)*, which 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*, which coordinates Federal agency Grid and middleware R&D and supports Grid applications
- *Networking Research Team (NRT)*, which coordinates Federal agency research programs in advanced networking technologies and protocols

### 4.2 DEFINITION OF THE 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 that 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 include:

- 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

Other areas of LSN concern are:

- International cooperation in the areas of optical network coordination and certificate authorities
- Outreach and cooperation with university and commercial communities
- Network science education and workforce issues

To address these concerns, LSN R&D pursues the following technical goals:

- End-to-end performance measurement including measurement 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 of LSN programs are:

- 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

### **4.3 LSN CG ACTIVITIES**

The LSN CG holds an Interagency Coordination Meeting each year 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 outside (or larger than) the stated scope of the LSN CG, such as cybersecurity

The LSN held its annual Interagency Coordination Meeting on November 3, 2003, with presentations and program descriptions from nine Federal agencies and several divisions within

NSF. The LSN Teams also presented their accomplishments and plans. This meeting served as the LSN CG's Special Meeting. Discussion among the meeting participants identified the following three overarching areas where most of the LSN agencies had significant efforts and interests that would benefit from cooperation:

- Network testbed infrastructure
- Network performance measurement
- Grid outreach

These areas are now described.

#### **4.3.1 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)

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 their recommendations.

#### **4.3.2 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.

### **4.3.3 Grid Outreach**

Many user and application communities are developing Grid capabilities, often focused on their particular needs. The LSN CG agencies are fostering the development of standardized Grid middleware resources, tools, architectures, and infrastructure to facilitate Grid applications. The LSN CG urged the LSN agencies to provide 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 those that are being developed
- Avoid balkanizing Grid resources by relying on standardized tools, middleware, and architecture rather than unique discipline-specific Grid capabilities
- Benefit from common policies, resource infrastructure, and the Global Grid Forum GGF.<sup>94</sup> 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.

### **4.3.4 Additional Topics for LSN and IWG Coordination**

The LSN CG discussion identified three additional topics of interest across several of the CGs that provide opportunities for future collaboration and coordination by the LSN CG and with other CGs. They are:

- Autonomic networking
- High speed transport
- Security

## **4.4 LSN MULTI-AGENCY ACTIVITIES**

### **4.4.1 Common Interests**

The LSN agencies identified a commonality of commitment and interest across a wide range of network research topics, summarized as follows:

---

<sup>94</sup> <http://www.gridforum.org/>

- LSN Agency Networking Programs and Interests
  - Networking Research
    - Basic technologies
    - 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

#### **4.4.2 Coordination of Network Performance Measurement**

The LSN agencies benefit from a strong initiative to coordinate end-to-end performance measurement across the Federal research networks (JETnets, including DOE/SC's ESnet, DOE/SC's UltraScience Net, NASA's NREN, NSF's EINs, DoD's DREN, Abilene, several RONS, and National LambdaRail.) (The complete list of JETnets appears in section 4.5.1.) They are developing a common set of performance measurement tools and capabilities, developing agreements to share performance data among the LSN networking realms, and opening their networks for other LSN networks to measure cross-border performance.

#### **4.4.3 Optical Networking Workshop**

At its November 3, 2004 meeting, the LSN CG identified a need to hold a workshop on optical networking. As a consequence, the NASA Research and Education Network (NREN) held a workshop on Optical Networking Testbeds August 9-11, 2004 at NASA Ames Research Center, Moffett Field, California. NREN worked with the LSN agencies to define the objectives, sessions, and key participants. LSN agency representatives participated in the Workshop Organizing Committee to assure that each agency's areas of interest were addressed.

#### **4.4.4 Coordination of Meetings**

At LSN CG meetings, the LSN agencies inform each other of upcoming agency meetings and workshops so that program managers from all LSN agencies can participate in and benefit from them.

#### **4.4.4.1 DREN Workshop on Network Security**

DREN held an interagency workshop on network security in early August, 2004. At this workshop, JET members provided information on their activities, tools, and implementation architectures for network security. The workshop enabled coordination between the Federal research network community and the larger community of Federal networks.

#### **4.4.4.2 Cooperation on Principal Investigator Meetings**

Individual LSN agencies have adopted a policy of opening Principal Investigator meetings to all other LSN agencies. The researchers benefit from coordinating research plans and results and Federal networking research program managers benefit from a broad view of all the Federal networking research programs.

### **4.5 LSN TEAM ACTIVITIES**

LSN Teams — JET, MAGIC, and NRT — coordinate activities in focused areas of interest with researchers, implementers, and commercial sector participants involved in networking research. The Teams also respond to specific taskings from the LSN CG. The accomplishments, plans, and coordination of the LSN Teams contribute strongly to the objectives of the LSN CG. The LSN Team activities and plans were presented at the November 3, 2003 LSN Interagency Coordination Meeting and are now described.

#### **4.5.1 Joint Engineering Team (JET)**

The JET provides for cooperation among the Federal agencies that maintain JETnets and those having an interest in coordinating with or using the JETnets. 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 (DIA), NASA, NRL, 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 network peering point in Chicago
- UCAID's Abilene

JET participants include representatives from:

- LSN agencies that operate JETnets: DoD/HPCMPO, DOE/SC, NASA, and NSF
- Other JETnets: Abilene, NLR, and StarLight

- Other LSN agencies – NIH/NLM, NIST, and NSA
- Other interested agencies – USGS
- Other organizations interested in research networking: ANL, ORNL, Indiana University, University of Illinois-Chicago, Cisco Systems, Juniper Networks, and Qwest

JET is focused on:

- Coordinating interconnections among the JETnets
- Providing a forum for collaboration on planning, architecture, implementation, operations, and maintenance among the JETnets
- Encouraging technical consensus and sharing of resources
  - Promoting interoperability and transparency among the JETnets
  - Fostering best practices for advanced networking implementations and infrastructure
  - Enhancing performance to the network edge and end users
  - Encouraging the deployment of interoperable advanced services such IPv6 and Ipv4 and IPv6 multicast
- Serving as a focal point to facilitate operations and resolve emerging problems and technical issues
- Supporting outreach and applications demonstrations to the research network user community (that is, acting as a communication vehicle for the technical community and providing a Federal voice for effective coordination with international research and engineering networks)

Current JET activities include:

- Improving access through Federal peering points and StarLight
- Coordinating and promoting traffic monitoring and performance measurement
- Developing best practices. An example is 9000 Byte Message Text Unit (MTU) size. Increasing the packet size allows transmittal of large amounts of information using fewer packets and results in higher reliability.
- New technology deployment to support the 1 Gigabit Ethernet (1GigE) protocol, 10 GigE, and all-optical networking
- Fostering cooperation on network security best practices
- Coordination with international research networks on connectivity and transparency and cooperation on developing Grid and security standards
- Coordination with MAGIC on network support for Grid applications

Looking forward, the JET will focus on:

- Cybersecurity at line speed
- Firewalls
- Dark fiber deployments and JETnet connectivity to RONS and the NLR
- 10 Gigabits per second (Gbps) Network Interface Cards (NICs)

#### **4.5.1.1 JET Roadmap Workshop**

The JET held an interagency workshop in April, 2004 at the Thomas Jefferson National Accelerator Facility (JLAB) in Newport News, Virginia, to develop a three-year roadmap for JET activities including coordination of performance measurement, coordination of Federal research network peering points, engineering and coordinating LSN network infrastructure, and coordinating with regional, national-scale, and international research networks.

#### **4.5.1.2 JETnets Annual Reviews**

Individual JETnets hold annual reviews of their agency-specific programs, plans, and accomplishments. Representatives from other JETnets are invited to participate in these meetings to promote openness, awareness, and opportunities for collaborating and coordinating. For example, several times a year the DOE/SC ESnet holds Joint Techs meetings in coordination with Internet2 meetings. At these meetings the ESnet Steering Committee (ESSC) invites other Federal networks and JET members to participate in their review of ESnet activities.

#### **4.5.2 Middleware and Grid Infrastructure Coordination (MAGIC) Team**

MAGIC provides coordination among Federal agencies with responsibility for middleware and Grid projects, individuals involved in middleware and Grid research, individuals involved in implementing or operating Grids, and users of Grids and middleware. MAGIC participants include representatives from the Federal agencies, the commercial sector, academia, and Federally-funded laboratories. Federal agency and Federal lab participants include ANL, DOE/SC, LANL, LBL, NASA, NIH, NIST, NOAA, NSF, PNNL, and UCAR. Non-Federal participants include The Boeing Company, Cisco Systems, Educause, Hewlett-Packard (HP), IBM, Internet2, Information Sciences Institute (ISI), Level3 Communications, Microsoft, the University of Chicago, the University of Illinois at Urbana-Champaign, and the University of Wisconsin.

MAGIC's responsibilities are to:

- Coordinate interagency middleware and Grid efforts
- Enhance and encourage interoperable Grid technologies and deployments
- Promote usable, widely deployable middleware tools and services
- Provide a forum for effective international coordination of these technologies

MAGIC focuses on Grid middleware and enabling Grid users. Topics of recent concern include:

- Trust models such as implementing policies and architectures for authentication, authorization, accounting, and revocation of authorization. Of particular concern is implementing trust across institutional and international boundaries and models that scale to large numbers of users.
- Intellectual property and open source software
- Program and project updates

MAGIC plans for the coming year are to:

- Collaborate with the larger Grid community, for example, commercial and international sectors
- Promote progress on key technical and policy issues
- Provide outreach to user communities
- Track Grid R&D progress and promote its visibility

### **4.5.3 Networking Research Team (NRT)**

NRT promotes collaboration and exchange of information on networking research. Specifically, its goals are to:

- Coordinate network R&D efforts across member agencies through joint participation in review panels and program reviews
- Conduct interagency workshops to identify R&D gaps in combined portfolios of member agencies
- Coordinate interagency PI meetings to enable researchers to leverage R&D efforts of member agencies
- Carry out tasks assigned by LSN

Members are DARPA, DOE/SC, NASA, NIH/NLM, NIST, and NSF. NRT plans to open participation to the university and commercial sectors to provide broader research and user participation.

NRT held a workshop on Research Needs for Large Scale Networks on May 13-14, 2003 in Lansdowne, Virginia, to address the following fundamental network security concerns:

- To what extent do non-technical obstacles such as economic factors, policy issues, and man-machine interfaces contribute to network security problems in large-scale networks?
- Is there solid evidence that today's large-scale networks are too insecure to serve their intended purposes?
- To what extent are persistent security problems the result of deployment obstacles?
- Would effective cybersecurity technology, had it been adopted, have solved the problems?

The workshop concluded that:

- Definitions, metrics, standards, and models of security are needed.
- Mathematical methods and theoretical research are needed to understand real-world security.
- Experimental research is needed.

In FY 2004, NRT will develop mechanisms for disseminating NRT information such as by using the NRT Web site.<sup>95</sup>

NSF's EINs and DOE/SC's UltraScience Net will provide high performance networks to support networking research.

Other key NRT topics are:

- Survivability of the Internet: A critical infrastructure assessment
- Building on network R&D of other agencies such as on open source issues
- Outreach to other groups with interests in networking research including JET, MAGIC, Canarie2 (Canadian research network), and EU research networks

## **4.6 LSN AGENCY ACTIVITIES**

Each LSN agency focuses on network research programs and activities that address the mission needs of the agency while cooperating with other LSN agencies to meet these needs and to expand their awareness and involvement in other networking areas in their interest. This section reviews the programs and activities of each LSN agency.

### **4.6.1 NIH/NLM**

The NIH/NLM program on Healthcare and the NGI funds testbeds that demonstrate next generation Internet capabilities by the healthcare community including:

- Quality of Service (QoS)
- Security and medical data privacy
- Nomadic computing
- Network management
- Infrastructure technology as a means for collaboration

In FY 2003, NLM initiated a program in Applications of Advanced Network Infrastructure Technology in Health and Disaster Management that 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 such as biometrics or smartcards
- Are nomadic or use GIS technology

The NLM Scalable Information Infrastructure Awards<sup>96</sup> support 11 testbed projects:

---

<sup>95</sup> <http://www.nitrd.gov/iwg/lsn/nrt>

<sup>96</sup> These programs are described at <http://www.nlm.nih.gov/research/siiawards.html>

- 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

## **4.6.2 NSF**

### **4.6.2.1 CISE Directorate**

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, fine-grained to episodic network services using optical, fiber, wireless, satellite, and hybrid network links.

FY 2003 awards are:

- The Networking Research Testbed (NRT) program that 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

Ten NRT program 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 is a complementary program to NRT. Some EIN characteristics 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

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, now called Circuit-switched High-speed End-to-End Transport Architecture (CHEETAH)
- Cyber defense technology experimental research (DETER) network
- PlanetLab: An overlay testbed for disruptive network services
- WAN-in-Lab
- NSF has a continuing NSF Middleware Initiative (NMI)<sup>97</sup> program. The 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, sometimes called the "large" 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. NMI awards are:
  - Disseminating and supporting middleware infrastructure: Engaging and expanding scientific Grid communities
  - Designing and building a national middleware infrastructure (NMI-2)
  - An integrative testing framework for Grid middleware and Grid environments
  - Extending integrated middleware to collaborative environments in research and education
  - Instruments and sensors as network services: Instruments as first class members of the Grid
  - Middleware for Grid portal development

---

<sup>97</sup> <http://www.nsf-middleware.org/>

- FY 2003 theme areas for the NSF Strategic Technologies for the Internet (STI)<sup>98</sup> program 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
  - Innovative access network technologies

The FY 2003 STI awards are for:

- A security architecture for IP telephony
- Network Startup Resource Center (NRSC)
- 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 future transport protocol for high performance network environments
- Media-aware congestion control
- Self-organizing spectrum allocation

#### **4.6.2.2 NSF International Connections**

NSF funds several international initiatives:

- The 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.
- The 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 the Network Access Point (NAP) of the Americas

---

<sup>98</sup> [http://www.cise.nsf.gov/funding/pgm\\_display.cfm?pub\\_id=5381&div=sci](http://www.cise.nsf.gov/funding/pgm_display.cfm?pub_id=5381&div=sci)

- The TransLight program supports additional optical networking connectivity to Canada (Canarie2 network), Prague, Stockholm, and London.

#### **4.6.2.3 NSF Information Technology Research (ITR) Program**

The NSF ITR Program<sup>99</sup> 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 ITR 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

#### **4.6.2.4 Extensible Terascale Facility (ETF) Connectivity**

Networks connect the 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.

### **4.6.3 DOE/SC**

#### **4.6.3.1 Application Drivers for the High-Performance Network Environment**

The needs of the DOE/SC scientific research community are growing:

- 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.

These changes result in increasing demands being placed on DOE/SC's network. For example, five years out, DOE/SC scientists are expected to need the network to:

- Transfer three petabytes per year

---

<sup>99</sup> <http://www.nsf.gov/home/crssprgm/itr/>

- 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 one to ten 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 its 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 and with other LSN agencies. Other DOE/SC priorities are program integration and priority setting for its facilities and research.<sup>100</sup>

#### **4.6.3.2 UltraScience Net**

The UltraScience Net<sup>101</sup> is a new DOE/SC 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 to StarLight, was put in place in early 2004. Linking of Sunnyvale (California), CERN, and LBNL is being done cooperatively with NLR. This network will support end-to-end scientific discovery. Towards that end, DOE/SC is supporting:

- Networking research
- Middleware research
- Connectivity with a high-performance production network
- Collaboratory pilots and testbeds

UltraScience Net technologies will be transitioned to a DOE/SC R&E, high-impact science network as the technology is shown to meet requirements. This R&E prototype will be developed in coordination with other LSN agencies. The prototype will support very high throughput data streams and network measurement and diagnosis.

#### **4.6.3.3 Networking Research**

DOE/SC networking research programs address 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. DOE/SC is also addressing current protocol limitations for transporting large data sets. New protocols are being developed to support large data transfers. Cybersecurity will be

---

<sup>100</sup> <http://www.sc.doe.gov/feature/ASCR.htm>

<sup>101</sup> <http://www.csm.ornl.gov/ultranet/>

needed at OC-12 to OC-192 speeds, requiring policy and technology advances. Research is also being carried out in QoS and dynamic provisioning services.

#### **4.6.3.4 Middleware Research**

DOE/SC middleware research is focused 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
- Facilitating the discovery and use of scientific data, computers, software, and instruments over the network in a controlled fashion

#### **4.6.3.5 ESnet**

DOE/SC 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 R&E prototype described in section 4.6.3.2 to a production environment and to tighten the coupling of the prototype and ESnet.

#### **4.6.3.6 Connectivity**

DOE/SC connectivity programs are developing IPv6 protocol implementation, multiplatform video conferencing services, and implementation of distributed network measurement and analysis tools. DOE/SC is coordinating with other LSN agencies and with the university and commercial sectors through:

- Inviting LSN agencies to ESnet Steering Committee meetings
- Tactical coordination with program managers in other LSN agencies with common interests, particularly in the NSF Cyberinfrastructure program

### **4.6.4 NASA**

#### **4.6.4.1 Adaptive Networking**

The NASA Research and Education Network (NREN) adaptive networking program is developing a tool called Resource Allocation, Measurement and Adjustment System (RAMAS) to reserve bandwidth for applications, either when requested or later. RAMAS has the following features:

- Passive monitoring and measurement
- Data capture of OC-3 and OC-12 data streams
- GUI-enabled 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, Asynchronous Transfer Mode (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. It identified the difficulty of optimizing TCP for satellite transmissions.

In FY 2004 RAMAS is developing:

- Upgrades to OC-48 and O-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
- Identification of performance bottlenecks for selected wireless applications

#### **4.6.4.2 Nomadic Networking: Portable Satellite Dishes**

NASA's program in nomadic networking is designed to enable NASA science and engineering in remote locations. It has three components: *ad hoc* networking, hybrid satellite and terrestrial networking, and mobile IP/IPv6. Planned FY 2004 activities include:

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

A Mobile Agents Field Experiment is 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.

NASA is also supporting a "ground-truthing" experiment to calibrate satellite-derived geology data with in-situ data taken in real time by on-site geologists. It uses wireless networking and 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.

#### **4.6.4.3 NASA Wide Area Network (WAN) Testbed**

NREN has a WAN testbed that connects five NASA Centers and two exchange points. It uses ATM and POS OC-12 and OC-3 circuits. Satellite connectivity is available at the NASA Glenn Research Center. The purpose of the testbed is to conduct research and develop 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 Transportable Earth Station (TES) platform.

#### **4.6.5 DARPA**

DARPA's Information Processing Technology Office (IPTO)<sup>102</sup> has a program in Network Modeling and Simulation. The program supports projects to:

- Develop modeling and simulation tools for on-line 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

One recent accomplishment is 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 used to design better and faster protocols. This work was co-funded with NSF.

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

DARPA held a Network Modeling and Simulation (NMS) PI meeting in January 2004 in Miami, Florida in coordination with other LSN agencies.

DARPA/IPTO FY 2004 plans also include work in Cognitive Networking and Self-Aware Collective Systems:

- *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. This research will focus on self-diagnosis, automatic adaptation to changing and hostile environments, reconfiguration in response to changes in the

---

<sup>102</sup> <http://www.darpa.mil/ipto/>

environment, intelligent negotiation for tasks, and resources and robustness under attack. It 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.

- The *Self-Aware Collective Systems* technology thrust will enable heterogeneous teams of individuals (for example, people, software agents, robots) and/or organizations (for example, coalition forces) to rapidly form and easily manage and maintain virtual alliances concerned with a specific task. This thrust comprises two efforts: Self-Aware Peer-to-Peer Systems and Collective Cognitive Information Processing for Improved Asset Performance:
  - *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 (for example, vehicles or other equipment).

## **4.6.6 DOE/NNSA**

### **4.6.6.1 Distance Computing**

The Distance Computing (DisCom) program is NNSA’s primary component of the LSN PCA. DisCom is an element 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 developments in distance computing. Distance Computing will extend the environments required to support high-end computing to remote sites. DisCom 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, BlueGene/L, and beyond. DisCom will support tools development for more reliable and persistent file-transfer mechanisms over the WAN for FY 2005 deployment.

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

#### **4.6.7 NIST**

NIST supports network-related programs in its Advanced Network Technologies Division (ANTD) and its Computer Security Division (CSD) within 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's goal is to provide the networking industry with the best test and measurement research. It supports modeling and analysis, testing, developing methods and tools, and prototypes and empirical studies. 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* including 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 healthcare industry
- *Infrastructure Protection* for securing core services, secure BGP,<sup>103</sup> 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.

##### **4.6.7.1 Resilient Agile Networking**

The Resilient Agile Networking (RAiN) program is developing, testing, and standardizing technologies for fault-tolerant, self-adaptive, and ad-hoc networks. A key challenge is to detect network failures and quickly adapt. RAiN projects include:

---

<sup>103</sup> BGP is a protocol for exchanging routing information between gateway hosts (each with its own router) in a network of autonomous systems. It generally is the protocol used between gateway hosts on the Internet.

- *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. NIST is developing an architecture for Forwarding and Control Element Separation (ForCES) to improve survivability of the Internet infrastructure.
- *Large Scale BGP Attack Modeling* to characterize the performance impact and benefits of various proposed approaches to secure BGP. NIST is modeling, analyzing, and providing visualization tools for global BGP data. This will enable analysis of alternative approaches for BGP router security. It contributes to characterizing BGP mechanisms and vulnerabilities and to developing standards.
- *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

#### **4.6.7.2 Self-Managing Systems**

NIST is conducting research and developing the test and measurement basis and standards foundations for future self-managing systems. This work is being performed in cooperation with DARPA, the Global Grid Forum, Distributed Management Task Force (DMTF), IETF, the Institute for Electrical and Electronic Engineers (IEEE), IBM, Sun Microsystems, Cisco Systems, 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 the industry in evaluating the behavior and performance of emerging first generation self-managing systems

#### **4.6.7.3 Information Security**

NIST programs in information security are focusing on cryptographic standards and applications:

- Secure encryption, authentication, non-repudiation, key establishment, and pseudo-random number generation algorithms
- Standards and guidance for e-Gov and e-Authentication
- PKI and DNS security standards, interoperability, assurance, and scalability
- Wireless and mobile devices and smart card security
- Quantum computing and quantum cryptography (in coordination with DARPA)

#### **4.6.7.4 NIST GLASS: GMPLS/Optical Simulation Tool**

NIST is developing a modeling tool to evaluate architectures and protocols for routing, signaling, and managing of GMPLS for optical networks and to support multi-level, multi-protocol schemes for traffic engineering, QoS, protection, and restoration. NIST is developing a modular simulation framework to support:

- Optical network restoration

- MPLS DiffServ protocol recovery
- GMPLS and optical Common Control and Measurement Plane (CCAMP)

NIST is also developing a protection and restoration tool to provide multi-level recovery schemes.

#### **4.6.7.5 NIST Priority Taskings**

NIST is responding to tasking to evaluate the economic consequences of IPv6. At the request of the White House, NIST organized a workshop on Spam in January 2004.

#### **4.6.8 NOAA**

NOAA's networking and advanced IT strategy is 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 that strategy through activities such as:

- *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

NOAA projects in computer and network security address:

- Encrypted file systems
- Firewalls and intrusion detection at Gigabit speeds
- Automated enterprise-wide patching
- Wireless security

The goal of NOAA's Next Generation Internet program is to use advanced networking technologies to enhance NOAA data collection and dissemination and to support the NOAA HPC architecture. To that end NOAA provides connectivity to seven NOAA sites and to radar sites providing near-real-time weather data. It is developing phased-array radars (NEXRAD) to collect near-real-time weather data. Near-term deployment of these radars will significantly increase data transport requirements.

NOAA is developing discipline-specific user toolsets to support collaboration and Grid applications. It recently demonstrated an advanced Intrusion Detection System (IDS). It plans to explore the use of Grid systems for data handling, computation, and collaboration. It will test deployment of IPv6 at three NOAA sites. NOAA is considering the impact of deploying optical networking on NEXRAD and to meet other NOAA needs.

NOAA is planning to implement additional applications as they become practical, including:

- Collaboratories
- Grid computing
- Storm-scale simulations with immersive environments
- Wireless: Data anywhere, anytime
- Remote operations

## **4.6.9 NSA**

### **4.6.9.1 Laboratory for Telecommunications Sciences (LTS)**

LTS programs continue to emphasize transmission of quantum communications through optical elements. A quantum channel on the same fiber as a SONET channel was demonstrated. This work is joint with DOE/SC.

In FY 2003, NSA demonstrated techniques for regionalized QoS management and developed a model for QoS pricing of Internet services. In FY 2004, NSA plans to join these two efforts to provide a viable pricing model and requisite enforcement mechanisms.

LTS also continued work on critical infrastructure protection issues for converged networks.

## **4.6.10 DoD/HPCMPO**

DoD/HPCMPO's Defense Research and Engineering Network (DREN)<sup>104</sup> maintains a production WAN to support DoD's HEC systems environment. DREN uses MCI to provide DS-3 through OC-768 links and to deploy GigE and 10 GigE services. This network supports IP, IPv6, and multicast with increased security. It supports both Continental United States (CONUS) and Outside the Continental United States (OCONUS) sites with OC-3 and OC-12 peering to OCONUS. MCI expanded the core DREN network to OC-192 in June 2004 and will expand to OC-768 when required. Services it provides include:

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

---

<sup>104</sup> <http://www.hpcmo.hpc.mil/Htdocs/DREN/dren-def.html>

DREN supports scientists and engineering technologists performing networking research throughout the United States. It connects to DoD's ATDnet, StarLight, and other experimental networks to support development of new technologies and services. As new technologies and services are developed and mature, they are transitioned to the operational DREN network.

DREN provides connectivity to remote sites including Alaska with OC-3 connectivity and Hawaii with Synchronous Transport Mode1 (STM1) services. DREN is coordinating with NASA on connectivity to Alaska and Hawaii.

DREN has had an IPv6 pilot program for several years and now deploys IPv6 native. It supports the Abilene IPv6 Transit Access Point (6TAP) testbed and the IPv6/MPLS end-to-end testbed that will provide users with secure, reliable, global data access by deploying:

- The latest available IPv6 and MPLS services
- Efficient routing using protocols such as BGP+, GMPLS, and Open Shortest Path First (OSPF)
- A range of QoS services
  - Priority, pre-emption, policy, authorization, and audit services
  - Voice, video, and data services support
- Dynamic network monitoring
- Scalable speeds
  - Network test and measurement capabilities
  - Passive monitoring with time-stamping

The DREN performance measurement program is now testing 10 GigE equipment. Supported by the measurement program, DREN is now able to achieve near wire-speed performance over 80 percent of the time across its OC-12 links.

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







































































































































































National Coordination Office for  
Information Technology Research and Development

Suite II-405  
4201 Wilson Blvd.  
Arlington, VA 22230  
(703) 292-4873

Email address: [nco@nitrd.gov](mailto:nco@nitrd.gov)

Web address: <http://www.nitrd.gov>