H.R. 4218, HIGH-PERFORMANCE COMPUTING REVITALIZATION ACT OF 2004

HEARING

BEFORE THE

COMMITTEE ON SCIENCE
HOUSE OF REPRESENTATIVES

ONE HUNDRED EIGHTH CONGRESS

SECOND SESSION

MAY 13, 2004

Serial No. 108–55

Printed for the use of the Committee on Science

Available via the World Wide Web: http://www.house.gov/science

U.S. GOVERNMENT PRINTING OFFICE

93–214PS WASHINGTON : 2004

For sale by the Superintendent of Documents, U.S. Government Printing Office
Internet: bookstore.gpo.gov Phone: toll free (866) 512–1800; DC area (202) 512–1800
Fax: (202) 512–2250 Mail: Stop SSOP, Washington, DC 20402–0001

COMMITTEE ON SCIENCE

HON. SHERWOOD L. BOEHLERT, New York, *Chairman*

RALPH M. HALL, Texas
LAMAR S. SMITH, Texas
CURT WELDON, Pennsylvania
DANA ROHRABACHER, California
KEN CALVERT, California
NICK SMITH, Michigan
ROSCOE G. BARTLETT, Maryland
VERNON J. EHLERS, Michigan
GIL GUTKNECHT, Minnesota
GEORGE R. NETHERCUTT, JR., Washington
FRANK D. LUCAS, Oklahoma
JUDY BIGGERT, Illinois
WAYNE T. GILCHREST, Maryland
W. TODD AKIN, Missouri
TIMOTHY V. JOHNSON, Illinois
MELISSA A. HART, Pennsylvania
J. RANDY FORBES, Virginia
PHIL GINGREY, Georgia
ROB BISHOP, Utah
MICHAEL C. BURGESS, Texas
JO BONNER, Alabama
TOM FEENEY, Florida
RANDY NEUGEBAUER, Texas
VACANCY

BART GORDON, Tennessee
JERRY F. COSTELLO, Illinois
EDDIE BERNICE JOHNSON, Texas
LYNN C. WOOLSEY, California
NICK LAMPSON, Texas
JOHN B. LARSON, Connecticut
MARK UDALL, Colorado
DAVID WU, Oregon
MICHAEL M. HONDA, California
BRAD MILLER, North Carolina
LINCOLN DAVIS, Tennessee
SHEILA JACKSON LEE, Texas
ZOE LOFGREN, California
BRAD SHERMAN, California
BRIAN BAIRD, Washington
DENNIS MOORE, Kansas
ANTHONY D. WEINER, New York
JIM MATHESON, Utah
DENNIS A. CARDOZA, California
VACANCY
VACANCY
VACANCY

CONTENTS

May 13, 2004

H.R. 4218, HIGH-PERFORMANCE COMPUTING REVITALIZATION ACT OF 2004

THURSDAY, MAY 13, 2004

HOUSE OF REPRESENTATIVES,
COMMITTEE ON SCIENCE,
Washington, DC.

The Committee met, pursuant to call, at 10:30 a.m., in Room 2318 of the Rayburn House Office Building, Hon. Sherwood L. Boehlert [Chairman of the Committee] presiding.

COMMITTEE ON SCIENCE
U.S. HOUSE OF REPRESENTATIVES

H.R. 4218, the High-Performance Computing Revitalization Act of 2004

Thursday, May 13, 2004
10:30 a.m.-12:30 p.m.
2318 Rayburn House Office Building (WEBCAST)

Witness List

Dr. John H. Marburger III
Director
White House Office of Science and Technology Policy

Dr. Irving Wladawsky-Berger
Vice President for Technology and Strategy
IBM Corporation

Dr. Rick Stevens
Director, Mathematics and Computer Science Division
Argonne National Laboratory

Dr. Daniel Reed
William R. Kenan, Jr. Eminent Professor
University of North Carolina at Chapel Hill

Section 210 of the Congressional Accountability Act of 1995 applies the rights and protections covered under the Americans with Disabilities Act of 1990 to the United States Congress. Accordingly, the Committee on Science strives to accommodate/meet the needs of those requiring special assistance. If you need special accommodation, please contact the Committee on Science in advance of the scheduled event (3 days requested) at (202) 225-6371 or FAX (202) 225-0891.
Should you need Committee materials in alternative formats, please contact the Committee as noted above.

HEARING CHARTER

COMMITTEE ON SCIENCE
U.S. HOUSE OF REPRESENTATIVES

H.R. 4218, High-Performance Computing Revitalization Act of 2004

THURSDAY, MAY 13, 2004
10:30 A.M.–12:30 P.M.
2318 RAYBURN HOUSE OFFICE BUILDING

1. Purpose

On Thursday, May 13, 2004, the House Science Committee will hold a hearing to examine federal high-performance computing research and development (R&D) activities and to consider H.R. 4218, the *High-Performance Computing Revitalization Act of 2004,* which would amend the *High-Performance Computing Act of 1991.*

The bill is timely because high-performance computing in the U.S. is at a turning point. The fastest computer in the world today is in Japan not the U.S., and several federal agencies are in the process of reformulating their high-performance computing programs, in part, in response to the challenge posed by Japan.

2. Witnesses

Dr. John H. Marburger, III is the Director of the White House Office of Science and Technology Policy (OSTP). Prior to joining OSTP, Dr. Marburger served as President of the State University of New York at Stony Brook and as Director of the Brookhaven National Laboratory.

Dr. Irving Wladawsky-Berger is Vice President for Technology and Strategy for IBM Corporation. Dr. Wladawsky-Berger previously served as co-chair of the President's Information Technology Advisory Committee (PITAC), and as a founding member of the Computer Sciences and Telecommunications Board of the National Academy of Sciences.

Dr. Rick Stevens is the Director of the Mathematics and Computer Science Division at Argonne National Laboratory (ANL). He is also a Director of the National Science Foundation (NSF) TeraGrid project, which aims to build the Nation's most comprehensive, open infrastructure for scientific computing.

Dr. Daniel Reed is the William R. Kenan, Jr. Eminent Professor at the University of North Carolina at Chapel Hill (UNC–CH). Previously, Dr. Reed served as Director of the National Center for Supercomputing Applications at the University of Illinois Urbana-Champaign, one of NSF's university-based centers for high-performance computing. Dr. Reed is a current member of PITAC.

3. Overarching Questions

The hearing will address the following overarching questions:

1. How does high-performance computing affect the international competitiveness of the U.S. scientific enterprise?
2. Are current efforts on the part of the federal civilian science agencies in high-performance computing sufficient to assure U.S. leadership in this area? What should agencies such as the NSF and the Department of Energy (DOE) be doing that they are currently are not?
3. Where should the U.S. be targeting its high-performance computing research efforts? Are there particular industrial sectors or science and engineering disciplines that will benefit in the near-term from anticipated high-performance computing developments?

4. Brief Overview

- High-performance computers (also called supercomputers or high-end computers) are an essential component of U.S. scientific, industrial, and military competitiveness. However, the fastest and most efficient supercomputer in the world today is in Japan, not the U.S. Japan was successful in producing a computer far ahead of the American machines in part because Japan focused on a type of computer architecture that the U.S. had ceased developing. Also, Japan focused a large amount of money on a single machine, while the U.S. funds a variety of computer development projects.

- Despite the recent technical success of the Japanese, most experts still rate the U.S. as highly competitive in high-performance computing. The depth and strength of U.S. capability stems in part from the sustained research and development program carried out by federal science agencies under an interagency program codified by the High-Performance Computing Act of 1991. That Act is widely credited with reinvigorating U.S. high-performance computing capabilities after a period of relative decline during the late 1980s.
- The Federal Government promotes high-performance computing in several different ways. First, it funds research and development (R&D) at universities, government laboratories and companies to help develop new computer hardware and software; second, it funds the purchase of high-performance computers for universities and government laboratories; and third, it provides access to high-performance computers for a wide variety of researchers by allowing them to use government-supported computers at universities and government labs.
- In recent years, federal agency efforts once again appear to have lost momentum as federal computing activities began focusing less on high-performance computing and more on less specialized computing and networking technologies.
- Responding to concerns that U.S. efforts to develop and deploy high-performance computers may have flagged, OSTP created an interagency task force—the High-End Computing Revitalization Task Force (HEC–RTF)—to examine federal high-performance computing programs and make recommendations for improvement. Dr. Marburger will release the task force report during his appearance before the Committee.
- On April 27, 2004, Representative Judy Biggert introduced H.R. 4218, the *High-Performance Computing Revitalization Act of 2004,* which would update the *High-Performance Computing Act of 1991* and, in particular, would require the High-Performance Computing R&D Program to "provide for sustained access by the research community in the United States to high-performance computing systems that are among the most advanced in the world in terms of performance in solving scientific and engineering problems, including provision for technical support for users of such systems." H.R. 4218 also requires the Director of OSTP to "develop and maintain a research, development, and deployment roadmap for the provision of high-performance computing systems for use by the research community in the United States." This and other provisions in the bill are designed to ensure a robust ongoing planning and coordination process so that the national high-performance computing effort is not allowed to lag in the future.

5. Major Issues Addressed in H.R. 4218

Assuring U.S. Researchers Access to the Most Advanced High-Performance Computing Systems Available.

What the Bill Does: The bill requires the High-Performance Computing Research and Development Program to "provide sustained access by the research community in the United States to high-performance computing systems that are among the most advanced in the world in terms of performance in solving scientific and engineering problems, including provision for technical support for users of such systems." The bill also specifically requires the NSF and the DOE Office of Science to provide U.S. researchers with access to "world class" high-performance computing systems.

Why That's Necessary: Beginning in the 1980s with the NSF supercomputer centers program, the Federal Government has been providing university researchers with access to the fastest computers. Today, university researchers are concerned that the Federal Government, and particularly NSF, may be moving away from a commitment to provide such access. While NSF has reiterated its intention to continue to provide access to the fastest computers through supercomputer centers, it has also said it will place greater emphasis on distributed collections of many computers (known as "grid computing"), which may not provide computing capability equal to that of the fastest supercomputers. At the same time, DOE has indicated it wants to expand its efforts to provide access to large, single-location machines, but it is not clear how much access DOE will be able to provide or whether its machines will be open to researchers in all fields as NSF-funded machines are.

Assuring Balanced Progress on All Aspects of High-Performance Computing.

What the Bill Does: The bill also requires the program to support all aspects of high-performance computing for scientific and engineering applications, including software, algorithm and applications development, development of technical standards, development of new computer models for science and engineering problem solving, and education and training in all the disciplines that support advanced computing.

Why That's Necessary: New supercomputers (hardware) alone won't help researchers. The development of advanced software and applications programs is essential to enable researchers to use the additional computing power.

Assuring an Adequate Interagency Planning Process to Maintain Continued U.S. Leadership.

What the Bill Does: The bill requires the Director of OSTP to "develop and maintain a research, development, and deployment roadmap for the provision of high-performance computing systems for use by the research community in the United States." This and other provisions in the bill are designed to ensure a robust ongoing planning and coordination process so that the national high-performance computing effort is not allowed to lag in the future.

Why That's Necessary: The *High-Performance Computing Act of 1991* codified an interagency planning process that remains in place today. However, the chief product of this process in recent years has been an annual review of activities undertaken by agencies, rather than a prospective planning document. A forward-looking process would enhance coordination between agencies and maximize the total benefit of federal investment.

6. Current Issues in High-Performance Computing

Is the U.S. Competitive?

The world's fastest computer, Japan's Earth Simulator, is designed to perform simulations of the global environment and to address scientific questions related to climate, weather, and earthquakes. NEC, a leading Japanese computer manufacturer, built the Earth Simulator for the Japanese government at a cost of at least $350 million. The first measures of the Earth Simulator's speed, taken in April 2002, determined that the Earth Simulator was significantly faster than the former record holder—the ASCI White System at Lawrence Livermore National Laboratory—and also used the machine's computing power with far greater efficiency.[1]

Twice a year, researchers at the University of Tennessee and the University of Mannheim (United Kingdom) compile a list of the world's 500 fastest supercomputers. The latest list became public on November 16, 2003 (see Table 2 in Appendix II).[2] The Earth Simulator is approximately twice as fast as the second place machine, the ASCI Q system (located at Los Alamos National Laboratory and built by Hewlett-Packard). Of the top twenty machines, eight are located at DOE national laboratories and two at U.S. universities.[3] IBM manufactured six of the top twenty machines and Hewlett-Packard manufactured five.

[1] For the fastest U.S. computers, typical scientific applications are usually only able to utilize 5–10 percent of the theoretical maximum computing power, while the design of the Earth Simulator makes 30–50 percent of its power accessible to the majority of typical scientific applications.

[2] The top 500 list is compiled by researchers at the University of Mannheim (Germany), Lawrence Berkeley National Laboratory, and the University of Tennessee and is available on line at http://www.top500.org/. For a machine to be included on this public list, its owners must send information about its configuration and performance to the list-keepers. Therefore, the list is not an entirely comprehensive picture of the high-performance computing world, as classified machines, such as those used by NSA, are not included.

[3] The two university machines are located at the Pittsburgh Supercomputing Center (supported primarily by NSF) and Louisiana State University's Center for Applied Information Technology and Learning. The remaining 12 machines include four in Europe, two in Japan, and one each at the National Oceanic & Atmospheric Administration, the National Center for Atmospheric Research, the Naval Oceanographic Office, and NASA.

What Types of High-Performance Computers Should the U.S. Develop?

The success of the Earth Simulator has caused a great deal of soul-searching in the high-performance computing community in the U.S. The Earth Simulator is built from custom-made components, and is based on a computer architecture that the U.S. had stopped pursuing in the 1990s. At that time, U.S. programs chose to favor the use of commercially available components for constructing high-performance computers. An advantage of this approach was that it made high-performance computers more cost-effective to develop, by leveraging development costs against a larger market.

Some computing experts have concluded that this strategy of relying largely on commercial needs to guide the development of supercomputer components has left U.S. academic researchers at a disadvantage. That's because certain kinds of research questions—such as those involved in climate modeling—are difficult to pursue on the kinds of computers that can be built with commercial components. The Japanese Earth Simulator, for example, is not based on a computer architecture that would be of widespread interest in the commercial market.

Federal agencies are in the process of reviewing their programs to decide which kinds of computer architecture R&D to pursue. H.R. 4218 is silent on this issue, but a decision on what kinds of computer architectures to pursue would be part of the planning required by the bill.

This question is significant in that NSF first became involved in offering supercomputer access because in the early 1980s foreign researchers often had more and better access to top supercomputers than U.S. researchers did. With the advent of the Earth Simulator, this may be true again for climate and earthquake researchers. Federal civilian agencies, particularly NSF, need to figure out how to help develop computers that will be useful to U.S. scientists in a wide variety of fields. The research needs of different scientific fields require distinct computer architectures, and so serving the entire user community will most likely require the development of a number of diverse computer architectures.

Supercomputers—regardless of the extent of their appeal in the commercial market—are still in the end manufactured private companies. In the U.S., the major producers of high-performance computers include IBM, Hewlett-Packard, and Silicon Graphics, Inc. and Cray. Leading Japanese manufacturers include NEC, Fujitsu, and Hitachi. In the past, Congress prevented federal research funds from being used to purchase Japanese supercomputers.

Where are the NSF and DOE Office of Science and Programs Headed?

NSF and the DOE Office of Science are the lead agencies responsible for providing high-performance computing resources for U.S. civilian research. (See Appendix II.) Both NSF and the DOE Office of Science are moving ahead in significant new directions. NSF recently signaled that it will place greater emphasis on developing grid computing resources. Meanwhile, DOE has indicated it will expand its efforts to provide access to large, single-location machines but has not yet implemented these plans. Both agencies are at a point of transition as they redefine their roles in providing access to U.S. researchers to high-performance computing resources.

NSF's support three large supercomputer centers,[4] which in FY03 served approximately 3,000 users, mostly from academia. (When the supercomputer center program started, there were five initial centers.) In addition to providing cyberinfrastructure, NSF's Computer and Information Sciences and Engineering Directorate supports roughly $70 million of research on hardware, systems architecture, and advanced applications.

In FY04, the DOE Office of Science initiated a new effort in the development of next-generation computer architectures (NGA). The program will emphasize the development of computer architectures that do not rely on commercial components or computing needs. The Department issued an initial request for proposals for the NGA program in March 2004. The NGA Program received $38 million in FY04, and the same amount is requested for FY05.

DOE also administers the National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory, which provides high-end computing resources to over 2,000 scientists annually. According to Department figures, 35 percent of NERSC users are university-based, but the majority are those are funded through DOE grants. The budget for NERSC is on an upward trend, up from $22 million in FY03 to $32 million in FY04, with $38 million proposed for FY05.

[4] The three NSF-supported centers are the San Diego Supercomputing Center at the University of California-San Diego, the National Center for Supercomputing Applications at the University of Illinois Urbana-Champaign, and the Pittsburgh Supercomputing Center, jointly run by Carnegie Mellon University and the University of Pittsburgh.

These increases reflect the Office of Science strategy to expand its role as a provider of high-performance computing resources.

Also, NSF and the Defense Advanced Research Projects Agency (DARPA) have jointly released a solicitation for software for high-performance computing (NSF/DARPA).[5]

7. Background

What Is High-Performance Computing?

High-performance computing—also called supercomputing, high-end computing, and sometimes advanced scientific computing—refers to the use of machines or groups of machines that can perform very complex computations very quickly. High-performance computers are, by definition, the most powerful computers in the world at a given moment in time. High-performance computers are used to solve highly complex scientific and engineering problems, or to manage vast amounts of data. Technologies improve so quickly that the high-performance computing achievements of a few years ago could now be handled by today's desktops.

The speed of high-performance computers is measured in "flops," a unit signifying a calculation each second. The prefix "Tera" signifies trillions, and thus a one Teraflop machine can execute a trillion calculations each second. The world's fastest machine, Japan's Earth Simulator, can execute 35 Teraflops, or 35 trillion calculations each second.

What Is High-Performance Computing Used For?

High-performance computers are often used to simulate physical systems that are difficult to study experimentally. Such simulations can be an alternative to actual experiments (e.g., for nuclear weapon testing and climate modeling), or can test researchers' understanding of a system (e.g., for particle physics and astrophysics). Industry researchers use high-performance computers to simulate how new products will behave in different environments (e.g., for development of new industrial materials). Other major uses for supercomputers include performing massive mathematical calculations (e.g., for codebreaking) and managing vast amounts of data (e.g., for government personnel databases).

Scientific Applications: High-performance computers are used to tackle a rich variety of scientific problems. Large-scale climate modeling examines possible future scenarios related to global warming. In biology and biomedical sciences, researchers perform simulations of protein structure and folding, and also model blood flows. Astrophysicists model planet formation and supernova, while cosmologists simulate conditions in the early universe. Particle physicists perform complex calculations involving the basic building blocks of matter. Geologists model stresses within the earth to study plate tectonics, while civil engineers simulate the impact of earthquakes.

National Defense Applications: The National Security Agency (NSA) is a major user and developer of high-performance computers for specialized tasks relevant to codebreaking (such as factoring large numbers). The DOE National Nuclear Security Administration (NNSA) is also a major user and developer of machines used in modeling nuclear weapons. The Department of Homeland Security uses high-performance computing to extract useful data from large amounts of information; to model the dispersal of plumes of biological, chemical, and radiological agents; and to identify pathogens using their DNA signatures. The Department of Defense uses high-performance computing to model armor penetration, and for weather forecasting. Many scientific applications may have future defense applications. For example, computational fluid dynamics studies could be used to model turbulence surrounding military aircraft.

Industrial Applications: The automotive industry uses high-performance computers for vehicle design and engineering. The movie industry uses massive computer animation programs to produce films. Pharmaceutical companies simulate chemical interactions to design new drugs. The commercial satellite industry manages huge amounts of data in generating maps. Financial companies and other industries use large computers to process immense and unpredictable Web transaction volumes, to mine databases for sales patterns or fraud, and to measure the risk in investment portfolios.

---

[5] The NSF/DARPA solicitation for research on software and tools for high-end computing is available on line at http://www.nsf.gov/pubs/2004/nsf04569/nsf04569.htm.

What Types of High-Performance Computers Are There?

There are a number of different ways to build high-performance computers, and different configurations are better suited to different problems. While there are many possible configurations, they can be roughly divided into two classes: big, single-location machines and distributed collections of many computers (this approach is often called grid computing). Each approach has its benefits—the big machines can be designed for a specific problem and are often faster, while grid computing is attractive in part because the purchase and storage cost is often lower than for a large specialized supercomputer.

At least since the mid-1990's, the U.S. approach to developing new capabilities has emphasized using commercially-available components as much as possible. This emphasis has resulted in an increased focus on grid computing, and has influenced the designs of large, single-location machines. The U.S. has favored supercomputer designs based on ever-larger numbers of commercially available processors, coupled with improvements in information sharing between processors.

Users thus have a number of options for high-performance computing, and must take into account the pros and cons of different configurations when deciding what sort of machine to use. Users must also design software to allow the machine to solve each problem most efficiently. For example, some problems, such as climate modeling and codebreaking, require a great deal of communication between computer components. Other applications, such as large-scale data analysis for high energy physics experiments or bioinformatics projects, can be more efficiently performed on distributed machines, each tackling its own piece of the problem in relative isolation.

What's the Status of Federal High-Performance Computing Capabilities?

In 1991, Congress passed the High-Performance Computing Act, establishing an interagency initiative (now called National Information Technology Research and Development (NITRD) programs) and a National Coordination Office for this effort. Eleven agencies or offices participate in the high-end computing elements of the NITRD program. Tables 1a and 1b in Appendix II show the funding level by agency for FY03, the most recent year for which budget data is available. (The overall FY05 budget request for NITRD is $2 billion, but the breakout for the high-performance computing component of that is not yet available.)

The total requested by all 11 agencies in FY03 for high-performance computing was $846.5 million. The largest research and development programs are at NSF, which requested $283.5 million, and the DOE Office of Science, which requested $137.8 million. Other major agency activities (all between $80 and $100 million) are at the National Institutes of Health (NIH), DARPA, the National Aeronautics and Space Administration (NASA), and NNSA. Different agencies concentrate on serving different user communities and on different stages of hardware and software development and application. (Tables 1a and 1b do not include the procurement costs for high-performance computers purchased by agencies, such as NNSA and the National Oceanic and Atmospheric Administration (NOAA), for computational science related to their missions.[6])

> *National Science Foundation:* In the mid-1980s, NSF established supercomputer centers to serve the academic community. These supercomputing centers provide researchers with access to high-performance computing capabilities and also with the technical support they need to use the facilities effectively. NSF also supports the development of the Extensible Terascale Facility (ETF), a nationwide grid of machines that can be used for advanced communications and data management. The ETF will be coming online in the next year, and a challenge for NSF will be managing the ETF to serve a wide array of users with different scientific computation needs while integrating the ETF with the supercomputing centers.
>
> *Department of Energy:* DOE has been a major force in advancing high-performance computing for many years. Both the Office of Science and the NNSA invest significantly in high-performance computing. Activities under the Office of Science include the Advanced Scientific Computing Research program, which funds research in applied mathematics, in network and computer sciences, and in advanced computing software tools. In FY04, the Office of Science initiated a new program on next-generation architectures (NGA) for high-performance computing. NNSA uses high-performance computers for simulations and weapons modeling through the Accelerated Strategic Computing Initiative (ASCI).

[6] For example, in FY03 NOAA spent $36 million on supercomputers—$10 million for machines for climate modeling and $26 million for machines for the National Weather Service.

Defense Advanced Research Projects Agency: DARPA has traditionally focused on hardware development, including research into new architectures. On July 8, 2003, DARPA announced it had selected Cray, IBM, and Sun Microsystems to participate in the second phase of its High-Productivity Computing Systems program. The goal of the program is to provide a new generation of economically viable, high-productivity computing systems for national security and industrial applications by the year 2010.

Other Agencies: NIH, NASA, and NOAA are primarily users of high-performance computing. NIH manages and analyzes biomedical data and models biological processes. NOAA uses simulations for weather forecasting and climate change modeling. NASA relies on high-performance computers for applications including atmospheric modeling, aerodynamic simulations, data analysis and visualization. Scientists at the National Institute of Standards and Technology collaborate with companies and universities to develop high-performance computing applications to address industrial problems. The NSA both develops and uses high-performance computing for a number of applications, including codebreaking. As a user, NSA has a significant impact on the high-performance computing market, but due to the classified nature of its work, the size of its contributions to High-End Computing Infrastructure and Applications and the amount of funding it uses for actual operation of computers is not public.

Interagency Coordination: The National Coordination Office (NCO) coordinates planning, budget, and assessment activities for the NITRD Program through a number of interagency working groups. The NCO reports to OSTP and the National Science and Technology Council. The NCO also manages the HEC–RTF, an interagency effort on the future of U.S. high-performance computing. The HEC–RTF is tasked with the development of a roadmap for the interagency research and development for high-end computing core technologies, a federal high-end computing capacity and accessibility improvement plan, and a discussion of issues relating to federal procurement of high-end computing systems.

8. Witness Questions

The witnesses were asked to address the following questions in their testimony:

Questions for Dr. Marburger

1. What are the Administration's views on the *High-Performance Computing Revitalization Act of 2004?*
2. Please describe the findings and recommendations of the High-End Computing Revitalization Task Force. How will these findings and recommendations be incorporated into the Networking and Information Technology Research and Development program that you oversee?
3. What are the respective roles of the National Science Foundation and the Department of Energy with regard to the provision of high-performance computing resources to university researchers?

Questions for Dr. Wladawsky-Berger

1. How does high-performance computing affect U.S. industrial competitiveness?
2. Are current efforts on the part of the federal civilian science agencies in high-performance computing sufficient to assure U.S. leadership in this area? What should agencies such as the National Science Foundation and the Department of Energy be doing that they are not already doing now?
3. Where are you targeting IBM's high-performance computing research efforts? Are there particular industrial sectors that will benefit in the near-term from anticipated high-performance computing developments?

Questions for Dr. Stevens

1. How does high-performance computing affect the international competitiveness of the U.S. scientific enterprise?
2. Are current efforts on the part of the federal civilian science agencies in high-performance computing sufficient to assure U.S. leadership in this area? What should agencies such as the National Science Foundation and the Department of Energy be doing that they are not already doing now?
3. Where should the U.S. be targeting its high-performance computing research efforts? Are there particular industrial sectors or science and engineering

disciplines that will benefit in the near-term from anticipated high-performance computing developments?

Questions for Dr. Reed

1. How does high-performance computing affect the international competitiveness of the U.S. scientific enterprise?
2. Are current efforts on the part of the federal civilian science agencies in high-performance computing sufficient to assure U.S. leadership in this area? What should agencies such as the National Science Foundation and the Department of Energy be doing that they are not already doing now?
3. Where should the U.S. be targeting its high-performance computing research efforts? Are there particular industrial sectors or science and engineering disciplines that will benefit in the near-term from anticipated high-performance computing developments?

APPENDIX I

SECTION-BY-SECTION ANALYSIS OF H.R. 4218, THE HIGH-PERFORMANCE COMPUTING REVITALIZATION ACT OF 2004

Sec. 1. Short Title

"High-Performance Computing Revitalization Act of 2004."

Sec. 2. Definitions

Amends section 4 of the High-Performance Computing Act of 1991 (HPC Act) to further elaborate on, or amend, the definition of terms used in the Act:

- "Grand Challenge" means a fundamental problem in science or engineering, with broad economic and scientific impact, whose solution will require the application of high-performance computing resources and multi-disciplinary teams of researchers
- "high-performance computing" means advanced computing, communications, and information technologies, including supercomputer systems, high-capacity and high-speed networks, special purpose and experimental systems, applications and systems software, and the management of large data sets
- "Program" means the High-Performance Computing Research and Development Program described in section 101
- "Program Component Areas" means the major subject areas under which are grouped related individual projects and activities carried out under the Program

Strikes the definition of "Network" that refers to the National Research and Education Network, which no longer exists as such.

Sec. 3. High-Performance Computing Research and Development Program

Amends section 101 of the HPC Act, which describes the organization and responsibilities of the interagency research and development (R&D) program originally referred to as the National High-Performance Computing Program—and renamed the High-Performance Computing Research and Development Program in this Act. Requires the program to:

- Provide for long-term basic and applied research on high-performance computing
- Provide for research and development on, and demonstration of, technologies to advance the capacity and capabilities of high-performance computing and networking systems
- Provide for sustained access by the research community in the United States to high-performance computing systems that are among the most advanced in the world in terms of performance in solving scientific and engineering problems, including provision for technical support for users of such systems
- Provide for efforts to increase software availability, productivity, capability, security, portability, and reliability
- Provide for high-performance networks, including experimental testbed networks, to enable research and development on, and demonstration of, advanced applications enabled by such networks
- Provide for computational science and engineering research on mathematical modeling and algorithms for applications in all fields of science and engineering
- Provide for the technical support of, and research and development on, high-performance computing systems and software required to address Grand Challenges
- Provide for educating and training additional undergraduate and graduate students in software engineering, computer science, computer and network security, applied mathematics, library and information science, and computational science
- Provide for improving the security of computing and networking systems, including research required to establish security standards and practices for these systems

Requires the Director of the Office of Science and Technology Policy (OSTP) to:

- Establish the goals and priorities for federal high-performance computing research, development, networking, and other activities
- Establish Program Component Areas that implement the goals established for the Program and identify the Grand Challenges that the Program should address
- Provide for interagency coordination of federal high-performance computing research, development, networking, and other activities undertaken pursuant to the Program
- Develop and maintain a research, development, and deployment roadmap for the provision of high-performance computing systems for use by the research community in the United States

Leaves substantially unchanged the provisions of the HPC Act requiring the Director of OSTP to:

- Provide an annual report to Congress, along with the annual budget request, describing the implementation of the Program, including current and proposed funding levels and programmatic changes, if any, from the previous year
- Consult with academic, State, and other appropriate groups conducting research on and using high-performance computing

Requires the Director of OSTP to include in his annual report to Congress:

- A detailed description of the Program Component Areas, including a description of any changes in the definition of activities under the Program Component Areas from the previous year, and the reasons for such changes, and a description of Grand Challenges supported under the Program
- An analysis of the extent to which the Program incorporates the recommendations of the Advisory Committee established by the HPC Act—currently referred to as the President's Information Technology Advisory Committee (PITAC)

Requires PITAC to conduct periodic evaluations of the funding, management, coordination, implementation, and activities of the Program, and to report to Congress once every two fiscal years, with the first report due within one year of enactment.

Repeals section 102 of HPC Act, the "National Research and Education Network," which requires the development of a network to link research and educational institutions, government, and industry. This network was developed but has since been supplanted by the Internet.

Repeals section 103 of the HPC Act, "Next Generation Internet," as this program is no longer in existence.

Sec. 4. Agency Activities

Amends section 201 of the HPC Act, which describes the responsibilities of the National Science Foundation (NSF) under the Program. Requires NSF to:

- Support research and development to generate fundamental scientific and technical knowledge with the potential of advancing high-performance computing and networking systems and their applications
- Provide computing and networking infrastructure support to the research community in the United States, including the provision of high-performance computing systems that are among the most advanced in the world in terms of performance in solving scientific and engineering problems, including support for advanced software and applications development, for all science and engineering disciplines
- Support basic research and education in all aspects of high-performance computing and networking

Amends section 202 of the HPC Act, which describes the responsibilities of the National Aeronautics and Space Administration (NASA) under the Program. Requires NASA to conduct basic and applied research in high-performance networking, with emphasis on:

- Computational fluid dynamics, computational thermal dynamics, and computational aerodynamics
- Scientific data dissemination and tools to enable data to be fully analyzed and combined from multiple sources and sensors
- Remote exploration and experimentation

- Tools for collaboration in system design, analysis, and testing

Amends section 203 of the HPC Act, which describes the responsibilities of the Department of Energy (DOE) under the Program. Requires DOE to:

- Conduct and support basic and applied research in high-performance computing and networking to support fundamental research in science and engineering disciplines related to energy applications
- Provide computing and networking infrastructure support, including the provision of high-performance computing systems that are among the most advanced in the world in terms of performance in solving scientific and engineering problems, and including support for advanced software and applications development, for science and engineering disciplines related to energy applications

Amends section 204 of the HPC Act, which describes the responsibilities of the Department of Commerce, including the National Institute of Standards and Technology (NIST) and the National Oceanic and Atmospheric Administration (NOAA), under the Program.

Requires NIST to:

- Conduct basic and applied metrology research needed to support high-performance computing and networking systems
- Develop benchmark tests and standards for high-performance computing and networking systems and software
- Develop and propose voluntary standards and guidelines, and develop measurement techniques and test methods, for the interoperability of high-performance computing systems in networks and for common user interfaces to high-performance computing and networking systems
- Work with industry and others to develop, and facilitate the implementation of, high-performance computing applications to solve science and engineering problems that are relevant to industry

Requires NOAA to conduct basic and applied research in high-performance computing applications, with emphasis on:

- Improving weather forecasting and climate prediction
- Collection, analysis, and dissemination of environmental information
- Development of more accurate models of the ocean-atmosphere system

Amends section 205 of the HPC Act, which describes the responsibilities of the Environmental Protection Agency (EPA) under the Program. Requires EPA to conduct basic and applied research directed toward the advancement and dissemination of computational techniques and software tools with an emphasis on modeling to:

- Develop robust decision support tools
- Predict pollutant transport and their effects on humans and on ecosystems
- Better understand atmospheric dynamics and chemistry

APPENDIX II

Table 1a: Fiscal Year 2003 Budget Requests for High End Computing by Agencies Participating in the National Information Technology Research and Development program (dollars in millions)

Agency	High End Computing: Infrastructure and Applications	High End Computing: Research and Development	Total for High End Computing
NSF	**215.2**	**68.3**	**283.5**
DOE Office of Science	**98.5**	**39.3**	**137.8**
DARPA	16.8	81.9	98.7
NIH	88.2	8.9	97.1
NASA	68.4	26.0	94.4
DOE/NNSA	41.4	39.5	80.9
NSA	--	31.9	31.9
NOAA	13.3	1.8	15.1
NIST	3.5	0.0	3.5
EPA	1.8	0.0	1.8
ODDR&E	--	1.8	1.8
Total:	**547.1**	**299.4**	**846.5**

Source: NITRD National Coordination Office Fiscal Year 2003 Blue Book. The Blue Book is released in August of each year, and thus the data on FY 2003 spending and FY 2004 budget requests levels has not yet been provided to the National Coordination Office.

Note: In addition to the research and development-type activities that are counted for the data included in this table and Table 1b, many agencies devote significant funding to the purchase and operation of high-performance computers that perform these agencies' mission-critical applications.

Acronyms: DARPA—Defense Advanced Research Projects Agency, DOE/NNSA—Department of Energy's National Nuclear Security Administration, EPA—Environmental Protection Agency, NASA—National Aeronautics and Space Administration, NIH—National Institutes of Health, NIST—National Institute of Standards and Technology, NOAA—National Oceanic and Atmospheric Administration, NSA—National Security Agency, NSF—National Science Foundation, ODDR&E—Office of the Director of Defense Research and Engineering.

Table 1b: Funding History from fiscal year 1992 to fiscal year 2003 of high-performance computing research and development programs at various agencies (dollars in millions)

	FY 1992	FY 1993	FY 1994	FY 1995	FY 1996	FY 1997	FY 1998	FY 1999	FY 2000	FY 2001	FY 2002	FY 2003 (Requests)
NSF	**127.00**	**133.90**	**139.10**	**150.00**	**140.32**	**129.20**	**132.90**	**224.70**	**289.80**	**311.70**	**291.50**	**283.50**
DOE Office of Science	**73.00**	**76.20**	**84.60**	**73.10**	**84.49**	**86.00**	**90.53**	**91.90**	**84.10**	**130.30**	**126.70**	**137.80**
DARPA	141.80	169.20	136.20	142.70	77.96	72.70	84.80	48.00	36.50	96.20	81.30	98.70
NIH	8.90	34.40	29.50	29.90	22.40	23.40	23.74	27.10	34.10	59.50	87.20	97.10
NASA	64.00	70.20	84.60	87.40	75.55	88.00	90.10	71.40	124.80	86.80	62.10	94.40
DOE NNSA									113.90	168.30	75.60	80.90
NSA		40.20	32.70	28.20	29.48	30.40	26.42	24.00	31.70	32.90	41.60	31.90
NOAA	1.80	9.40	9.80	2.80	3.30	4.30	4.30	8.80	13.20	12.00	15.60	15.10
NIST	0.90	0.90	0.90	3.60	5.59	4.00	3.99	3.50	3.50	3.50	3.50	3.50
EPA	4.50	6.10	5.90	10.50	8.70	5.60	5.38	4.20	3.90	3.50	1.80	1.80
ODDR&E									2.00	2.00	2.00	1.80
VA					3.00	1.00						
Totals	**421.90**	**540.50**	**523.30**	**528.20**	**450.79**	**444.60**	**462.16**	**503.60**	**737.50**	**906.70**	**788.90**	**846.50**

Source: NITRD National Coordination Office Blue Books, Fiscal Years 1992 to 2003.

Acronyms: DARPA—Defense Advanced Research Projects Agency, DOE/NNSA—Department of Energy's National Nuclear Security Administration, DOE/SC—Department of Energy's Office of Science, EPA—Environmental Protection Agency, NASA—National Aeronautics and Space Administration, NIH—National Institutes of Health, NIST—National Institute of Standards and Technology, NOAA—National Oceanic and Atmospheric Administration, NSA—National Security Agency, NSF—National Science Foundation, ODDR&E—Office of the Director of Defense Research and Engineering, VA—Department of Veterans Affairs.

Program History: Figures from FY 1992–1995 reflect the funding for the High-Performance Computing Systems and the Advanced Software Technology and Algorithms Programs. Figures from FY 1996–1999 reflect the funding for the High-End Computing and Computation Program. Figures from FY 2000–2003 reflect the funding for the High-End Computing Infrastructure and Applications and Research and Development Programs.

Table 2: The top twenty machines of the TOP500 List of the World's Fastest Supercomputers (full list available on line at http://www.top500.org/).

Rank	Site Country/First Year of Operation	Computer / Number of Processors Manufacturer
1	Earth Simulator Center Japan/2002	Earth-Simulator / 5120 NEC
2	Los Alamos National Laboratory United States/2002	ASCI Q - AlphaServer SC45, 1.25 GHz / 8192 HP
3	Virginia Tech United States/2003	X 1100 Dual 2.0 GHz Apple G5/Mellanox Infiniband 4X/Cisco GigE / 2200 Self-made
4	NCSA United States/2003	Tungsten PowerEdge 1750, P4 Xeon 3.06 GHz, Myrinet / 2500 Dell
5	Pacific Northwest National Laboratory United States/2003	Mpp2 Integrity rx2600 Itanium2 1.5 GHz, Quadrics / 1936 HP
6	Los Alamos National Laboratory United States/2003	Lightning Opteron 2 GHz, Myrinet / 2816 Linux Networx
7	Lawrence Livermore National Laboratory United States/2002	MCR Linux Cluster Xeon 2.4 GHz - Quadrics / 2304 Linux Networx/Quadrics
8	Lawrence Livermore National Laboratory United States/2000	ASCI White, SP Power3 375 MHz / 8192 IBM
9	NERSC/LBNL United States/2002	Seaborg SP Power3 375 MHz 16 way / 6656 IBM
10	Lawrence Livermore National Laboratory United States/2003	xSeries Cluster Xeon 2.4 GHz - Quadrics / 1920 IBM/Quadrics
11	National Aerospace Laboratory of Japan Japan/2002	PRIMEPOWER HPC2500 (1.3 GHz) / 2304 Fujitsu
12	Pittsburgh Supercomputing Center United States/2001	AlphaServer SC45, 1 GHz / 3016 HP
13	NCAR (National Center for Atmospheric Research) United States/2003	pSeries 690 Turbo 1.3 GHz / 1600 IBM
14	Chinese Academy of Science China/2003	DeepComp 6800, Itanium2 1.3 GHz, QsNet / 1024 Legend
15	Commissariat a l'Energie Atomique (CEA) France/2001	AlphaServer SC45, 1 GHz / 2560 HP
16	HPCx United Kingdom/2002	pSeries 690 Turbo 1.3GHz / 1280 IBM
17	Forecast Systems Laboratory - NOAA United States/2002	Aspen Systems, Dual Xeon 2.2 GHz - Myrinet2000 / 1536 HPTi
18	Naval Oceanographic Office (NAVOCEANO) United States/2002	pSeries 690 Turbo 1.3GHz / 1184 IBM
19	Government United States/2003	Cray X1 / 252 Cray Inc.
20	Oak Ridge National Laboratory United States/2003	Cray X1 / 252 Cray Inc.

Chairman BOEHLERT. The hearing will come to order. I want to welcome everyone here today to discuss an issue that has been of continuing interest to this committee, high-performance computing.

I first became interested in this issue back in the early '80s, when I sat right at the end of the first row as a junior Member, when Ken Wilson, a Nobel laureate in physics, who was then at Cornell, testified that he and his students sometimes had to go overseas to get access to the fastest computers.

Prompted by those concerns, and by concerns about the health of the U.S. computing industry, this committee helped provide the impetus for the National Science Foundation Supercomputer Center program, which I think everyone here would agree has been a resounding success.

Indeed, spawned in part by those centers, there has been a supercomputing revolution in this country. High-performance computing has become an everyday part of scientific research in both academia and industry. Computation has become a third way of pursuing scientific questions, along with theory and experimentation.

And while the computing industry doesn't look much like it did in the early '80s—thank God for that—revolutions often leave bodies in their wake. U.S. computing capability has continued to advance, and we often hear that today's desktop computers have the power that was once limited to the highest-end models. It never ceases to amaze me that my 12-year-old grandson can hold some game of his in his hands that has greater capacity than what I was initially exposed to when Sperry Univac developed something way back when.

But we can't take that success for granted, and indeed, there are signs of trouble ahead. The Japanese Earth Simulator was a wake-up call that our leadership is being challenged and that we, perhaps, had put too many of our eggs in pursuing computer architectures with commercial applications. And we are starting once again to hear concerns from academia that they may not have continuing access to the fastest machine. That sounds an alarm.

This concern is provoked, in part, by the somewhat mixed signals being sent both by NSF and the Department of Energy about how they will proceed in the future. I am also concerned that we not have a situation in which NSF and DOE both run to catch this particular ball, and end up with it falling between them.

The antidote to all of this is, in part, to re-invigorate the interagency process we put together in the High-Performance Computing Act of 1991. I particularly wish to congratulate Mrs. Biggert and Mr. Davis for introducing a bill that would do just that. We plan to move this bill forward swiftly.

We hope that the revived process and clearer focus called for in the bill will ensure an integrated, adequately funded supercomputing effort among the federal agencies that will help the computing industry develop new machines and will help academic researchers gain access to them.

I hope our distinguished witnesses today will help us figure out how we can accomplish these goals and what else we should be doing, and I hope that Dr. Marburger will be able to assure us that we will be investing the necessary resources in high-performance computing which now undergirds all of science and engineering.

With that, let me yield the remainder of my time to Mrs. Biggert, the chair of our Energy Subcommittee, to talk about her bill.

[The prepared statement of Mr. Boehlert follows:]

PREPARED STATEMENT OF CHAIRMAN SHERWOOD BOEHLERT

I want to welcome everyone here today to discuss an issue that has been of continuing interest to this committee, high-performance computing.

I became interested in this issue back in the early '80s, in the first years I served on this committee, when Ken Wilson, a Nobel laureate in physics who was then at Cornell, testified that his students sometimes had to go overseas to get access to the fastest computers.

Prompted by those concerns, and by concerns about the health of the U.S. computing industry, this committee helped provide the impetus for the National Science Foundation (NSF) supercomputer center program, which I think everyone here would agree has been a resounding success.

Indeed, spawned in part by those centers, there has been a supercomputing revolution in this country. High-performance computing has become an everyday part of scientific research in both academia and industry; computation has become a third way of pursuing scientific questions, along with theory and experimentation.

And while the computing industry doesn't look much like it did in the early '80s—revolutions often leave bodies in their wake—U.S. computing capability has continued to advance, and we often hear that today's desktop computers have the power that was once limited to the highest-end models.

But we can't take that success for granted, and indeed there are signs of trouble ahead. The Japanese Earth Simulator was a wake-up call that our leadership is being challenged and that we perhaps had put too many of our eggs in pursuing computer architectures with commercial applications. And we are starting once again to hear concerns from academia that they may not have continuing access to the fastest machines.

This concern is provoked, in part, by the somewhat mixed signals being sent both by NSF and the Department of Energy (DOE) about how they will provide access in the future. I'm also concerned that we not have a situation in which NSF and DOE both run to catch this particular ball and end up with it falling between them.

The antidote to all of this is, in part, to re-invigorate the interagency process we put together in the *High-Performance Computing Act of 1991.* I want to congratulate Mrs. Biggert and Mr. Davis for introducing a bill that would do just that. We plan to move the bill forward swiftly.

We hope that the revived process and clearer focus called for in the bill will ensure an integrated, adequately funded supercomputing effort among the federal agencies that will help the computing industry develop new machines and will help academic researchers gain access to them.

I hope our distinguished witnesses today will help us figure out how we can accomplish those goals and what else we should be doing, and I hope that Dr. Marburger will be able to assure us that we will be investing the necessary resources in high-performance computing, which now undergirds all of science and engineering.

With that, let me yield the remainder of my time to Mrs. Biggert, the chair or our Energy Subcommittee, to talk about her bill.

Ms. BIGGERT. Thank you, Mr. Chairman, and thank you for yielding me time, and thank you for holding this hearing today.

When we think of how computers affect our lives, we probably think of the work we do on our office desktop machines, or maybe the Internet surfing we do in our spare time. We don't normally think of the enormous contribution that supercomputers, also called high-performance computers, make to the world around us.

You can't have world class science if you don't have world-class computers, and that's why my bill, H.R. 4218, allows U.S. researchers access to the high-performance computing systems that are among the most advanced in the world. To facilitate broader and easier access, H.R. 4218 also provides technical support for those users.

Keeping high-performance computing strong in this country requires coordination of our R&D efforts. Unfortunately, the interagency planning progress has lost the vitality it once had. Congress must find a way to invigorate that process. My bill does so by requiring the White House Office of Science and Technology Policy to direct an interagency planning process and develop and maintain a roadmap for the research, development, and deployment of high-performance computing resources.

The report Dr. Marburger has brought with him today is an excellent beginning, and I commend the High-End Computing Revitalization Task Force for making this valuable contribution. It is clear from the report that we have a lot of catching up to do, but now, we have a map for the first part of our journey.

There is more to supercomputing than building big machines. We need to have a balanced approach that includes software, algorithm, and applications development, development of technical standards, and education and training. H.R. 4218 requires the relevant federal agencies to support all these aspects of high-performance computing.

We could not imagine the kind of problems that the supercomputers of tomorrow will be able to solve, but we can imagine the kind of problems we will have if we fail to provide researchers in the United States with the computing resources they need to remain world-class.

I look forward to today's testimony on this important issue, and yield back.

[The prepared statement of Mrs. Biggert follows:]

PREPARED STATEMENT OF REPRESENTATIVE JUDY BIGGERT

When we think of how computers affect our lives, we probably think of the work we do on our office desktop machines, or maybe the Internet surfing we do in our spare time. We don't normally think of the enormous contribution that supercomputers—also called high-performance computers—make to the world around us.

World-class computers are essential for doing world-class science. My bill, H.R. 4218, ensures that the U.S. research community has access to high-performance computing systems that are among the most advanced in the world, and provides technical support for users of these systems.

Keeping high-performance computing strong in this country requires support at the federal level. Unfortunately, interagency planning process has lost the vitality it once had. Congress must find a way to reinvigorate that process. My bill does so by requiring the White House Office of Science and Technology Policy to develop and maintain a research, development, and deployment roadmap for the provision of high-performance computing resources to the U.S. research community.

The report Mr. Marburger has brought with him today is an excellent beginning and I commend the Task Force for making this valuable contribution. It's clear from the report that we have a long way to go, but now we have a map for the first part of our journey.

We know it's not enough to simply buy big machines. We need to have a balanced approach that includes software, algorithm and applications development; development of technical standards; education, and training. I note that my bill provides support for all these aspects of high-performance computing.

As we meet in this chamber today, we cannot imagine the kinds of problems that the supercomputers of tomorrow will be able to solve. But we can imagine the kind of problems we will have if we fail to provide researchers in the United States with the computing resources they need to remain world-class. I look forward to hearing today's testimony on this important issue.

Thank you.

Chairman BOEHLERT. Thank you very much. Mr. Davis.

Mr. DAVIS. Mr. Chairman, thank you very much. I am pleased to join you in welcoming our witnesses in this hearing that we are having on H.R. 4218, the *High-Performance Computing Revitalization Act of 2004,* which Congresswoman Biggert and I have introduced.

I look forward to working with you on this bill. The need for the legislation we are considering arises from what I would characterize as a weakening of the planning mechanisms for the program established in the High-Performance Computing Act of 1991. The annual program plan required by the 1991 statute is no longer delivered to Congress at the time of the President's budget submission, and it now serves as, more often, an overview of past results than as a description and rationale for funding priorities going forward.

Another strong indicator for breakdown in the planning process is the special task force that was created last year to assess federal efforts to deploy and develop high-end computing systems, partly in response to the concern that the U.S. was falling behind in this technology.

This matter clearly should have been an important agenda item, and subsequently addressed in a comprehensive way, under the normal interagency planning and coordinating process that was established by the 1991 Act.

The High-Performance Computing Revitalization Act has specific provisions that attempt to elevate the priority of high-end computing under this program. It also seeks to strengthen the process for allocating program priorities, and improving program implementation by requiring formal biennial reviews by the President's Information Technology Advisory Committee.

Today, the Committee will hear from the President's science advisor, and from outside experts who have been asked to review the bill and provide their comments and recommendations. I am interested, obviously, in your views on whether the current priorities and resource allocations of interagency programs are properly balanced, and whether the current agency roles are effective.

In my District, we are particularly proud of Oak Ridge National Lab as it leads the supercomputing efforts of the Department of Energy. Oak Ridge and its partners will receive a $25 million grant from the Department of Energy for a supercomputer to be housed in a new 170,000 square foot facility and supported by a staff of 400.

I am thrilled that East Tennessee will be the new home of the world's fastest computer. I appreciate the attention of our—the attendance of our witnesses, and I look forward to our discussion. I yield back the remainder of my time.

[The prepared statement of Mr. Davis follows:]

PREPARED STATEMENT OF REPRESENTATIVE LINCOLN DAVIS

Mr. Chairman, I am pleased to join you in welcoming our witnesses to this hearing on H.R. 4218, the *High-Performance Computing Revitalization Act of 2004,* which Congresswoman Biggert and I have introduced.

The need for the legislation we are considering today arises from what I would characterize as a weakening of the planning mechanism for the program established in the *High-Performance Computing Act of 1991.* The annual program plan required by the 1991 statute is no longer delivered to Congress at the time of the President's budget submission, and it now serves as more of an overview of past results than as a description and rationale for funding priorities going forward. Another strong indicator of a breakdown in the planning process is the special task force that was created last year to assess federal efforts to develop and deploy high-end computing systems, partly in response to concerns that the U.S. was falling behind in this technology. This matter clearly should have been an important agenda item, and subsequently addressed in a comprehensive way, under the normal interagency planning and coordination process established by the 1991 Act.

The High-Performance Computing Revitalization Act has specific provisions that attempt to elevate the priority of high-end computing under the program. It also seeks to strengthen the process for allocating program priorities and improve program implementation by requiring formal biennial reviews by the President's Information Technology Advisory Committee.

Today, the Committee will hear from the President's Science Advisor and from outside experts who have been asked to review the bill and provide their comments and recommendations. I am interested in their views on whether the current priorities and resource allocations of the interagency program are properly balanced and whether current agency roles are effective.

In my district, we are particularly proud of Oak Ridge National Laboratory as it leads the supercomputing efforts for the Department of Energy. Oak Ridge and its partners will receive a $25 million grant from the Department of Energy for a supercomputer to be housed in a new 170,000 square foot facility and supported by a staff of 400. I am thrilled that East Tennessee will be the new home of the world's fastest computer.

I appreciate the attendance of our witnesses, and I look forward to our discussion.

Chairman BOEHLERT. Thank you very much, Mr. Davis.

[The prepared statement of Mr. Smith follows:]

PREPARED STATEMENT OF REPRESENTATIVE NICK SMITH

I'd like to thank Chairman Boehlert and Ranking Member Gordon for holding this hearing to examine the Federal Government's role in the development of high-performance computing capabilities. I would also like to thank the distinguished witnesses for joining us here today.

Supercomputers allow us to make discoveries and develop new products more quickly and at a much lower cost than we would have thought imaginable even 10 years ago. I welcome Dr. John H. Marburger, III, Dr. Irving Wladawsky-Berger, Dr. Rick Stevens, Dr. Daniel Reed here today and look forward to learning more about the current uses, issues and relevance in the development of high-performance computers

As the Chairman of the Research Subcommittee, I am especially interested in the much needed continuous investment at all stages of the technology pipeline, from initial investigation of new concepts to technology demonstrations and products. With no initial, speculative research, this becomes a problem with no gain or success. With the current lack of technology demonstrations, new research ideas are much less likely to grow beyond anything but an idea. Continuous investment is needed in all contributing sectors and agencies including but not limited to the financial investment and support. Universities, national laboratories, private sector corporations and vendors need to share in every aspect of the effort to develop high-performance computers that will better the U.S. both economically by providing jobs, but also by gaining respect among the international community.

In my home state of Michigan, the auto industry is the source of a lot of jobs, but I don't think anyone back home will be too concerned if supercomputer impact modeling puts a few crash test dummies out of work.

Supercomputers are vitally important to our technological and economic competitiveness globally, so it is obviously disturbing that Japan's Earth Simulator is faster and more efficient than anything in the United States. The best hope for the U.S. to maintain its edge against rising global competition is by fostering and expanding our most prized intellectual asset: innovation. Over the past 30 years, innovation has given the U.S. and the rest of the world wave after wave of technological advancement and generated millions of high-skilled jobs. If we want to ensure that successive waves of innovation begin in the U.S., and that U.S. workers are first to benefit from "the next big things," we have to have necessary innovation infrastructure in place. I'm glad that we are talking about this issue today, but I hope that we don't rush to judgment on how the Federal Government can "fix" the problem.

According to an April, 2003 report, IBM it is developing, in conjunction with Lawrence Berkeley National Laboratory and Argonne National Laboratory, a system that will perform at twice the level as the Earth Simulator by 2005. In addition, the Department of Energy has contracted with IBM to develop two systems, ASCI Purple and Blue Gene/L, that together will be able to perform 460 trillion operations per second. The Earth Simulator's peak capability is 40 trillion operations per second.

There may be some need to adjust how the Federal Government supports high-end computing to address areas of need for specific industries or types of research. Still, America's supercomputing capabilities are technologically competitive and I hope that as we move forward with this dialogue that we focus on ways to build on that strong track record.

Again, I would like to thank the Chairman and Ranking Member for holding this hearing.

[The prepared statement of Mr. Costello follows:]

PREPARED STATEMENT OF REPRESENTATIVE JERRY F. COSTELLO

Good morning. I want to thank the witnesses for appearing before our committee to discuss federal research and development activities in support of high-performance computing and the *High-Performance Computing Revitalization Act of 2004* recently introduced by my colleagues Congresswoman Biggert and Congressman Davis. Supercomputers are an essential component of U.S. scientific, industrial, and military competitiveness. Users of these computers are spread throughout government, industry, and academia.

Within my home state of Illinois, the University of Illinois has the Center for Supercomputing Research and Development (CSRD). The CSRD conducts research in supercomputing and parallel processing and has developed the Cedar panel processing system to demonstrate that this technology is practical across a wide range of applications.

As the U.S. develops new high-performance computing capabilities, continued coordination among agencies and between government and industry will be required. The bill introduced by my colleagues seeks to improve coordination and accomplish the goal of developing new capabilities efficiently so that all of the scientific, governmental, and industrial users have access to the high-performance computing hardware and software best suited to their needs.

I am interested to know about the current state of U.S. competitiveness in supercomputing. Further, I am interested to know if adequate research programs are currently in place for the development of future supercomputing systems that will meet the needs of most science and engineering fields.

I thank the witnesses for appearing before our committee and look forward to their testimony.

[The prepared statement of Ms. Johnson follows:]

PREPARED STATEMENT OF REPRESENTATIVE EDDIE BERNICE JOHNSON

Thank you, Mr. Chairman, for calling this hearing to examine the very important issue of High-Performance Computing. I also want to thank our witnesses for agreeing to appear today.

We are here to examine the role the Federal Government can play in high-performance computing research and development activities. There has been much discussion on whether the United States is losing ground to foreign competitors in the production and use of supercomputers and whether federal agencies' proposed paths for advancing our supercomputing capabilities are adequate to maintain or regain the U.S. lead.

As we all know, a high-performance computer, also called a supercomputer, is a broad term for one of the fastest computers currently available. Such computers are typically used for number crunching, including scientific simulations, (animated) graphics, analysis of geological data (e.g., in petrochemical prospecting), structural analysis, computational fluid dynamics, physics, chemistry, electronic design, nuclear energy research, and meteorology.

Supercomputers are state-of-the-art, extremely powerful computers capable of manipulating massive amounts of data in a relatively short time. They are very expensive and are employed for specialized scientific and engineering applications that must handle very large databases or do a great amount of computation, among them meteorology, animated graphics, fluid dynamic calculations, nuclear energy research and weapon simulation, and petroleum exploration.

High-performance computers are gaining popularity in all corners of corporate America. They are used to analyze vehicle crash tests by auto manufacturers, evaluate human diseases and develop treatments by the pharmaceutical industry and test aircraft engines by the aero-space engineers.

It is quite evident that supercomputing will become more important to America's commerce in the future. I look forward to working with this committee on its advancement. Again, I wish to thank the witnesses for coming here today to help us conceptualize this goal.

[The prepared statement of Ms. Jackson Lee follows:]

PREPARED STATEMENT OF REPRESENTATIVE SHEILA JACKSON LEE

Mr. Chairman,

'hank you for convening this timely and provocative hearing. It seems that almost everything we do in Science, in Research and Development, is all critically dependent on computers and information analysis. American leadership in everything from Space exploration, to drug design, to defense could be jeopardized by losing the edge in computing speed and efficiency. The startup of the Earth Simulator in Japan two years ago served as a wake-up call that perhaps we are lagging in this critical field.

We need to take a close look at the possible effects of our investments, or lack of investments, in supercomputing technology. What might be the long-term effects of giving up leadership in supercomputing? Will that loss trickle-down and lead to us falling behind in chip manufacturing, software design, or education of the next-generation engineers and computer scientists? Will our industries and perhaps even defense become dependent on a foreign power?

I hope not. It is in the American spirit to strive for excellence. The *High-Performance Computing Act of 1991* was meant to set us on a course to retain our leadership in computing in an array of scientific and engineering fields. Unfortunately, that initiative is falling into disarray. The Administration's proposed budget for FY 2005 actually cuts the coordinated R&D program by one percent, at a time when our economy is still struggling to rebound and federal investments in growth industries are absolutely critical.

To accent the lack of "vitality" in our high-end computing endeavors, the President's budget description includes a new "High-End Computing *Revitalization* Task Force" with members from around various federal R&D agencies.

We are at a cross-roads here. Japan has recently taken a lead in the supercomputing race—we can either celebrate their progress and find ways to capitalize on their investments, or we can be spurred on to greatness on our own. This committee, with excellent leadership from the Chairman and Ranking Member, has never been afraid to take on such far-reaching questions. I welcome this fine panel of experts to guide us through this dialogue, and thank them for taking the time to be here today.

I look forward to the discussion. Thank you.

Chairman BOEHLERT. And now, for our very distinguished panel, and I want to thank all of you for being resources to this committee. You help us learn, and then hopefully, we can follow and lead.

Dr. John H. Marburger III, Director of the White House Office of Science and Technology Policy. Dr. Marburger, welcome back. Dr. Irving Wladawsky-Berger, Vice President for Technology and Strategy, IBM Corporation. Doctor. And for the purposes of an introduction, the Chair now recognizes the gentlelady from Illinois, Ms. Biggert.

Ms. BIGGERT. Thank you, Mr. Chairman. It is my pleasure to introduce one of our witnesses, Dr. Rick Stevens. Dr. Stevens is the Director of the Mathematics and Computer Science Division at Argonne National Laboratory, which is located in my District, as if you didn't know, because I mention it all the time, but——

Chairman BOEHLERT. The whole world knows.

Ms. BIGGERT. He is also a Director of the National Science Foundation TerraGrid project, which aims to build the Nation's most comprehensive open infrastructure for scientific computing. And I think it is safe to say that he is probably one of the smartest residents of my District, and it is an honor for me to be able to congratulate him publicly today.

As Mr. Davis mentioned just yesterday, the DOE announced that Oak Ridge National Laboratory and Argonne, in partnership with IBM, Cray, and Silicon Graphics, had won a peer-reviewed competition to develop the next generation architectures for high-performance computers. So congratulations, Dr. Stevens, for leading your team at Argonne in this successful collaborative effort, and also congratulations to Dr. Wladawsky-Berger from IBM.

So welcome, Dr. Stevens.

Dr. STEVENS. Thank you.

Chairman BOEHLERT. Thank you very much, and congratulations, Dr. Stevens. You have a very effective advocate here in Washington. She is sitting to my left. And for the purposes of an introduction, the Chair recognizes Mr. Miller of North Carolina.

Mr. MILLER. Thank you, Mr. Chairman. I am pleased to introduce Professor Daniel A. Reed, who now resides in North Carolina. He is the Director of the Renaissance Computing Institute—is it pronounced RENCI—an interdisciplinary center spanning the University of North Carolina–Chapel Hill, Duke University, and North Carolina State University.

Before that, he was the Director of the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, where he also led the National Computational Science Alliance, a consortium of roughly 50 academic institutions and national laboratories that is developing the next generation of software infrastructure for scientific computing; was one of the principal investigators and chief architect of the NSF TerraGrid.

Professor Reed is also the former head of the Department of Computer Science at the University of Illinois, which is one of the oldest and most highly ranked computer science departments in the country, although I assume he will bring the ranking with him now to my alma mater.

He is the William R. Kenan, Jr. Eminent Professor at the University of North Carolina–Chapel Hill, where he conducts interdisciplinary research in high-performance computing.

Chairman BOEHLERT. Thank you very much, and Dr. Stevens and Dr. Reed, it must comfort you some to know that you have Mr. Miller and Ms. Biggert here constantly reminding us of the excellence with which you do your work.

This is a wonderful panel, and I also want to have a couple of words to say about one who is not here today. That is Mr. Robert Bishop, Chairman and Chief Executive Officer of Silicon Graphics, Inc. He has coined a phrase that I think neatly sums up the task before us, and this is his phrase: "In order to out-compete economically in the 21st Century, America will have to out-compute its international competitors."

Mr. Bishop had come to Washington from the West Coast to testify at a hearing that unfortunately had to be cancelled because of the schedule of the House. He could not join us today, but he is a valuable resource, also, as all of the panel members are, and we appreciate his good words.

We will start with Dr. Marburger.

Dr. MARBURGER. Thank you, Mr. Chairman.

I welcome this opportunity to discuss high-performance computing and the Administration's views on the *High-Performance Computing Revitalization Act of 2004.* And I ask that my full written statement be included in the record.

I have a short oral presentation.

Chairman BOEHLERT. And without objection, the full statements of all the panelists will be included in their entirety in the record. We would ask that you try to summarize—or not be arbitrary in the time allocated, but we give you a guideline, five to seven minutes.

Thank you.

STATEMENT OF DR. JOHN H. MARBURGER, III, DIRECTOR, WHITE HOUSE OFFICE OF SCIENCE AND TECHNOLOGY POLICY

Dr. MARBURGER. Thank you.

Information technology does underlie many of the most technological developments of our time. It plays an enabling role in all of the President's priorities—winning the war on terrorism, securing our homeland, and strengthening the economy. Consequently, networking and information technology R&D continues to be one of this Administration's highest interagency R&D priorities. Our Office of Science and Technology Policy is actively engaged in interagency coordination of this area.

The *High-Performance Computing Act of 1991* laid the foundations for the multi-agency networking and information technology R&D program, which we call NITRD, which represents the Federal Government's combined R&D efforts in this field. This program remains a priority of this Administration, and is flourishing today.

In the *High-Performance Computing Revitalization Act of 2004,* the Committee has provided a timely update of this important legislation, while preserving the original legislation's intent and scope. I share your enthusiasm for and commitment to high-performance computing, and I am pleased to convey the Administration's support for this bill, the *High-Performance Computing Revitalization Act of 2004,* in its current form.

I would like to take this opportunity to mention some Administration initiatives related to high-end computing, or supercomputing, which has been and continues to be a high priority area within the broader NITRD program. The President's Fiscal Year 2004 and 2005 budgets stressed the importance of high-end computing, as did a priority guidance memo that was sent out, or will be sent out soon—actually, in the previous year, for Fiscal Year 2005. This is a document that the OMB Director and I send to the heads of science and technology agencies every year to outline our top multi-agency R&D priorities, and NITRD has been a priority ever since I have been in Washington.

We emphasize high-end computing, because the technical activities requiring it are growing, creating a need for advanced computational capabilities that has never been greater. Decisions made years ago that were sensible at the time led to a dependence largely on bundled clusters of commercial, off-the-shelf processors. The promise of high aggregate performance at relatively low cost made the choice of these systems highly attractive. However, while these systems are effective for some classes of applications, many others, including certain applications relevant to national security analyses, are poorly served by these commercial, off-the-shelf based solutions. Addressing this problem, however, is costly, beyond the resources of all but a few federal agencies, and virtually all private sector enterprises.

In the 1990s, due to the limited market for high-end computing systems and the dramatic expansion of the market for low and mid-range systems, the U.S. computer industry focused primarily on the hardware and software needs of business applications and smaller scale scientific and engineering problems, and as a result, the flow of R&D needed to maintain high-end computing technologies in the U.S., and the human capital required to sustain its cutting edge, have failed to keep up with the opportunities for development.

With these concerns in mind, my office, OSTP, created a task force under the auspices of the National Science and Technology Council, and made up of agency experts in high-end computing. This High-End Computing Revitalization Task Force, with an unpronounceable acronym, was asked to develop a forward-looking plan for federal high-end computing. And I am pleased to provide the Committee today with the Task Force's report, *The Federal Plan for High-End Computing,* which you have, I think everyone here has it.

In it, the Task Force addresses the needs of major federal science and technology areas for high-end computing, articulating and synthesizing the urgent problems facing high-end computing, and providing proposed solutions for addressing them.

These include detailed roadmaps for investments in key R&D areas, which include hardware, software and systems. Importantly, the report also includes a recommendation that future so-called leadership class systems—leading edge high capability computers capable of tackling heretofore unsolvable computational problems—be treated as national resources for use by all of the agencies that participate in the systems development, and those agencies' constituents. I provided more information on the Task Force's findings and recommendations in my extensive written testimony.

The recommendations will certainly not be implemented overnight. They will require a dedicated effort by all the relevant agencies, and OSTP is committed to facilitating this effort. Some benefits of the Task Force's work are already evident, primarily as a result of the high level of interagency cooperation in preparing the report. To cite just one example, three agencies, NSF, Department of Energy's Office of Science, and the Department of Defense, have combined forces to initiate the High-End Computing University Research Activity, a pilot program aimed at funding basic research in different theme areas related to high-end computing. Joint planning has led to two closely coordinated solicitations. The agencies' involvement in the Task Force was a key factor in the development of this program, and a sign of the future benefits we can expect from this important effort.

I commend the Task Force for developing this report and for their commitment to continue the work they have begun, by making high-end computing a continued vigorous interagency activity that fully captures the synergies evident in this report, and I look forward to working with all of the agencies this year, to see that the Task Force's recommendations are considered in the preparation of the agencies' Financial Year '06 budget requests. Addressing the issues facing the Nation's high-end computing enterprise will require a sustained and coordinated effort. The Task Force's report constitutes an important first step.

And Mr. Chairman, I think this hearing itself is another important step. Thank you very much for the opportunity to address you on this issue today.

[The prepared statement of Dr. Marburger follows:]

PREPARED STATEMENT OF JOHN H. MARBURGER, III

Mr. Chairman and Members of the Committee, I am pleased to meet with you today to discuss high-performance computing and share with you the Administration's views on the *High-Performance Computing Revitalization Act of 2004.* Networking and information technology (IT) research and development (R&D) continues to be one of this Administration's highest interagency R&D priorities, and the Office of Science and Technology Policy (OSTP) is actively engaged in interagency coordination of this area.

Advancements in IT underlie many of the most important technological developments of our time. The influence of IT is truly pervasive, having a profound impact on the way we work, learn, do business, and communicate. IT plays an enabling role in all of the President's priorities: winning the war on terrorism, securing the homeland, and strengthening the economy. Its impact in this last area has been particularly profound, with tremendous increases in productivity, in particular, serving to reshape the economy. Virtually all aspects of commerce today have felt the impact of IT, from product development to supply-chain management. Federally-funded R&D underpins these advances.

The NITRD program

For all of these reasons, the multi-agency Networking and IT R&D (NITRD) program, which represents the Federal Government's combined R&D efforts in this field, has been and remains a priority of this Administration. As such, it has been featured in each of President Bush's budget requests to Congress. The R&D aspects of the Budget are in turn shaped in part by the memorandum that the Office of Management and Budget (OMB) Director and I send to the heads of agencies with science and technology responsibilities every year, outlining our top multi-agency R&D priorities. Agencies take this memo into account when crafting their budget submissions. The commitment to the NITRD portfolio signaled in these memos is reflected in the funding increases this program—one of the more mature R&D programs in the federal portfolio—has realized. The increases to the NITRD portfolio total 14 percent, to over $2 billion, since President Bush took office in 2001.

A formal interagency working group, which exists under the National Science and Technology Council's (NSTC's) Committee on Technology, coordinates interagency efforts related to the NITRD program. The NSTC is a Cabinet-level council that advises the President on science and technology. It is chaired by the President or Vice President, though that responsibility is typically delegated to the OSTP Director. It is the principal means to coordinate science and technology matters within the federal research and development enterprise.

The Interagency Working Group on NITRD is made up of experts from 12 different agencies with responsibilities for R&D in networking and IT. The group meets regularly and has established seven reporting categories in order to focus on particular areas of emphasis within the overall NITRD portfolio. These Program Component Areas (PCAs) cover the following areas: (1) high-end computing infrastructure and applications, (2) high-end computing research and development, (3) human computer interaction and information management, (4) large-scale networking, (5) software design and productivity, (6) high-confidence software and systems, and (7) social, economic and workforce issues related to IT. Coordinating groups associated with these PCAs meet regularly to determine research needs, coordinate activities, and review progress.

Every year, the NITRD "blue book"—a supplement to the President's Budget—outlines the activities and funding levels for each of the seven areas listed above. This document provides more detailed descriptions of NITRD program activities and more specific budgetary information than is present in the overall Budget. The FY 2005 blue book will be available this summer.

The President's Information Technology Advisory Committee (PITAC), which is made up of private sector representatives with expertise in IT, provides expert, outside advice to the NITRD program. President Bush announced his intention to appoint the current 24 members of PITAC to their positions in May of last year. They have since tackled the important issue of the role of IT in the health care system, and are embarking on an examination of the Nation's cyber security R&D activities. A future activity will address issues related to computational science, a field that focuses on scientific simulation.

The High-Performance Computing Revitalization Act of 2004

Both the NITRD program's and PITAC's foundations are found in the *High-Performance Computing Act of 1991* (P.L. 102–194). The Act, which was subsequently updated with the *Next Generation Internet Act of 1998* (P.L. 105–305), defines an interagency program for the Nation's networking and IT R&D activities. It required the formation of goals and priorities for high-performance computing, which was defined broadly to mean "advanced computing, communications, and information technologies.. . ." It required establishment of an advisory committee to provide outside advice to the program, and identified specific agency activities.

The program that developed from this legislation—the NITRD program—is flourishing today. In the *High-Performance Computing Revitalization Act of 2004,* the Committee has provided a timely update of this important legislation while preserving the original legislation's intent and scope. I share your enthusiasm for and commitment to high-performance computing and I am pleased to convey the Administration's support for the *High-Performance Computing Revitalization Act of 2004,* in its current form.

High-end computing within the NITRD program

High-end computing—or supercomputing, as it is sometimes referred to—is an important element of the NITRD program. Certain of today's important and unsolved scientific and engineering problems can be answered only with high-end computers employing hundreds to thousands of times more computational power than is available in today's systems. These unsolved problems include important national security challenges in areas such as cryptanalysis and image processing of satellite and other data, as well as important scientific and technological questions related to the analysis of complex systems such as aircraft, the atmosphere, and biological systems.

Two PCAs exist to support interagency coordination of high-end computing within the NITRD program, one on Infrastructure and Applications, and the other on R&D. Together, they encompass advances in hardware, software, architecture, and application systems; advanced concepts in quantum, biological, and optical computing; algorithms for modeling and simulation of complex physical, chemical, and biological systems and processes; and information-intensive science and engineering applications.

A number of agencies with active interest in high-end computing participate in coordination: the National Science Foundation (NSF), the National Institutes of Health (NIH), the National Aeronautics and Space Administration (NASA), the Department of Defense (DOD), which includes the Defense Advanced Research Projects Agency (DARPA), the National Security Agency, and the Office of the Director, Defense Research and Engineering, the Department of Energy (DOE) (both the Office of Science and the National Nuclear Security Administration), the National Institute of Standards and Technology (NIST), the National Oceanic and Atmospheric Administration, and the Environmental Protection Agency (EPA).

High-end computing has been and continues to be a high-priority area within the NITRD program. The President's FY 2004 and 2005 Budgets stressed the importance of high-end computing, as did the OSTP/OMB FY 2005 guidance memorandum I referred to earlier.

NSF's and DOE's provision of high-end computing resources to academic researchers

I understand that the Committee is particularly interested in better understanding the provision of high-end computing resources by DOE and NSF to university researchers. NSF remains the largest provider of supercomputing resources to academic researchers, though need continues to outstrip demand. In addition to

NSF-funded scientists and engineers, users include large numbers of NIH-, NASA-, and DOE-funded scientists and engineers.

NSF support for high-performance computing will continue to advance a broad range of science and engineering areas, with emphasis on the support of university-based science and engineering research and education. Moreover, the national community has identified a pressing need to create a state-of-the-art cyberinfrastructure that integrates and makes broadly accessible state-of-the-art high-performance compute nodes, research instruments that generate research data, data storage and management resources, visualization tools that advance capabilities to interpret and analyze data, and new tools for collaboration.

Responsive to this need, NSF's focus on cyberinfrastructure will continue to advance high-performance computing while broadening the scope of facilities and services supported to create new science and engineering knowledge. In addition, NSF will continue, through education, outreach and training as well as development of "services" to make this new cyberinfrastructure available to and usable by a wider range of the national research and education community.

NSF-funded high-performance computing centers include the San Diego Supercomputing Center, the National Center for Supercomputing Applications, and the Pittsburgh Supercomputing Center. These Centers are partnering in the Teragrid effort that integrates their leading edge high-end computing facilities with complementary resources at the California Institute of Technology, Argonne National Laboratory, Indiana University, Purdue University, the University of Texas, and Oak Ridge National Laboratory; the resources are connected by a high-performance backbone network (40 gigabytes/second). NSF's Middleware Initiative is developing software to support distributed applications including collaboration and grid computing.

NSF builds on a wide range of collaborations among universities, federal partnerships (including DOE and DOE Labs), and other sectors. Access to these facilities is available to university researchers through application to the centers. Accounts tailored to development, mid- and high-range needs, educational use, and for Southeastern Universities Research Association and Experimental Program to Stimulate Competitive Research applicants are available. The Partnerships for Advanced Computational Infrastructure and Teragrid facilities allocated more than 169,000,000 CPU (central processing unit) hours to users in FY 2003. Upgrades, both in progress and planned, will significantly increase available CPU hours.

NSF continues significant investments in high-end computing; NSF plans $70 million in FY 2005 for high-end computing facilities. This investment is complemented by significant investments in education, outreach and training, which increase the number and diversity of the user communities, as well as investments in application codes, software, and new technologies for the next generation of computing.

DOE's Office of Science operates several high-end computing facilities, including (1) the National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory, which is the flagship high-end computing facility for the Office of Science; (2) the Center for Computational Sciences (CCS) at the Oak Ridge National Laboratory; and (3) the Environmental Molecular Sciences Laboratory (EMSL) at Pacific Northwest National Laboratory. All are managed as unclassified open facilities in support of the DOE Office of Science mission. University researchers who are working on applications that are relevant to the broad science mission of the Office of Science can apply for access to these facilities, which is granted on a competitive peer-reviewed basis. For example, up to seven percent of NERSC resources are available to researchers for mission-relevant work that is not directly supported by the Office of Science.

An exception to the requirement for mission relevance is DOE's *Innovative and Novel Computational Impact on Theory and Experiment* (INCITE) program at NERSC. The goal of the program is to provide ten percent of the computational resources at NERSC in very large allocations to a small number of computationally intensive large-scale research projects selected based on their ability to make high-impact scientific advances. The INCITE program specifically encouraged proposals from universities and other research institutions.

In FY 2004, 52 proposals were submitted, with more than 60 percent coming from academic researchers, requesting a total of more than 130 million hours of supercomputer processor time. The three awards in FY 2004 amount to ten percent of the total computing time available on NERSC's current IBM supercomputer.

The Office of Science yesterday announced an award for their "Leadership-class System," a $25 million investment in FY 2004. The request for applications for acquisition of this leadership-class system specified that "Proposed activities should be designed to support computational science applications research areas relevant to the mission of the Office of Science, as well as those of other federal agencies." University researchers—regardless of which federal agency supports their work—will be granted access to this leadership-class computational resource, again on a competitive peer-reviewed basis.

Challenges facing the high-end computing enterprise

The challenges facing high-end computing today are significant. Decisions made years ago-sensible at the time—led to a dependence largely on bundled clusters of commercial-off-the-shelf (COTS) processors. The promise of high aggregate performance at relatively low cost made the choice of these systems highly attractive. However, we now know that while these systems are effective for some classes of applications, many others—including certain applications relevant to national security considerations—are poorly served by COTS-based solutions. Addressing this problem, however, is costly—prohibitively so—for all but a few federal agencies and virtually all private-sector enterprises.

In the 1990s, due to the limited market for high-end computing systems and the dramatic expansion of the market for low and mid-range systems, the U.S. computer industry focused primarily on the hardware and software needs of business applications and smaller scale scientific and engineering problems. As a result, the flow of R&D needed to maintain high-end computing technologies in the U.S., and the human capital required to sustain its cutting edge, have failed to keep up with opportunities for development.

The High-End Computing Revitalization Task Force

With these concerns in mind, OSTP initiated the organization of a task force, under the auspices of the NSTC, made up of agency experts in high-end computing. This High-End Computing Revitalization Task Force (HECRTF) was given a specific charge based on the issues outlined in the President's FY 2004 Budget, which said:

> *"Due to its impact on a wide range of federal agency missions ranging from national security and defense to basic science, high-end computing—or supercomputing—capability is becoming increasingly critical. Through the course of 2003, agencies involved in developing or using high-end computing will be engaged in planning activities to guide future investments in this area, coordinated through the NSTC. The activities will include the development of an interagency R&D roadmap for high-end computing core technologies, a federal high-end computing capacity and accessibility improvement plan and a discussion of issues (along with recommendations where applicable) relating to federal procurement of high-end computing systems. The knowledge gained from this process will be used to guide future investments in this area. Research and software to support high-end computing will provide a foundation for future federal R&D by improving the effectiveness of core technologies on which next-generation high-end computing systems will rely."*

Specifically, the Task Force was asked to develop a forward-looking plan for high-end computing with the following three components: (1) an interagency R&D roadmap for high-end computing core technologies, (2) a federal high-end computing capacity and accessibility improvement plan, and (3) recommendations relating to federal procurement of high-end computing systems.

I am pleased to provide the Committee with the Task Force's report, the *Federal Plan for High-End Computing.* In its report, the Task Force addresses the needs of major federal science and technology areas for high-end computing, articulating and synthesizing the urgent problems facing high-end computing.

The Task Force lays out detailed roadmaps for investments in key R&D areas, which include hardware, software, and systems. They emphasize the importance of addressing the increasing gap between the theoretical peak performance and the sustained system performance of high-end computers—a problem that has plagued the massive multi-processor systems currently in use. Their report also emphasizes the need for procurement of "early access" systems that will enable the development of more robust systems and help identify failed approaches before full-scale procurements take place.

The report also addresses issues related to the acquisition, operations, and maintenance of high-end computing systems by agencies, including so-called "leadership class" systems—leading-edge, high-capability computers capable of tackling heretofore unsolvable computational problems. The Task Force recognized that the costs associated with the development of leadership systems are beyond the reach of almost any agency working alone. At the same time, the Task Force emphasized that the need is great: demand for high-end computing capabilities surpasses the resources available in every agency, and some of the smaller agencies, such as EPA and NIST, rely on the resources of other agencies to meet their need. To address this, the Task Force recommends that future leadership systems be treated as national resources, for use by all of the agencies that participate in the system's development (and those agencies' constituents). They suggest specific mechanisms by which agencies that lack the resources to develop high-end computing systems can partner with larger agencies for access to existing systems.

Additional sections of the report address procurement issues, which are currently hampered by the diversity of agency needs for high-end systems and their practices governing procurement of them. The Task Force suggests the initiation of several pilot projects related to procurement to address this. These include the development of improved suites of benchmarks that better mirror applications, an evaluation of the total cost of ownership of several similar systems, and the development of a common solicitation and use of a single suite of benchmarks for procurement, using lessons learned from the first two pilot projects.

Finally, the report describes interagency mechanisms through which to coordinate implementation of various aspects of the plan.

It is important to recognize that benefits of the Task Force's work have already begun to accrue, with the high level of interagency cooperation already leading to tangible results. For example, three agencies—NSF, DOE's Office of Science and DOD—have combined forces to initiate the High-End Computing University Research Activity, a pilot program aimed at funding basic research in different "theme" areas related to high-end computing. Joint planning has led to two closely coordinated solicitations. With software as the theme for 2004, NSF recently issued a program solicitation (that also incorporates DARPA interests) for research on "Software and Tools for High-End Computing." This program, for which the anticipated funding of $7 million was provided by both NSF and DARPA, will support "innovative research activities aimed at building complex software and tools (on top of the operating system) for high-end architectures." A second solicitation, from DOE's Office of Science but also with DARPA interest and funding, is focused on "Operating/Runtime Systems for Extreme Scale Scientific Computation." The agencies' involvement in the HECRTF was a key factor in the development of these programs, and a sign of the future benefits we can expect from this important effort.

I commend the Task Force for developing their report and for their commitment to continue the work that they have begun by making high-end computing a continued, vigorous interagency activity that fully captures the synergies evident in their report. I look forward to working with all of the agencies this year to see that the Task Force's recommendations are considered in the preparation of agencies' FY 2006 budget requests. Addressing the issues facing the Nation's high-end computing enterprise will require a sustained and coordinated effort. The Task Force's report constitutes an important first step.

BIOGRAPHY FOR JOHN H. MARBURGER, III

John H. Marburger, III, Science Adviser to the President and Director of the Office of Science and Technology Policy, was born on Staten Island, N.Y., grew up in Maryland near Washington D.C. and attended Princeton University (B.A., Physics 1962) and Stanford University (Ph.D., Applied Physics 1967). Before his appointment in the Executive Office of the President, he served as Director of Brookhaven National Laboratory from 1998, and as the third President of the State University of New York at Stony Brook (1980–1994). He came to Long Island in 1980 from the University of Southern California where he had been a Professor of Physics and Electrical Engineering, serving as Physics Department Chairman and Dean of the College of Letters, Arts and Sciences in the 1970's. In the fall of 1994 he returned to the faculty at Stony Brook, teaching and doing research in optical science as a University Professor. Three years later he became President of Brookhaven Science Associates, a partnership between the university and Battelle Memorial Institute that competed for and won the contract to operate Brookhaven National Laboratory.

While at the University of Southern California, Marburger contributed to the rapidly growing field of nonlinear optics, a subject created by the invention of the laser in 1960. He developed theory for various laser phenomena and was a co-founder of the University of Southern California's Center for Laser Studies. His teaching activities included "Frontiers of Electronics," a series of educational programs on CBS television.

Marburger's presidency at Stony Brook coincided with the opening and growth of University Hospital and the development of the biological sciences as a major strength of the university. During the 1980's federally sponsored scientific research at Stony Brook grew to exceed that of any other public university in the northeastern United States.

During his presidency, Marburger served on numerous boards and committees, including chairmanship of the governor's commission on the Shoreham Nuclear Power facility, and chairmanship of the 80 campus "Universities Research Association" which operates Fermi National Accelerator Laboratory near Chicago. He served as a trustee of Princeton University and many other organizations. He also chaired the highly successful 1991/92 Long Island United Way campaign.

As a public spirited scientist-administrator, Marburger has served local, State and Federal Governments in a variety of capacities. He is credited with bringing an open, reasoned approach to contentious issues where science intersects with the needs and concerns of society. His strong leadership of Brookhaven National Laboratory following a series of environmental and management crises is widely acknowledged to have won back the confidence and support of the community while preserving the Laboratory's record of outstanding science.

Chairman BOEHLERT. Thank you very much, Dr. Marburger. Dr. Wladawsky-Berger.

STATEMENT OF DR. IRVING WLADAWSKY-BERGER, VICE PRESIDENT FOR TECHNOLOGY AND STRATEGY, IBM CORPORATION

Dr. WLADAWSKY-BERGER. Good morning, Mr. Chairman. I genuinely appreciate the opportunity to be here with you.

I was asked to comment on three questions, and have done so at length in the testimony I have submitted for the record. All three questions go to the heart of some very critical issues of competitiveness, the role of government, and our own strategy for high-performance computing.

I have given considerable thought to issues like this in the course of my 30 plus years in the IT industry. During that time, I have been associated in one way or another with high-performance computing, and based on that experience, I am convinced that supercomputers are more important now than they have ever been.

In response to the Committee's first question, let me say, as unambiguously as I can, that supercomputers are essential to overall U.S. leadership in a global marketplace, and in particular, to U.S. industrial competitiveness. I say that for two reasons. First, the increasing importance of Grand Challenge applications, such as those originally posed by the *High-Performance Computing Act of 1991.* And second, the fact that we are becoming an increasingly integrated information-based society, subject to unremitting change and relentless competitive pressures.

The Grand Challenges that the HPC Act envisioned us tackling were profound, among them, the prediction of global climate change, new improved drugs, and understanding the formation of galaxies, the nature of new materials, and the structure of biological molecules.

Thanks to the combined efforts of industry, academia, and government, the U.S. established a strong position of global competitive leadership in high-performance computing. We did so with machines that by today's standards are rudimentary. Ten years ago, for example, the number one ranked machine on the world's Top 500 list of supercomputers performed 125 billion calculations per second. Today, it would not even make the list. And we did it, not with machines alone, but by building and exploiting an HPC infrastructure of skills, applications, and R&D, as well as government, university, and industry collaborations. All in all, that infrastructure has ensured sustainable leadership for the long-term.

Today, the Grand Challenges are grander still, both in their complexity and in the opportunity they present. Life science, for example, is an entirely new Grand Challenge for supercomputing, one that can revolutionize health care in this country and the rest of the world. It is a Grand Challenge we cannot afford to ignore. Our country's continuing commitment to high-performance computers will make it possible to address the new Grand Challenges and continue to lead the world in these crucial areas.

My second reason for believing that high-performance computing is more important than ever is that we are becoming an increasingly integrated information-based society. Omnipresent communications keep the world online and in touch 24 hours a day. Billions of devices are being connected to the Internet. Microprocessors are turning up in everything, from oil drilling rigs to home appliances. Open standards are integrating all this technology and enabling it to amass and transmit colossal volumes of information.

At IBM, we call this emerging state On Demand. On Demand describes an information-based society with everything and everyone connected using open standards, and with computing power, storage, and networking essentially unlimited. Even at that accelerated integration, it came as no surprise to us that 400 chief executive officers in a recent IBM survey cited their ability to respond to change as an absolute priority. Those CEOs are looking for the ability to take all that information created by customers and competitors and process it in real time.

In the On Demand world, real-time applications and mountains of information make supercomputing not an option, but an essential requirement. Information without real-time analysis and insight cannot deliver the competitive advantage required by an On Demand society. That is why we believe that supercomputing is rapidly becoming part of the modern computing fabric.

Supercomputers are the high-leverage tools that can mean the difference between success and failure in a hypercompetitive global economy. American institutions, from business to government to healthcare and education, need these tools to continue competing and winning.

In an information-based society, supercomputers must be ubiquitous, and that means that they must become more and more affordable. The same forces that drove down the cost of PCs, bandwidth, memory, and IT in general are making high-performance computing more affordable. And as a result, high-performance computing is crossing the boundary between the lab and the rest of society, becoming ever-present in the IT infrastructure and essential for our institutions to innovate.

Innovation remains the key to competing in a changing global economy. Research is a critical driver of innovation, and is needed more than ever, given the environment of change and opportunity that we face. And that is why we strongly support H.R. 4218 and its objective of enhancing U.S. leadership in high-performance computing.

Let me address the question of U.S. leadership. I believe that leadership includes many factors; hardware and software, of course, but also skills, applications, and application development tools, training, research, and development, and the many other factors that make supercomputing valuable.

Our nation needs the federal agencies, civilian and defense, to focus on all these factors to ensure continued success and leadership. The U.S. should also increase its application capability in a cost-effective manner, focusing on the importance of commercially viable technologies. Agencies are questioning whether they can meet their highly specific mission needs using commercial technology, technology which tends to be less expensive, and therefore, more likely to spread through society.

I believe strongly that they can. Industry stands ready to partner with federal agencies to understand and help solve their critical application needs. Thank you, and I look forward to answering your questions.

[The prepared statement of Dr. Wladawsky-Berger follows:]

PREPARED STATEMENT OF IRVING WLADAWSKY-BERGER

Good morning, Mr. Chairman and Members of the Committee and thank you for inviting me to be with you today. My name is Irving Wladawsky-Berger. I am Vice President, Technology and Strategy at the IBM Corporation. I genuinely appreciate the opportunity to offer you our perspective on the questions before the Committee.

Having been associated with high-performance computing for more than 30 years, I think it is important to share with you the fundamental shift we see happening in supercomputing and the role it will play in determining our nation's position in the global economy.

First, I'd like to thank Representative Biggert for her leadership on the important issue of high-performance computing and express my appreciation to all of you for considering H.R. 4218 today. It is critical that our nation support the basic tenets of this bill to: 1) assure U.S. researchers access to the most advanced high-performance computing systems available; 2) assure balanced progress on all aspects of high-performance computing; and 3) assure an adequate interagency planning process to maintain continued U.S. leadership.

Second, I think a little historical perspective may be helpful.

There was a time when many in the U.S. feared that we would lose leadership in this critical area to the Japanese IT industry. Instead, thanks to the combined efforts of industry, academia and government, the U.S. established a strong leadership position in high-performance computing.

Why was this so important?

Because we needed supercomputing to address such grand challenges as:

- Enhancing military systems
- Building more energy-efficient cars and airplanes
- Designing better drugs
- Forecasting weather and predicting global climate change
- Improving environmental modeling, and
- Understanding the formation of galaxies, the nature of new materials, and the structure of biological molecules.

Our leadership in high-performance computing technology allowed us to maintain our leadership internationally in these areas, and we did so with machines that are rudimentary by today's standards. Ten years ago, for example, the number one ranked machine on the world's Top 500 list of supercomputers performed 125 billion calculations per second. Today that computer would not even make the list.

I believe that supercomputing is even more important today than it was in the 1990s when we established our leadership. And if anything it is even more important now that we not only maintain but extend our leadership.

The same economic and social forces that are making PCs, the Internet, wireless and other technologies ubiquitous are transforming the high-performance computing segment.

Supercomputers have become so much less expensive and so much more powerful that they can now be applied in areas where they were never before affordable. In effect, the country's continuing commitment to this technology is making it possible to address new grand challenges. It is imperative that we do so.

- EPA, for example, will use a powerful new supercomputer to assess the risks to human health and the environment posed by exposure to chemical and air pollution and other agents.
- The State University of New York at Buffalo will use high-performance computers at our Deep Computing On Demand center to study human proteins and target drugs for cancer, Alzheimer's, AIDS, multiple sclerosis and other diseases.

Life sciences clearly represents an entirely new set of Grand Challenges for supercomputing with the potential to revolutionize health care in this country and the rest of the world. We cannot afford to ignore it.

We created our Blue Gene supercomputer initiative—ironically using the same chips found in game-players—to tackle the Grand Challenge of protein folding. But there are other milestones we must reach—including the simulation of drug interactions with human cells—that are beyond today's systems. Today, we ultimately test new drugs on human beings. We know the cost and human suffering inherent in this process can be reduced dramatically over time with very sophisticated high-performance simulations leveraging many petaflops of computing power.

But supercomputing is no longer limited to the "classic" Grand Challenges.

At IBM we have described an emerging state of business called On Demand. This is fundamentally what happens when we become an information-based society with everything and everyone connected using open standards, and with computing power, storage and networking essentially unlimited.

Real-time applications and unprecedented amounts of data are creating an environment in which supercomputing is a requirement. Real-time transactions and data without real-time analysis and insight are no longer enough. We see this already in areas as diverse as fraud detection and customer relationship management. We believe supercomputing is rapidly becoming an essential part of the modern computing fabric.

Omnipresent communications keeps the world online and in touch 24-hours a day. Some experts believe that by 2006 the number of devices attached to the Internet—everything from PCs, smart phones and set-top boxes to RFID tags, home appliances and automobiles—will approach ten billion; the number of users will approximate one billion; the number of online buyers a half-billion; and the total amount of commerce $5.5 trillion. Indeed, the price/performance ratio of microprocessors has made them so affordable that they can be integrated in huge numbers into everything from oil well drilling rigs and home appliances to vending machines and automobiles. Adidas is even putting them in running shoes.

Open standards are integrating all this technology and enabling it to amass and transmit information. The availability of information on such a scale and timeframe leads to decisions, decisions to actions, actions to change and change to the need for response. The pace of change will only accelerate and its magnitude will only increase with the constant proliferation and integration of technology.

Given that prospect, it is not surprising that in a recent IBM survey of 400 chief executive officers worldwide, the ability to respond to change was cited as a major need. Those CEOs were calling for the ability to take all that information created by customers and competitors and process it in real time. More and more, it is important to solve complex problems that are critical to competing in a global marketplace that demands the highest quality products offered at attractive prices with the best possible customer service.

Supercomputers are an excellent tool to collect and analyze data; simulate and model problems; and create real-time solutions. The power of supercomputers helps industry and the scientific community to innovate and create solutions faster and at less cost.

It is only with high-performance computing that we can hope to do the real-time information analysis that will enable us to respond faster and more effectively to the developing challenges and growing opportunities all our institutions face. Examples include gathering data to meet security challenges, developing everything from airplanes to health-related items, meeting customer needs, simulating drug reactions in the body, and tracking climate and weather to better understand the environmental challenges of the modern world.

Supercomputers can permit just about all of society's institutions—not just the research community—to understand change better and to act with precision. But to make supercomputers more ubiquitous and increasingly helpful in a wide range of problems in business, health care, education, national security and every other aspect of society, those supercomputers must be affordable.

High-performance computing is crossing the boundary between the lab and the rest of society and is on the road to becoming a ubiquitous and conventional part of the IT infrastructure. As such, it should continue to be a driver of economic growth, a strategic tool for our scientific and business communities, and a strong pillar of our competitiveness in a changing, often turbulent, global marketplace.

The United States must ensure that it will have the high-performance computing assets needed in order to prosper in a constantly changing environment. Clearly, that requires aggressive research, performed at a level commensurate with the environment of change that we face, including the application of high-performance computing to produce real innovation.

We need to foster an environment of innovation much the way the *High-Performance Computing Act of 1991* and the Federal High-Performance Computing and Communications (HPCC) program did when they gave scientists, engineers and industry leaders increased access to high-performance computers, thus building the user community and advancing science.

Innovation has always been the strong suit of the United States. Today, innovation remains the key to maintaining our ability to compete in a changing global economy where technology, science and education are becoming widespread among developing as well as developed nations. And the most advanced technologies—like supercomputing—remain the key to innovation and competitive advantage.

Let me turn now to the specific questions posed by the Committee.

How does high-performance computing affect U.S. industrial competitiveness?

Supercomputing today is more important than ever, especially with the massive amounts of data we collect, analyze and use as well as the increasing complexity of our world. This is true given the competitive environment we live in, with constant growth in Asia and the European Union. This is equally true at the level of the individual firm, where customers have become far more demanding in terms of responsiveness, quality and price. U.S. businesses recognize the value of high-end computing, and want the benefit of affordable access to these tools.

High-end computing has become the third node of science and engineering. By bridging theory and experimentation with computing and simulation, American industry is able to address some of the most complex, computationally intensive problems. Application areas extend from aircraft and automobile design to fusion reactor and accelerator design to materials science to petroleum exploration. High-end computing extends the amount of science and engineering that can be supported by available computational resources.

Supercomputing is the preferred tool of analysis for the sheer mass of available digital data created by advances in processing capability and inexpensive communications. New applications include the processing of streaming data, analysis of video and audio data, real-time security scanning and new areas such as information-based medicine.

Consider what two of our customers are already doing:

- Locus Pharmaceuticals is using supercomputing to develop novel small molecule therapeutics for viral diseases like AIDS.
- General Motors is installing the industry's fastest supercomputer based on our own POWER4 technology to promote greater global collaboration, improve validation testing and reduce product-development costs. They expect it to shorten some vehicles' time-to-market by as much as four years.

High-performance computing is making it possible to provide high-powered analytical capability to traditional commercial applications. In the past, such systems have been focused on the management of data and planning models. Today, we are adding real-time operational capabilities, permitting analysis of the data and response to changing external situations.

Delivering these capabilities sooner rather than later will be vital to U.S. industry's ability to compete in an economic and regulatory environment that is changing and often uncertain. Fortunately, a new business model for delivering high-end computing to U.S. industry is emerging, effectively widening the application base and reducing costs. Specifically, I am referring to the offering of supercomputing power to customers over the Internet, helping to free them from the fixed costs and management responsibility of owning a supercomputer. In this model, a business is able to avoid technological risk as well as the financial risk associated with supercomputer ownership. That is especially important for companies with short-term projects or those with variable needs for supercomputing power.

Let me reiterate that price/performance plays a major role in making supercomputing a prime tool for competitive advantage. In that regard, scalable systems based on common components make it possible to reach a large user base, help reduce the cost and risk of development, and support a wide range of applications.

Cooperation among application and systems developers is key to achieving sustained performance improvements. This is true in both the business world and in the academic and scientific arenas. In universities, where individual investigators lead small research teams and are funded by research grants, a system's price is a major factor in determining which projects proceed and at what rate.

As you know, however, technology by itself is not enough. Our competitiveness will also depend on fostering a broad set of sophisticated skills to match the sophistication and capability of the technology. Our analysis indicates a growing need for many special skills like technical and scientific solutions architects, business transformation consultants, software engineers and application portfolio managers. Highly skilled personnel are critical to the success of the IT industry which in turn is necessary for the economy's competitiveness.

That is one reason IBM invests heavily in training and professional development. This year we will invest over $750 million to help our employees build skills, including more than $200 million for "hot" skills. $400 million (53 percent) will be spent in the U.S. This investment will ensure that our employees have the skills that customers need in today's highly competitive IT world.

Are current efforts on the part of the federal civilian science agencies in high-performance computing sufficient to assure U.S. leadership in this area? What should agencies such as the National Science Foundation and the Department of Energy be doing that they are not already doing?

The current efforts of federal civilian agencies are a good start, but are not enough to meet present demands. This is why we support the bill under consideration and its objectives of: 1) assuring U.S. researchers access to the most advanced high-performance computing systems available; 2) assuring balanced progress on all aspects of high-performance computing; and 3) assuring an adequate interagency planning process to maintain continued U.S. leadership. I believe that these steps will help the U.S. to advance high-performance computing and maintain our position of leadership.

That leadership is based on a many factors. They include: sustainability, meeting application needs, developing algorithms, enhancing skills and creating test beds and partnerships between government, industry and universities. By these measures, there is no question that the U.S. continues to lead the world in high-performance computing.

However, to meet the challenges and complexity of the world today, supercomputing must both meet the "classic" Grand Challenges and become ubiquitous in the solution of a wide variety of problems. There must be a concerted effort to do the necessary research and to move even faster than before if we are to maintain our leadership. In the final analysis, it is the cumulative presence of a variety of leadership characteristics, including skills, technologies (both hardware and software), application development, training methodologies, research, development, engineering, and manufacturing capabilities that will advance high-performance computing. Agencies must focus on all of these components to ensure success.

My fundamental view is that the U.S. should increase its application capability in a cost-effective manner. The roadmap developed to meet these needs must be based on commercially viable technologies that can be optimized for application-specific needs.

The government agencies must work with the research communities and the private sector to define supercomputing applications and technology solutions. The Federal Government should not attempt to dictate market trends and architectural paths for industry. Rather, the government as a partner with industry should specify its critical needs and work with industry to meet them. These partnerships are critical.

Where are you targeting IBM's high-performance computing research efforts? Are there particular industrial sectors that will benefit in the near-term from anticipated HPC developments?

IBM's research strategy revolves around solving complex scientific and business problems more quickly and at lower costs. We continue to aggressively evolve and improve our product line by developing advanced microprocessors which we then use to build scalable families of products. We are also conducting considerable research to overcome obstacles to high degrees of parallelism.

We are doing a number of things to advance our systems, such as:

- Studying cost effective, uniprocessor building blocks that take advantage of the ability to run multiple system activities at the same time (concurrency, i.e., interconnecting main memory, storage, various caches, and then processor execution units and algorithms and application software).

- Recognizing that sustained system performance is more than just hardware, but includes also application development performance and application execution performance.
- Bringing evolutionary technological improvements to current systems with functional integration at the chip package level to provide differentiation.
- Continuing to perform research into the most difficult problems in silicon semiconductor technology and performance.
- Exploring open standard software as a critical aspect of future research and performance.

Our strategy requires that we pursue application-driven design through partnerships with the national labs, universities and government agencies. We are working to satisfy a spectrum of customer performance and price needs, so naturally we maintain continued partnerships with the technical and scientific community. We are engaged in a number of studies to combine new processor architectures with innovative high-performance networks.

Our strategy is based on the following beliefs:

1. HPC systems and applications are crucial, since they will continue to drive advancement in the computer industry. It is not an issue of just technology and hardware. Advancement depends on servers, software, storage, communications and a business model for low-cost delivery of high-performance computing.
2. Petaflop performance will advance in response to the needs of the scientific community, and growing application complexity requires adaptable high-performance computing systems. It is critical to listen to users and then focus on and develop the applications that meet their needs.
3. Architecture should scale up and scale out. We have pioneered both these models. We are committed to sustainable models and long-term viability as well as to ensuring that our customers have the greatest performance for the least amount of money.
4. Simulation and modeling are key to solving 21st century problems.
5. Partnerships between government, universities and industry are critical.

Therefore, our research strategy involves working closely with the Federal Government in general and not solely with the agencies within the jurisdiction of the Science Committee, like the Department of Energy's Office of Science and the National Science Foundation, but very actively with other agencies, such as the Department of Defense and the Department of Energy's Defense programs. In this regard, I believe that the National Institutes of Health should place greater focus on the power that supercomputing could provide for further advances in the life sciences.

We view each of our government collaborations as an opportunity to undertake Grand Challenge applications and address the most complex problems of our times. Our view is that we should leverage our systems expertise in these arenas. These partnerships are valuable to industry, universities and government and we all benefit in unique ways. For a company like IBM, for instance, these projects are relevant to our commercial business and we can leverage this opportunity for learning and importing these new ideas into our products.

Industrial sectors that will benefit include: the life sciences, aircraft and automotive manufacturers, pharmaceutical companies, petroleum companies, and consumer products businesses.

Conclusion

It is critical that high-performance computing in the United States advance to meet the challenges of our complex world. Meeting our applications needs, the needs of our scientists and our businesses, and the skill demands of the 21st century will help us to advance high-performance computing and keep the U.S. at the keen edge of innovation.

H.R. 4218 will help us accomplish this goal. Its emphasis on a mix of leadership, partnerships, powerful and affordable systems, and a strong focus on basic research will keep the U.S. competitive and help us maintain the innovative spirit that has made us global leaders in technology and the most prosperous society on Earth.

BIOGRAPHY FOR IRVING WLADAWSKY-BERGER

Dr. Irving Wladawsky-Berger has responsibility for key IBM initiatives that are critical to the future of the IT industry. In that capacity, he leads IBM's company-wide e-business on demand initiative. The next major phase of the Internet and e-business, e-business on demand helps customers fuse their business processes with advanced IT capabilities to achieve whole new dimensions in productivity and innovation. "On Demand businesses" are more responsive in real-time to any threat or opportunity, more focused on their own core expertise, better able to implement a variable cost structure and—being built on a resilient IT infrastructure—more available to their constituents.

In conjunction with this, Dr. Wladawsky-Berger leads IBM's participation in the movement toward open standards and open source software like Linux; and guides the company's Next Generation Internet efforts. In addition, he collaborates very closely on IBM's Grid and Autonomic Computing efforts to make the Internet a self-managing, distributed computing platform capable of delivering computing services on demand.

Dr. Wladawsky-Berger's role in IBM's Internet and e-business activities began in December 1995 when he was charged with the dual objectives of formulating IBM's overall strategy in the emerging Internet opportunity, and developing and bringing to market leading-edge Internet technologies that could be integrated into IBM's mainstream business.

He began his IBM career in 1970 at the Company's Thomas J. Watson Research Center where he started technology transfer programs to move the innovations of computer science from IBM's research labs into its product divisions. After joining IBM's product development organization in 1985, he continued his efforts to bring advanced technologies to the marketplace, leading IBM's initiatives in supercomputing and parallel computing including the transformation of IBM's large commercial systems to parallel architectures. He has managed a number of IBM's businesses, including the large systems software and the UNIX systems divisions.

Dr. Wladawsky-Berger is a member of the University of Chicago Board of Governors for Argonne National Laboratories and the Technology Advisory Council for BP International. He was co-chair of the President's Information Technology Advisory Committee, as well as a founding member of the Computer Sciences and Telecommunications Board of the National Research Council. He is a Fellow of the American Academy of Arts and Sciences. A native of Cuba, he was named the 2001 Hispanic Engineer of the Year.

Dr. Wladawsky-Berger received an M.S. and a Ph.D. in physics from the University of Chicago.

May 13, 2004

The Honorable Sherwood Boehlert
Chairman, Science Committee
2320 Rayburn Office Building
Washington, DC 20515

Dear Congressman Boehlert:

Thank you for the invitation to testify before the U.S. House of Representatives Committee on Science on May 13th for the hearing entitled ***H.R. 4218, the High-Performance Computing Revitalization Act of 2004***. In accordance with the Rules Governing Testimony, this letter serves as formal notice of the federal funding I currently receive in support of my research.

I received no federal funding directly supporting the subject matter on which I testified, in the current fiscal year or either of the two proceeding fiscal years.

Sincerely,

Irving Wladawsky-Berger

Chairman BOEHLERT. Thank you very much, Doctor. Dr. Stevens.

Dr. STEVENS. Good morning, Mr. Chairman.

Chairman BOEHLERT. Microphone, please.

Dr. STEVENS. Good morning, Mr. Chairman, Members of the Committee, and especially Representative Biggert. I think——

Chairman BOEHLERT. Now, wait a minute. Proceed, Doctor.

STATEMENT OF DR. RICK STEVENS, DIRECTOR MATHEMATICS AND COMPUTER SCIENCE DIVISION, ARGONNE NATIONAL LABORATORY

Dr. STEVENS. I thank you for granting me this opportunity to comment on the future path of high-performance computing research in the U.S.

I would like to start by thanking Representatives Biggert and Davis for introducing H.R. 4218 to reauthorize the High-Performance Computing Act. This is a very critical bill. This bill, like its predecessor, will have a considerable impact on science in the U.S.

What I want to do right now is make a couple of points, summarize two important activities happening at NSF and DOE, and then provide a few recommendations.

My first point is that high-performance computing is a critical technology for the Nation. It is needed by all branches of science and engineering, and it is a critical policy tool for government leaders. More important, its availability is a pacing item for much of science. Without increased access to high-performance computing, certain activities and scientific inquiries will slow down. For me personally, scientific computing is the most important thing that I can think of to work on. My most recent interests are in evolution of bacteria and studying epilepsy in children. High-performance computers are essential for both of those activities.

Second point. The United States is the undisputed leader in the development of high-performance computing technologies, hardware, software, et cetera. But we are also the undisputed leader in education and training for high-performance computing, and because we are training the next generation, we are setting the direction not just for the U.S., but for the world. That direction must be to use this leadership to improve our scientific productivity over the long-term, and the impact that will have on our economics.

Third point. In addition to the computing hardware and software, high-performance computing environments today are required to be connected to many other kinds of resources. Databases and instruments are two very important things. Grid computing, as Irving had mentioned, is a mechanism for doing that that enables us to tie high-performance computing to experimental technologies in life science, medicine, nanoscience, and physics, and to use these systems to analyze the large volumes of data that will come out of those endeavors.

Fourth point. While we are maintaining our leadership in science and technology, we have to have a rigorous research activity to improve performance and usability of these systems. Performance cannot be measured simply by a benchmark on the Top 500 list. It needs to be measured by real applications and real results. We have fallen away from that recently.

Fifth point. Maintaining our international leadership in science and technology requires that the United States dramatically improve its performance in deploying large-scale systems for civilian science and engineering. We have made dramatic progress in deploying these systems for defense. We have not kept up in the civilian sector.

The NSF has embarked on a large-scale project known as the TerraGrid to connect resources across the Nation for serving the university community. In this way, grid computing will provide the power of entire laboratories to individual researchers, regardless of their location. NSF and the Department of Energy should collaborate to ensure that grid technology is broadly deployed and uses standard protocols and interfaces.

Secondly, DOE has recently started the development of national leadership computing capability, the recent announcement on Tuesday. By deploying the highest performance open computers possible, these leadership computing systems will enable researchers to push the scientific envelope and create next-generation software for critical applications in areas of interest to the Nation, including global climate modeling, fusion energy research, life sciences, nanoscience, astrophysics, and chemistry. DOE and the National Science Foundation should collaborate in the development and deployment of these scale systems for the future.

Let me try to summarize with three high level recommendations. First of all, we need to aim high. The U.S. should aim for nothing less than world leadership in high-performance computing. We need to develop the most capable computer systems in the world, make them work, make them work well, and make them available to the broad national scientific community.

DOE and NSF should have a focused research and development program to achieve breakthrough level computing performance on a set of representative applications that are critical for the next ten years of scientific process. Examples such as I gave before, bioinformatics, computational biology, nanoscience, environment, climate, complex device modeling, et cetera.

By focusing on achieving performance breakthroughs on real applications, instead of benchmarks or abstract peak performance, many new ideas may be brought to bear on the problem, and novel application-specific systems may be developed that will provide new ideas for next-generation general purpose systems.

Second recommendation. We must learn from our mistakes. We learned from the original High-Performance Computing bill in '91 that sometimes it doesn't work well to have different agencies working on different parts of the problem, one responsible for hardware, one responsible for software, one responsible for applications, and no one responsible for integrating these systems into a coherent whole, and making them available to users. We must not make that mistake this time around.

Third recommendation. We must connect high-performance computing to the future. We recognize that some of the biggest scientific impacts in the future may not come from the same directions as they have in the past. In particular, we are in the midst of a revolution in biology as a result of access to large-scale computers, data systems, and high-throughput experimental technologies. This revolution will have a far-ranging impact on our science, our society, our security, and our health. So, how to engage the NIH is one of the critical questions facing those in government that manage advanced computing programs and those of us in academia and research laboratories who try to do that work.

Each institute has a potential need for high-performance computing. There are 27 institutes in NIH. We need to somehow find a way to engage them. NIH needs broad access to significant amounts of capacity computing, but they also need access to the most capable computer systems for those areas that are ready, like lung and heart modeling, neuroscience, infectious disease modeling, and cancer.

In conclusion, Mr. Chairman, thank you for your time and this committee's support for the U.S. scientific enterprise, support that has created a system capable of fueling sustained economic growth while fostering an open environment for scientific discovery. I would be happy to answer any questions.

[The prepared statement of Dr. Stevens follows:]

PREPARED STATEMENT OF RICK STEVENS

Good morning, Mr. Chair and Members of the Committee. Thank you for granting me this opportunity to comment on the future path of high-performance computing research. I am Rick Stevens, Director of the Argonne National Laboratory's Mathematics and Computer Science Division and founding director of the Computation Institute and professor of computer science at the University of Chicago. I am also the current director of the NSF TeraGrid project. I am a researcher in scientific and high-performance computing.

I have prepared remarks addressing your questions regarding the reauthorization of the *High-Performance Computing Act of 1991.*

- **How does high-performance computing affect the international competitiveness of the U.S. scientific enterprise?**

During the past several decades high-performance computing has become a critical capability for U.S. science and engineering research. The quantity and quality of scientific projects that rely on high-performance computing either for simulations or for data analysis are increasing rapidly worldwide.

In some areas of research—such as materials science, genomics, astrophysics, climate modeling, high-energy physics, plasma physics, and cosmology—scientific progress can be linked directly to sustained availability of high-performance computing systems. In these areas U.S. researchers are competing directly with their international peers based on the level of computing capability they can bring to bear on a problem.

Therefore, it is reasonable to state that U.S. international scientific competitiveness is directly affected by high-performance computing.

In addition, emerging economies such as India and China will eventually (perhaps greatly) exceed the United States in the total number of employed scientists and engineers. To maintain our leadership in important science and technology areas, we will need to make our scientists as productive as possible. One way to do so is to *extend our leadership* in high-performance computing and extend our ability to apply high-performance computing to emerging areas such as nanotechnology, biotechnology, engineering, and environmental research—areas where rapid technological progress is possible and where the economic benefits of this rapid progress will have near-term impact.

Most university-based U.S. scientists have access through peer-reviewed proposals to the NSF and DOE high-performance computer systems, which are among the most powerful in the world. Access to high-performance computing (HPC) systems by non-university-based researchers varies depending on agency, with some agencies such as NNSA, NASA, and DOD providing considerable access and other such as EPA and NIH providing less access.

- **Are current efforts on the part of the federal civilian science agencies in high-performance computing sufficient to assure U.S. leadership in this area? What should agencies such as the National Science Foundation and the Department of Energy be doing that they are not already doing now?**

The current efforts of the civilian science agencies are commendable but inadequate to ensure sustained and broad U.S. leadership. These efforts are also inadequate to meet the demonstrated current demand from U.S. scientists. Current demand is approximately three times the current capacity.

The United States has arguably the best science funding system in the world. The diversity of funding agencies and the mixture of basic research supported by the NSF and mission research supported by DOE, NASA, NIH, EPA, and NIST have enabled a rich national research portfolio, in fact the richest portfolio of any nation. However, this diversity of funding sources and programs also means that there are occasional missed opportunities and lack of coordination.

Coordination is particularly important when developing computing and data infrastructures (e.g., Grids) and the systems software necessary to integrate computing, databases, instruments and other resources into a coherent scientific resource for the community. Without explicit roles and responsibilities and the associated funding, doing the right thing is often impossible.

In the past, there have also been difficulties in the "technology pipeline" handoff. For many years the DOD and recently the NNSA have played a leading role in developing new HPC architectures. DARPA played a major role in the 1980s and 1990s in developing parallel computing systems. During this same time NSF, DOE, and NASA were responsible for deploying systems for civilian science users and for developing systems software, applications, and networking. However, no single agency or set of agencies was explicitly responsible for deploying "at scale" the most advanced systems for general scientific use. As a consequence the final integration of software, hardware, and applications necessary to make full use of the advanced capabilities was often left undone: usability suffered, users suffered, and science was not well served.

Historically it has been assumed (until recently) that the best way to provide HPC capabilities to the research community was to fund the basic architecture research at universities and occasionally companies, fund some of the enabling software research at labs and universities, and fund the applications, but to rely on the commercial marketplace to move the ideas and technology from the research stage to the product stage for hardware and to have the commercial market complete the software environments necessary to make the machines usable.

Our experience of the past 5–10 years indicates that this strategy is not adequate to maintain leadership in high-performance computing. While there is some commercial demand for high-performance systems, this demand tends to focus on the lower-end of these systems and to be concerned mainly with achieving low-cost capacity cycles.

The research community has a need for capacity, and its demand can generally be met by low-end commercial offerings. However, the research community also requires purpose-built "high-capability" systems. It is these purpose-built capability systems that are the drivers for scientific progress. Like special-purpose instruments—space telescopes, electron microscopes, particle accelerators, and Mars rovers—they capture the scientific imagination, and entire communities are built around them. Unfortunately, there is not a high commercial demand or, in some cases, even any commercial demand, for these systems.

As we push the frontiers on computer technology, it is likely that there will be a partial divergence between those systems that are ideally suited for classes of large-scale scientific computation and those systems that are best suited for general-purpose business computing.

When the scientific community can leverage commodity technologies, commodity components and commodity software, it should. Where these technologies are not adequate for the task, then appropriate technologies should be developed and put to use.

NSF and DOE should work together and with other agencies, particularly with DARPA, to plan large-scale development and deployment of future scientific computing systems aimed at creating a sustained series of advances in computer performance delivered to real scientific applications.

Applications science communities need fundamental improvements in supercomputer performance and scalability. However, we should not aim to achieve a one-time performance record but to begin multiple activities that can be sustained over many hardware generations (5–10 years). These sustained efforts will enable us to understand which applications are best suited for which types of architectures and to optimize them.

Important problems in predicting regional impacts of global warming, modeling pollution transport, understanding the evolution of molecular machines, predicting new drug targets, developing novel materials, and even developing new computational devices require orders of magnitude more computing power than is currently available to academic and laboratory scientists. It is unlikely that one type of high-performance computing architecture will be sufficiently effective on all applications areas. Therefore, it is important to have a diversity of HPC systems under development and to engage the applications community to evaluate each class of system to determine which combinations of algorithms and architectures are best suited for each problem domain and to provide some risk management, in case some ideas turn out not to work. I therefore further suggest that

DOE and NSF work together to develop and deploy a series of the most capable systems in the world for civilian science. These systems should span a range of architectural ideas, and vendors should balance price/performance against applications specificity.

As leading agencies for supporting civilian computational science, NSF and DOE should work together to ensure that the United States designs, builds, and deploys a comprehensive integrated computing and data infrastructure (i.e., a National Science Grid) that is usable by all U.S. scientists regardless of institutional affiliation. NSF has already made an excellent start in this direction with programs such as the National Middleware Initiative (NMI) and the Extended Terascale Facility (i.e., TeraGrid). DOE has developed numerous technologies in the SciDAC and National Collaboratories program that are directly relevant to this infrastructure. NASA also has much to contribute through its Information Power Grid project. However, more needs to be done to ensure that U.S. researchers can access resources supported from multiple agencies in a convenient and secure fashion and with standard protocols and standard tools. Agencies also need to focus on enabling applications communities to exploit this shared infrastructure to reduce overhead, improve productivity, and facilitate sharing and collaboration. Therefore, I suggest the following.

NSF and DOE should work together to construct a National Science Grid.

The National Science Grid would further the democratization of U.S. science by empowering individual researchers—regardless of their location—with the power of entire institutions. This effort will teach us much about how to improve scientific productivity and will lead to commercial benefits as well. It is also in this National Science Grid that we must deploy next-generation supercomputers.

- **Where should the U.S. be targeting its high-performance computing research efforts? Are there particular industrial sectors or science and engineering disciplines that will benefit in the near-term from anticipated high-performance computing developments?**

High-performance computing research should be targeted at four major goals.

1. **Developing Multiple Generations of New Systems.** It should produce multiple new "purposebuilt" architectures that are optimized for large-scale scientific computing. Each of these systems should target particular classes of applications such that the total of all classes cover the important and known applications areas. Areas of importance include systems that address both regular and irregular problems, data-intensive problems, and problems that require interactivity. These systems should reach for performance goals of three to four orders of magnitude beyond current systems over the next ten years.
2. **Develop Systems Software Needed to Make Next-Generation Systems Highly Usable.** Scalable systems software is needed that enables the largest systems to run reliably, with high-throughput I/O, advanced scheduling, secure access, scalability, and extensibility. Systems software research should be open source and cross-platform wherever possible to provide maximum benefit to the community.
3. **Develop Next-Generation Environments for Scientific Problem Solving.** Advanced software environments for scientific computing are needed that improve our ability to solve large-scale problems. Creating these environments will require research in new types of languages such as automated reasoning systems, new language implementation techniques and compilers, visualization and interactive analysis methods, collaboration tools, and data management technologies.
4. **Invest in Fundamental Research.** Accelerated research is needed in fundamental methods and algorithms for scientific problem solving. This research should include novel theoretical formulations of problems and methods that trade computation for storage or that might be applicable for new types of computational devices (e.g., field programmable gate arrays or cellular automata).

A number of scientific and engineering areas can benefit from increased access to high-performance systems in the near-term and new architectures aimed specifically at them in the long-term. These include climate modeling, materials science and nanoscience, molecular modeling, phylogeny and molecular evolution, genomics analysis, computational astrophysics and cosmology, computational chemistry and drug design, theoretical physics, plasma physics, and computational modeling of the heart, lungs, and nervous system. I believe that the interaction between NSF, DOE, and NIH will be a particularly important and fruitful area for collaboration in the near-term and the long-term.

In summary:

1. *HPC is a critical technology for the Nation.* It is needed by all branches of science and engineering and is a critical policy tool for government leaders. Its availability is a pacing item in many areas of science.
2. *The United States is the undisputed world leader in the development of HPC technologies, including hardware, software, and applications.* The United States also leads the world in education and training for HPC.
3. In addition to computing hardware and software, *HPC environments today include advanced networking, Grid computing, and data-intensive computing,* in addition to classical simulation and modeling. New high-throughput experimental technologies in life science and medicine, nanoscience, and physics, as well as large-scale imaging and sensing networks, are highly dependent on increased access to HPC for data analysis and acquisition.
4. *Maintaining our international leadership in science and technology requires that the United States maintain a vigorous research and development pro-*

gram in HPC in universities, laboratories, and private industry. These R&D programs should set their sights on the most aggressive performance and usability goals possible.

5. *Maintaining our international leadership in science and technology requires that the United States dramatically improve its performance in deploying large-scale systems for civilian science and engineering research and make these systems available to all qualified users in the U.S. scientific community regardless of institutional affiliation or funding source.*
6. The NSF has embarked on a large-scale project known as the "TeraGrid" to deploy, via the Grid, high-performance computing to the civilian science community. Grid computing connects multiple distributed large-scale computing resources with high-performance storage, leading-edge visualization resources, scientific databases, and instruments to create a unified computing environment for science. In this way Grid computing will provide the computing power of entire laboratories to individual researchers regardless of their location. *NSF and DOE should collaborate to ensure that Grid technology is broadly deployed and uses standard protocols and interfaces.*
7. DOE has begun development of a national leadership computing capability that will provide unprecedented computing performance to all areas of science and engineering. By deploying the highest-performance open computers possible, these leadership-computing systems will enable researchers to push the scientific envelope and create next-generation software for critical applications in areas of interest to the Nation, including global climate modeling, fusion energy, life sciences, nanoscience, astrophysics, and computational chemistry. *DOE and NSF should collaborate in the development and deployment of leadership-class HPC systems.*

Recommendations

1. **Aim high.** The U.S. should aim for nothing less than world leadership in HPC. We need to develop the most capable computer systems in the world, make them work, and make them available to the broad national scientific community.

The DOE and the NSF should have a focused research and development program to achieve breakthrough-level computing performance on a set of set of representative applications that are critical for the next ten years of scientific progress. Examples of such areas include bioinformatics and computational biology, computational nanoscience, environmental and climate modeling, complex device modeling, and multi-scale multi-physics applications in astrophysics and advanced industrial processes.

By focusing on achieving performance breakthroughs on real applications, instead of benchmarks or abstract peak performance, many new ideas may be brought to bear on the problem, and novel application-specific systems may be developed that will provide new ideas for next-generation general purpose systems.

2. **Learn from our mistakes.** The original HPCC (1991) program showed that it doesn't work well to have different agencies responsible for hardware development, software, and applications and no agency responsible for integration and broad deployment. We should charge NSF and DOE with this broad mission: NSF because of its strong connection to university science and DOE because of its experience in developing large-scale user facilities and technology integration.

We as a nation should pursue multiple computer development paths, including public and private partnerships and novel architectures, while increasing the level of expectations for usability of deployed computing environments. The key goal is that there should be a number of projects each managed by a single agency responsible for making usable resources from the technology developed across the broad national effort.

3. **Connect HPC to the future.** We recognize that some of the biggest scientific impacts in the future may come from different directions from those in the past. The NIH has the largest non-defense research budget in the world and funds the vast majority of life science and biomedical research in the United States. It is widely recognized that bioinformatics and computational biology are revolutionizing both basic biology research, and research of direct clinical importance. I therefore recommend that NIH be considered as a partner with NSF and DOE in the future responsibility of applications science for our national HPC program.

How to effectively engage NIH is one of the critical questions facing those in government that manage advanced computing programs. NIH is a large organization with many institutes. Each institute has a potential need for HPC and could be a target of partnerships with agencies with established programs and with existing HPC infrastructures. NIH needs broad access to significant amounts of capacity computing, as well as access to the most capable systems for those areas of research that are ready to exploit these systems (e.g., neuroscience, heart and lung modeling, infectious disease). We are in the midst of a revolution in biology as a result of access to large-scale computers, data systems, and high-throughput experimental techniques. This revolution will have far-ranging impact on our science, our security, our economy, and our health.

In conclusion, Mr. Chair, I thank you for your time and this committee's support for the U.S. scientific enterprise, support that has created a system capable of fueling sustained economic growth while fostering an open environment of discovery and wonder. I would be happy to answer any questions that you may have.

BIOGRAPHY FOR RICK STEVENS

Professor Rick Stevens is Director of the Mathematics and Computer Science Division at Argonne National Laboratory and co-founder and Director of the University of Chicago/Argonne Computation Institute, which was created to provide an intellectual home for large-scale interdisciplinary projects involving computation at the two institutions. He is internationally recognized for his work in high-performance computing, collaborative and visualization technologies, and computational science, including computational biology. He has a broad set of research interests best characterized by the idea that advanced computing and communications technology is a primary enabling tool for accelerating scientific research. His research has focused on a range of strategies for increasing the impact of computation on science, from architectures and applications for petaflops systems to Grid computing to advanced visualization and collaboration technology for improving scientific productivity of distributed teams. He is currently Director of the NSF TeraGrid project and formerly was chief architect of the National Computational Science Alliance. He has a long-standing interest in applying computing to problems in the life sciences and has been systematically focusing his energies in this direction during the past decade. He is Professor of Computer Science at the University of Chicago, where he teaches and supervises graduate students in the areas of systems biology, collaboration and visualization technology, and computer architecture.

ARGONNE NATIONAL LABORATORY
MATHEMATICS AND COMPUTER SCIENCE DIVISION
9700 SOUTH CASS AVENUE, ARGONNE, ILLINOIS 60439-4844

TELEPHONE 630-252-3378
FAX: 630-252-6333
EMAIL: STEVENS@MCS.ANL.GOV

May 6, 2004

The Honorable Sherwood Boehlert
Chairman, Science Committee
2320 Rayburn Office Building
Washington, DC 20515

Dear Congressman Boehlert:

Thank you for the invitation to testify before the U.S. House of Representatives Committee on Science on May 13 for the hearing entitled *H.R.4218, High-Performance Computing Revitalization Act of 2004*. In accordance with the Rules Governing Testimony, this letter serves as formal notice of the federal funding I currently receive in support of my research; each of the following is renewable.

- $291K, ANI-EIA-012-523, NSF via the University of Texas, Date-Intensive Challenge: The Instrumented Oilfield of the Future, FY04
- $450K, ACI 9619018, NSF via National Computational Science Alliance at University of Illinois, Program for Partnerships for Advanced Computational Infrastructure, FY04
- $900K, ACI-022509, NSF via National Computational Science Alliance at University of Illinois Urbana/Champaign, Extensible Network Services for the Access Grid, FY04
- $6,680K, ACT-012296, NSF via National Computational Science Alliance at University of Illinois at Urbana-Champaign, Extensible TeraGrid Facility, FY04
- $180K, BAA 02-103/VMS, NIH via the University of Chicago, Advanced Biomedical Tele-Collaboration (ABC) Testbed in Surgery, Anesthesia, and Emergency Medicine, FY04
- $500K, KJ-01, FWP 57541, DOE-MICS, SciDAC Middleware to Support Group to Group Collaboration, FY04
- $220K, KJ-01, FWP 57542, DOE-MICS, SciDAC National Collaboratory to Advance the Science of High-Temperature Plasma Physics for Magnetic Fusion, FY04
- $180K, EIA 103723, NSF, Framework for Developing Complex Applications on High-End Petaflops Class Machines, FY04
- $273K, AI-02-031, NIH via the University of Chicago, Resource Center of Excellence - Microbial Informatics Resource Core, FY04

Best Regards,

Rick L. Stevens
Director, Mathematics and Computer Science
Division Argonne National Laboratory and
Professor, Department of Computer Science,
University of Chicago

Chairman BOEHLERT. Thank you very much, Dr. Stevens. Dr. Reed.

STATEMENT OF DR. DANIEL A. REED, WILLIAM R. KENAN, JR. EMINENT PROFESSOR, UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

Dr. REED. Good morning, Chairman Boehlert, and Members of the Committee.

As Representative Miller mentioned, I am chair of a new institute in the Carolinas looking at applications of high-performance computing across a range of disciplines, but especially life sciences. I am delighted to be here to discuss H.R. 4218. I would also like to express my appreciation to Representative Biggert for her sponsorship and leadership with this bill. I believe it is critical. I also chair the Community Input Workshop, which provided input to the agency process for producing the HECRTF report to which Dr. Marburger alluded.

In response to your questions regarding the HPC Computing Revitalization Act, I would like to make three points today. The first, related to international competitiveness, is that high-performance computing, as you noted in your opening remarks, has emerged as a third element of a research portfolio that complements theory and experiment. In cosmology, where experiments are not possible, it allows researchers to explore models of the universe's origins. In climatology, it allows rapid analysis of humanity's long-term effects on the environment. And in biology, it enables researchers to study the effects of genetics, pathogens, and particulates on respiratory disease.

Legend says that Archimedes remarked on discovery of the lever "Give me a place to stand, and I can move the world." Today, science and computational science have become largely synonymous, and high-performance computing is the intellectual lever that helps assure U.S. competitiveness in an increasingly competitive world.

However, there is one unique aspect that I think we must realize about high-performance computing that distinguishes it from our investments in other research instruments, and that is its universality as an intellectual amplifier. Powerful new telescopes advance astronomy, but not material science. Powerful new particle accelerators advance high-energy physics, but not genetics. High-performance computing is universal as a tool that advances research discovery in all of the sciences.

That brings me to my second point, the current status of our efforts in coordinated solutions. Because all research domains do benefit from high-performance computing, but none is solely defined by it, the high-performance computing endeavors often lack the cohesive community of advocates that might be found in an individual discipline, and this has often led to, in my judgment, an overdependence on market forces to shape what emerges as technologies to advance science.

During the past three years, at least six community reports have highlighted the need for more integrated approaches, and in this regard, I applaud the Committee for capturing these recommendations in the HPC Revitalization Act. My only recommendation beyond the basic Act would be to consider mechanisms to aid the transfer of technologies that are promising and applicable to the nature of science and to commercial practice. There are substantial costs to develop high-performance computing systems tailored to science, and the limited markets associated with those sometimes mean that government mechanisms may be necessary to help sustain those developments, again, for systems tailored to science.

More generally, I believe an integrated interagency initiative, as envisioned by the HPC Act, should clearly articulate the scope of each agency's responsibilities, and as Dr. Stevens noted, as part of a broad computing ecosystem. It should include verifiable metrics for interagency collaboration and progress that are coupled to national priorities.

Now, there has been a lot of debate about the relative roles of the National Science Foundation and the Department of Energy in providing access to high-end computing systems. In my judgment, this debate misses the critical point. The collaborative commitments of both are necessary to sustain scientific research. Both need to deploy and maintain world-class computing systems in support of scientific discovery, again, in an integrated infrastructure, the supports, data management, storage, network, and workforce development.

Finally, to echo something that Dr. Stevens said, the biological triumphs of the last decade are due in no small part to biological insight, but also to the judicious applications of computing technologies. Hence, I believe it is critical that the National Institutes of Health should also lead by supporting computing research, and by working with the other agencies to deploy in the integrated infrastructure in support of biomedical research.

This brings me to my third and final point, where we go from here. Today, the lack of high-performance computing systems designed for important scientific and national problems unnecessarily constrains our innovation. Integrated vehicle designs with lifetime warranties are within reach. Personalized medicines tailored to the individual genetics of particular individuals are also on the horizon. To make these opportunities a reality, however, we must develop new high-performance computing systems that better support the needs of critical applications as a focus initiative that focuses on sustained, not simply peak performance.

In addition, we must recognize that we must make these systems easier to use and more productive, particularly for commercial domains in support of national competitiveness. There is no silver bullet that will eliminate our current problems. Rather, the challenge is in sustaining an integrated interagency research, development, and deployment initiative that is reflective of national needs and opportunities.

Today, high-performance computing is reaping the rewards of yesterday's research. We must seed tomorrow's crop of research ideas today, lest I fear we will tomorrow subsist on wild berries, rather than the fruits of today's research.

So in conclusion, let me say that I strongly support H.R. 4218 and its vision for high-performance computing, and I would be happy to take questions.

[The prepared statement of Dr. Reed follows:]

PREPARED STATEMENT OF DANIEL A. REED

Good morning, Chairman Boehlert and Members of the Committee. Thank you very much for granting me this opportunity to comment on appropriate paths for scientific computing. I am Daniel Reed, Director of the Renaissance Computing Institute (RENCI), a collaborative activity of the University of North Carolina at Chapel Hill, Duke University and North Carolina State University. I am the former Director of the National Center for Supercomputing Applications (NCSA) at the University of Illinois, one of three NSF-funded high-end computing centers. I am also a researcher in high-performance computing.

In response to your questions regarding the *High-Performance Computing Revitalization Act of 2004,* I would like to make three points today regarding high-performance computing.

1. International Competitiveness

High-performance computing has emerged as the third element of the research portfolio, complementing theory and experiment. Computing breathes life into the underlying mathematics of theoretical models, allowing us to understand nuanced predictions and to shape experiments more efficiently. Computing also allows us to capture and analyze the torrent of experimental data being produced by a new generation of scientific instruments and sensors, themselves made possible by advances in computing and microelectronics.

Legend says that Archimedes remarked, on the discovery of the lever, "Give me a place to stand, and I can move the world." Today, computing pervades all aspects of science and engineering. "Science" and "computational science" have become largely synonymous, and high-performance computing is the intellectual lever that helps assure U.S. scientific leadership in an increasingly competitive world.

High-performance computing plays a special and important role as an intellectual lever by allowing researchers and practitioners to bring to life theoretical models of phenomena when economics or other constraints preclude experimentation. Computational cosmology, which tests competing theories of the universe's origins by computationally evolving cosmological models, is one such example. Given our inability to conduct cosmological experiments (we cannot create variants of the current universe and observe its evolution), computational simulation is the only feasible way to conduct experiments.

High-performance computing also enables researchers to evaluate larger or more complex models and to manage larger volumes of data than would be possible on conventional computer systems. Although this may seem prosaic, the practical difference between obtaining results in hours, rather than weeks or years, is substantial—it qualitatively changes the range of studies one can conduct. For example, climate change studies, which simulate thousands of Earth years, are only feasible if the time to simulate a year of climate in a few hours. Moreover, conducting parameter studies (e.g., to assess sensitivity to different conditions such as the rate of fluorocarbon or CO_2 emissions) is only possible if the time required for each simulation is small.

Finally, high-performance computing allows us to couple models to understand the interplay of processes across interdisciplinary boundaries. Understanding the environmental and biological bases of respiratory disease or biological attack requires coupling of fluid dynamics models to model airflow and inhalants, whether smoke, allergens or pathogens, materials models to surface properties and interactions, biophysics models of cilia and their movements for ejecting foreign materials, and deep biological models of the genetic susceptibility to disease. The complexity of these interdisciplinary models is such that they can only be evaluated using high-performance computers.

The breadth of these examples highlights a unique aspect of high-performance computing that distinguishes it from other scientific instruments—its universality as an intellectual amplifier. Powerful new telescopes advance astronomy, but not materials science. Powerful new particle accelerators advance high energy physics, but not genetics. In contrast, high-performance computing advances all of science and engineering, because all disciplines benefit from high-resolution model predictions, theoretical validations and experimental data analysis. As new scientific discoveries increasingly lie at the interstices of traditional disciplines, high-performance computing is the research integration enabler.

Although this universality is the intellectual cornerstone of high-performance computing, it is also its political weakness. Because all research domains benefit from high-performance computing, but none is solely defined by it, high-performance computing lacks the cohesive, well-organized scientific community of advocates found in other disciplines. In turn, this has led to over-dependence on market forces to shape the design and development of high-performance computing systems, to our current detriment.

Fueled by weapons research and national security concerns, until the 1980s, the U.S. government's high-performance computing needs could substantively influence the commercial market and assure U.S. supremacy in high-performance computing. Scientific and government high-performance computing needs are now a much smaller fraction of the overall computing market, with concomitantly less economic influence.

With the explosive growth of the computing industry and the internationalization of information technology, we are in danger of losing our international competitive advantage in high-performance computing, with serious consequences for scientific research and industrial competitiveness. This economic milieu has had profound effects on all aspects of high-performance computing—research and development, marketing, procurement and operation.

This brings me to my second point: the current status of our efforts.

2. Current Status and Coordinated Solutions

Not only has high-performance computing enriched and empowered scientific discovery, as part of a larger information technology ecosystem, it has also been responsible for substantial economic growth in the United States. **Because of this success, information technology and high-performance computing are increasingly international activities, with associated competition for intellectual talent and access to world-class computing resources.**

In an era of constrained federal budgets and fierce international competition, we cannot afford wasted or duplicative efforts. The great strength of the U.S. research system is its diversity—many research ideas can be explored, with funding opportunities at multiple agencies. In computing, this diversity also creates leaks in the pipeline from basic research to deployment and commercial infrastructure, and many promising ideas are lost. The pipeline from basic research, through advanced prototyping and evaluation, to either research infrastructure or commercial development, requires tactical and strategic coordination across agencies.

Hence, we must encourage cross-agency collaboration and coordination, while leveraging the unique missions and attributes of each agency. Only via such interagency coordination can we maintain international leadership in high-performance computing. This belief is supported by broad community consensus. **During the past three years, at least six community reports[1] have highlighted the limitations of current approaches and have recommended an integrated, interagency initiative in high-performance computing.**

I applaud the Committee for capturing the central elements of these recommendations in the High-Performance Computing Revitalization Act, namely the need to (a) train a new generation of high-performance computing users and researchers, (b) conduct basic research and advanced prototyping for high-performance computing, and (c) develop and deploy high-performance systems that match scientific needs. **In addition to these goals, I recommend that the HPC Act also include mechanisms to aid the transfer of promising technologies to commercial practice.** The substantial engineering costs to develop high-performance computing systems and their limited market means that government incentives or support may prove necessary to sustain development of high-performance systems that can meet national scientific and security needs.

I believe an interagency initiative in high-performance computing should be based on the following principles:

[1] NSF Workshop on Computation as a Tool for Discovery in Physics, September 2001—www.nsf.gov/pubs/2002/nsf02176; Report on High-Performance Computing for the National Security Community, July 2002—www.hpcc.gov/hecrtf-outreach/bibliography/200302_hec.pdf; Blueprint for Future Science Middleware and Grid Research and Infrastructure, August 2002—www.nsf-middleware.org/MAGIC/default.htm; Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure, January 2003—http://www.cise.nsf.gov/sci/reports/toc.cfm; DOE Science Networking Challenge, June 2003—gate.hep.anl.gov/may/ScienceNetworkingWorkshop/; DOE Science Case for Large Scale Simulation, June 2003—www.pnl.gov/scales/; Community Workshop on the Roadmap for the Revitalization of High-End Computing, June 2003—www.hpcc.gov/hecrtf-outreach

1. *An integrated strategic plan that articulates the responsibilities, scope and financial scale of each agency's responsibilities.*
2. *Regular deployment and support of the world's highest performance computing facilities for open scientific use, as part of a broad ecosystem of supporting infrastructure, including high-speed networks, large-scale data archives, scientific instruments and integrated software.*
3. *Coordination and support for national priorities in science, engineering, national security and economic competitiveness.*
4. *Vendor engagement to ensure technology transfer and economic leverage.*
5. *Verifiable metrics of interagency collaboration, community engagement and technical progress that are tied to agency funding.*

The National Science Foundation (NSF) and the Department of Energy (DOE) are the primary supporters of physical science and engineering research, whereas the National Institutes of Health (NIH) fund the majority of life science and biomedical research. Each of these and other federal civilian science agencies has a unique, though critical role in the computing technology pipeline.

There has been much debate about the relative roles of NSF and DOE in providing access to high-performance computing for scientific research. This debate misses the critical point—the coordinated actions of both agencies are needed to ensure U.S. competitiveness, and both should be charged with deploying and operating systems with the highest possible capability.

Reflecting its role as a basic research agency, NSF should support advanced systems research, including new architectures, software and tools and advanced algorithms. This research is the well spring of tomorrow's computing systems and infrastructure and the educational opportunity for a new generation of high-performance computing researchers. **Concurrently, NSF should continue to develop and support leading edge computing and data management systems, both for open community access and to support its Major Research Equipment (MRE) projects.**

Investments in "computing as science" (i.e., basic research in next generation computing technologies) and "computing for science" (i.e., deployment of computing infrastructure as a scientific enabler) are complementary, with qualitatively different time scales and needs. Given the rapid rates of change in computing technologies, high-performance computing infrastructure must be sustained at adequate levels for long periods and renewed regularly if it is to remain relevant to research facilities that have 10-20 year operational lifetimes.

Many high-performance computing research ideas can only be validated by constructing large-scale prototypes. In the 1970s and 1980s, the U.S. funded several research and development efforts in high-performance computing, and we continue to harvest insights from these experiments. Today, there are few, if any such projects, with concomitant loss of experience and insight. Hence, **DOE should lead advanced prototyping and deployment of next-generation high-performance computing systems, coupled to its scientific facilities and laboratory mission. This advanced prototyping and development should harvest basic research ideas from the DOE and NSF portfolios for national deployment.**

Finally, as quantitative biology and biomedicine expand to include tools and techniques from the physical and mathematical sciences, the National Institutes of Health (NIH) must also assume a leadership role in computational science and high-performance computing. **The biological research triumphs of the past decade were due in no small measure to a combination of biological insight and judicious application of new computing technology. Equally importantly, the biomedical discoveries of this decade, with concomitant cost savings and improved treatments, will depend critically on the deep integration of biology, medicine, software, algorithms and hardware. Hence, NIH should also lead by supporting both computing research and the creation of a national infrastructure for biomedical data sharing, computational modeling and distributed collaboration that is interoperable with that being deployed by NSF and DOE.**

While we debate appropriate actions, our international competitors are moving ahead. As part of the Sixth Framework, the European Union plans to deploy a pan-European Grid as a baseline infrastructure in support of scientific research. In the U.S., we are developing a set of loosely connected Grids without a common framework or strategic funding plan. Similarly, Japanese investment in the Earth System Simulator, the world's fastest computing system, is well known.

This leads me to my third and final point: research needs and opportunities.

3. Actions

The explosive growth of commodity clusters has reshaped the high-performance computing market. Although this democratization of high-performance computing has had many salutatory effects, including broad access to commodity clusters across laboratories and universities, it is not without its negatives. Not all applications map efficiently to the cluster programming model of loosely coupled, message-based communication, and it is difficult for vendors to make a profit developing systems tailored for scientific research. Hence, some researchers and their applications have suffered due to lack of access to more tightly coupled supercomputing systems. Second, **an excessive focus on peak performance at low cost has limited research into new architectures, programming models, system software and algorithms.** The result has been the emergence of a high-performance "monoculture" composed predominantly of commodity clusters and small symmetric multiprocessors (SMPs).

In the 1990s, the U.S. high-performance computing and communications (HPCC) program supported the development of several new computer systems. In retrospect, we did not recognize the critical importance of long-term, balanced investment in hardware, software, algorithms and applications. Achieving high-performance for complex scientific applications requires a judicious match of computer architecture, system software, tailored algorithms and software development tools. We have substantially under-invested in the research needed to develop a new generation of architectures, programming systems and algorithms. The result is a paucity of new approaches to managing the increasing disparity between processor speeds and memory access times (the so-called von Neumann bottleneck).

Hence, **we must target exploration of new systems that better support the irregular memory access patterns common in scientific and national defense applications. In turn, promising ideas must be realized as advanced prototypes that can be validated with scientific codes.** In addition, we must recognize that new programming models and tools are needed that simplify application development and maintenance. The current complexity of application development unnecessarily constrains use of high-performance computing, particularly for commercial use. Finally, increases in achieved performance over the past twenty years have been due to both hardware advances and algorithmic improvements; we must continue to invest in basic algorithms research. This critical cycle of prototyping, assessment, development and deployment must be a long-term, sustaining investment, not a one time, crash program.

Opportunities abound for application of high-performance computing in both science and industrial sectors. Integrated vehicle designs with lifetime warranties, based on coupled electrical, mechanical and power train models, are within reach. Higher resolution cosmological models would allow testing of competing theories of the evolution of the universe, with sufficient resolution to simulate galaxy formation. Personalized medicines, tailored to minimize toxicity and maximize efficacy based on individual genetics, are possible based on drug chemistry models. All require a new generation of high-performance computing systems that can deliver high sustained performance for a suite of coupled models.

There is no "silver bullet" that will eliminate current problems and ensure continued U.S. preeminence in high-performance computing. Rather, the challenge is creating and sustaining an integrated, interagency research, development and deployment program that is reflective of national needs and opportunities. **Today, high-performance computing is reaping the rewards of yesterday's research investment. We must seed tomorrow's crop of research ideas today, else tomorrow we will subsist on wild berries.**

In conclusion, Mr. Chairman, let me thank you for this committee's longstanding support for scientific discovery and innovation. Thank you very much for your time and attention. I would be pleased to answer any questions you might have.

BIOGRAPHY FOR DANIEL A. REED

Professor Daniel A. Reed is Director of the Renaissance Computing Institute (RENCI), an interdisciplinary center spanning the University of North Carolina at Chapel Hill, Duke University and North Carolina State University. He was previously Director of the National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign, where he also led National Computational Science Alliance, a consortium of roughly fifty academic institutions and national laboratories that is developing next-generation software infrastructure of scientific computing. He was also one of the principal investigators and chief architect for the NSF TeraGrid. Professor Reed is also the former head of the Department of Computer Science at the University of Illinois, one of the oldest and most highly ranked computer science departments in the country. He holds the Chancellor's Eminent Professorship at the University of North Carolina at Chapel Hill where he conducts interdisciplinary research in high-performance computing.

THE UNIVERSITY OF NORTH CAROLINA
AT
CHAPEL HILL

Daniel A. Reed
Chancellor's Eminent Professor
Phone: (919) 962-1796
FAX: (919) 962-1799

Carolina Renaissance Computing Institute
Campus Box 3175, Sitterson Hall
Chapel Hill, NC 27599-317
Email: Dan_Reed@unc.edu

May 13, 2004

The Honorable Sherwood Boehlert
Chairman, Committee on Science
2320 Rayburn Office Building
Washington, DC 20515

Dear Congressman Boehlert:

Thank you for the invitation to testify before the U.S. House of Representatives Committee on Science on May 13 for the hearing entitled H.R. 4218, *High-Performance Computing Revitalization Act of 2004*. In accordance with the Rules Governing Testimony, this letter serves as formal notice of the federal funding I currently receive in support of my research.

Please see attached table containing federal awards over the past three years.

Sincerely,

Daniel A. Reed
Chancellor's. Eminent Professor
Director, Renaissance Computing Institute
University of North Carolina at Chapel Hill

Attachment

Grant Number	Federal Agency/Source	Title	Project Dates	Total Project Amount
N00014-03-1-0765	Office of Naval Research	National Center for Advanced Secure Systems Research	05/05/2003 - 04/27/2004	4,195,814
DE-FC02-01ER41205	Department of Energy	National Computational Infrastructure for Lattice Gauge Theory	08/15/2001 - 08/14/2004	$ 273,273
ARMY RAYTHEON AA19	Raytheon Company	NCSA Participation in High-Performance Computing Modernization Program at Army	09/23/1996 - 08/17/2004	3,278,490
NSF ACI-0219597	National Science Foundation	ITR: Intelligent High-Performance Computing on Toys	09/01/2002 - 08/31/2004	$ 400,000
NSF EIA 9972884 EQ	National Science Foundation	Intelligent Information Spaces: A Testbed to Explore and Evaluate	09/15/1999 - 08/31/2004	$ 2,000,000
DE-FC02-01ER25488	Department of Energy	High-End Computer System Performance: Science and Engineering	09/15/2001 - 09/14/2004	$ 801,000
NSF CMS 01-17853	National Science Foundation	NEESGRID: A Distributed Virtual Laboratory for Advanced Earthquake Experiment	08/01/2001 - 09/30/2004	6,501,869
NSF ACI 01-22296	National Science Foundation	The TERAGRID: Cyberinfrastructure for 21ST Century Science and Engineering	10/01/2001- 9/30/2004	36,045,500
NSF EOA 9975248	National Science Foundation	CADRE: A National Facility for High-Performance I/O Characterization	07/15/1999 - 06/30/2003	$ 1,200,000
R71700H-29200099	William Marsh Rice University (Department of Energy Prime)	Los Alamos Computer Science Institute	05/03/1999 - 09/20/2003	$ 1,164,909
R38144-79200003	William Marsh Rice University (National Science Foundation Prime)	NGS: GrADS: Efficient Script-Based Application Development for Networked High Performance Computing Environment	10/01/2002 - 09/20/2003	$ 105,000
NSF ACI-9619019	National Science Foundation	Enabling Technology Team C: I/O and Virtual Environments Software	10/01/1997 - 09/30/2003	$ 800,000
NSF ACI 96-19019COOP	National Science Foundation	National Computational Science Alliance - PACI	10/01/1997 - 9/30/2003	192,846,587
DOE LLNL B505214	Lawrence Livermore National Lab (Department of Energy Prime)	Wide-Area, Adpative I/O Systems for Data and Visualization Corridors	07/29/1999 - 07/20/2002	$ 702,819
PC253664	California Institute of Technology (Department of Energy Prime)	Accelerated Strategic Computing Initiative	10/01/1997 - 11/30/2002	$ 515,658
R36504-29200099	William Marsh Rice University (National Science Foundation Prime)	Grid Appplication Development Software (GrADS)	10/01/1999 - 12/31/2002	$ 445,935
NSF 02-168	Oklahoma University (National Science Foundation Prime)	Linked Environments for Atmospheric Discovery	09/01/2003 - 08/31/2008	$ 2,921,835

NSF CCR-0331654	William Marsh Rice University (National Science Foundation Prime)	Virtual Grid Application Development Software (VGrADS)	09/01/2003 - 08/31/2008	$ 677,320

DISCUSSION

Chairman BOEHLERT. Thank you. Excuse me, thank you very much. Thank all of the panelists.

Let me start by commending Dr. Marburger for identifying high-performance computing as a top tier issue for the science agencies and for convening the White House-led interagency task force to develop a revitalization plan. The Task Force in its report constituted a textbook example of how OSTP can constructively guide federal science programs. It is exactly what Congress had in mind when the Science Committee created OSTP back with the *Science and Technology Policy Act of 1976.* So, Dr. Marburger, once again, by everyday performance, you distinguish yourself and the important post you hold.

Let me ask each of the witnesses, what is the most important thing the Federal Government ought to be doing that it isn't doing, or isn't doing enough of, in the area of high-performance computing, and think about that for a moment?

Then I will ask Dr. Marburger to tell us if those gaps are reflected in the HECRTF report, and how they will be addressed.

Let us go—Dr. Wladawsky-Berger. Of Big Blue.

Dr. WLADAWSKY-BERGER. When we look at progress in high-performance computing, the biggest steps happen when we are working together between industry, the research community in universities and national labs, for advanced applications, because it is usually by pushing that envelope of advanced applications that we learn what works, what doesn't work, that we also learn how to make systems usable, because the technology by itself is not usable. You need to add considerable software. You need to add a lot in application development tools. You need to develop applications and algorithms, and the bigger the system, the harder it is to develop all those additional facilities.

I think I want to second the statements that Rick Stevens and Dan Reed made that in today's world, the grandest of challenges are in life sciences. I mean, I am a physicist by training. I also come from near territory of Representative Biggert, having gotten my degree at the University of Chicago, which is sort of nearby. And you know, physics drove a lot of high-performance in the 20th Century. In the 21st Century, it is life sciences, and the applications are incredibly, and their potential is absolutely incredible, but they require a lot more work, a lot more pilots, a lot more development of applications than we have today.

So, that—I would say that would be my number one priority.

Chairman BOEHLERT. Thank you. Dr. Stevens.

Dr. STEVENS. My assessment of the number one priority is that we need to have a sustained development activity over multiple generations of hardware, and we need to have multiple paths. So, we have to work on multiple kinds of architectures, and we need to do it over multiple generations. Each generation of hardware might take two or three years to develop, a year to manufacture, a year to install, so you are—we are talking about a 10 to 15 year horizon, not a three year horizon.

We have to commit to a multi-year program. Part of that program needs to be large-scale deployment, so the scientists that can be involved in that multi-generation development activity have an expectation of being able to do science on those systems when they are developed.

Right now, as much as this HECRTF plan is actually a huge step forward, it still is relatively silent on what is necessary to deploy these systems to advanced sciences. It is mainly about developing the systems, somewhat silent about the deployment strategy. Once you deploy this hardware, you need to develop the application software, as Irving has—Wladawsky-Berger has mentioned. And life science and nanotechnology are my two thoughts about which areas are most important for the future.

So, to summarize, multi-year, multi-generation deployment, development and deployment, with the software applications bundled.

Chairman BOEHLERT. Well, you know, given the focus on biology, should NIH be playing a larger role?

Dr. STEVENS. Absolutely.

Dr. REED. Absolutely.

Dr. STEVENS. There is jurisdictional issues with respect to maybe this committee, but NIH needs to be engaged at the highest levels, and both vertically and horizontally across that agency to advance. And there has been a number of reports in the last five years that have tried to lay out a roadmap for NIH's participation. We just need to see some of that executed.

Chairman BOEHLERT. Thank you. Dr. Reed.

Dr. REED. I would echo what both of my friends have said. I think the challenge is in sustaining investment and in sustaining understanding. My grandmother used to tell me that good judgment comes with experience, and experience comes from bad judgment. And part of what that means in any discipline is, and particularly true of high-performance computing, where one is looking at the interaction of complex systems with application domains, is you really have to turn the crank multiple times. You have to do the R&D, you have to deploy a generation of systems, and gain experience with those systems with real applications, take those insights, and feed them back in to successive generations of improved designs. And that ability to move through multiple generations means that we have to sustain the investment across the R&D and the deployment in order to make those systems really effective.

We tend to start things, but not finish things, and the notion that we can solve a problem in a couple of years with a crash initiative and then declare victory and move on doesn't solve this kind of problem. And in terms of investment and driving problems, I agree that biology and biomedical research is one of the great untapped opportunities, and I really believe that NIH has to be a player in this.

I would also say that, though, it is not a black and white thing, because a lot of the discoveries, as biology becomes quantitative, are from interdisciplinary interactions and insights from the physical sciences and mathematics, and that fusion of research collaborations in interdisciplinary ways is a place that will make a lot of biological discovery happen, and high-performance computing is absolutely critical to that.

Chairman BOEHLERT. Dr. Marburger, your comment.

Dr. MARBURGER. Well, thank you very much.

First of all, I would like to point out that Congress has created some pretty heavy machinery to accomplish some of the things that my colleagues here on the panel have pointed to. Certainly, the need for sustained investment for getting that experience is, I think, satisfied by the existence of the National Information Technology R&D Initiative itself, and the establishment of PITAC [President's Information Technology Advisory Committee], this expert FACA [Federal Advisory Committee Act] group, and the existence of a coordinating office that reports up under OSTP. I believe that our engagement with this issue during the past two years has been productive and will be increasingly productive, and I must say that the attention by Congress to this issue has been very important in sustaining our attention to it.

So, I do agree with the need for sustained effort first of all, and secondly, for effective coordination across agencies. And I would like to say a word about that, and particularly the participation by NIH. NIH does participate in our interagency effort. And I believe that in the future, as a follow-on activity to the preparation of this report, which does indeed focus more on development than on deployment, I believe that a focus on real life problems and applications, as recommended by my colleagues here, will have the effect of engaging NIH more effectively in future deliberations. Because there is a great—high-performance computing comes in different colors and varieties. There are different architectures, and different kinds of hardware structures related to the different types of applications. And I believe that NIH is currently getting a lot of mileage out of the existing high-end computing architectures that are available, these clusters of existing, off-the-shelf microprocessors.

They are enormously powerful for some types of bioinformatics, so NIH has a lot to chew on with the existing state of supercomputing, and they are doing great things with it. A focus on real life applications of other kinds of architectures can enable all agencies to learn about the potential for increased investment and application relevant to their field in a way that our previous focus on development of architectures would not.

So, I think there are some interrelationships among the various recommendations that have been made by our panelists that are all likely to be addressed under the bill that you have proposed, and the structure that has been created and will be strengthened in the bill.

Chairman BOEHLERT. Thank you very much, Doctor. Mr. Davis.

Mr. DAVIS. Thank you, Mr. Chairman. The Department of Energy has begun a national leadership class supercomputer that will be enormous—I think enormous value to those in the business community, as well as research institutions.

And I certainly applaud these efforts, and look forward to America regaining the leadership in this role. However, Dr. Stevens, when you testified, you indicated that perhaps maybe because there is not any particular agency that would be responsible for providing leadership and guidance, that perhaps we should choose one, or one of the different agencies that would actually provide oversight.

The question is should responsibility for providing access to and support on specialized high-end computing systems be assigned to a particular agency, and would you recommend which one?

Dr. STEVENS. So, multiple agencies have high demand. The two agencies that leap to mind as the candidates for having this leadership role are the National Science Foundation and the Department of Energy. Both of them have complementary skills to bring to the table.

The NSF funds many thousands of researchers across the country and universities, and is in touch with the pulse of what academic research is doing, and they need to play a very strong role in providing capability systems to that community.

On the other hand, DOE has the skill and the organizational structure to field large development projects, large-scale construction projects, large-scale research instruments, and to manage them as national user facilities.

These two skills need to be combined. Neither agency, I think, should have the sole responsibility for serving the country. That is a single point of failure. I don't think we want to have that risk in this endeavor. On the other hand, there needs to be much more coordination between those two agencies, and a linkage of how they are going to provide access.

For example, the National Science Foundation is prohibited by current policy from providing access to researchers at FFRDCs [Federally Funded Research and Development Center] for access to the high-performance computers at the NSF Supercomputer Centers. DOE, up until recently, had a more of an internal focus on its allocation of computing resources. In order for NSF to be that lead, that policy for FFRDC access would have to change. In order for DOE to be the single lead, DOE's policy for assigning time to mission applications would have to take a second seat to a peer review process similar to that applied at the experimental facilities like the Light Sources that awards time based on merit, not mission.

So, I think both agencies need to have this role. I think what is really limiting progress is not really the politics of interagency role, as much as resources available for large-scale deployment.

And if I could just point out that the Japanese Earth Simulator was not really a technological enterprise that somehow beat us. It was primarily a resource deployment issue. That machine cost on the neighborhood of $400 million to deploy. At the time at which the Japanese made that commitment, the largest systems the U.S. was deploying in that same timeframe were on the order of $100 million per system. So scale of deployment is really the issue here, not so much agency politics.

Mr. DAVIS. Would someone else like to respond?

Dr. MARBURGER. There are interagency issues associated with the guidelines that exist in their operations, and these are on the table in the discussions that we have in the OSTP-sponsored interagency working groups. I believe that continued focus on these issues will bear some fruit. I was very pleased when the Department of Energy, for example, did open its computing facilities much more broadly to scientists receiving their support from other agencies.

But as this committee is fully aware, there is some controversy regarding the opening up of NSF resources to scientists who are not at universities, and working—the FFRDCs refers to the Department of Energy National Laboratories, of course.

Mr. DAVIS. Any other—first of all, I am sorry.

Dr. REED. I think the challenge, really, is in strategic planning across the agencies, and looking at acquisitions and deployments as a rolling, sustained activity. We have to move to a model where we can make long-term plans about the infrastructure that we deploy in support of national scientific discovery, because the uncertainty about when a new machine will appear at appropriate scale has long-term ramifications, and to hark back to what several of us said before about the deep integration now of computing as an enabler for science, if one looks at the timescale for other large-scale scientific facilities we build, we have multi-year planning processes and operational lifetimes that may be measured in 10, 20, 30 years. We don't have, at the moment, that kind of strategic planning, acquisition, and deployment process for high-end computing and the associated ecosystem of infrastructure that supports it.

That creates a lot of uncertainty, not only among the people who are primarily—use computing in the narrow sense, but in the way that computing infrastructure supports the broader scientific enterprise. The data management, the analysis, the collaboration support that go with those other instruments is now inextricably intertwined with this infrastructure as well, and it has to be part of a larger planning process.

Chairman BOEHLERT. Thank you very much. The gentleman's time has expired. Ms. Biggert.

Ms. BIGGERT. Thank you, Mr. Chairman. Just one further step in that question. Dr. Marburger, if you could wave a magic wand—I know you would love to, but that hasn't been invented yet, and—what are the two or three changes that you would make now to strengthen the interagency coordination?

Dr. MARBURGER. That is a dangerous question for me to answer. I think first of all, we have unprecedented cooperation among the agencies. There has been tension about strategies. People were concerned about competition for resources, and perhaps losing control over their own assets, and losing leverage over certain types of applications or designs. And much of those tensions have dissipated during the past year, during the activities of this committee.

The most important thing is to maintain engagement at a sufficiently high level within the agencies to make budget decisions and resource allocation decisions. And the first thing that I would just like to do is to make sure that the relevant agencies are engaged at a sufficiently high level. I think they are, but that it is important to maintain that focus of leadership at the top.

Another magic wand would be simply to try to get everybody at the same level of awareness of some of the broader scale technical issues, like the differences among these types of computing facilities, the parallel versus the vector, and so forth. That would make it easier to discuss some of these things. I do believe that some differentiation in roles is absolutely essential. I think it is appropriate for the National Science Foundation to focus on connectivity, as it is doing, the Department of Energy to focus on major facilities, as it is doing, and then other agencies to be users of those capabilities in appropriate ways.

I wish I could reduce the barriers to resource flow among the agencies. That is a very—that has always been a serious problem for coordinating science programs across agencies. There are semi-permeable membranes to the flow of funds and resources across the agencies. Fortunately, there is no barrier to agencies planning together, and trying to overcome these membranes, as it were, that separate them, to overcome them from the top down, as they plan their activities.

These are just some thoughts that come to mind.

Ms. BIGGERT. If DOE provides new computer resources to academic researchers not associated with DOE, will that complement or partially substitute for what NSF is—currently provides through its computer centers?

Dr. MARBURGER. I think that the facilities that DOE tends to provide are unique. It is necessary for NSF-supported researchers to have access to them. The model of the Synchrotron Light Sources and other accelerator-based user facilities at the Department of Energy Laboratories is a very good one for supercomputing. It is one that we have in mind and would like to support. So, operating these new facilities as if they were accelerators is a good model. NSF does provide connectivity, currently, that the Department of Energy laboratories take good advantage of. And that can continue.

So, I don't see any insuperable barriers to the model that is being proposed here. There are some minor difficulties, but I believe they can be worked out.

Ms. BIGGERT. Thank you. And before I proceed to other questions with the panel, I just wanted to state for the record that I did not pack this panel, that I did not know that everyone was from the Midwest at the time, but I am very glad that you are all here. This is for Dr. Stevens and Dr. Reed.

How do you anticipate that academic researchers would react to DOE taking on a greater role in providing university researchers with access to the high-performance computers?

Dr. STEVENS. I think they would react positively under the following condition, if they had confidence that the allocation on those systems was a peer review process of the highest standard, number one.

Number two, if the systems provided were really unique, that is, one of the challenges that we have been talking about here is the need for multiple architectures. For example, the recent announcement for DOE actually talks about two or three architectures ultimately being deployed. If those were pushed to the extremes over the next decade, they would be quite different from each other.

The Cray system and the IBM system are on two very different paths. If they were pushed to the extreme and at very large scale, much larger than say, what could be deployed and supported at a university site, then they would be truly national resources. They would be unique. They would complement what university groups can have, and with appropriate peer review, I think those are the combinations for success.

Chairman BOEHLERT. Thank you. The gentlelady's time has expired.

Ms. BIGGERT. Could Dr. Reed answer——

Chairman BOEHLERT. Dr. Reed, did you want to respond?

Ms. BIGGERT. Yes.

Dr. REED. Just briefly.

Chairman BOEHLERT. I want to give North Carolina equal time.

Dr. REED. Well, I do—I agree with what Rick Stevens said. I think that the other aspect to bear in mind is that computing is part of this ecosystem that connects instruments, and they are being developed by multiple agencies, and so the notion that one agency provides sole access, I think, has to recognize the ground truth that we need to bring this broad infrastructure together, and that is important.

Chairman BOEHLERT. Thank you very much. Ms. Lofgren. Or Ms. Woolsey, I am sorry. Zoe was here.

Ms. WOOLSEY. She was here, but I was here before her.

Chairman BOEHLERT. Well, all right. And you are still here.

Ms. WOOLSEY. Pay attention, Mr. Chairman.

Chairman BOEHLERT. So you get——

Ms. WOOLSEY. I get 10 minutes.

Chairman BOEHLERT. No.

Ms. WOOLSEY. Hers and mine. First of all, thank you, panel. As usual, the panelists that, even if they are all from the Midwest, are always wonderful, that are picked by this committee, and by Ms. Biggert herself.

I would like to acknowledge Dr. Wladawsky-Berger, and—for being named the 2001 Hispanic Engineer of the Year.

Dr. WLADAWSKY-BERGER. Thank you.

Ms. WOOLSEY. And being a native of Cuba. And I want to thank you for using your brilliance and vision and intellect here in our country, in the United States.

Dr. WLADAWSKY-BERGER. Thank you.

Ms. WOOLSEY. Thank you very much. Says something about immigration, doesn't it, folks?

With hardware and software being a main emphasis in this nation, and with our need for paying more attention—paying the right amount of attention to supercomputing, I want to ask you if—about the role of the telecom industry, the wireless industry, the fiber-optics industry. I want to know if we are putting enough—investing enough in that industry, because wouldn't it—I mean, I think I am right that it would certainly impede the get-along with supercomputing, yes, if we hold it up through that industry.

So, I am asking you, are we doing enough with the infrastructure, the telecom infrastructure, with research and development in that area, and are we funding this research appropriately? Just any one of you, starting with Dr. Wladawsky——

Dr. WLADAWSKY-BERGER. Well, if I may start, I mean, I can't say enough about the importance of broadband to everything, to our security, to our economic competitiveness, to healthcare, to education. What the Internet taught us is how much more valuable all this technology is when all the pieces are connected with each other using open standards, than in the old days, not too long ago, when they were all separate and they were not connected.

And as for trying to take these supercomputing capabilities, and make them available everywhere, from the very largest to scientists, to others in a more commercial world, the more we have broadband, line-connected and wireless, the much more valuable it is going to be.

Let me just give one little example. We are working, in IBM, with a small company that is developing some very innovative approaches to detecting skin cancers. They have some tools that are noninvasive that analyze the skin, but then that information gets transmitted over broadband in real time to some commercial supercomputer centers—in this case we partner with them—it analyzes that, gets back the answer in real time, and now the combination of the supercomputers with the broadband is helping whole new applications that you just would never have been able to do otherwise.

Ms. WOOLSEY. Well, all right. Are we doing enough in that direction? I mean, I feel like we——

Dr. STEVENS. Well——

Ms. WOOLSEY. Go ahead, Dr. Stevens.

Dr. STEVENS. Let me try to take a stab at it.

Let me just make an observation. The Earth Simulator. If you want to use that, you fly to Japan, and you sit—you go into a building, and you sit there with your Japanese colleagues and type directly at the machine. That machine is not on the network, okay.

If Japan decided to connect that machine to, say, a high-performance particle accelerator to analyze protein structures, they decide not to do that. In the U.S., deploying systems that way would be crazy, right. The NSF has recognized this in the TerraGrid project. Now, let us just play this picture forward a little bit in time. Today, we are deploying systems that are in the order of 10 to 100 teraflops. In five years, we will be deploying systems that are a petaflop, 10 to the 15th operations per second. If we are lucky, five years beyond that in the exaflops and so forth.

Ms. WOOLSEY. And you really think I know what that means, don't you?

Dr. STEVENS. It is this really big number.

Ms. WOOLSEY. Mrs. Biggert does. That is why she——

Dr. STEVENS. It is really big. But——

Ms. WOOLSEY.—gets—yeah——

Dr. STEVENS.—here is the point. The point is that in supercomputing, if you want to move the data between these machines, you need networks that will keep up. That means very soon, we will need terabit per second networks. Today, we do not have an aggressive R&D program to develop or deploy terabit networks to support interconnecting these supercomputing resources. So, in that sense, we are falling behind.

Ms. WOOLSEY. Dr. Reed.

Dr. REED. So let me just amplify those issues. I think there are several reasons why the answer to your question is yes, we need to do more. One has to do with the connectedness of individuals, and in a knowledge economy, our challenge is to allow people to work together, and that means exploiting the best intellectual talent across the country, regardless of location. And networking, high-speed networking, broadband networking is the way to do that.

It is also true that managing the large data volumes, whether it be for business and commercial applications or for scientific applications, we are a long way from where we need to be. The other part of the computer revolution that has produced large volumes of data, moving those to people for efficient analysis is a remaining challenge.

And so, how we break down those barriers of time and space, and connect everyday things to allow information to flow efficiently, we need—yes, we need an integrated program that couples that with the other aspects of computing.

Chairman BOEHLERT. The gentlelady's time has expired.

Ms. WOOLSEY. Thank you.

Chairman BOEHLERT. The Chair recognizes the distinguished gentleman from Michigan, Dr. Ehlers.

Mr. EHLERS. Thank you, distinguished Chairman. I would like to just talk about the hardware of that, and try and get a better understanding of that.

The—several of you commented on the need for major advances in hardware. Dr. Stevens, you said that you saw things such as global warming or drugs, et cetera, require "orders of magnitude more computing power." And I would like to get a better handle on that. For example, the life sciences, nanotechnology problems, can they be handled on the grid systems, such as the NSF TerraGrid or something similar? Are we talking about orders of magnitude of improvement in other ways? Where are we going in this whole field? And I am somewhat familiar with what the Japanese did, and recognize they are approaching their limits. Are we going to jump ahead, and how are we going to do it? And our—do we need further improvements in bandwidth and interconnection as well with this, or are we talking about more centralized computer facilities that you can access with ordinary broadband? A whole series of questions pop in my mind. I am not articulating them very well. Well, let us just go down the line from right to left, and get your comments. My right to your left. To my left.

Dr. REED. Sir, there are a whole series of problems that we can, in some sense, see solutions from here, but we can't get there at the moment. And let me give you an example of one in which I am involved now, that captures in a biological and biomedical sense a flavor of that.

I am involved with a group of researchers that span biology, chemistry, physics, and medicine that are trying to build a model, a virtual model, of a lung, to understand the effects of smoking, cystic fibrosis, cancer, how the interactions of it, at a physical science level, what is really a computational fluid dynamics model of air flow at the large, gross level in the lungs, down through intermediate structures, and how particles interact with surfaces, down to the bottom, where you are looking at a biophysics problem in understanding how cilia and mucus help eject materials, and then at the very bottom, the genetic basis of human variation.

That kind of interdisciplinary problem is the—if we can solve a problem like that and build an integrated model, we can get some deep insights into the effects of environment on health, the genetic susceptibility to various kinds of disease. But we don't have the computing capability to solve that problem right now. We can model small pieces of that problem. We are one, maybe two orders of magnitude from where we would need to be to be able to solve that problem.

So, the thing that I think—and I have said this a couple times—I think is really important in this domain is that there is no one single solution to this problem. If you look at the broad range of problems, we need leadership class computing systems, because there are some classes of applications that can only be solved with very tightly coupled, single-site systems. There are other kinds of critical problems that coupling distributed data archives and instruments with some intermediate but still high-performance computing capability will let us solve, and then there are others where even more mundane systems coupled together in the right ways give distributed groups of people the ability to solve problems.

But there is absolutely no doubt that there are science and economic benefits that we can see from where we are, if we had another order or two of magnitude and capability in high-performance systems, even in the centralized case.

Mr. EHLERS. Now, Dr. Stevens, one answer from that, from Dr. Reed was one or two orders of magnitude. Would you agree with that or are you looking further into the future in——

Dr. STEVENS. I like to look further in the future, of course. We need the one or two orders of magnitude to solve the problem Dan is talking about, but of course, as soon as we can solve that problem, we will want to ask——

Mr. EHLERS. Yeah.

Dr. STEVENS.—deeper questions, like gee, if we can build a virtual lung, why can't we build a virtual human, and now understand what happens, instead of doing drug testing on people, we can do drug testing, say, for drug interactions or whatever, on this virtual human, and maybe we can build virtual children, because we don't tend to do drug testing on children today, even though it is an important problem for the pharmaceutical industry. So, there are all kinds of things that I think we will find that we want to do, beyond this one or two orders of magnitude.

Mr. EHLERS. Now, let me be a little more specific. Obviously, we are—we don't have unlimited financial resources at the Federal Government. Where should our efforts go in order to get those one, two, or three orders of magnitude?

Dr. STEVENS. So——

Mr. EHLERS. What approach should we be taking?

Dr. STEVENS. In the near-term, we need to exploit the architectures that we know that work, and we need to scale them up to the practical limits of that technology. So in the case of vector processors, the recently announced DOE program is a good start. In terms of these embedded designs, system on a chip designs, which the IBM Blue Gene machine is another one. We know practically where we can take that, and it will scale over maybe another couple of orders of magnitude.

To go beyond that, we need to do fundamental R&D in some new technologies, okay, and here is a couple of technologies that we need to work on. One is that we need to make hardware more flexible. And what does that mean? It means right now, the hardware that we use to build these computers is sort of fixed at the factory. One idea is to make that hardware less fixed at the factory, so that each application can reconfigure the hardware to be more efficient. That is one idea we need to test, and if we can find that works in the small, we need to see if it can work in the large.

Another idea is optics, improving the ability to go to optics directly onto the chips, so that we don't have to use copper wires in the middle of these machines any more. We can do many thousands of optical fibers off of a single chip. That will give us enormous flexibility in terms of network topologies and improving bandwidth.

Finally, we are going to reach limits with lithography. We are at 90, 60 nanometers currently, and within the next decade, we will be down to feature sizes that start to approach single molecule sizes, and so we need to leverage research in, say, molecular transistors, to figure out how can we make these systems several orders of magnitude smaller than they are now, and still get performance at reasonable power densities.

So those are some examples.

Mr. EHLERS. And before we go to the next one, I would just point that is why we need more money for the National Institute of Standards and Technology, to help with the lithographic process. So, Mr. Chairman——

Chairman BOEHLERT. Amen, amen. Next—Mr.——

Dr. WLADAWSKY-BERGER. Let me talk about efforts we have going at IBM, and as my colleagues have said throughout, there is no one single architecture that works on everything, so we have some programs that are aimed at building the highest performance microprocessors you can, and aggregate them in large numbers.

There is another program, which is Blue Gene, where we want to aggregate them in huge numbers, in fact, the Blue Gene that is going to Lawrence Livermore Lab in 2005 will have 2,000—65,000 microprocessors, and to do that, you want to use low-cost microprocessors that don't use too much power, so you can aggregate them in large numbers. And that is a very good example of the innovation ahead of us.

How can you design the most powerful supercomputers possible at the most affordable cost possible? The approach we are taking is to use essentially commercial components, and then add a lot of value around them, so you can aggregate them in larger and larger and larger numbers. Let us remember that human beings, like all organisms, are built out of commodities, cells, but by the time you get to higher organisms, let alone human beings, I don't think we are commodities. I think some very exquisite things we hope have happen to be able to aggregate all of those components.

That is a lot of the excitement of future designs, to push orders of magnitude into the future, that will take tremendous R&D, and it would also take a lot of understanding of the applications.

Chairman BOEHLERT. The gentleman's time has expired. Ms. Jackson Lee.

Ms. JACKSON LEE. Thank you very much, Mr. Chairman, and let me commend you for a series of very effective and important hearings. And if I might, just very briefly indulge me. Yesterday, Mr.

Chairman, I was detained during the hearing of H.R. 4107, the Assistance to Firefighters Act. I was in judiciary markup with a number of my own bills before the Committee. And I just wanted to take a moment before I pose questions to those gentlemen to—first of all, say to you that I look forward to working with you on this legislation, because I am on the Homeland Security Committee with you, and I know your interest and your commitment.

I want to raise two points, and in my study of the bill, I am still studying it, I am a chauvinist on Homeland Security. I believe it is an important aspect of our work, but I am also concerned that our firefighters, who all of us know are probably best served by being—keeping those fire grants in the U.S. Fire Administration. I raise that point, and hope that we will continue to work through that issue.

And the other point would be that we clarify the very valued aspect of the legislation dealing with volunteer firefighters, and maintain, however, the credibility of things like—in my community, we have things like meet and confer. I think that makes us feel better than maybe if we hear some other words, but the whole concept of collective bargaining, you coming from New York, I know you fully appreciate and understand that, but we can work as partners together on this. The first people I called in, being able to get home to Houston after 9/11 were my firefighters, and we huddled in a meeting for a long period of the day, and so I know that they are eager to work with H.R. 4107, and this legislation.

I am eager to work with you on this as well, and wanted to make mention of that, and wanted to give my apology for being detained in——

Chairman BOEHLERT. Thank you. I, too, was detained in an Intelligence Committee meeting of—dealing with the abuse allegations in Iraq, so those are other subjects for other discussions, but I will be glad to work with gentlelady, who has two minutes and 40 seconds left.

Ms. JACKSON LEE. Thank you very much, Mr. Chairman. Let me just say that the statement, I think, that is most clear is that high-performance computing in the United States is at a turning point, because we all know that Japan has the either smarter one or the faster one. Dr. Marburger, if you can just say to me what that does to NASA, what that does to our educational desires in that area, and can we get the Administration's full support in helping us with the request for increased funding, not only to get equal with Japan but to get ahead of them, particularly as relates to producing more of our scientists who can engage us in this research.

Dr. MARBURGER. This Administration does place high-performance computing very high on its priorities. We intend to continue to follow this. We strongly support the bill. It has been introduced by this committee, and we look forward to working together to get the resources necessary to maintain our leadership in the computing area.

Ms. JACKSON LEE. Are we disadvantaged at NASA by not having a computer of that level?

Dr. MARBURGER. Are we disadvantaging NASA? I would say that the NASA programs are the world's leaders in the areas for which they have responsibility, and that that leadership position of NASA is not currently in jeopardy.

Ms. JACKSON LEE. And what about Homeland Security, which is one of the crux of our concerns?

Dr. MARBURGER. I do not believe that Homeland Security is jeopardized by any current program or proposal. I think the Homeland Security computing needs are being addressed. There are foreseeable applications in the future, not currently being conducted, that could benefit from the types of computing architectures being discussed here.

Ms. JACKSON LEE. But the Administration is supportive of increased funding to help us develop this technology?

Dr. MARBURGER. The Administration supports adequate funding for maintaining our leadership in all of these areas.

Ms. JACKSON LEE. All right. Then we probably have a disagreement there. I think we need increased funding. Dr. Wladawsky-Berger. Help me out. How are we being disadvantaged by not having the technology that we need, or being competitive with Japan in terms of the type of supercomputer that we need, high-performing computer?

Dr. WLADAWSKY-BERGER. Let me say, I believe Japan right now has the fastest computer, but in reality, I really do believe the U.S. is way ahead of the rest of the world——

Ms. JACKSON LEE. Good news.

Dr. WLADAWSKY-BERGER.—in the use of supercomputing, and in particular, in the widespread use of supercomputing. Now, it is not enough, and the reason it is not enough is because the opportunities are so much bigger throughout society, whether it is applied to healthcare, to education, to economic development, to financial services, and of course, to national security. There is probably no problem that cannot be made better by the judicious use of information analysis and simulation, and that is why we believe we need to do so much more, because we all believe, I think, this is the key to competitiveness and national security.

Ms. JACKSON LEE. Thank you.

Chairman BOEHLERT. The gentlelady's time has expired.

Ms. JACKSON LEE. Thank you.

Chairman BOEHLERT. Mr. Sherman.

Mr. SHERMAN. Thank you, Mr. Chairman. I want to thank Dr. Marburg—I am going to mispronounce your name—for

—Marburger, for his speech of December 3, where he focused on the important provisions that our committee wrote dealing with nanotechnology and the importance of looking at the societal implications.

I think that we, as a species, are faced with three related technologies, supercomputing, nanotechnology, and genetic engineering, that I would refer to as a reverse Pandora's box. You remember Pandora's box. Every evil was in that box, and one embodiment of hope. I think these three technologies offer us the reverse. Every kind of hope, and one or two unspeakable evils.

The concern I have is—and I have expressed this to my colleagues on the Committee—the creation of new intelligent life forms through either of two paths, perhaps converging paths. One would be through artificial intelligence, supercomputing and the related software, in effect, a new silicon life form, if you will, although I am told that supercomputing levels, ultimately, you will be using a different substrate than silicon.

And the other would be through genetic engineering. I have asked some of my constituents whether they think their kids will compete successfully on the LSAT with an 800 pound being with four 50 pound brains. Some of the more confident have told me their kid will still do better on the LSAT. I am not so sure. I have met their kids.

Anyway, I—these technologies will interrelate. Supercomputing will obviously help genetic engineering. Nanotechnology and the biosciences may allow us to reverse engineer the human brain, to turn our artificial—to turn our computers into artificial intelligence should we decide to do that, a big if.

And I know that there is a tendency to think that artificial intelligence is separate from supercomputing, because while supercomputing might grind out more calculations, it doesn't come with the new software architecture, but I think what we have seen is that if you get enough computer power, you can do amazing things with really weak software, and/or barely adequate software. So, I don't think that we can regard new software as separate from new hardware. The two will work together.

Dr. Marburger, I know you speak for the President at—how close are we to a machine that has reached a level of intelligence where it would be entitled to the minimum wage?

Dr. MARBURGER. Not very. We are quite far from that. In terms of just the numbers of components measured by neurons, for example, the interconnectivity of the human brain far exceeds anything that we can currently build or foresee in the immediate—in the foreseeable future with computer hardware. But we have three experts here who are closer to this field than I am, and I think we should hear from them, if you——

Mr. SHERMAN. I—let me ask that not in terms of—I know it is not in the foreseeable future. It won't happen during the Kerry Administration. Sorry, I had to say that. But do we expect this, and keep—it is so hard to predict, because you are predicting an accelerating process, while the Internet connects, and a growing number of scientists working on a growing number of projects, using new tools. The computers get smarter, the—you build one on the other. But I will ask all three panelists. Are we talking 25 years, 50 years, 100 years?

Dr. WLADAWSKY-BERGER. Well, let me start—the reality, I think, is we don't know. Now, I think that at least in the foreseeable future, the real danger is not that advanced computers will have evil intent, but they could frustrate the—a lot by just not working well, because of the complexity of managing the incredibly large infrastructure. I mean, look at your PC. It may or may not be evil, but God, how many times does it frustrate you, because it is not——

Mr. SHERMAN. Well——

Dr. WLADAWSKY-BERGER.—doing what you want it to do.

Mr. SHERMAN. Yeah. I think there is frustration in our foreseeable future. I—before we get to Dr. Stevens, because I know my time is about to elapse, if it hasn't already. But one argument is made that even if there was a self-aware computer, we would have total control, because it couldn't act in the physical world without human beings running around doing its bidding, and I would simply say that I know several people that would give hands to the Devil in return for a good stock tip.

Dr. WLADAWSKY-BERGER. Yeah. Let me just add that one of the hallmarks of good research is to anticipate problems. It is not just—I am sorry—to create things, but to anticipate the negative implications of what we are creating, whatever it is. Right now, we are all very worried about the complexity of managing and programming, and you are bringing up some farther out issues, and one of the reasons we all are so strong in supporting fundamental research is because that is how you anticipate problems and start working on their solutions, way ahead of the time those problems hit us.

Mr. SHERMAN. Dr. Stevens, I don't know—if the Chairman will indulge me, I would like your response as well.

Dr. STEVENS. Well, I mean, it is a fascinating topic. And my personal view is that I would be much more concerned with near-term issues associated with large-scale computing, either this frustration issue, or the use of large-scale data systems to collect information and—that may be used for purposes that the people whose information it is is not in agreement with, whether that is for privacy or other purposes.

So, I think we are on a path to build a large-scale cybernetic structure on this planet. That is the destiny of where connecting millions of computers and devices will go. We have no idea how to program that system in a way that would exhibit intelligent behavior currently. And as you pointed out, we have demonstrated through projects like the Deep Blue at IBM that relatively straightforward algorithms can, in fact, exceed human performance in very constrained activities.

I would like to see some of those activities to be used for good purposes, and to apply simpler intelligences to, sort of, you know, be used instead of troops in battle or whatever, that may provide benefit in the near-term before we achieve dramatic intelligence.

Just a final comment.

Mr. SHERMAN. I would also point out, though, that our—we will have the ability at some point to reverse engineer the human brain, and that that ability will be enhanced by the supercomputing capacity——

Dr. STEVENS. We—absolutely——

Mr. SHERMAN.—and the increased capacity we have for brain scans.

Dr. STEVENS. Absolutely. And in fact, it will not be possible to reverse engineer the brain, or any large, complex biological system, without advanced computing, okay. That is clear. Right now, if you had to estimate what is the most intelligent device we can build, it is roughly between a worm and an insect in terms of what it can do.

Chairman BOEHLERT. On that closing note——

Mr. SHERMAN. Thank you.

Chairman BOEHLERT. The gentleman's time has expired. Thank you for tickling our fancy, so to speak, and giving us food for thought for the future, and thank all of the witnesses for your very productive testimony and for being resources to this committee. The hearing is adjourned.

[Whereupon, at 12:15 p.m., the Committee was adjourned.]

Appendix:

Additional Material for the Record

I

108TH CONGRESS
2D SESSION

H. R. 4218

To amend the High-Performance Computing Act of 1991.

IN THE HOUSE OF REPRESENTATIVES

APRIL 27, 2004

Mrs. BIGGERT (for herself, Mr. DAVIS of Tennessee, Mr. BOEHLERT, and Mr. JOHNSON of Illinois) introduced the following bill; which was referred to the Committee on Science

A BILL

To amend the High-Performance Computing Act of 1991.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. SHORT TITLE.

This Act may be cited as the "High-Performance Computing Revitalization Act of 2004".

SEC. 2. DEFINITIONS.

Section 4 of the High-Performance Computing Act of 1991 (15 U.S.C. 5503) is amended—

(1) in paragraph (2), by inserting "and multidisciplinary teams of researchers" after "high-performance computing resources";

(2) in paragraph (3)—

(A) by striking "scientific workstations,";

(B) by striking "(including vector supercomputers and large scale parallel systems)";

(C) by striking "and applications" and inserting "applications"; and

(D) by inserting ", and the management of large data sets" after "systems software";

(3) in paragraph (4), by striking "packet switched"; and

(4) by amending paragraphs (5) and (6) to read as follows:

"(5) 'Program' means the High-Performance Computing Research and Development Program described in section 101; and

"(6) 'Program Component Areas' means the major subject areas under which are grouped related individual projects and activities carried out under the Program.".

SEC. 3. HIGH-PERFORMANCE COMPUTING RESEARCH AND DEVELOPMENT PROGRAM.

Title I of the High-Performance Computing Act of 1991 (15 U.S.C. 5511 et seq.) is amended—

(1) in the title heading, by striking "**AND THE NATIONAL RESEARCH AND EDU-**

CATION NETWORK" and inserting "**RESEARCH AND DEVELOPMENT**";

(2) in section 101—

(A) the section heading, by striking "**NATIONAL HIGH-PERFORMANCE COMPUTING**" and inserting "**HIGH-PERFORMANCE COMPUTING RESEARCH AND DEVELOPMENT**";

(B) in subsection (a)—

(i) in the subsection heading, by striking "NATIONAL HIGH-PERFORMANCE COMPUTING" and inserting "HIGH-PERFORMANCE COMPUTING RESEARCH AND DEVELOPMENT";

(ii) by striking paragraphs (1) and (2) and inserting the following: "(1) The President shall implement a High-Performance Computing Research and Development Program, which shall—

"(A) provide for long-term basic and applied research on high-performance computing;

"(B) provide for research and development on, and demonstration of, technologies to advance the capacity and capabilities of high-performance computing and networking systems;

"(C) provide for sustained access by the research community in the United States to high-performance computing systems that are among the most advanced in the world in terms of performance in solving scientific and engineering problems, including provision for technical support for users of such systems;

"(D) provide for efforts to increase software availability, productivity, capability, security, portability, and reliability;

"(E) provide for high-performance networks, including experimental testbed networks, to enable research and development on, and demonstration of, advanced applications enabled by such networks;

"(F) provide for computational science and engineering research on mathematical modeling and algorithms for applications in all fields of science and engineering;

"(G) provide for the technical support of, and research and development on, high-performance computing systems and software required to address Grand Challenges;

"(H) provide for educating and training additional undergraduate and graduate students in software engineering, computer science, computer and network security, applied mathematics, library and information science, and computational science; and

"(I) provide for improving the security of computing and networking systems, including Federal systems, including research required to establish security standards and practices for these systems.";

(iii) by redesignating paragraphs (3) and (4) as paragraphs (2) and (3), respectively;

(iv) in paragraph (2), as so redesignated by clause (iii) of this subparagraph—

(I) by striking subparagraph (B);

(II) by redesignating subparagraphs (A) and (C) as subparagraphs (D) and (F), respectively;

(III) by inserting before subparagraph (D), as so redesignated by subclause (II) of this clause, the following new subparagraphs:

"(A) establish the goals and priorities for Federal high-performance computing research, development, networking, and other activities;

"(B) establish Program Component Areas that implement the goals established under subparagraph (A), and identify the Grand Challenges that the Program should address;

"(C) provide for interagency coordination of Federal high-performance computing research, development, networking, and other activities undertaken pursuant to the Program;"; and

(IV) by inserting after subparagraph (D), as so redesignated by subclause (II) of this clause, the following new subparagraph:

"(E) develop and maintain a research, development, and deployment roadmap for the provision of high-performance computing systems under paragraph (1)(C); and"; and

(v) in paragraph (3), as so redesignated by clause (iii) of this subparagraph—

(I) by striking "paragraph (3)(A)" and inserting "paragraph (2)(D)";

(II) by amending subparagraph (A) to read as follows:

"(A) provide a detailed description of the Program Component Areas, including a description of any changes in the definition of or activities under the Program Component Areas from the preceding report, and the reasons for such changes, and a description of Grand Challenges supported under the Program;";

(III) in subparagraph (C), by striking "specific activities" and all that follows through "the Network" and inserting "each Program Component Area";

(IV) in subparagraph (D), by inserting "and for each Program Component Area" after "participating in the Program";

(V) in subparagraph (D), by striking "applies;" and inserting "applies; and";

(VI) by striking subparagraph (E) and redesignating subparagraph (F) as subparagraph (E); and

(VII) in subparagraph (E), as so redesignated by subclause (VI) of this clause, by inserting "and the extent to which the Program incorporates the recommendations of the advisory committee established under subsection (b)" after "for the Program";

(C) in subsection (b)—

(i) by redesignating paragraphs (1) through (5) as subparagraphs (A) through (E), respectively;

(ii) by inserting "(1)" after "ADVISORY COMMITTEE.—";

(iii) in paragraph (1)(C), as so redesignated by clauses (i) and (ii) of this subparagraph, by inserting ", including funding levels for the Program Component Areas" after "of the Program";

(iv) in paragraph (1)(D), as so redesignated by clauses (i) and (ii) of this subparagraph, by striking "computing" and inserting "high-performance computing and networking"; and

(v) by adding at the end the following new paragraph:

"(2) In addition to the duties outlined in paragraph (1), the advisory committee shall conduct periodic evaluations of the funding, management, coordination, implementation, and activities of the Program, and shall report not less frequently than once every two fiscal years to the Committee on Science of the House of Representatives and the Committee on Commerce, Science, and Transportation of the Senate on its findings and recommendations. The first report shall be due within one year after the date of enactment of this paragraph."; and

(D) in subsection (c)(1)(A), by striking "Program or" and inserting "Program Component Areas or"; and

(3) by striking sections 102 and 103.

SEC. 4. AGENCY ACTIVITIES.

Title II of the High-Performance Computing Act of 1991 (15 U.S.C. 5521 et seq.) is amended—

(1) by amending subsection (a) of section 201 to read as follows:

"(a) GENERAL RESPONSIBILITIES.—As part of the Program described in title I, the National Science Foundation shall—

"(1) support research and development to generate fundamental scientific and technical knowledge with the potential of advancing high-performance computing and networking systems and their applications;

"(2) provide computing and networking infrastructure support to the research community in the United States, including the provision of high-performance computing systems that are among the most advanced in the world in terms of performance in solving scientific and engineering problems, and including support for advanced software and applications development, for all science and engineering disciplines; and

"(3) support basic research and education in all aspects of high-performance computing and networking.";

(2) by amending subsection (a) of section 202 to read as follows:

"(a) GENERAL RESPONSIBILITIES.—As part of the Program described in title I, the National Aeronautics and Space Administration shall conduct basic and applied research in high-performance computing and networking, with emphasis on—

"(1) computational fluid dynamics, computational thermal dynamics, and computational aerodynamics;

"(2) scientific data dissemination and tools to enable data to be fully analyzed and combined from multiple sources and sensors;

"(3) remote exploration and experimentation; and

"(4) tools for collaboration in system design, analysis, and testing.";

(3) in section 203—

(A) by striking subsections (a) through (d) and inserting the following:

"(a) GENERAL RESPONSIBILITIES.—As part of the Program described in title I, the Secretary of Energy shall—

"(1) conduct and support basic and applied research in high-performance computing and networking to support fundamental research in science and engineering disciplines related to energy applications; and

"(2) provide computing and networking infrastructure support, including the provision of high-performance computing systems that are among the most advanced in the world in terms of performance in solving scientific and engineering problems, and including support for advanced software and applications development, for science and engineering disciplines related to energy applications."; and

(B) by redesignating subsection (e) as subsection (b);

(4) by amending subsection (a) of section 204 to read as follows:

"(a) GENERAL RESPONSIBILITIES.—As part of the Program described in title I—

"(1) the National Institute of Standards and Technology shall—

"(A) conduct basic and applied metrology research needed to support high-performance computing and networking systems;

"(B) develop benchmark tests and standards for high-performance computing and networking systems and software;

"(C) develop and propose voluntary standards and guidelines, and develop measurement techniques and test methods, for the interoperability of high-performance computing systems in networks and for common user interfaces to high-performance computing and networking systems; and

"(D) work with industry and others to develop, and facilitate the implementation of, high-performance computing applications to solve science and engineering problems that are relevant to industry; and

"(2) the National Oceanic and Atmospheric Administration shall conduct basic and applied research on high-performance computing applications, with emphasis on—

"(A) improving weather forecasting and climate prediction;

"(B) collection, analysis, and dissemination of environmental information; and

"(C) development of more accurate models of the ocean-atmosphere system."; and

(5) by amending subsection (a) of section 205 to read as follows:

"(a) GENERAL RESPONSIBILITIES.—As part of the Program described in title I, the Environmental Protection Agency shall conduct basic and applied research directed toward advancement and dissemination of computational techniques and software tools for high-performance computing systems with an emphasis on modeling to—

"(1) develop robust decision support tools;

"(2) predict pollutant transport and the effects of pollutants on humans and on ecosystems; and

"(3) better understand atmospheric dynamics and chemistry.".

○

PREPARED STATEMENT OF BOB BISHOP

TO OUT-COMPETE IN THE 21ST CENTURY, U.S. INDUSTRY MUST OUT-COMPUTE

BOB BISHOP
CHAIRMAN AND CHIEF EXECUTIVE OFFICER, SGI

HOW DO U.S. COMPANIES DEPLOY HIGH-PERFORMANCE COMPUTING AND HOW DOES IT AFFECT U.S. INDUSTRIAL COMPETITIVENESS?

The role of HPC in U.S. industry today is to solve technical problems quickly, gain insight into design alternatives, and bring safe and secure products and services to the market early, thus creating competitive advantage and improving the quality of our daily lives. HPC stimulates global competition, then helps companies compete in fiercely competitive markets.

HPC can also be seen as a nerve center within the corporate setting and a conduit to cross-functional thinking. It brings together specialists from different fields who, by interaction with each other, rapidly improve their understanding, insight and problem solving in matters of great complexity. HPC eliminates stovepipe thinking.

Managers and specialists leverage each other's knowledge in such an environment, asking multiple "what if" questions, evaluating countless scenarios while accelerating cooperative decision-making along the way. As a consequence, enterprise level strategy and tactics are broadened and strengthened.

HPC in U.S. industry today is not compute-only activity conducted in glass-house isolation. HPC centers are connected via high-speed lines to other geographically dispersed decision centers both inside and outside of the enterprise. HPC may also direct-connect with laboratory instruments, sensor networks, satellite feeds or real-time video signals. In fact, it is increasingly common to find rich media from multiple sources "fused" into a single image, overlaying locally generated graphics, and effectively granting "X-ray vision" to all participants in the HPC session. In this way, HPC becomes a tool for superior decision-making.

Increasingly, HPC drives a creative food chain, from innovation to operations, and increasingly delivers interactive real-time solutions. Speed and innovation are critical in the corporate race for global success.

Leading U.S. industries have aggressively adopted HPC to improve their productivity and competitiveness. Defense, aerospace, automobile, chemical, pharmaceutical, medical, energy and media lead the way. Other U.S. industries are adopting HPC at a more modest rate.

Worldwide deployment of HPC is found in similar industry sectors, especially in Japan, Germany, France, and the UK. China and India are beginning to rapidly adopt HPC as well. The U.S. remains by far the most predominant supplier of HPC products however, for both hardware and software. Japan is the only other significant HPC equipment maker.

Leading-edge developments and break-through ideas in modern industries require high levels of modeling, simulation, visualization, and life-cycle data management. Biotechnology, nanotechnology and material science, for example. Vast amounts of intellectual property and future wealth are created in the process. Competitors strive to out-gun each other with in-house HPC capability, and win the right to patent, copyright and trade mark their knowledge.

HPC must be understood however, not as a single technology, but as an ecosystem of multiple technologies, each with its own set of issues and challenges: fast processors, complex memory hierarchies, interconnect fabrics, massive storage facilities, high-fidelity visualization, networking, and multi-layered software, to name but a few. A single weak factor will likely reduce the overall effectiveness of any HPC installation, dramatically.

HPC buyers must judiciously balance and combine HPC sub-system technologies appropriate to the real-world problems that they are attempting to solve. Even then, buyers need to continuously stay abreast of updates and developments, keeping their facilities relevant and at the leading edge, if they wish to survive.

Attracting and retaining talent to run an effective HPC facility is difficult for most U.S. corporations, especially with the recent dearth of computer science graduates emerging from advanced engineering schools.

Perhaps this is because corporate IT spending in the recent past has been dominated by business-process applications á la enterprise resource planning (ERP), customer relationship management (CRM), Internet deployment and mobile computing.

Such applications have improved the background context in which all corporations must operate. However, spending in these areas has not helped the core HPC user, except in the few cases where commercial technologies can be successfully re-purposed within the HPC mission. For example, Internet technology is useful for everyone, technical and commercial, as is the PC, the PDA and the cell phone. These latter devices however, are mostly used as access mechanisms to remote HPC resources, and do not constitute HPC technology in its own right.

The annual spending of U.S. corporations on business-process applications is one hundred-fold greater than that spent on engineering and scientific applications. With few exceptions, computer vendors are therefore attracted to the commercial side.

To help spread the adoption of HPC within U.S. private industry more broadly, and to help ensure more U.S. government and U.S. industry interchange in the future evolution of this critical capability, the Washington, DC-based organization "Council on Competitiveness" has recently begun a High-Performance Computing Initiative. I am privileged to serve on this Council's Executive Committee, and would encourage the Chairman of the House Committee on Science and its Members to be in contact with this effort. The Initiative is gathering data that will provide a timely and accurate profile of key HPC users, application areas and bottlenecks experienced in U.S. industry today.

This data will also highlight the multitude of factors that determine private industry HPC deployment in the U.S., including application software availability, ease-of-use, total-cost-of-ownership for equipment and personnel, and return-on-investment to the buyer.

As for U.S. computer vendors, in the absence of significant HPC volume procurements by corporations, it is difficult for them to focus solely on industry HPC markets. Hence U.S. computer vendors generally concentrate their product developments on the larger business-process markets, positioning their HPC activities as a minor sideline. Alternatively, they will repackage their commercial machines for technical purposes. Neither approach however, will allow HPC to reach its full potential. The market requires U.S. Government HPC procurement in steady volume to sustain strong U.S. HPC capability. This U.S. Government additional volume is especially critical to the health and survival of the few computer vendors that remain alive and dedicated to HPC today.

WHAT ARE SOME OF THE CURRENT HPC EFFORTS OF THE FEDERAL CIVILIAN SCIENCE AGENCIES? ARE THEY SUFFICIENT TO ENSURE U.S. LEADERSHIP IN HPC?

Recent events have conspired to raise alarm that the U.S. HPC industry has fallen behind its foreign rivals. For example, the powering on of Japan's Earth Simulator in March 2002, was a "Sputnik-like event," overshadowing all HPC machines on the planet. As of today, this machine is still at the head of the Top 500 Supercomputing Sites, as last published in November 2003. The machine is optimized for geoscience applications, and is front-ended by three Onyx machines supplied by Silicon Graphics Inc (SGI) that convert its numerical output to interactive immersive high-fidelity visualization. You can't drink from a firehose!

The ES–40 (Earth Simulator-40 Teraflops) price-tag exceeded $300 million, excluding the elegant new buildings in which it is housed. It was paid for by the Japanese Government and built by NEC along the lines of its SX–6 machine, a clustered-vector architecture in its sixth generation.

This is an outstanding example of government-industry cooperation in open science, but not necessarily a good example of HPC innovation or good HPC architecture. It is certainly a shining example of what money can buy. However, the recently installed ALTIX supercomputers at Tokyo University's Earthquake Research Center run several earthquake codes at similar speeds as the ES–40 at a much lower price. The two ALTIX machines, supplied by SGI, incorporate the latest Itanium 2 processor technology from Intel, the Linux operating system, and SGI's fourth generation global shared memory NUMAflex architecture.

Within the U.S. Government, the National Weather Service (NWS), the National Center for Atmospheric Research (NCAR), and the Geophysical Fluid Dynamics Laboratory (GFDL) belonging to National Oceanographic and Atmospheric Administration (NOAA) are already heavy HPC users. All of these centers however, would benefit greatly from additional HPC capability, given the importance of weather in our daily lives and given the difficulty of weather science. Severe weather continues to wreak havoc in many areas of the U.S., and the cost of more accurate weather modeling and forecasting capability pales in comparison to the damages caused by unforeseen weather events. The cost of hurricane evacuation alone on the Atlantic seaboard exceeds $1 million per coastal mile, or $100 million in the case of a hurricane that cannot be predicted to come ashore within one hundred miles. A 50 percent improvement in forecast accuracy would lower this cost by $50 million, provided it could be accomplished in a timely manner; enough to recover the cost of HPC equipment in a single event, and more importantly, saving lives along the way.

The key to solving problems in weather, climate and environmental science is HPC. Nature can only be accurately described and computed from equations that take account of complex non-linear interactions between multiple natural systems, i.e., rivers, lakes, oceans, mountains, forests, dust, pollution, cloud cover, snow cover, ice, polar regions, etc. Such equations of motion are so interconnected and intertwined that they can only be managed when all aspects are held in the global shared memory of a large HPC machine and computed simultaneously.

We have a similar experience at NASA's Goddard Space Flight Center and at the NASA Ames Research Center. Both are heavily committed to HPC and are driving their climate modeling programs to higher performance through extensive use of leading edge HPC. NASA Ames has in fact tuned their 512-pprocessor ALTIX machine to world record-breaking memory bandwidth performance (the first machine in HPC history to break one terabyte-per-second, as measured by the STREAMS Triad benchmark). Both NASA facilities will require much more HPC capability however, to achieve the Administration's recently announced Code T program consisting of a permanent Moon-colony and manned space flight to Mars. There is an opportunity here for NASA to build Moon and Mars simulators, along the lines of the Japanese Earth Simulator. Such simulators would be less difficult however, given that neither the Moon nor Mars has an active weather or tectonic system like the Earth.

There is also the need to design and simulate a new generation of spacecraft for the long voyages entailed. Moreover, since NASA's three space shuttles will most likely stop flying by the year 2010, the design of new generation space vehicles should begin very soon.

Human and Health Services (HHS) is yet another federal civilian science agency that must strongly encourage the deployment of HPC. Rapid recognition of pathogens and viruses and the development of their counter-acting vaccines is critical to public health. The recent global outbreaks of SARS, Ebola, Avian flu, and West Nile disease maybe an indication of worse to come. Rapid government response will only be achieved through HPC centers and laboratories that are globally connected.

Bio-terrorism is an additional threat for HHS to manage. Crisis management will ultimately require real-time modeling and simulation of toxin dispersion at the resolution of city streets and office buildings, at least in the top one hundred population centers of the U.S. These issues and others overlap with the newly formed Department of Homeland Security (DHS), which itself must become HPC capable to be fully effective.

The U.S. Department of Energy (DOE) has extensive experience in HPC, although mostly for weapons design and nuclear stockpile stewardship. HPC deployment however, is recently gaining momentum within DOE's Open Science program, and this is a very encouraging trend for the U.S. HPC community as a whole. DOE will play a critical role in guiding the Nation's future energy infrastructure and building alternative energy technologies. It also has extensive experience with environmental remediation. These are grand challenge problems that require significant HPC resources.

Generally speaking, there is a clearer recognition across the federal civilian agencies today that personal computers do not deliver the true horse-power of HPC machines, no matter how many units are networked together. One thousand bicycles do not make a truck! However, the low entry price of commodity clusters is often attractive for certain engineering and scientific applications, especially when these applications entail little inter-communications between the elements of the cluster. Even then, commodity clusters are only effective if there are no real-time interactivity requirements. Surprisingly however, the long-term total cost of ownership of a commodity cluster can be higher than expected if the full cost of maintenance, software licensing and system administration is taken into account.

Finally, the recent formation of a High-End Computing Revitalization Task Force (HECRTF) has been very helpful in building knowledge and momentum around the importance of HPC to both U.S. industry and the U.S. Federal Government. There is now a greater interagency discussion on the topic, and private industry is being heavily consulted. We are eagerly awaiting the outcome of this effort. Nothing will encourage more future spending by the U.S. computer vendors on HPC research and development however, than a strong increase in U.S. federal HPC procurement and deployment.

SUMMARY OF SGI's HIGH-PERFORMANCE COMPUTING RESEARCH EFFORTS

SGI regularly spends 13 percent of its annual revenues on research and development. This entire amount is spent on high-performance computing, high-performance storage, and high-performance visualization.

SGI dedicates its R&D efforts to system-level architectures utilizing industry standard components where appropriate. The unique combination of system-level architectures built with standard high-volume off-the-shelf commodity components, yields an overall price/performance balance that is very attractive to the HPC user. Full-custom products are generally too expensive, and full-commodity products lack the required performance or productivity. The blended use of custom/commodity by SGI is illustrated below:

	Custom Architecture	Commodity Component
Computing	NUMAflex interconnect fabric Global shared memory Single system image	Intel Itanium 2 processors Standard DIMMs Linux-64 operating system
Storage	Integrated DAS, NAS, SAN Heterogeneous connect Global shared files	LSI Logic disk drives StorageTek tapes
Visualization	Scalable graphic compositor Visual area networking	ATI graphic chips Standard PCs, laptops, and PDAs

SGI is aggressively focused on the technical, engineering and scientific marketplace. Problems in this space require large numbers of processors, large amounts of memory, and large amounts of I/O bandwidth, all tightly coupled with each other. SGI servers scale-up the number of processors, the amount of memory and the level of I/O bandwidth independently. To date, SGI has shipped HPC machines with 1,024 processors and with four terabytes of globally shared main memory. Current R&D efforts within SGI are aimed at scaling systems to 128 thousand processors and to one petabyte of main memory, globally shared among all processors. This is an ultra-scale machine, and one that is within reach by SGI in the 2007–2008 time frame, in partnership with the appropriate funding agency.

Furthermore, it is SGI's intention to integrate scalar, vector, streaming, and special-function processors directly onto the shared memory architecture of this machine. The most appropriate processor elements will then be brought into action on-the-fly, while the user's application code is being executed. This "multi-paradigm" concept will therefore embrace the best features of several architectures that are in the marketplace today. The machine will reconfigure itself in a dynamic manner to best suit the application as it runs.

With respect to our R&D efforts in storage and data life-cycle management, CXFS from SGI is a very successful shared-file heterogeneous-connect storage area network (SAN) in the market today. It will be extended to run over a wide area network, and thus enable nationwide single-level file addressing (SAN over WAN).

With respect to SGI's R&D efforts in visualization, our work involves the interactive visualization of massive data sets stored in global shared memory, using the diverse compute elements of the multi-paradigm architecture. We will bring high-fidelity visualization to the Linux environment in the near future.

And with respect to SGI's R&D efforts in software, we will assist the Open-Source community scale its Linux-64 operating system to accommodate as large a number of processors in a single system image configuration as possible. We will also help bring high-level scientific programming tools into the market and application program interfaces (APIs) that improve the ease-of-use of HPC equipment in general.

SGI's goal is to maintain its position as HPC thought leader and the leading supplier of real-time big data machines on the planet.

BIOGRAPHY FOR BOB BISHOP

Bob Bishop has served as Chairman and Chief Executive Officer for SGI since 1999. He joined the company in 1986 as founding president of SGI's World Trade Corporation and was responsible for all company activities outside North America until 1995.

Prior to joining SGI, Bishop held senior positions with Apollo Computer, Inc., from 1982 to 1986 and Digital Equipment Corporation from 1968 to 1982.

Bishop is an elected member of the Swiss Academy of Engineering Sciences, serves on the international advisory panel for the Multimedia Super Corridor in Malaysia, and is a member of the Executive Committee for the Council on Competitiveness in Washington, D.C.

He earned a B.Sc. (First Class Honors) in mathematical physics from the University of Adelaide, Australia, and an M.Sc. from the Courant Institute of Mathematical Sciences at New York University.