

[SPEAKERS](#)

[CONTENTS](#)

[INSERTS](#)

[Page 1](#)

[TOP OF DOC](#)

74-543PS  
2002  
*INNOVATION IN INFORMATION  
TECHNOLOGY: BEYOND FASTER COMPUTERS  
AND HIGHER BANDWIDTH*

HEARING

BEFORE THE

SUBCOMMITTEE ON RESEARCH  
COMMITTEE ON SCIENCE  
HOUSE OF REPRESENTATIVES

ONE HUNDRED SEVENTH CONGRESS

FIRST SESSION

JULY 31, 2001

Serial No. 107-18

Printed for the use of the Committee on Science

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

[Page 2](#)

[PREV PAGE](#)

[TOP OF DOC](#)

COMMITTEE ON SCIENCE

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

LAMAR S. SMITH, Texas  
CONSTANCE A. MORELLA, Maryland  
CHRISTOPHER SHAYS, Connecticut  
CURT WELDON, Pennsylvania  
DANA ROHRBACHER, California  
JOE BARTON, Texas  
KEN CALVERT, California  
NICK SMITH, Michigan

ROSCOE G. BARTLETT, Maryland  
VERNON J. EHLERS, Michigan  
DAVE WELDON, Florida  
GIL GUTKNECHT, Minnesota  
CHRIS CANNON, Utah  
GEORGE R. NETHERCUTT, JR., Washington  
FRANK D. LUCAS, Oklahoma  
GARY G. MILLER, California  
JUDY BIGGERT, Illinois  
WAYNE T. GILCHREST, Maryland  
W. TODD AKIN, Missouri  
TIMOTHY V. JOHNSON, Illinois

[Page 3](#)

[PREV PAGE](#)

[TOP OF DOC](#)

MIKE PENCE, Indiana  
FELIX J. GRUCCI, JR., New York  
MELISSA A. HART, Pennsylvania  
J. RANDY FORBES, Virginia

RALPH M. HALL, Texas  
BART GORDON, Tennessee  
JERRY F. COSTELLO, Illinois  
JAMES A. BARCIA, Michigan  
EDDIE BERNICE JOHNSON, Texas  
LYNN C. WOOLSEY, California  
LYNN N. RIVERS, Michigan  
ZOE LOFGREN, California  
SHEILA JACKSON LEE, Texas  
BOB ETHERIDGE, North Carolina  
NICK LAMPSON, Texas  
JOHN B. LARSON, Connecticut  
MARK UDALL, Colorado  
DAVID WU, Oregon  
ANTHONY D. WEINER, New York  
BRIAN BAIRD, Washington  
JOSEPH M. HOEFFEL, Pennsylvania  
JOE BACA, California  
JIM MATHESON, Utah  
STEVE ISRAEL, New York

[Page 4](#)

[PREV PAGE](#)

[TOP OF DOC](#)

DENNIS MOORE, Kansas

MICHAEL M. HONDA, California

Subcommittee on Research

NICK SMITH, Michigan, *Chairman*

LAMAR S. SMITH, Texas

CURT WELDON, Pennsylvania

GIL GUTKNECHT, Minnesota

FRANK D. LUCAS, Oklahoma

GARY G. MILLER, California

JUDY BIGGERT, Illinois

W. TODD AKIN, Missouri

TIMOTHY V. JOHNSON, Illinois

FELIX J. GRUCCI, JR., New York

MELISSA A. HART, Pennsylvania

SHERWOOD L. BOEHLERT, New York

EDDIE BERNICE JOHNSON, Texas

BOB ETHERIDGE, North Carolina

STEVE ISRAEL, New York

LYNN N. RIVERS, Michigan

JOHN B. LARSON, Connecticut

BRIAN BAIRD, Washington

JOE BACA, California

DENNIS MOORE, Kansas

[Page 5](#)

[PREV PAGE](#)

[TOP OF DOC](#)

MICHAEL M. HONDA, California

RALPH M. HALL, Texas

**SHARON HAYS** *Subcommittee Staff Director*

**PETER HARSHA** *Republican Professional Staff Member*

**JIM WILSON** *Democratic Professional Staff Member*

**DIANE JONES** *Professional Staff Member*

**NATALIE PALMER** *Staff Assistant*

## C O N T E N T S

July 31, 2001

Witness List

Hearing Charter

Opening Statement by Nick Smith, Chairman, Subcommittee on Research, U.S. House of Representatives

Opening Statement by Bob Etheridge, Member, Subcommittee on Research, U.S. House of

## Representatives

### Witnesses:

Ruzena Bajcsy, Ph.D., Assistant Director, Computer and Information Science and Engineering, National Science Foundation; Chair, Interagency Working Group on Information Technology R&D

[Page 6](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Oral Statement

Prepared Statement

Biography

Hans-Werner Braun, Ph.D., Research Scientist, San Diego Supercomputing Center

Oral Statement

Prepared Statement

High Performance Wireless Research and Education Network

Funding Disclosure

Vitae

Helen Berman, Ph.D., Director, The Protein Data Bank; Board of Governors Professor of Chemistry, Rutgers, The State University of New Jersey

Oral Statement

Prepared Statement

Biography

Funding Disclosure

William Blake, Vice President for High-Performance Technical Computing, Compaq Computer Corporation

Oral Statement

Prepared Statement

Slide Presentation

Biography

Carol Wideman, Chief Executive Officer and Founder, Vcom3D

Oral Statement

Prepared Statement

Funding Disclosure

[Page 7](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Biography

Discussion

HEARING ON INNOVATION IN INFORMATION TECHNOLOGY: BEYOND FASTER COMPUTERS AND HIGHER BANDWIDTH

TUESDAY, JULY 31, 2001

House of Representatives,  
Subcommittee on Research,  
Committee on Science,  
Washington, DC.

The Subcommittee met, pursuant to call, at 2:03 p.m., in Room 2318 of the Rayburn House Office Building, Hon. Nick Smith [Chairman of the Subcommittee] presiding.

Committee on Science  
Subcommittee on Research  
U.S. House of Representatives

[Page 8](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Washington, DC 20515

*Hearing on*

Innovation in Information Technology: Beyond Faster Computers and Higher Bandwidth

*Tuesday, July 31, 2001*

Witness List

Ruzena Bajcsy, Ph.D.

Assistant Director, Computer and Information Science and Engineering,

National Science Foundation;

Chair, Interagency Working Group on Information Technology R&D

Hans-Werner Braun, Ph.D.

Research Scientist,

San Diego Supercomputing Center

Helen Berman, Ph.D.

[Page 9](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Director, The Protein Data Bank;

Board of Governors Professor of Chemistry,

Rutgers, The State University of New Jersey

William Blake

Vice President for High-Performance Technical Computing,

Compaq Computer Corporation

Carol Wideman

Chief Executive Office and Founder,

Vcom3D

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 (three days requested) at (202) 225-6371 or FAX (202) 225-0891.

[Page 10](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Should you need Committee materials in alternative formats, please contact the Committee as noted above.

HEARING CHARTER

SUBCOMMITTEE ON RESEARCH

COMMITTEE ON SCIENCE

U.S. HOUSE OF REPRESENTATIVES

Innovation in Information Technology:

Beyond Faster Computers and Higher Bandwidth

TUESDAY, JULY 31, 2001

2:00 P.M.–4:00 P.M.

2318 RAYBURN HOUSE OFFICE BUILDING

## I. Purpose

On Tuesday, July 31, 2001, at 2:00 p.m. the Subcommittee on Research of the House Committee on Science will hold a hearing to examine the impact federal investment has had on promoting innovation in information technology and fostering a variety of sophisticated applications that infuse information technology into areas such as education, scientific research, and delivery of public services. The hearing will also examine the limits of current technology and highlight research questions and technological applications that require additional investment.

[Page 11](#)

[PREV PAGE](#)

[TOP OF DOC](#)

## 2. Background

### The Federal IT R&D Effort

In a Research Subcommittee hearing on June 26, 2001, *Reinventing the Internet: Promoting Innovation in IT*, witnesses described the impact of the federal investment in Information Technology (IT) Research and Development (R&D) on innovation and the overall economy. The witnesses stressed the importance of federally sponsored, long-term, fundamental IT research in enabling the private sector to focus their efforts on shorter-term, more product-driven research and development. These views echoed the recommendations of a 1999 report from the President's *Information Technology Advisory Committee (PITAC)*, *Information Technology in Research: Investing in Our Future*. Specifically, PITAC recommended that increased future Federal investment include an emphasis on long-term, high-risk research goals and that the government develop a strategic initiative in long-term information technology R&D. Further, PITAC recommended that the federal government pay particular attention to four priority research areas including software, scalable information infrastructure, high-end computing and the socioeconomic impact of IT in the overall research agenda.

The current federal IT R&D program encompasses a number of agencies including the National Science Foundation (NSF), Department of Defense (DoD), Department of Energy (DoE), the National Aeronautics and Space Administration (NASA), the Environmental Protection Agency (EPA), the National Institutes of Health (NIH), the National Institute for Standards and Technology (KIST), and the National Oceanic and Atmospheric Administration (NOAA). Coordination of the federal IT R&D investment is facilitated by the White House's Office of Science and Technology Policy through the Interagency Working Group on Information Technology R&D (IWG), which meets regularly to evaluate the current state of research and to define areas which require additional investment.

[Page 12](#)

[PREV PAGE](#)

[TOP OF DOC](#)

The Interagency Working Group consists of representatives from each of the agencies involved in IT R&D and is chaired by Dr. Ruzena Bajcsy, who currently serves as the head of the National Science Foundation's Computer and Information Science and Engineering directorate. The IWG serves to provide a practical link between legislative directives and the evolving needs of the IT field.

## NSF IT Research

Though the IWG has no official authority over the agencies represented by the group, participants voluntarily mobilize the recommendations of the IWG through their agency budget requests, extramural programs, and IT initiatives. The National Science Foundation, for example, has a number of programs that constitute its proposed FY 2002 budget of \$643 million for investment in IT research. A major component of this investment is the interdisciplinary Information Technology Research (ITR) initiative, a Foundation-wide program aimed at advancing high performance computing capacity through the support of long-term, innovative research. The goal of the NSF ITR initiative is to amplify the benefits of IT in all areas of science and engineering and spur progress across the national economy and society. The ITR initiative provides IT funding in addition to that of the core research programs in NSF's Computer and Information Sciences and Engineering, Engineering, and Math and Physical Sciences directorates.

The ITR program involves seven comprehensive and complementary focus areas including: large-scale networking; high-end computing; high-end computation and infrastructure; high-confidence software and systems; human-computer interaction and information management; software design and productivity; and social, economic and workforce implications of IT. In FY 2000, the NSF Information Technology Research program stressed fundamental research; in the second year additional applications in science were added, and in the third year, the program will expand research in multidisciplinary areas, focusing on fundamental research at the interfaces between fields and disciplines. The Fiscal Year 2002 Budget Request reflects the interdisciplinary nature of the current ITR program.

[Page 13](#)

[PREV PAGE](#)

[TOP OF DOC](#)

74543a.eps

As illustrated in the table above, the NSF ITR initiative is supported by a number of NSF directorates including Biological Sciences (BIO), Computer and Information Sciences and Engineering (CISE), Engineering (ENG), Geosciences (GEO), Mathematical and Physical Sciences (MPS), Social, Behavioral and Economic Sciences (SBE), Polar Programs (OPP), and Education and Human Resources (EHR). Each directorate determines the IT needs of its own community and develops programs appropriately. The Computer and Information Sciences and Engineering (CISE) directorate receives the largest ITR budget and supports a variety of research projects on next-generation cyberinfrastructure, human augmentation, research at the interface of biology and information technology, and computer security. In response to the PITAC recommendation to increase funding for high-risk research projects, CISE devotes 10 percent of its ITR budget to such research. Each NSF directorate determines focus areas for its ITR budget, which are communicated externally through program announcements and other dissemination methods.

A quick overview of the ITR priorities from across the Foundation illustrates the diverse and comprehensive nature of the NSF ITR program. For example, the ITR focus areas for the Biological Sciences Directorate (BIO) include genome-enabled sciences, molecular computing, biological databases, development of real-time information networks, complex modeling, interactive systems, and informatics. The Geosciences ITR emphases include modeling to understand the interactions taking place in the Earth system, development of tools for knowledge discovery and interpretation of large-scale data sets,



development of infrastructure to study geospatial data, and extension of local networking and computing capabilities in support of large-scale modeling and database activities in the geosciences. The Engineering Directorate supports computational simulation and modeling of complex materials, structures and processes and the development of high-end computing tools to accelerate the design of next generation IT manufacturing techniques. Mathematical and Physical Sciences focuses its ITR support on algorithm development, statistical analysis, optimization theory, network design, quantum and optical computing, the development of nano-devices, and advanced computational methods. The Social, Behavioral and Economic Sciences Directorate emphasizes the need for fundamental research on social, economic and workforce issues associated with computational social science and also supports applications of technology that facilitate improved learning at a variety of levels. Finally, NSF's Office of Polar Programs spends its ITR dollars on priorities such as the development of remote operation capabilities and accessible information systems for polar data. The NSF ITR program promotes interdisciplinary collaboration that ultimately yields more complex and meritorious results than is possible through traditional discipline-specific endeavors.

[Page 14](#)

[PREV PAGE](#)

[TOP OF DOC](#)

### 3. Witnesses

While the NSF ITR program lists an extensive array of research priorities, a sampling of actual research projects supported by NSF is most illustrative of the contributions the ITR initiative has made to advanced technology, research science, education, and other areas. Computer scientists and other researchers solve problems that once imposed limits on research and technology—and the applications of both—while they help define the research questions of tomorrow and develop strategies for finding answers and new possibilities enabled by IT. This hearing will highlight a sampling of work on the frontiers of scientific research, illustrating how key technological problems are solved through NSF-supported research. The hearing will also show how technological advances drive a myriad of creative applications that promise to one day impact not only other researchers but also policy makers, educators, manufacturers, merchants, and many others. The witnesses will also speak to the continuing needs for federal investment in IT as well as the particular areas of focus that need special attention to drive subsequent generations of technologies and applications.

The witnesses will discuss their work in developing high technology solutions for scientific and commercial enterprises including, for example: the installation of reliable wireless networks to provide ubiquitous web access to rural, remote, and emergency response communities; the expansion of a powerful, international scientific database that facilitates highly technical exploration of protein structure and function; and the creation of supercomputers to facilitate powerful computational functions for a variety of research efforts. In addition, two of the witnesses will also discuss their work on developing and adapting technological advances to meet the specific needs of the education community by using, for example, a wireless network to provide access to three remote Native American reservations, or by developing software that features high quality 3D characters that use sign language to promote web accessibility and new learning options for deaf and hard-of-hearing students.

[Page 15](#)

[PREV PAGE](#)

[TOP OF DOC](#)

The following witnesses will address the Subcommittee:

*Dr. Ruzena Bajcsy*, Chair of the Interagency Working Group and Assistant Director, NSF, Computer and Information Science and Engineering. Dr. Bajcsy is also a robotics researcher at the University of Pennsylvania.

*Dr. Hans-Werner Braun*, a Research Scientist at the San Diego Supercomputing Center. Dr. Braun conducts research on wireless networks to explore ways to increase bandwidth, especially in remote and rural locations, and to enable rapid installation and portable networks that can be used in emergency management situations and in non-static scientific research environments. Dr. Braun's work is funded through the NSF ITR initiative.

*Dr. Helen Berman*, the Director of the Protein Data Bank and Board of Governors Professor of Chemistry and Professor of Chemistry at Rutgers, The State University of New Jersey. Dr. Berman's research is focused, in part, on developing biologically-relevant information databases that utilize sophisticated data curation and data query technologies to ensure accuracy, ease of use, and interoperability with other related databases. Dr. Berman's work is funded through the NSF Biological Sciences Directorate, Division of Biological Infrastructure.

*Ms. Carol Wideman*, the CEO and founder of Vcom3D. Ms. Wideman's company is currently developing the foundations of real-time 3D graphics and simulation capabilities and has utilized this technology to offer Internet access and educational products to the deaf and hard-of-hearing. Ms. Wideman's work was supported by NSF through the Small Business Innovative Research program.

[Page 16](#)

[PREV PAGE](#)

[TOP OF DOC](#)

*Mr. Bill Blake*, the Vice President for High-Performance Technical Computing at Compaq Computer Corporation. Mr. Blake's division works closely with the academic and private sector communities to develop and provide high-end computers to meet the computational needs of a variety of researchers. Compaq developed high end computers for research efforts including the Human Genome Project at Celera Genomics and the terascale computer at the Pittsburgh Supercomputing Center, an NSF-funded ITR/Major Research Equipment project.

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

1. What are the basic Information Technology research questions that must be answered if the United States is to remain economically and technologically strong and how is the Federal government mounting a coordinated and balanced effort to drive innovations and discoveries at the current technological limits?
2. How does NSF balance the need to be fiscally responsible with the need to support high-risk, high-stakes projects? How does NSF ensure that it supports novel and innovative approaches without bias toward existing platforms, infrastructure, or data streaming protocols?
3. How is current IT R&D moving various fields beyond their current limitations and toward new

capabilities? What other technologies will likely be enabled by the advances derived both directly and indirectly from this research and where is additional emphasis required?

4. One goal of the federal IT R&D program has been to provide academic researchers with access to leading edge scientific computers. The approach being taken in the program for very high performance computing is to employ scalable computer architectures using thousands of commodity processors. Are there classes of problems in science and engineering research that are not well suited to this type of computer architecture, and if so, should federal resources be provided for developing and acquiring alternative types of high performance computers?

[Page 17](#)

[PREV PAGE](#)

[TOP OF DOC](#)

5. How should the federal high performance computing investment be balanced between software and hardware development? What is the potential role that open source software could play in the high performance computing world and what should be the role of the Federal government, if any, in open source software development?

Chairman **SMITH**. The Subcommittee on Research will now be in session. I am looking for my first page. Well, as everyone knows, last year we passed an information technology bill of—the Chairman at that time, Mr. Sensenbrenner. We are now looking at new legislation and trying to determine, with expert advice, such as the witnesses we have today, where we go in Federal involvement and Federal policy directions that we apply to information technology.

This is the second time that the Subcommittee has met to discuss Federal support of information technology research. On June 26, we met to examine the Federal investment in IT research and development and to try to understand how that investment is divided among agencies and research priority areas. That hearing covered the Federal information technology oversight structure. And the recommendation of both the President's Information Technology Advisory Council or PITAC, and the Interagency Working Group on Information Technology Research and Development. Both oversight groups recommended increased support for long-term, high-risk, and high-potential IT research.

Today's hearing will focus more closely on the way the National Science Foundation uses the recommendations of these oversight groups and other members of the community to prioritize its IT R&D efforts. NSF provides most of its support for IT through Computer Information Sciences and Engineering directorate. And through the CISE, NSF supports research in such areas as computational infrastructure, networking, and data storage.

[Page 18](#)

[PREV PAGE](#)

[TOP OF DOC](#)

In addition, the Information Technology Research Program, or ITR, is an agency-wide IT priority that provides additional funding for information technology across disciplines and directorates that I think is so important if we are going to make the right choices and come up with the best possible solutions to how we use taxpayers' money, how we coordinate our efforts within and without government. The ITR program

recognizes that every discipline and directorate at NSF is, to some degree, dependent on applications of IT research.

Our witnesses today will describe how NSF support has enabled them to push the limits of what is possible, creating new applications that benefit other researchers, communities, and the Nation as a whole. New research in wireless, high-quality Internet connections is allowing children in the most remote rural locations of our country to have real-time access to today's leading researchers. It is enabling communities struck by disaster to coordinate relief efforts when phone and fiberoptic networks are down. High-performance supercomputing is enabling better weather predictions, and reliable flight patterns to make better use of our airways, and more efficient manufacturing. And IT enabled adaptive technologies are allowing disabled children, some for the first time in their life, to participate fully in traditional school programs.

I hope our witnesses today will give us your best analysis of some of the mistakes that government has made as we try to support IT research that sometimes has some political bias. But I hope you will also comment on other advances in IT R&D and highlight the gaps in our current knowledge and our current capacity.

As this Subcommittee moves ahead with legislation to authorize the Federal role in supporting IT research, I hope our witnesses can also give their impressions of NSF's role in the process. Where does NSF do things right and where does NSF have areas that it can improve on?

[Page 19](#)

[PREV PAGE](#)

[TOP OF DOC](#)

We owe much of our current prosperity to the gains in productivity enabled by innovations in IT—gains that were, in turn, enabled by past investments in information technology research. Our future productivity, it seems to me—in fact, our future prosperity are tied to the success which is derived from our investments in information technology support. I think the witnesses are a great asset to how we develop our final information technology bill. I look forward to your testimony. And would turn to Mr. Etheridge for his comments.

[The prepared statement of Mr. Smith follows:]

#### PREPARED STATEMENT OF CHAIRMAN NICK SMITH

Good afternoon and welcome to this hearing of the Subcommittee on Research, *Innovation in Information Technology: Beyond Faster Computers and Higher Bandwidth*. This is the second time the Subcommittee has met to discuss federal support of information technology research. On June 26th, we met to examine the federal investment in IT research and development and to understand how that investment is divided up among agencies and research priority areas. That hearing covered the federal IT oversight structure and the recommendations of both the President's Information Technology Advisory Council, or PITAC, and the Interagency Working Group on Information Technology R&D. Both oversight groups recommended increased support for long-term, high risk and high potential IT research.

Today's hearing will focus more closely on the way the National Science Foundation uses the

recommendations of oversight groups and other members of the community to prioritize its IT R&D funding. NSF provides most of its support for IT through its Computer and Information Sciences and Engineering directorate (CISE). Through CISE, NSF supports research in such areas as computational infrastructure, networking, and data storage. In addition, the Information Technology Research program, or ITR, is a agency-wide IT "priority" that provides additional funding for IT across disciplines and directorates. The ITR program recognizes that every discipline and directorate at NSF is, to some degree, dependent upon applications of IT research.

[Page 20](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Our witnesses today will describe how this NSF support has enabled them to push the limits of what's possible, creating new applications that benefit other researchers, communities, and the Nation as a whole. New research in wireless, high-quality Internet connections is allowing children in the most remote and rural locations in our country to have real-time access to today's leading researchers. It's enabling communities struck by disaster to coordinate relief efforts when phones and fiberoptic networks are down. High-performance supercomputing is enabling better weather predictions, reliable flight patterning, and more efficient manufacturing. And IT enabled adaptive technologies are allowing disabled children, some for the first time in their life, to participate fully in traditional school programs.

I hope our witnesses will also comment on other advances in IT R&D and highlight the gaps in our current knowledge and capacity. As this subcommittee moves ahead with legislation to authorize the federal role in supporting IT research, I hope our witnesses can also share their impressions of NSF's role in the process. Where does NSF do things right? Where might NSF work to improve its programs?

We owe much of our current prosperity to the gains in productivity enabled by innovations in IT—gains that were, in turn, enabled by past investments in IT research. Our future productivity and prosperity are equally tied to the investments we make today supporting this innovative research. I thank the witnesses for appearing today to discuss their research with us and I look forward to your testimony.

Mr. **ETHERIDGE**. Thank you, Mr. Chairman. And let me just tell you how pleased I am that you have called this hearing today to explore how federally sponsored research related to information technology is addressing societal needs and helping to advance the progress in science and engineering—critical area, as we look to the 21st century. And I join you in welcoming our witnesses here this afternoon.

[Page 21](#)

[PREV PAGE](#)

[TOP OF DOC](#)

The Subcommittee held a hearing last month, as the Chairman has already indicated, to look broadly at the interagency IT R&D program. This was really the kickoff for a process to gather information needed to update the authorizing legislation for the program. I believe this has been a successful program. It puts in place a process to develop the coordinating—the coordinated research plan and budget for the research activities carried out by the participating agencies.

And it has been succeeded—it has also succeeded in developing the computing and networking

infrastructure needed to develop leading-edge research and to drive the technology forward for the benefits of society. And I won't go into all the details. The Chairman has touched on them. We all know how much it has benefited us in recent years.

Today, we will have the opportunity to review some specific applications of the technology that illustrates the kind of benefits that are possible as we push the wall back, as we move toward the—into the 21st century.

I invite our witnesses to comment on any aspects of the current IT R&D program that could or should be strengthened. We are interested in your observations on both this programmatic content and its research emphasis, as well, and in the way that it is ministered also, to make it easier to get it out to our colleagues who are working on it.

The technical advances that lead to today's information based society evolved from past federal-sponsored research in partnership with industry, as well as at the university level. I look forward to our discussions today on how we can ensure that the storehouse of basic knowledge continues to be replenished. Let me repeat that again. It is critical that we replenish it. You know, so many times we think about drawing it down and forget that we have to replenish it on a regular basis. And if we don't, we are in trouble. And if we replenish it, it, in turn, will enable the development of future generations of information technology products and services that will benefit not only society in this country, but people around the world.

[Page 22](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Mr. Chairman, again, I thank you for calling this hearing and join you in welcoming our distinguished guests this afternoon.

Chairman **SMITH**. And without objection, any additional opening comments by any other member of the Committee will be printed in the work—in the record. I would like to introduce our panelists at this time. Dr. Ruzena Bajcsy, welcome again, and it is good to see you. Dr. Bajcsy is Chair of the Interagency Working Group on Information Technology Research and Development, and Assistant Director of the National Science Foundation. You are Director—or Assistant Director of NSF's Computer and Information Science and Engineering Directory. And Dr. Bajcsy is also a Robotics Researcher at the University of Pennsylvania. So welcome.

Second witness is Dr. Hans-Werner Braun, a Research Scientist at the San Diego Supercomputing Center. And Dr. Braun conducts research on wireless networks to explore ways to increase bandwidth, especially in remote and rural locations, and to enable rapid installation and portable networks that can be used in emergency management situations and in non-static scientific research environments. Mr. Braun's work is funded through the NSF, ITR initiative.

Next, we have Helen Berman. Dr. Berman is the Board of Governors Professor of Chemistry and the Director of the Protein Data Bank at Rutgers University. And Dr. Berman's research is focused on developing biologically relevant information databases and is funded through NSF's Biological Science Directorate's Division of Biological Infrastructure.

The next witness is Mr. Bill Blake. And Bill Blake is the Vice President for High-Performance Technical Computing at Compaq Computer Corporation. And Mr. Blake's division works closely with the academic and private sector communities to develop and provide high-end computers to meet the computational needs of a variety of researchers. Compaq developed high-end computers for research efforts, including the Human Genome Project at Solera Genomics, and the Terascale Computer at the Pittsburgh Supercomputing Center, an NSF funded ITR major research equipment project.

[Page 23](#)

[PREV PAGE](#)

[TOP OF DOC](#)

And our final witness is Mrs. Carol Wideman—Ms. Carol Wideman, the CEO and founder of—is it Vcom3? Vcom3D, maybe. Ms. Wideman's company is currently developing the foundation of real-time 3D graphic simulation capabilities and utilized this technology to offer Internet access and educational products to the depth. Dr. Wideman's work was supported by NSF through the Small Business Innovative Research Program.

And with that, as you know, your written testimony will be all entered into the record and we would ask you, as best as you can, to try to limit your verbal comments to 5 minutes. And with that, Dr. Bajcsy, we will start with you.

STATEMENT OF DR. RUZENA BAJCSY, Ph.D., ASSISTANT DIRECTOR, COMPUTER AND INFORMATION SCIENCE AND ENGINEERING, NATIONAL SCIENCE FOUNDATION; CHAIR, INTERAGENCY WORKING GROUP ON INFORMATION TECHNOLOGY R&D

Dr. **BAJCSY**. Thank you, Mr. Chairman. I am truly delighted to be here and I am most grateful for this hearing. As you mentioned, the genesis of the NSF Information Technology Research Program can be traced to the publication of the PITAC Information Technology Research Investing in Our Future in 1999, when the Committee concluded that the current Federal support of IT research is seriously inadequate, as Mr. Etheridge also mentioned, that we need to replenish our knowledge.

The NSF arose to the plate and put into its program a research agenda. And in fiscal year, as you see on your slide—fiscal year 2000 and 2001, and requested in 2001, the numbers for NSF information technology initiative budget.

[Page 24](#)

[PREV PAGE](#)

[TOP OF DOC](#)

As urged by PITAC, NSF took the lead on the overall Federal initiative. In its calls for proposals, NSF encouraged proposals for basic, long-term, high-end risky projects. NSF also sought greater involvement of not only the first-tier universities, but second and third-tier research universities. An additional desire was that much of the targeted research would also have eventual trickle-down effects resulting in useful technologies for the general public. In addition, the NSF solicitation emphasized support for Principal Investigators whose proposals would not only advance research, but materially assist in training the next generation workforce. As you know, this is a great concern to all of us.

In ITR's first year, NSF was overwhelmed by the response from the community, receiving a total of about 2,000 proposals. And subsequently, we were able to fund only about 11 percent of the deserving proposals. We divided the proposals into small, medium, and large. The details are in my oral testimony.

At NSF, preparation of the solicitation, as well as the management and evaluation of the proposals, is currently carried out by a coordinating committee, which is composed partially from my directorate and partially from all the other directorates.

The current geographic distribution of fiscal year 2000 is on your screen and I can elaborate later at your convenience, if you wish.

In fiscal 2001, the second year of the ITR program, the call for proposals emphasized more IT applications. Approximately 45 percent of the funding was distributed to other directorates, in addition to mine, to computer science topics. The response from the community was as substantial as that received in fiscal year 2000, with NSF receiving again about 2,000 proposals. And, again, we were able only to fund about 11 percent.

[Page 25](#)

[PREV PAGE](#)

[TOP OF DOC](#)

And the geographic distribution of the proposals submitted in 2001, you can see there. And I want to call your attention that there are still some states that are empty, have no dots. And this is really a concern to us.

So if I evaluate what we have accomplished, the last 2 years of experience with ITR having truly transformed both internal NSF IT—as you mentioned, Mr. Chair, this is, indeed, in my opinion, a very gratifying experience—it is a true success—as well as the external scientific community. NSF has been able to encourage and award more interdisciplinary, high-risk, and long-term research grants by having the ability to make larger awards and longer awards, 3 to 5 years, than the average NSF grants are.

Perhaps, most importantly, NSF has changed the scale of problems that are now being investigated. This is of—am I over?

Chairman **SMITH**. That signals that you have got about 60 seconds left.

Dr. **BAJCSY**. I see. Okay. Well, this is of enormous importance—this new scale. But what I want to say is now that NSF has its second year of ITR, the agency is beginning to face the problem of how to monitor and evaluate progress. In some ways, NSF functions like a taxpayer-funded venture capitalist organization, with funding coming from the U.S. citizens, as opposed to more narrow-focused private sources.

Chairman **SMITH**. Dr. Bajcsy—

[Page 26](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Dr. **BAJCSY**. Yes.



Chairman **SMITH**. Let us wait.

Dr. **BAJCSY**. Okay. Ah. Okay. All right. So I would like emphasize that we view NSF as a venture capitalist organization, with funding coming from the U.S. citizens as opposed to more narrowly focused private sources. As a result, NSF, without the constant pressure to report higher numbers each quarter, can afford to invest in broader, long-term, and more scientifically adventurous projects because the agency is not forced to be short term profit-oriented.

It is clear that IT is an essential economic driver in the United States, as well as the whole world. However, there are many unsolved research problems. As example, just think of the interhuman, man machine, human computer interfaces, especially when you think about the multi-dimensional data or visual and audio data.

While the other problem—we simply have not focused, for example, in assembling the right hardware/software/algorithms for reliable and secure end-to-end multimedia communication. Our user interfaces are not very good.

While the ability of NSF to invest in leading-edge IT R&D has increased significantly, thanks to Congress, current funding levels are still not meeting the necessary challenges to keep the IT industry expanding at its current rapid rate.

[Page 27](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Now, a different class of problems is involved in NSF's support for computational science, which provides computer simulations for physical processes. To enable work in this area, NSF supports the partnership, the PACI.

As a result of such activities, NSF's future points toward a more holistic view of the evolving cyberinfrastructure, which will consist not only of high-performance computers, but also will embrace broadband connectivity with large databases and distributed instruments. The evolution of the cyberinfrastructure is a natural outgrowth of PACI, since PACI, by definition, is a distributed partnership.

Finally, I would like to say, as is well-known, NSF's mission is to support basic research for all sciences. This differs significantly from other Federal agencies whose missions are strictly agency-oriented. Because of this division of our mission, NSF basic research is a pillar for all mission-oriented agencies involved IT R&D and, I should say, has the smallest budget. Thank you.

[The prepared statement of Dr. Bajcsy follows:]

PREPARED STATEMENT OF RUZENA BAJCSY

Information Technology Research

Recent dramatic advances in computing, communications, and in collecting, digitizing, and processing information are having a major impact today not only in the world of science, but on the everyday experiences of the average U.S. citizen. These advances are undeniable indicators that the horizons of

Information Technology (IT) are much broader, and its impacts on society far larger, than were anticipated even a few short years ago. In order to respond to the need for continuing rapid advancements in IT, NSF has established the Information Technology Research (ITR) Program.

[Page 28](#)

[PREV PAGE](#)

[TOP OF DOC](#)

ITR is focused on the analysis and synthesis of information systems. It includes hardware, system architectures, operating systems, programming languages, communication networks, and other systems that acquire, store, process, transmit, and display information. It also considers closely connected and/or distributed multi-processor computers and possible new modes of computation such as quantum, molecular, or DNA computing. While in the past, we have dealt with systems consisting of 100s of computers, today we have systems consisting of 1,000 to 10,000 computers—and tomorrow, a million or more processors may be connected in one system. This change of scale challenges our current understanding of systems and points out the compelling need for new methodologies, both theoretical and experimental, to address the complexity of these new systems. Ranging from high end computing to work in human-computer interfaces, ITR research is geared toward answering these complex scientific questions—as well as toward creating systems that are easily accessible to and useful for the average citizen.

The genesis of NSF's ITR Program can be traced in part to the publication of *Information Technology Research: Investing in Our Future*, the groundbreaking 1999 report by the President's Information Technology Advisory Committee (PITAC). In this report, the Committee concluded that current Federal support for IT research is seriously inadequate. The PITAC recommended that funding Increases for IT R&D across participating Government agencies should be (in \$ millions):

96

74543b.eps

For NSF the total investment into ITR and High End Acquisition are as follows:

[Page 29](#)

[PREV PAGE](#)

[TOP OF DOC](#)

49

74543c.eps

As urged by the PITAC, NSF took the lead on the overall Federal initiative. NSF's FY 2000 solicitation focused on eight areas, five of which were suggested by the PITAC report:

Applications in Computational Sciences (ACS)

Human Computer Interfaces (HCI)

Information management (IM)

Scalable Information Infrastructure (SII)

Software (SW)

Education, Workforce (EWF)

Revolutionary Computing (RC), including DNA and quantum computing

Social Impact of IT (SOC)

In its call for proposals, NSF encouraged proposals for basic, long term, high end, and sometimes risky research projects. NSF also sought greater involvement of second- and third-tier research universities. An additional desire was that much of the targeted research would also have eventual "trickle-down" effects resulting in useful technologies for the general public. In addition, the NSF solicitation emphasized support for Principal Investigators (PIs) whose proposals would not only advance research but materially assist in training the next generation workforce. In ITR's first year, NSF was overwhelmed by the response from the community, receiving a total of approximately 2,000 proposals with the consequence of being able to support only 11% of the proposals.

[Page 30](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Proposals were divided into three categories:

Small: <\$500K over 3–5 years (28% of total awarded proposals)

Medium: >\$500K up to \$5 Million over 3–5 years (62% of total awarded proposals)

Large: \$5M over 3–5 years (3 awards were made).

At NSF, preparation of the solicitation as well as the management and evaluation of the proposals is currently carried out by a Coordinating Committee, half of whose members are from the Computer and Information Science and Engineering (CISE) Directorate and half of whom are from the other NSF directorates. Guided by one full-time manager, this Committee makes recommendations for awards and for large and medium size proposals. The Assistant Director of CISE has final approval.

The current geographic distribution of NSF's FY 2000 ITR awards is shown on the following map:

74543d.eps

Large awards appear in red, while green designates the smaller awards.

Examples of large awards include promising high-end research to simulate how blood flows. Such studies can be of great significance in designing next-generation artificial hearts and valves, crucial to maintaining life while an appropriate donor heart is located (Ghattas). Another large, high end award is funding a comprehensive biogeometric study using geometric modeling to assist in new drug design (Edelbrunner).

[Page 31](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Additional examples of recent ITR awards granted by NSF include:

A project whose goal is to build a robot assistant to help the elderly in their homes (Dunbar-Jacob, University of Pittsburgh)

A study of the global spread of the Internet and impact on nations (Kraemer, UC Irvine)

A project that is focused on rewriting the air traffic Control System (Leveson, MIT)

Developing a computer system to answer spoken questions (Picone, Mississippi State)

R&D focused on creating displays touch-sensitive systems for the blind (Beebe, Univ. of Wisconsin)

New ways of teaching stroke victims how to move (S. Schaal, USC)

Significant small grants have been awarded to scientists whose research is focused on tracking how molecules move in biochemical solutions; rethinking how we model the chemical bonds that make materials strong; and enabling oceanographers to sail "virtual ships" to study mid-ocean ridges.

In FY 2001, the second year of the ITR program, the call for proposals emphasized more IT Applications (approximately 45% of funding) in addition to Computer Science topics. The response from the community was as substantial as that received in FY 2000, with NSF receiving approximately 2,000 proposals. With so many fine proposals received, NSF will be able to grant awards to only 11–15% of the proposals received.

[Page 32](#)

[PREV PAGE](#)

[TOP OF DOC](#)

The structure of the ITR awards, namely, their division into small, medium and large awards has been the same as in the FY 2000 call, and NSF's Coordinating structure also remains the same.

Programmatically, NSF called for new proposals in the following categories:

1. *Systems Design and Implementation*. This category included Software; Human-Computer Interfaces; Revolutionary Computing; and Fundamental IT Models.
2. *People and Social Groups Interacting with Computers and Infrastructure*. The subheadings of this category were: Social and Economic Implications of Information Technology; Universal Access; Information Technology in the Social and Behavioral Sciences Community; Expanding Infrastructure; and IT Workforce Education.
3. *Information Management*. Areas of special emphasis in this category were: Content and Access Research; Environmental Informatics; Geosciences Modeling and Representation; Informatics in Biological Sciences; Informatics in Mathematical and Physical Sciences; and Information Research, Repositories, and Testbeds.
4. *Applications in Computational Science and Engineering and Related Applications of Information*

*Technology.* This category includes: Algorithms and Data Models; Advanced Computation in Biological Sciences; Advanced Computation in Mathematical and Physical Sciences; Computational Geosciences; Applications in Engineering Sciences; Scientific Data Analysis; Computer Simulation and Visualization.

[Page 33](#)

[PREV PAGE](#)

[TOP OF DOC](#)

5. *Scalable Information Infrastructure for Pervasive Computing and Access.* Areas of special emphasis within this area include: Remote Access and Control of Experimental Facilities; Scalability, Security, Privacy and Integrity of the Information Infrastructure; Sensors and Sensor Networks; Tether-free Communication; Tether-free Networking Technology; and Teleimmersion.

The Geographic distribution of proposals submitted in FY 2001 is illustrated in the figure below.

74543e.eps

Again, red designates proposals for large awards, while green designates proposals for smaller awards.

The final awards have not yet been made at this time.

Evaluation:

The last two years of experience with ITR have truly transformed both internal NSF IT activities as well as the external scientific community. NSF has been able to encourage and award more interdisciplinary, high risk, and long term research grants by having the ability to make larger (>\$500K to \$3M per year) awards with longer (3 to 5 years) duration than the average NSF grants. In addition, NSF brought into the NSF Principal Investigators (PI) family researchers who, for various reasons, had not been able to participate before. These PIs are not only from traditional, first-tier research institutions but also from second- and third-tier institutions, significantly broadening NSF's outreach into the scientific community.

[Page 34](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Perhaps most importantly, NSF has changed the *scale* of problems that are now being investigated. This is of enormous importance since perhaps the largest obstacle in making progress in IT is that many software/hardware/algorithmic problems that were investigated by only one or two professors did not scale up. Such work is crucial to the future of high-end computing and computation.

Internally, NSF's ITR Program, encouraged by a Coordinating group comprised of members from every NSF directorate, has enabled NSF to build a more cohesive understanding of IT and its applications to science and engineering across the whole of NSF. In turn, this has influenced the CISE core program which now more accurately reflects the current problems in IT that have come about because of this change.

Now that NSF is in its second year of ITR, the agency is beginning to face the problem of how to monitor and evaluate progress in the awards it is granting. In some ways, NSF functions like a taxpayer-funded venture capitalist organization, with funding coming from the U.S. citizen as opposed to more narrowly-

focused private sources. As a result, NSF, without the constant pressure to report higher "numbers" each quarter, can afford to invest in broader, longer-term, and more scientifically adventurous projects because the agency is not forced to be short-term profit-oriented. At the same time, however, NSF must function as responsible stewards of the taxpayers' money and must be accountable for how the organization allocates investments, focusing not only on the science but on the quality of the product.

Above and beyond ITR, proposal evaluation at NSF is peer reviewed, in this case via panels and sometimes via individual written reviews. For larger grants (>\$500K over 3 years) I personally oversee the reviews and panel recommendations. In discussion with the coordinating committee and/or a subset of the committee, we make the final decisions together.

[Page 35](#)

[PREV PAGE](#)

[TOP OF DOC](#)

The Future:

It is clear that IT is an essential economic driver in the U.S. as well as the whole world. However, there are many unsolved research problems that remain obstacles to the continuation of the IT growth. A case in point is the current state of human-computer interface with respect to the personal computer (PC). Nearly all current users—from the sophisticated high end scientist to the average second-grade student—are still communicating with PCs and computing systems by means of keyboards, and assembling research and documentation by means of one or two main word-processing programs whose code and function have not evolved significantly in years. Opportunities for other kinds of human-computer interfaces, while occasionally available in the lab or to the general public in rudimentary, unreliable forms, are not widely available to the general public. This is due primarily to a lack of quality research in this area. We simply have not focused, for example, on assembling the right hardware /software/algorithms for reliable and secure end-to-end multimedia communication. Our user interfaces—even keyboards and wordprocessors—are fragile, difficult to use, and not optimized for universal use.

Sensors and sensory networks offer another challenge as a potential growth area in our IT economy, yet there are many difficult problems involved with making such networks reliable and ubiquitous in addition to being easy to use. We need theoretical and experimental work in creating and understanding large distributed systems that are reliable, adaptive, and secure. Security and privacy of our systems is, in fact, of paramount importance. It seems that almost weekly we are subjected to system failures and virus attacks against which we have only begun to construct adaptive defenses. Such critical areas of research as this still suffer from a lack of adequate investment, and we must increase our ability to fund leading-edge research in this crucial arena.

[Page 36](#)

[PREV PAGE](#)

[TOP OF DOC](#)

While the ability of NSF to invest in leading-edge IT R&D has increased significantly, current funding levels are still not meeting the necessary challenges to keep the IT industry expanding at its current rapid rate. Furthermore, the NSF IT research investment, concurrently with pushing the boundaries of understanding Computer Science and Engineering and its applications, is also responsible for helping to

train our future workforce, since NSF primarily funds U.S. universities where such training takes place. For this reason, we must continue to increase our efforts in this regard, or we risk not developing the scientists and researchers we will need for the future.

A different class of problems is involved in NSF's support for Computational Science, which provides computer simulations for physical processes as diverse as weather forecasting, design of large airplanes, air traffic control, protein folding, and other biological large systems simulations. To enable work in these areas, NSF supports the Partnership of Advanced Computational Infrastructure (PACT), which, though it does not possess the most powerful compute capabilities in the U.S. (when compared to the high end facilities of the DOE Labs, for example), does have the most advanced facility open to all researchers—and most importantly to students. PACT partners develop software for high end facilities, train the future workforce for high performance computing, and disseminate research across the university science communities while serving civilian high performance computing users as well. An excellent example of this kind of interface is NSF's interaction with the National Institutes of Health (NIH), whose research community is using the PACI facilities as their first preference for high-end biomedical and bioinformatics R&D.

As a result of such activities, NSF's future points toward a more holistic view of the evolving cyberinfrastructure which will consist not only of high performance computers but also will embrace broadband connectivity with large databases and distributed instruments. The evolution of the cyberinfrastructure is a natural outgrowth of PACI, since PACI by definition is a distributed partnership. The most important aspect of this development will be the realization of grid computing and the harnessing of our resources to significantly advance our computational capabilities, whether local or geographically distributed.

[Page 37](#)

[PREV PAGE](#)

[TOP OF DOC](#)

As is well-known, NSF's mission is to support basic research for all sciences. This differs significantly from other Federal agencies whose missions are strictly agency-oriented. Because of this division of our missions, NSF basic research is a pillar for all mission-oriented agencies involved with IT R&D. Accordingly, NSF frequently participates in coordinated, cooperative, and joint programs with other agencies. Looking toward the future, NSF's investment portfolio demonstrates an increasing emphasis on high-end computing; research that will eventually "trickle down" and touch the lives of all U.S. citizens; research that will involve the sometimes neglected second- and third-tier research universities as well as their students and graduate students, all of whom represent the future of IT in this country; and long-term high-end research that is not likely to be supported by U.S. businesses due to its long time horizons.

NSF, other participating agencies, and indeed, all U.S. citizens are living in the midst of a massive information revolution. To ensure that we continue to benefit from the fruits of IT, NSF has embarked on a broad, exciting, and ongoing Information Technology Research Program as recommended by the PITAC and further enhanced by the creativity and ideas shared by the entire scientific community. Ultimately, the goal of ITR is ubiquitous connectivity and access for all, including the elimination of the persistent digital divide. that separates the "haves" from the "have nots." In order to achieve this overarching goal, we will need to solve many deep, high end research issues in architecture, networking, sensor IT, I/O devices, and

system security, reliability, and error recovery while developing or evolving entirely new methods of computing and communication. We must also not lose sight of the need to support our undergraduates, graduates, and the workforce itself, preparing them to serve as eventual leaders in the IT revolution. These are indeed challenging times. And NSF is doing its part to help lead the United States and every U.S. citizen toward a happier and more productive new century.

[Page 38](#)

[PREV PAGE](#)

[TOP OF DOC](#)

## BIOGRAPHY FOR RUZENA BAJCSY

Dr. Ruzena Bajcsy ("buy chee") was named the Assistant Director for the Computer Information Science and Engineering Directorate (CISE) on December 1, 1998. As head of NSF's CISE directorate, Dr. Bajcsy manages a budget of approximately \$300 million annually. Dr. Bajcsy is the sixth person to be named to this position since the directorate was created in 1986. She comes to the National Science Foundation (NSF) from the University of Pennsylvania where she was computer science and engineering professor.

Dr. Bajcsy is a pioneering researcher in machine perception, robotics and artificial intelligence. She is a professor both in the Computer and Information Science Department and in the Mechanical Engineering and Applied Mechanics Department and is a member of the Neuroscience Institute in the School of Medicine. She is also director of the university's General Robotics and Active Sensory Perception Laboratory, which she founded in 1978.

Dr. Bajcsy has done seminal research in the areas of human-centered computer control, cognitive science, robotics, computerized radiological/medical image processing and artificial vision. She is highly regarded not only for her significant research contributions but also for her leadership in the creation of a world-class robotics lab, recognized world wide as a premiere research center. She is a member of the National Academy of Engineering as well as the Institute of Medicine. She is especially known for her wide-ranging, broad outlook on the field and cross-disciplinary talent and leadership, successfully bridging such diverse areas as robotics and artificial intelligence, engineering and cognitive science.

[Page 39](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Dr. Bajcsy received her Master's and Ph.D. degrees in electrical engineering from Slovak Technical University in 1957 and 1967, respectively. She received a Ph.D. in computer science in 1972 from Stanford University, and since that time has been teaching and doing research at Penn's Department of Computer and Information Science. She began as an assistant professor and within 13 years became Chair of the department. Prior to the University of Pennsylvania, she taught during the 1950s and 1960s as an instructor and assistant professor in the Department of Mathematics and Department of Computer Science at Slovak Technical University in Bratislava. She has served as advisor to more than 20 Ph.D. recipients.

Chairman **SMITH.** Dr. Bajcsy, you went slightly over your 5 minutes. But inasmuch as you are leaving us—and, at least, I would like to express the Committee's appreciation for your testimony over the past few years before this Committee and your work at the National Science Foundation. I understand you are returning next month back to the University of Pennsylvania. So——



Dr. **BAJCSY**. That is correct. Thank you very much.

Chairman **SMITH** [continuing]. Our thank-yous and our best wishes. And I am going to go a little out of order at this time. Mr. Braun, to have the gentleman from Texas, Mr. Smith, ask you a question that he is going to request that you respond in writing to. Is that right?

Mr. **SMITH OF TEXAS**. Well, actually, Mr. Chairman, I was first going to thank you for having the hearing. But, second, just to ask unanimous consent that I be allowed to address a written question to Mr. Braun that he could respond to over the next week or 10 days. And it was a question in regard to the availability of broadband in rural areas. And I am happy to submit that to him in writing and get a response later so as not to hold up the hearing.

[Page 40](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Chairman **SMITH**. Well, sir, we can somewhat command the authorization of the Committee to you to submit that question. Of course, it is up to Mr. Braun——

Mr. **SMITH OF TEXAS**. I would be——

Chairman **SMITH** [continuing]. Whether he answers you or not.

Mr. **SMITH OF TEXAS**. I am sure he would be happy to. But—and I won't be able to return after these votes, and that is why I wanted to ask you for that, Mr. Chairman.

Chairman **SMITH**. Without objection, it is so ordered. Does that leave us 5 minutes on this—this leaves us 9 minutes on this vote. So, Mr. Braun, we are going to ask you to proceed with your 5-minute presentation. Make it 6 and that will give us 4 minutes to go vote, if that is okay with you gentleman.

Mr. **BRAUN**. I don't see the slide yet. Oh. There it goes.

#### STATEMENT OF DR. HANS-WERNER BRAUN, Ph.D., RESEARCH SCIENTIST, SAN DIEGO SUPERCOMPUTING CENTER

Mr. **BRAUN**. Chairman Smith, and, members of the Subcommittee, it is an honor and a privilege having been invited to testify before you today regarding activities related to my NSF-funded High-Performance Wireless Research and Education Network. I am Principal Investigator on HPWREN and have served as a Principal Investigator for several NSF projects in the past. One such project has been the NSFNET backbone, while I was staffed at the University of Michigan, which was the Internet backbone at the threshold between the Internet being largely a government research project and its commercialization.

[Page 41](#)

[PREV PAGE](#)

[TOP OF DOC](#)

While the government supported NSFNET backbone partnership was key for driving the Internet

evolution, aspects of understanding of usages and needs for ubiquity at high-performance levels lagged behind. I worked on a number of NSF projects addressing such needs. Among them, this HPWREN project. My thoughts are presented in the written testimony. Permit me to just focus here on some of the key highlights.

The main thing is that we are trying to position this as an interdisciplinary collaboration not yet, yet another wireless network. We are working specifically with researchers and scientists and education people in trying to build a system surrounding the specific needs. Specifically, we do—we undertake research to understand application performance requirements while building a prototype network to demonstrate feasibility of wide areas of wireless high-performance network access.

We are focusing on rural access areas for research and education application with this emphasis on interdisciplinary collaboration, that I mentioned before, while specifically focusing also on fixed or ad hoc installations with—meaning installations that are more or less permanent with installations that have been created essentially in real time. And I have an example of that in a minute.

Sorry that the forms are so small, by the way. But the project is funded by the National Science Foundation and led by the San Diego Supercomputer Center and the Scripps Institution of Oceanography. The science applications currently collaborating with us are earth-grade people, two observatories, multiple field stations, and a couple of Indian reservations, including one of the fire stations in an Indian reservation. This has already resulted in a follow-on project in conjunction with Hewlett Packard and the Southern California Tribal Chairman's Association and San Diego County. But Hewlett Packard has made a large award to the group of Indian tribes to create even more of a digital village activity.

[Page 42](#)

[PREV PAGE](#)

[TOP OF DOC](#)

We also started collaboration with crisis management and other agencies. What you see here is the topology which spans about 50 miles in each direction, which is interconnected by mountaintops, essentially, by wireless links into those research and science locations, like the Palomar Observatory, Mt. Laguna Observatory, and a couple of Indian Reservations. The three reservations that you see actually required relay installation on Indian land which were done in collaboration with Native Americans. They have helped us installing it.

What you see here is a depiction of the interconnection between—along—I am sorry—along the 45 megabit per second backbone starting at UCSD via a commercial microwave tower and Mt. Woodson, North Peak, a commercial tower at Stephenson Peak, and the Mt. Laguna Observatory. North Peak in here is done in collaboration with the San Diego's Sheriff's Department and they are kind of very nice to us and we were able to get on there without having to pay them a fee.

A couple of quick images. This is an illustration of earthquake sensors in the desert where there is no universal service, no phone line, no electricity. They have to generate the power on their own and send the data back via wireless networks because there is no choice. This is an example of a vertical and horizontal array.

This is a view onto the Mt. Laguna Observatory, where you can see four domes that are connected in a

very mountainous setting at 6,100 feet that we connected at 45 megabits per second.

Same with the Mt. Palomar Observatory, which is doing now supernova research. Finding those very quickly—they have to find them within the first 48 hours of their decay to go from there to Hubble or other big telescopes to do their research.

[Page 43](#)

[PREV PAGE](#)

[TOP OF DOC](#)

With regard to the Indian reservations, I mentioned before, we needed to relay towers on Indian land. You see this here in the upper right corner where, in this case, there was also no electricity available and we had to generate the electricity locally out of solar power which power gel cell batteries and from there, the radios. So this is a totally self-contained unit on the mountaintop on Indian land.

To research some of the weather conditions and how they impact signal strength, we installed a couple of video cameras, which then later on has tickled the interest of some of the crisis management agencies, in terms of seeing events. For example, the forestry people would be interested in like a couple of cameras that triangulate in on events. And those are Internet-accessible, like you can pan and tilt and zoom them across the Internet.

Something else we investigated are researchers in the field, where a Yagi antenna is simply connected to a wavelength card in a laptop, which allows someone to connect to a mountaintop. And we used this technology later for a demonstration with the California Department of Forestry where we had a relay point, in the middle, in the upper left corner, that powers a radio out of a cigarette lighter in a firefighter vehicle and had two antennas connected, one to the mountaintop in the upper right corner, and one to Jim Garra, the Battalion Chief from San Diego County, in the lower right corner, being able to access the CDF web site at

---

Chairman **SMITH**. Mr. Braun, we are going to ask you, when we return, to try to wrap it up in about another 30 seconds, after we return. So, in order for us to make this vote, we are in recess for the next two votes until we get back here.

[Page 44](#)

[PREV PAGE](#)

[TOP OF DOC](#)

[Recess]

Chairman **SMITH**. Before I hit the gavel, sort of my apologies for having—this is our last week before we go on recess on Friday. The Subcommittee on Research will reconvene and Mr. Braun for your—

Mr. **BRAUN**. Well, I am almost done anyway. It is just one slide left. I think I am finished—pretty well was finished with this slide. And the only remaining one is we did a demonstration with ecology researchers at the Santa Margarita Ecological Reserve who are having the 4,500-acre ecological reserve and—to be implemented multiple links and demonstrated multiple telemetry instruments and a camera to the public there.

These are just examples of what we have been doing, and I can talk at length about many of those. But

besides that, I will just thank you for your attendance and inviting me to this Subcommittee hearing.

[The prepared statement of Mr. Braun follows:]

## PREPARED STATEMENT OF HANS-WERNER BRAUN

High Performance Wireless Research and Education Network

Table of Contents:

1. Written testimony by

[Page 45](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Hans-Werner Braun

High Performance Wireless Research and Education Network

University of California, San Diego

2. Appendices

a. Frank Vernon, UCSD Scripps Institution of Oceanography geophysicist and HPWREN Co-PI

b. Geneva Lofton-Fitzsimmons, UCSD American Indian Outreach Initiative

c. Greg Aldering, LBNL, Palomar Observatory scientist

d. Jared Aldern, Warner Unified School District

e. John Helly, UCSD–SDSC scientist

f. Lorraine Orosco, San Pasqual Band of Indians

g. Pamela Arviso, Two Directions, Inc.

h. Paul Etzel, SDSU Astronomy Department

i. Robert Pozos, SDSU Biology professor

j. Robert Smith and Doretta Musick, Pala Band of Mission Indians

k. Sedra Shapiro, SDSC Field Station

l. Srinivas Sukumar, Hewlett-Packard Company

3. Oral presentation slides

4. Posters

5. Disclosure letter

[Page 46](#)

[PREV PAGE](#)

[TOP OF DOC](#)

6. Curriculum vitae

## INTRODUCTION

My thanks to Chairman Smith, Ranking Member Johnson, and members of the Subcommittee on

Research for the opportunity to discuss my activities related to the NSF funded High Performance Wireless Research and Education Network—referred to as HPWREN. I am Principal Investigator for HPWREN and have previously served as a Principal Investigator for several NSF projects in the past. One such project has been the NSFNET backbone, which was the Internet backbone at the threshold between the Internet being largely a government research project and its commercialization.

Several HPWREN collaborators have contributed materials that I would like to bring to your attention. Though I will mention these throughout my testimony to you, I would also encourage you to review the written letters, as they contain valuable information related to current and future network applications in rural areas.

I believe that the involvement of the federal government in the evolution of the Internet is as crucial as ever. While the federal government has historically played a key role in driving the network performance edge, significant areas remain underdeveloped, including the sophistication of Internet applications and national network ubiquity fulfilling demanding performance requirements. For example, in rural America, even in technologically advanced areas such as San Diego, the notion of high performance quickly falls apart outside major populated areas, where even cell phone systems often turn into an illusion of reachability. However, the technology needs of rural areas—of perhaps no immediate business case to commercial service providers—should not be underestimated. Stimulating data communications needs and solutions today can pay off significantly over time.

[Page 47](#)

[PREV PAGE](#)

[TOP OF DOC](#)

## PROJECT DESCRIPTION

Measuring approximately 50 miles by 50 miles, the HPWREN project aims to demonstrate ways in which a high performance network can be created and used by network applications of remote communities. HPWREN fosters research and education opportunities for Americans that need and deserve the same access as those of us who live in urban areas.

HPWREN directly impacts the ways in which area field scientists—such as ecologists, astronomers, and earthquake researchers conduct their studies. HPWREN also provides rural Native American communities in the area with connectivity for education opportunities such as interactive computer classes and remote tutoring programs, and has been directly stimulating such activities. We first work with the network users to define their needs—and then build the network accordingly. This attempts a model shift from the predominantly existing practice of driving the needs from a network-centric view, towards people being able to first define their needs and the network being developed to best fit the situation. It is a partnership that begins with an outreach by the network developer to the scientist and educator users of the network.

## CURRENT AND FUTURE APPLICATIONS

The HPWREN project builds and investigates high performance networks for research and education applications in rural San Diego County, working directly with people responsible for those applications. The research applications specifically include geophysicists who need access to real-time earthquake data in remote deserts and mountains, astronomers needing to detect supernovae within their first hours of decay

and near-earth asteroids before they come too close, and ecologists who expect to interact with their field stations from home institutions while also incorporating real-time data into classroom curricula. Our education facet focuses on rural Native American Learning Centers, which serve both tribal and non-tribal members of the local communities.

[Page 48](#)

[PREV PAGE](#)

[TOP OF DOC](#)

In addition to the research and education applications, we are also investigating ad-hoc advanced network development and experimentation, while collaborating with local crisis management agencies. For example, we recently participated in a demonstration with firefighters of the California Department of Forestry—building, demonstrating, and tearing down a high-performance network within a few hours—and under bad weather conditions. This activity demonstrated that such transient high performance network infrastructures are entirely feasible.

The HPWREN research focuses on users and applications in need of high performance networking in these rural, and often quite remote, areas. Considerations include who needs what network performance where and when, and how such information can be used to evolve critical infrastructures? How are performance parameters to be defined that understandably communicate provided services? And how can the needs of new and revolutionary networking applications be assessed and integrated into evolving network architectures? How should the infrastructure deal with legitimate applications able to absorb all available resources by themselves.

## IMPORTANCE OF FEDERAL FUNDING TO NETWORK TECHNOLOGY ADVANCEMENTS

The issue of government funding needs to be separated into component issues. At least one is where funding is coming from, another is how the funding is being used. About 15 years ago the NSFNET began to trigger the evolution of the Internet from a government research and development environment towards today's large and commercialized infrastructure. NSF funds were critical to accomplish this, but, once successful, private companies raised issues about government competition. My belief is that the Internet would not have become as quickly successful without the early investment made by the National Science Foundation. After the initial seeding of new initiatives with government funding, the private market should be encouraged to evolve projects, with NSF funding eventually refocusing on the next "new, high risk things."

[Page 49](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Even for HPWREN, there have already been opportunities of leveraging and partnerships, such as seen by the five million dollar Hewlett-Packard award to the Southern California Tribal Chairman's Association—more than twice the amount of the NSF HPWREN funding. The Hewlett-Packard grant aims to build upon the prototypes created with the NSF funding for learning centers in three reservations, and to create digital village settings within all 18 reservations of the County. The hope is that the evolution of these Hewlett-Packard activities—which are largely led by the reservations themselves—will eventually create a solid infrastructure that allows the tribes to become self-sufficient.

In addition, from the indications I have, including attached support materials, the high speed connections for the science applications we are integrating, would not have happened without NSF's HPWREN investment. This is in large part due to funding. However, the human element of scientists collaborating in an interdisciplinary way with network researchers who are trying to understand, assess, and deliver on their needs, should not be underestimated. In support of this, the measurement infrastructure included in the HPWREN project, and a direct result of prior NSF funded activities, allows us to accurately determine how network resources are being consumed, and we can use such results in turn to evolve future networking environments, while making performance data publicly available.

## HIGHLIGHTS FROM THE APPENDICES

I would like to suggest your attention to some highlights I found particularly interesting in the testimony support letters provided to you from collaborators of the HPWREN project in the appendices.

[Page 50](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Frank Vernon, UCSD–SIO geophysicist

If a significant earthquake occurs in our region, we can quickly deploy additional sensors in the epicentral region, evaluate the data, and continue to adapt the station locations to maximize the scientific data return. This is especially important when some or all of the sensors need to be placed at remote sites. The ability to adapt environmental monitoring systems to significant transient events will be of great benefit to all field sciences.

Geneva Lofton-Fitzsimmons, UCSD American Indian Outreach Initiative

. . .the wireless Internet has provided a reliable educational resource, allowing students to do research and other projects.

Greg Aldering, LBNL Staff Scientist, astronomy

This digital image subtraction involves numerous steps to align the images and account for blurring by the Earth's atmosphere, and requires the equivalent power of 50 desktop computers to keep up with the data. Because the amount of data is so large (50 billion bytes per night), the image archive even larger (presently 8 trillion bytes and growing), and the computations so extensive, it is critical that the imaging data be transferred to a large computing center (in this case NERSC, the National Energy Research Scientific Computing Center at LBNL) as quickly as possible. Since such extensive computing facilities could not be maintained at the observatory, the only alternative to a fast data link would be to write the data to tape and ship it, at the cost of delaying supernova discoveries by several days.

[Page 51](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Jared Aldern, Warner Unified School District

Most importantly, several of our high school students were able to take part in the installation of the equipment. These students gained the satisfaction of providing a service to their community and valuable experience working with high-tech equipment, and their eyes were opened onto a whole new world of endeavors by the opportunity to collaborate with world-class scientists and technicians.

Anyone who cares about the future of rural America—and who realizes how tightly the fate of the countryside is tied to that of urban and suburban America—will do well to consider the state of the nations technological infrastructure in all areas.

John Helly, UCSD researcher

A particular challenge in this work, and an aspect that makes it especially interesting to other parts of the world, is the general lack of information within San Diego County regarding surface water flow.

Lorraine Orosco, San Pasqual Band of Indians

HPWREN has given the Tribes of San Diego a vision of the possibilities of using this technology to build stronger communities.

Pam Arviso, Two Directions, Inc.

[Page 52](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Thanks to the HPWREN project, our labs have been connected to the high-speed Internet for almost nine months.

Please convey to the National Science Foundation the sincere appreciation of the Native American community from North San Diego County, California for providing us with this very important technology.

Paul Etzel, SDSU astronomer

...and a consensus is now building between SDSU and other CSU astronomers to build a much larger 3.0 meter class telescope to operate in a completely Robotic mode over Internet2 via HPWREN.

Robert Pozoz, SDSU biologist

Presently, the College of Sciences at San Diego State University and the Super Computer Center at University of California San Diego are jointly developing interactive educational programs for rural American Indian reservations. The educational programs will cover the physiology and clinical implications dealing with diabetes and obesity.

Robert Smith and Doretta Musick, Pala Band of Mission Indians

We have seen a lot of progress, where students have raised their math levels from D's and F's to A's and B's. Parents are pleased, not only for what its doing for their children, but also how it has helped their



community.

[Page 53](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Sedra Shapiro, SDSU Field Station Programs

SDSU efforts to collect and disseminate environmental monitoring data to a variety of users are significantly enhanced by the HPWREN wireless network and an integrated, real-time data management and delivery system. Each of these efforts involves significant IT research challenges including those associated with networking remote sensor arrays, integrating diverse monitoring platforms, acquiring data in real-time, and archiving it continuously. It is important to demonstrate on a regional scale that multidisciplinary environmental monitoring is both practical and scalable.

Srinivas Sukumar, Hewlett-Packard E-Inclusion Executive

...we realize that it is possible because of a creative public-private partnership. Without public funding for Internet infrastructure development—like those programs that bring education to remote schools over the Internet—the Tribal Digital Village would not exist. We respectfully request your coordination and support of increased funding of these important initiatives. Together, we can make a difference.

## THE FUTURE OF WIRELESS NETWORKS

Performance metrics to consider include more than speed, but also real ubiquity and ease of a speedy and cost effective setup. It includes technology being an enabler for science and education activities. For example, astronomers are interested in remotely steerable telescopes and robust image databases. Seismologists need more and more earthquake sensors with reliable real-time access and quick deployability. Ecologists need real-time access to telemetry data in remote areas so that they can incorporate information into databases and classroom curricula. Educators in rural locations need communication capabilities to maintain education opportunities via distance education and tutoring. The HPWREN project attempts to be an early enabler for such applications at a time—and in an area—where there is currently no commercial business case.

[Page 54](#)

[PREV PAGE](#)

[TOP OF DOC](#)

## CLOSING REMARKS

Again, I would like to thank you for your attention to these considerations, and emphasize that much work remains to be done to ensure equal network access for both urban and rural research and education communities, and to make, a real difference where ever-increasing performance needs require to be fulfilled. Initially it may benefit science and education, but eventually it needs to extend to all people interested in the future of this society.

It has been an honor and a privilege to testify before you today, and I welcome your questions.

## APPENDICES

74543f.eps

[Page 55](#)

[PREV PAGE](#)

[TOP OF DOC](#)

74543g.eps

74543h.eps

74543i.eps

74543j.eps

74543k.eps

74543l.eps

74543m.eps

74543n.eps

74543o.eps

74543p.eps

74543q.eps

74543r.eps

74543s.eps

[Page 56](#)

[PREV PAGE](#)

[TOP OF DOC](#)

74543t.eps

74543u.eps

74543v.eps

74543w.eps

74543x.eps

74543y.eps

74543z.eps

74543aa.eps

74543bb.eps

74543cc.eps

74543dd.eps

74543ee.eps

[Page 57](#)

[PREV PAGE](#)

[TOP OF DOC](#)

74543ff.eps

74543gg.eps

74543hh.eps

74543ii.eps

74543jj.eps

74543kk.eps

74543ll.eps

74543mm.eps

74543nn.eps

74543oo.eps

Chairman **SMITH**. Well, thanks. The thanks is ours. Dr. Berman.

STATEMENT OF HELEN BERMAN, Ph.D., DIRECTOR, THE PROTEIN DATA BANK; BOARD OF GOVERNORS PROFESSOR OF CHEMISTRY, RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY

Dr. **BERMAN**. Good afternoon. My name is Helen Berman and I am a Board of Governors Professor of Chemistry at Rutgers University, and I am also the Director of the Protein Data Bank and Nucleic Acid Database. These are archives that contain information about the three-dimensional structures of biological molecules that include DNA, RNA, and proteins. These are the molecules of life that are found in all organisms, including bacteria, yeast, plants, flies, mice, healthy, as well as diseased, human beings.

I appreciate the opportunity to speak before you today to discuss how innovations in computer technology have revolutionized the way we do biology.

Exactly 30 years ago this summer, 1971, the producers and potential users of structural data agreed that a single formal electronic international archive of structures should be established. This was an extremely visionary step since, at that time, there were less than a dozen structures, and each was very simple and very small. Now, in the year 2001, the PDB, now managed by a consortium called the Research Collaboratory for Structural Bioinformatics, contains almost 16,000 such structures, including the manufacturing site of proteins, the ribosome.

The incredible growth in both the number of complexity, which we could only imagine 30 years ago, is the result of tremendous advances in technology in protein chemistry, in robotics, in imaging, and high-performance computing.

The way in which the data are organized and distributed to the worldwide community has also changed dramatically as computer technology has improved. In the beginning, data was distributed on punched cards, magnetic tape, and the processed data were distributed via the U.S. mail on magnetic media. The data were loosely structured in a free-text type format.

Today, the data are submitted to the PDB electronically via the Internet. The development of software has made it possible to use modern database technologies that were originally designed for business applications to organize and archive the biological data. Advanced query and visualization tools make it possible to analyze and view individual structures in real time. New algorithms for structure and sequence comparison, the ability to crosslink with other data resources—all of these things have made it possible to obtain a more complete picture of all that is known about the systems under study.

Today, there are more than 100,000 structure files downloaded. In 1 year, there are more than 1.5 million transactions at the primary PDB site alone. Scientists in academia, government, and industry access the PDB. These scientists use these data to plan new experiments, analyze their results, and compare them with others. These analyses are aimed primarily at trying to understand how these molecules work, how they interact, and they contain key information for drug discovery.

Significantly, the PDB has emerged as a powerful teaching tool and is used by students in classrooms around the world.

The complexity of the data provided many, many challenges for computational biology. Methods for visualizing molecules, predicting structures, comparing structures, using the sequence information that came out—has come out as part of the genome projects—all of this has provided a fertile test bed for people doing computational methods and has—and for developing advanced methods for data mining.

[Page 60](#)

[PREV PAGE](#)

[TOP OF DOC](#)

The challenge for the PDB is that the PDB does not stand alone. It must interoperate with other bases in order for us to really understand how biological molecules work. Our long-term goal is to define biology at a molecular level. To help, PDB has led the way in developing a dictionary that will give the terms that will allow other databases to interoperate with us. We will also have increasing demand for rapid query and the Internet must be able to handle the volume of requests that will be needed to download these very complex files.

The PDB is run by a consortium consisting of three institutions: Rutgers, San Diego Supercomputer Center, and NIST. The roles and responsibilities of each institution are consistent with its strengths and with its capabilities.

According to a memorandum of understanding, three agencies fund the PDB: the National Science Foundation, the National Institutes of Health, and the Department of Energy. A Cooperative Agreement between Rutgers and NSF outlines the responsibilities of the parties involved in running the PDB. We have a fair amount of oversight by various kinds of advisory committees. We are very happy with having this kind of involvement by so many agencies, however, there are problems in preparing budgets for such a complex organization involving three institutions and three agencies.

The PDB is a strategic resource in modern biology. Historically, it has enjoyed a synergistic relationship with computer and information science. Now, with the massive amounts of data being produced, it is critical that the PDB continue to forge this relationship with information technology. Harnessing the data within the PDB and combining it with data from other resources will enable biology to continue to advance at a breathtaking speed.

[Page 61](#)

[PREV PAGE](#)

[TOP OF DOC](#)

To do this, it is critical that the PDB have long-term fiscal stability so it can continue to be a key international resource in this post-genomic era. Thank you very much for letting me speak.

[The prepared statement of Dr. Berman follows:]

#### PREPARED STATEMENT OF HELEN M. BERMAN

Good afternoon. My name is Helen Berman. I am a Board of Governors Professor of Chemistry at Rutgers University and the Director of the Protein Data Bank (PDB) and the Nucleic Acid Database (NDB). The PDB and NDB are archives containing information about the three dimensional structures of biological molecules including DNA, RNA and proteins. These are the molecules of life that are found in all organisms

including bacteria, yeast, plants, flies, mice and healthy as well as diseased humans (*Figure 1*).

I appreciate the opportunity to speak before you today to discuss how innovations in computer technology have revolutionized how we do biology.

## History of the PDB

Exactly thirty years ago, the producers and potential users of structural data agreed that a formal electronic international archive of structures should be established. Thus, the Protein Data Bank was born at Brookhaven National Laboratory with distribution sites in Cambridge, England and later in other sites around the world. This was a visionary step since at the time there were less than a dozen structures and each one was very simple and small. Now in 2001, the PDB (now managed by a consortium called the Research Collaboratory for Structural Bioinformatics (RCSB)) contains almost 16,000 such structures including the manufacturing site of proteins-the ribosome (*Figure 2*).

[Page 62](#)

[PREV PAGE](#)

[TOP OF DOC](#)

This incredible growth in both number and complexity (which we could only begin to imagine thirty years ago) is the result of tremendous advances in technology in many fields. With the advent of cloning, proteins can be expressed rapidly. Crystals of proteins are required in order to perform the x-ray diffraction experiments that lead to the structures. Obtaining a high quality single crystal of a protein is often the rate-limiting step in determining the protein structure. Robots have now been developed to aid in crystallization. Advances in imaging technology, the availability of intense x-ray sources at the synchrotron sites, and advanced computer technologies have allowed some structures to be determined in a matter of days rather than months and years as before. Thus, when the sequences of genomes including the human genome were determined the concept of determining structures in a factory-like or high throughput mode on a genomic level became truly feasible rather than a vague possibility.

The way in which the data in the PDB are organized and distributed to the worldwide community has changed dramatically as computer technology has improved. In the beginning data were submitted on punched cards or on magnetic tapes and the processed data were distributed via the U.S. mail on magnetic media. The data were loosely structured with free text information describing how the experiment was done and what the results might mean.

## The Modern Day PDB

Today data are submitted to the PDB electronically via the Internet. The development of well-defined software accessible data dictionaries and controlled vocabularies has made it possible to use database technologies (originally designed for business applications) to organize and archive the biological structural data. Advanced query and visualization tools make it possible to analyze and view individual structures in real time over the World Wide Web. New algorithms for structure and sequence comparison coupled with the organization of all the data in a database enable scientists to look at characteristics of groups of structures in order to derive new knowledge about the principles that underlie protein folding. The ability to crosslink with other data resources has made it possible to obtain a more complete picture of all that is known about the systems under study. The availability of new types of storage media has enabled the PDB

to keep pace with the increasing volume of data and to ensure that there are redundant copies of this irreplaceable data. The availability of the World Wide Web makes it possible to provide many mirror sites that serve the global community of scientists.

[Page 63](#)

[PREV PAGE](#)

[TOP OF DOC](#)

## Usage of the PDB

Every day 100,000 structure files are downloaded. In one year there are more than 1.5 million transactions on the primary PDB site alone. Scientists in academia, government and industry access the PDB. These scientists use the data to plan new experiments, analyze their own results and compare them with others. These analyses are aimed primarily at trying to understand how these molecules work and how they interact. The data within the PDB provide information essential to understanding biology at its most basic level as well as provide the key elements for drug discovery. Today all that is required to obtain this information is a computer with access to the Internet.

Significantly, the PDB has also emerged as a powerful teaching tool and is used by students in classrooms around the world.

## The PDB and Computational Methods

The complexity of the data has provided many challenges for computational biology. In the beginning, methods for visualizing molecules and understanding their individual characteristics were the subject of many new types of algorithm development. Methods for predicting structures from first principles became a computational grand challenge. When the number of structures increased, methods for comparing structures became the subject of important research. As more and more DNA and protein sequences became available, the possibility of predicting new structures using both sequence and structure data became the subject of a great deal of research. Finally, the complex data contained in PDB entries including free text, well-defined data items describing experimental and structural features, and the three-dimensional coordinates collectively provides a fertile test bed for advanced methods of data mining.

[Page 64](#)

[PREV PAGE](#)

[TOP OF DOC](#)

## Challenges for the PDB

The number and complexity of structures will continue to grow. People will continue to exploit the information contained within the PDB and so new tools will continue to be developed to allow data exploration. But the PDB does not stand-alone. It must inter-operate with other databases that contain information about the chemistry and biology and activity of these molecules. The ability to combine and analyze all of these data is required in order to meet the long-term goal of defining biology at a molecular level. How to accomplish this type of inter-operation is a considerable challenge. To help do this, the PDB has led the way in developing a comprehensive dictionary for the data contained within it. Because of this, it has been possible to create specifications for an application programming interface (API) which will allow

other databases to inter-operate with it. Similar work is now ongoing in other disciplines. When these data dictionaries and supporting APIs are fully developed the exploitation of the massive amount of biological data will be possible.

It is also clear that as the PDB grows and the information contained within it becomes ever more valuable and sought after, there will be an increasing demand for rapid query. The Internet must be able to handle the volume of requests and the need to download large complex files. Thus, the PDB must have adequate resources to service this demanding level of access.

## Management and Funding of the PDB

The PDB is managed by a consortium consisting of three institutions: Rutgers, The State University of New Jersey; San Diego Supercomputer Center, University of California, San Diego; and the National Institute of Standards and Technology. Each institution has roles and responsibilities consistent with its strengths and capabilities. The Rutgers site handles incoming data processing and the development of data dictionaries; the San Diego site is responsible for distribution and query; the NIST site maintains the physical archive and plays a role in data uniformity. This type of distributed management would never be possible without the type of communication tools that we now have.

[Page 65](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Since its inception, the PDB has maintained the single worldwide archive of structural information and has worked closely with the international community so that it would continue to be a global resource.

According to a memorandum of understanding; three agencies fund the PDB: the National Science Foundation, the two National Institutes of Health, and the Department of Energy. It is important that these agencies all participate because each brings different scientific interests and expertise. A Cooperative Agreement between Rutgers and the NSF outlines the responsibilities of the parties involved in running the PDB. There are different types of review mechanisms to help ensure that the PDB continues to fulfill its mission in the best possible way. Progress Reports are submitted every six months for review by the agencies. In addition, the PDB has an international Advisory Committee consisting of senior scientists from academia and industry with expertise in x-ray crystallography, nuclear magnetic resonance, computer science and education. This committee provides the type of critical feedback that allows the PDB to continue to evolve and improve.

The budget preparation, however, presents problems. Each year Rutgers prepares a budget that includes SDSC as a subcontract. Because NIST is a federal agency itself, it cannot have a subcontract for personnel and instead receives its funds as an interagency transfer. The differences in the way that each agency deals with the funding process means that there is considerable paperwork involved each year in preparing the budgets and having them actually funded. Streamlining this process would allow our resources to be focussed more fully on continuing to develop and maintain a world class PDB.

[Page 66](#)

[PREV PAGE](#)

[TOP OF DOC](#)



## Conclusion

The PDB is a strategic resource in modern biology. Historically it has enjoyed a synergistic relationship with computer and information science. Now with the massive amount of data being produced it is critical that the PDB continue to forge this relationship with information technology. Harnessing the data within the PDB and combining it with data from other resources will enable biology to continue to advance at breathtaking speed.

To do this it is critical that the PDB have long term fiscal stability so that it can continue a key international resource in this post genomic era.

74543u3.eps

74543v3.eps

## BIOGRAPHY FOR HELEN M. BERMAN

### Personal

Address: Rutgers, The State University of New Jersey, Wright-Rieman Laboratories, Department of Chemistry, 610 Taylor Road, Piscataway, NJ 08854-8087; Telephone: (732) 445-4667; FAX: (732) 455-4320; E-Mail: berman@rcsb.rutgers.edu

Birthdate: May 19, 1943, Chicago, Illinois

[Page 67](#)

[PREV PAGE](#)

[TOP OF DOC](#)

### Education

A.B. 1964 Barnard College, Degree with Honors in Chemistry

Ph.D. 1967 University of Pittsburgh, Advisor: G.A. Jeffrey

### Appointments

1967-1969 National Institutes of Health Postdoctoral Trainee, Univ. of Pittsburgh, Pittsburgh, PA

1969-1989 Research Associate, Assistant Member, Associate Member, Member, Senior Member Fox Chase Cancer Center, Institute for Cancer Research, Philadelphia, PA

1972-1994 Research Collaborator, Department of Chemistry, Brookhaven National Laboratory, Upton, NY

1982-1989 Director, Research Computer Facilities, Fox Chase Cancer Ctr., Philadelphia, PA

1985-1993 Adjunct Professor of Chemistry, Department of Chemistry, University of Pennsylvania, Philadelphia, PA

1992-1997 Adjunct Professor of Crystallography, University of Pittsburgh, Pittsburgh, PA

1989-1999 Professor II, Department of Chemistry, Member, Waksman Institute, Rutgers University, New Brunswick, NJ

2000- Board of Governors Professor of Chemistry, Rutgers University, New Brunswick, NJ

### Professional Societies

American Chemical Society, American Crystallographic Association, American Society for Biochemistry and Molecular Biology, Biophysical Society, Sigma Xi

[Page 68](#)

[PREV PAGE](#)

[TOP OF DOC](#)

## Honors

Board of Governors Professor of Chemistry; Fellow of the Biophysical Society; 2000 Distinguished Service Award, Biophysical Society; Outstanding Woman Scientist Award, Assoc. for Women in Science (AWIS), NY Chapter; Fellow, American Association for the Advancement of Science; Haddow Fellow, Institute of Cancer Research; New Jersey Woman of Achievement

## Professional Service

Member, Molecular and Cellular Biophysics Study Section, 1980–1984; Adhoc Member, General Medicine Council of NIH, 1986, 1990; Member, NAS Committee on the Status of Crystallography, 1975; Member, USA National Committee for Crystallography, 1981–1993; Chairman, Small Molecule Special Interest Group, American Crystallographic Association, 1985; Program Chairman, American Crystallographic Association Annual Meeting, 1987; President, American Crystallographic Association, 1988; Member, Council for Scientific Society Presidents, 1988; Consultant, Franklin Institute of Science, 1988; Member, Board of Scientific Counselors, National Center for Biotechnology Information, National Library of Medicine, 1990–1994; Member, Council of the Biophysical Society, 1991–1994; Member, Advisory Committee for the Biological Sciences, National Science Foundation, 1992–1996; Director, Nucleic Acid Database, 1992–; Member, NRC Committee in National Needs for Biomedical and Behavioral Research Personnel, 1993–1994; Editorial Board, Journal of Biological Chemistry, 1994–; Chair, Publications Committee, Biophysical Society 1994–1996; Member, Nomenclature Committee, International Union of Biochemistry and Molecular Biology, 1995–; Chair, Database Committee, International Union of Crystallography, 1996–; Director, Protein Data Bank, 1998–; Editorial Board, Biochemistry, 1999–.

[Page 69](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Number of Publications to Date: 145

Selected Publications (1998–):

"The Nucleic Acid Database: A Resource for Nucleic Acid Science," H.M. Berman, C. Zardecki, and J. Westbrook, *Acta Cryst.*, D54:1095–1104, 1998.

"An Analysis of the Relationship Between Hydration and Protein-DNA Interactions," J. Woda, B. Schneider, K. Patel, K. Mistry, and H.M. Berman, *Biophysical J.*, 75:2170–2177, 1998.

"Hydration of the Phosphate Group in Double-Helical DNA," B. Schneider, K. Patel, and H.M. Berman, *Biophysical J.*, 75:2422–2434, 1998.

"The Crystal Structure of an Autoprocessed Ser221Cys-subtilisin E-Propeptide Complex at 2.0 Å Resolution," S.C. Jain, U. Shinde, Y. Li, M. Inouye, and H.M. Berman, *J. Mol. Biol.*, 284:137–144, 1998.

"Patterns of Hydration in Crystalline Collagen Peptides," R.Z. Kramer and H.M. Berman, *J. Biomol. Struct. Dyn.*, 16:367–380, 1998.

"The Past and the Future of Structure Databases," H.M. Berman, *Curr. Opin. Biotech.*, 10:76–80, 1999.

[Page 70](#)

[PREV PAGE](#)

[TOP OF DOC](#)

"À la Mode: a Ligand and Monomer Object Data Environment I. Automated Construction of mmCIF Monomer and Ligand Models," L. Clowney, J. Westbrook, and H.M. Berman, *J. Appl. Crystallogr.*, 32:125–133, 1999.

"The Nucleic Acid Database: A Research and Teaching Tool," H.M. Berman, C. Zardecki, and J. Westbrook, in *Handbook of Nucleic Acid Structure* (ed. S. Neidle) Oxford Univ. Press, p.77–92, 1999.

"Nucleic Acid Hydration," H.M. Berman and B. Schneider, in *Handbook of Nucleic Acid Structure* (ed. S. Neidle) Oxford University Press, p.295–310, 1999.

"Protein-DNA Interactions: A Structural Analysis," S. Jones, P. van Heyningen, H.M. Berman, and J.M. Thornton, *J. Mol. Biol.*, 287: 877–896, 1999.

"Sequence Dependent Conformational Variations of Collagen Triple-Helical Structure," R.Z. Kramer, J. Bella, P. Mayville, B. Brodsky, and H.M. Berman, *Nat. Struct. Biol.*, 6:5, 1999.

"Water-Mediation of Hydrogen Bonds in Collagen Triple-Helical Structure," H.M. Berman and R.Z. Kramer, in *Perspectives in Structural Biology* (eds. M. Vijayan, N. Yathindra, A.S. Kolaskar) Universities Press, 14:169–178, 1999.

"The Protein Data Bank (PDB)," H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne, *Nucleic Acids Research*, 28:235–242, 2000.

[Page 71](#)

[PREV PAGE](#)

[TOP OF DOC](#)

"An Overview of the Structures of Protein-DNA Complexes," N.M. Luscombe, S.E. Austin, H.M. Berman, and J.M. Thornton, *Genome Biology*, 1(1):1–37, 2000.

"Integrin-Collagen Complex: a Metal-Glutamate Handshake," J. Bella and H.M. Berman, *Structure*, 8:8121–8126, 2000.

"The Protein Data Bank and the Challenge of Structural Genomics," Helen M. Berman, T.N. Bhat, Philip E. Bourne, Zukang Feng, Gary Gilliland, Helge Weissig, & John Westbrook, in *Nat. Struct. Biol.*, 7:957–959, 2000.

"Staggered Molecular Arrangement of a Collagen-like Peptide With a Single Charged Pair," R.Z. Kramer, M. Venugopal, J. Bella, P. Mayville, B. Brodsky, and H.M. Berman, *J. Mol. Biol.*, 301:1191–1205, 2000.

"Quality Controls in Databanks for Molecular Biology," E.E. Abola, A. Bairoch, W.C. Barker, S. Beck, D. A. Benson, H. Berman, G. Cameron, C. Cantor, S. Doubet, T.J.P. Hubbard, T.A. Jones, G.J. Kleywegt, A.S. Kolaskar, A. Van Kuik, A.M. Lesk, H.W. Mewes, D. Neuhaus, F. Pfeiffer, L.F. TenEyck, R.J. Simpson, G. Stoesser, J.L. Sussman, Y. Tateno, A. Tsugita, E.L. Ulrich, J.F.G. Vliegthart, *BioEssays*, 22:1024–1034, 2000

"The PDB Data Uniformity Project," T.N. Bhat, Philip Bourne, Zukang Feng, Gary Gilliland, Shri Jain, Veerasamy Ravichandran, Bohdan Schneider, Kata Schneider, Narmada Thanki, Helge Weissig, John Westbrook and Helen M. Berman, *Nucleic Acids Research*, 29:214–218, 2001.

[Page 72](#)

[PREV PAGE](#)

[TOP OF DOC](#)

"Protein-RNA Interactions: A Structural Analysis," Susan Jones, David T. A. Daley, Nicholas M. Luscombe, Helen M. Berman, and Janet M. Thornton, *Nucleic Acids Research*, 29:943–954, 2001.

"Checking Nucleic Acid Crystal Structures," U. Das, S. Chen, M. Fuxreiter, A.A. Vaguine, J. Richelle, H.M. Berman & S.J. Wodak, *Acta Cryst.*, D57:813–828, 2001.

74543pp.eps

Chairman **SMITH**. Dr. Berman, thank you. Mr. Blake.

STATEMENT OF WILLIAM BLAKE, VICE PRESIDENT FOR HIGH-PERFORMANCE TECHNICAL COMPUTING, COMPAQ COMPUTER CORPORATION

Mr. **BLAKE**. The risks of technology. Thank you. Good afternoon. My name is Bill Blake. I represent Compaq Computer Corporation and I am here to talk about—excuse me.

Chairman **SMITH**. Just tell us what those lines mean, Mr. Blake.

Mr. **BLAKE**. Okay, sir. We—probably, if we can look—can you see the monitors on the side? Those seems to be working. If we can use those, that will—

[Page 73](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Chairman **SMITH**. Well, we want you to know, this is one of the most sophisticated technological hearing rooms in Congress.

Mr. **BLAKE**. Okay. What I would like to start with is a very quick overview of the important trends in high-performance computing, supercomputing, if you wish, that have been going on over the past 10 to 20 years. And this chart is really an attempt to highlight some of the important changes that have occurred over time.

If you look at the chart, the intent is to look at computer speed, the high-performance aspects of a single computer—and that is the vertical axis. And as we go higher on the axis, it indicates a shorter and shorter cycle time or faster and faster single processor. And on the horizontal axis, as we move out to the left, it indicates more and more processors being put into a single system. And the reason for this particular chart is that in the early days of supercomputing, the focus was on taking a single processor, or small number of processors, and making it process information as fast as possible.

Over time, there were some very different approaches to supercomputing which was to take hundreds of thousands of processors and, in effect, build a supercomputer out of a massive parallel array. Both of those approaches hit limitations. On the supercomputer side, in the attempt to make single processors or small numbers of supercomputers faster and faster, certain physical limits were reached—the ability to cool the machine, the ability to use reasonable manufacturing materials. It became a more exotic process. And in the massive parallel case, it simply became very difficult to program a machine to do effective work.

What happened over time, and the current generation of machines, really drifted toward the center of the curve where very fast microprocessors arranged in, in effect, multiprocessors that could be used in the commercial marketplace, as servers or database machines, were put together in a relatively small number called clusters. And it is these fast microprocessors in servers arranged in a cluster that are delivering the very high levels of performance at sites like the Department of Energy and the Pittsburgh Supercomputing Center.

[Page 74](#)

[PREV PAGE](#)

[TOP OF DOC](#)

The important technological trend to track here is that we have, with the help of leading appliers of high-performance computing, pushed the scaling limits of these, what would be off-the-shelf technologies, with improved interconnects, with high bin, with memory systems, to the point of solving extremely difficult problems, like the processing of the genome——

Chairman **SMITH**. Mr. Blake, I am going to interrupt you. And I am going to have to excuse myself for the next 20 minutes for a meeting, and I will return. The gentleman from New York will take the Chair and preside.

Mr. **GRUCCI** [presiding]. Please continue.

Mr. **BLAKE**. Thank you. So the current generation of supercomputers and the investments made, both at the resource level through Federal funding of massive computing at the national labs and terascale computing sponsors with the NSF, have provided us with an ability to scale up commercial technology in very large configurations to, in effect, go beyond in performance where classic supercomputers were able to reach.

Another important, very important, trend is these machines are increasingly being applied to the simulation and modeling of product developments and scientific research activities in a way that simply has not been able to be achieved with small-scale computing. Simulation application in a specific area, like biology, and advanced computer techniques to provide the scaling, are creating results that have heretofore been unachievable.

An important way of summarizing this is that modeling and simulation on these large-scale computers is actually creating an important third component of scientific method where science has relied upon theory and validation of theory. Through experiment, simulation and modeling now is becoming a third component of the scientific process used in industry for creating virtual prototypes of new products, but used in research for exploring phenomenon that in the laboratory sometimes is impossible to produce. Thank you.

[The prepared statement of Mr. Blake follows:]

#### PREPARED STATEMENT OF WILLIAM BLAKE

Very High Performance Computing: Making correct decisions and new discoveries in "zero time"

Term: HPTC

Definition: High Performance Technical Computing (HPTC) is the use of computers to produce information by simulating the behaviour of products and processes, or by algorithmic analysis of very large quantities of data.

*Supercomputing: "A supercomputer changes a compute-bound problem into an I/O bound problem"*  
*Seymour Cray*

*Cluster Supercomputing: "When you've got too much weight in your wagon for your pair of horses to pull, you don't get bigger horses, you get a bigger team of horses" Adm. Grace Hopper*

High Performance Technical Computing (HPTC) is the use of computers to produce information by simulating the behaviour of products and processes, or by algorithmic analysis of very large quantities of data.

The information resulting from HPTC is used to advance understanding in science and technology, in the development of products and services, and in defense and national security applications.

The availability of information when that information is needed to take decisions is a critical aspect of zero-latency decision-making. Increasingly, information needed to take decisions in the development and delivery of new products and processes comes from the result of simulation, or the analysis of large quantities of data ("What happens if. . .?", "What would be the consequences of. . .?")

The competitive advantage that accrues through the ability to make better decisions is the ultimate driving force behind the growth of computer performance.

Demands on the verisimilitude of simulation (the degree to which a simulation provides a true semblance of

actuality) are becoming ever stricter; margins for error, ever finer. The consequence is the need to deliver more accurate simulation results faster, and to extract more accurate and reliable information from ever-larger quantities of data.

This insatiable demand for computational performance in all its dimensions (rate of performance of arithmetic and logical operations, latency and bandwidth of access to data throughout the computer memory hierarchy) drives the development of ever faster and more powerful computing machines. In this manner, application of HPTC is the motor that drives the evolution of computer technology.

[Page 77](#)

[PREV PAGE](#)

[TOP OF DOC](#)

74543qq.eps

74543rr.eps

## Understanding in Science and Technology

Scientific and technological research in academia, government and industry has traditionally been the largest application of HPTC.

The ability of modern computers to perform realistic numerical simulation of natural phenomena in a reasonable time has elevated HPTC to the status of the third leg of scientific research, beside those of theory and experiment.

HPTC has also become an ineluctable tool to extract understanding and order from the rapidly increasing volume of data from experimental techniques in such diverse fields as high-energy physics, remote sensing, astronomy, and genetics research.

74543ss.eps

## Development of Products and Services

*The production of information through numerical simulation of product behavior and manufacturing processes is the most rapidly growing application of HPTC.*

[Page 78](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Zero-latency decision making in new product development is a goal the world's manufacturing industries are striving for, and is a major driver towards virtual product development, in which all critical product development decisions are taken, and their impact understood and verified, before manufacturing begins.

Take new vehicle development in the automotive industry as an example. One of the most expensive and time-consuming steps in traditional new vehicle development is the building and testing of material prototypes. The industry would like to eliminate material prototyping entirely through numerical simulation ("virtual prototyping").

74543tt.eps

One of the most important of these simulations is vehicle crashworthiness. New vehicle development is driven on ever tighter schedules, and getting timely input on whether the foot room of a new passenger vehicle remains intact during a crash, so that remaining design can proceed apace, is a critical item of information, requiring high-performance computers that do not fail. Auto manufacturers are currently investing heavily in HPTC equipment to enable faster and more accurate crashworthiness results. Web-based access to simulation results, and web-based engineering simulation environments for working engineers are becoming the norm, not only in crash simulation, but for the whole integrated product development process: *everything to the Internet*. Japanese and European Auto manufacturers are using HPTC as key to product quality and doubling their compute capacity each year—estimated to be four times the peak application capacity available to U.S. manufacturers today.

With few exceptions, application programs used for simulation by the manufacturing industry are provided by independent, third-party software vendors (ISVs). Only in the aerospace and petroleum industries does competitive advantage continue to justify the high cost of maintaining proprietary applications, and even here, the future trend is increasingly to third party applications.

[Page 79](#)[PREV PAGE](#)[TOP OF DOC](#)

A further example of a product and a service for which HPTC is critical, and in which Compaq is playing an increasingly important role, is numerical weather forecasting. Timely, reliable, and accurate weather forecasts are of immense economic importance. Forecast products are distributed in a variety of forms and media to a multitude of customers such as airlines, farmers, military planners, shipping companies, vacationers, or commuters. These are typical e-Business products, increasingly delivered over the Internet. Getting them to customers reliably, on time is critical.

Weather forecasts themselves are the result of numerical simulation that predicts the evolution of present weather into the future using HPTC. These simulations use the fastest available computers in order to make the most accurate possible prediction in time to get the forecast products to their customers. Non-stop in the simulation process is key to meeting timeliness requirement: since its establishment in 1980, the European Centre for Medium Range Weather Forecasting has never failed to deliver a forecast.

Prime Influencers—the role of the DoE National Labs, Pittsburgh Supercomputing Center and Celera Genomics

The DoE Accelerated Strategic Computing Initiative

The DoE Stockpile Stewardship program required an unprecedented level of computing power to perform the necessary high fidelity simulations of complex physical phenomenon, levels 10000 times greater than the largest supercomputers available six years ago.

[Page 80](#)[PREV PAGE](#)[TOP OF DOC](#)



The DoE Defense Program National Labs challenged industry to scale computing performance using commercial off the shelf components and not resort to expensive "purpose built supercomputers". This decision was key to allowing the technology to become available for scientific and industrial users in significantly smaller subsets of the systems used at the DOE yet preserve the important performance improvements.

74543uu.eps

74543vv.eps

74543ww.eps

The DoE ASCI program developed a highly effective way to influence the computer system vendors through a cost sharing vehicle called the ASCI Pathforward program. Compaq participated in this program with Sandia National Laboratory as the program manager with the objective to dramatically accelerate Compaq's plans to build clusters with very high bandwidth and low message passing latencies. The result was highly effective as the 4 year \$5M investment by DoE resulted in Compaq's introducing a new family of supercomputers, called the AlphaServer SC, with the key performance characteristics shaped by the Pathforward program.

How did the U.S. DoE/DoD performance ratios impact Compaq's product roadmap?

In a word, directly! Previous Alpha microprocessors had the highest peak floating point calculation rate but. . .

[Page 81](#)

[PREV PAGE](#)

[TOP OF DOC](#)

The new EV6 Alpha improves the bandwidth per floating point operation by a factor of 8 (represented on the "ASCI ratios" as increasing from 0.25 MB/FLOP to 2.0 MB/FLOP.

Previous AlphaServers delivered leadership performance but. . .

The new SMP Alphaservers have the highest bandwidth and lowest latencies in the industry.

Previous Compaq commercial clusters offered the lowest message passing latency, especially in parallel database applications, but. . .

The DoE pathforward project accelerated development of Terascale interconnects and operating system improvements for scaling to thousands of processors.

The Pittsburgh Supercomputing Center TCS Project

The Terascale Computing System

The TCS takes advantage of the improvements in Compaq products driven by the DoE while driving additional developments needed to support the development of the largest open science supercomputer facility in the world.

The intent of the work on the TCS is not to just enable calculations to be done much faster, but to deliver a level of computing that will transform the research paradigm in several fields. The TCS will be 12 times the computational power and 40 times the memory of PSC's last major machine. Yet, subsets of this system will be available to any of the U.S. research sites that are partnered with PSC.

[Page 82](#)

[PREV PAGE](#)

[TOP OF DOC](#)

One of the areas of system development on the TCS is a increase of the cluster interconnect to support 800 SMP nodes, up from 128 in Compaq's standard AlphaServer SC.

An important additional area of TCS focus will be solutions to increased fault tolerance in a system with over 2000 processors.

The race to complete the human genome and Celera Genomics

Celera Genomics was the first life science industrial company to apply computing power at the scale of a national laboratory to accelerate its work. The impact of Celera's ability to pull in the completion date for sequencing and assembling the human genome is well known.

Celera's requirements in high performance computing added the necessity of around the clock production capability while scaling to supercomputing levels. Celera pioneered a type of supercomputing that may best be called high data throughput computing with computationally intense components for processing programs such as BLAST while assembling massive databases with the results of assembly operations. The need for extremely high file system throughput and the ability to scale capacity continuously are key characteristics of the Celera system.

Why clustering matters

Simplifies systems management

[Page 83](#)

[PREV PAGE](#)

[TOP OF DOC](#)

— Single system image

— Optimum performance across many users and jobs

Delivers scalable performance

— Message passing fabric enables parallel scalability

— Easy addition of compute nodes

Improves data management

— Cluster file system simplifies disk I/O

— High performance through parallel I/O

Makes big systems more available

— Fast fail over with full data integrity

Supercomputing Challenges

Management of the computing "complex"

Availability of "super scale" data services

[Page 84](#)

[PREV PAGE](#)

[TOP OF DOC](#)

— Recovery, backup, queue management, diagnostics

Raw performance of individual nodes

Economic models that support research and investments

Lack of skilled resources

— Vanishing breed of computer scientists

The role of open source software development

Open source software development is an important way to support collaborative development of key system components, in a way unencumbered by royalty payments and many licensing restrictions. There must be a balance between the benefits of community-developed software and the ability of individual developers and companies to invest and take risks resulting in approaches that can be protected as proprietary products and inventions.

The initial results of commercial efforts to deliver open source products while developing a business plan solely on the returns of providing service and support are not promising.

Summary Comments

Consider High Performance Computing as a proxy for the wealth of a nation

Traditional measures such as electric generation per capita were useful in the industrial era. Today, in an interconnected world where "time to discovery" and "time to market" are key, a better measure of wealth might be Interconnect bandwidth and access coupled with high performance computing capability per capita.

[Page 85](#)

[PREV PAGE](#)

[TOP OF DOC](#)

There is significant evidence that the Federal Government, through agencies such as the Department of Energy and the National Science Foundation, can play a crucial role in the advancement of the U.S. lead in High Performance Computing.

74543xx.eps

74543yy.eps

74543zz.eps

74543a3.eps

74543b3.eps

74543c3.eps

74543d3.eps

74543e3.eps

74543f3.eps

74543g3.eps

74543h3.eps

[Page 86](#)

[PREV PAGE](#)

[TOP OF DOC](#)

74543i3.eps

74543j3.eps

74543k3.eps

74543l3.eps

74543m3.eps

## BIOGRAPHY FOR WILLIAM BLAKE

Bill Blake is the vice president of Compaq's worldwide High Performance Technical Computing business with responsibility for the development and marketing of HPTC solutions including the AlphaServer SC supercomputers. Under his direction, Compaq's HPTC business has grown at a 19.5% CAGR for the past five years, reaching \$1.3B in 2000 according to the industry analyst International Data Corporation. In this role he is Compaq's Executive partner to key supercomputing customers such as the U.S. DoE National Labs and Celera Genomics, Inc.

His group is also responsible for the development of compilers, software development tools, and parallel processing software with groups in the U.S., Ireland and Switzerland. Prior to Compaq, Bill had been involved with product development with Digital Equipment Corporation since 1973. His experience spans the development of VLSI components for video graphics, network controllers and high performance

interconnects to engineering applications of artificial intelligence software. Bill pioneered the development of CMOS cell-based design and logic synthesis CAD work at Digital, technology used in many VLSI products including the current AlphaServers. Bill was also responsible for the strategic alliance with Encore that introduced the MEMORY CHANNEL to Digital, and interconnect key to clustered systems, and the current strategic alliance with Quadrics Supercomputing World for the interconnect used in the AlphaServer SC.

[Page 87](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Trained at the undergraduate and graduate level in Electrical Engineering, Bill is a member of the IEEE, ACM, American Association of Artificial Intelligence, Scientific Advisory Board of Cluster Solutions, SA (Switzerland) and a recent member of the Open MP Board of directors.

Mr. **GRUCCI**. Thank you very much for your testimony. And, at this time, I would ask Ms. Wideman if she would proceed. Thank you.

STATEMENT OF MS. CAROL WIDEMAN, CHIEF EXECUTIVE OFFICER AND FOUNDER,  
VCOM3D

Ms. **WIDEMAN**. Chairman Smith, members and staff of the Committee, I am Carol Wideman, President and CEO of Vcom3D, a small business in Orlando, Florida, that develops innovative technologies for online learning and information accessibility. It is my privilege to discuss the importance of Federal Government investment in information technology research and its value to the United States economy.

While fast computers and high bandwidth connections are of great importance, it is the development of software technologies and online applications that make these computers and Internet connections valuable to American citizens. A case in point is the recent creation of web-based three-dimensional characters that communicate in sign language.

The acquisition of language skills is often delayed for deaf and hard-of-hearing children. For many, reading and writing English are frustrating experiences. This lack of written English language skills excludes many deaf learners from independent study, including participation in online learning environments and other digital media. This results in missed opportunities to develop key technology, communication, collaboration, and knowledge-building skills alongside their hearing peers. Adding captions, which are English text, does not provide access to these individuals.

[Page 88](#)

[PREV PAGE](#)

[TOP OF DOC](#)

SigningAvatar software converts annotated English text into real-time, 3-D graphics representations of sign language. The application runs in an Internet browser, and SigningAvatar content can be embedded in web pages and stored on the Internet server or on a CD-ROM. Notice that the SigningAvatar video interpreter web page, displayed on the screen and on the wall-mounted monitors, allows us to view the character from different angles and change the background color.

In applications where synchronization with videos not required, we may also have the option to change the signing speed and to select a different character. We have chosen Maria for our starring role during this testimony.

[Video presentation follows]

A quotation. "If I am not for myself, who will be for me? Now, a girl wearing hearing aids gets out of a car. My name is Lindsey Dorrish [ph]. I am a 15-year old girl and I am hearing impaired. I go to a school in Oakland, California called College Preparatory School. I really like school here. It is so fun because I have been accepted here and I really like challenging myself for the academics here."

[End of video presentation]

Large segments of the education and training market are not well served by existing web-based media and authoring technologies. One proven method to provide motivation and engage learners is to include computer-generated interactive characters that provide immediate feedback to the user.

[Page 89](#)

[PREV PAGE](#)

[TOP OF DOC](#)

According to research by Lester in 1997, the presence of a lifelike character in a learning environment can have a strong positive effect on a student's perception of their learning experience. Through Federal funding, Vcom3D extended the technology developed for signing by adding limp-synched voice and goal-oriented behaviors. This enables our characters to not only motivate the learner, but to use demonstration by example, mentoring and role-playing simulations. We are also creating authoring tools to enable subject matter experts, with no previous background in computer programming, graphics, or animation, to create learning content simply by using a simple characterization markup language. Interactive characters may then be embedded in web pages or CD-ROM educational software titles and viewed on standard PCs.

In addition to offering improved opportunities for exploratory learning, online training reduces cost for government agencies and corporations, by reducing travel.

Let us tune into the Popyon [ph] demonstration on the screen and on the wall-mounted monitors for an interactive learning experience in which we will not only learn about butterflies, but have the opportunity to draw our own butterfly.

[Video presentation follows]

"You chose the monarch butterfly. The males are bright orange and black. The females are orange, brown, and black. Would you like to help me draw a butterfly now? Great. We can draw a butterfly together. First, I will draw the butterfly's body. You draw the left wing for our butterfly first and then I will draw the ones on the right. That one is really bright. Isn't it? Say, I really like that color. Blue is such a great color. Wonderful. Now, I will draw the other wing. Then we can watch our beautiful butterfly fly."

[Page 90](#)

[PREV PAGE](#)

[TOP OF DOC](#)

[End of video presentation]

There remain numerous technological innovations that could improve the lives of our population, but which are not being addressed due to lack of short-term commercial payoff. Both speech recognition and speech synthesis are areas in which Federal funding of information technology could result in important advancements.

Additional information regarding research needs and potential applications of these technologies is included in my written testimony. Increasingly, private and institutional investors demand short-term high return on capital investments in emerging companies. Federal funding of innovative research and information technology provides seed capital for key technical innovations that are likely to be passed over in competition for venture capital.

The history of SBIR funding is filled with success stories that have enhanced the lives of underserved segments of our population, while creating high-tech jobs and valuable products for Americans. SigningAvatar assistive technology is just example of IT project that has not only created a valuable new technology, but has led to commercial products that are successful in the marketplace. Thank you very much.

[The prepared statement of Ms. Wideman follows:]

74543n3.eps

74543o3.eps

[Page 91](#)

[PREV PAGE](#)

[TOP OF DOC](#)

74543p3.eps

74543q3.eps

74543r3.eps

74543s3.eps

74543t3.eps

## BIOGRAPHY FOR CAROL J. WIDEMAN

### Address:

Vcom3D, Inc., 3452 Lake Lynda Drive, Suite 260, Orlando, FL 32817; Tel: 407-737-4695; Fax: 407-737-6821; E-mail: carolw@vcom3d.com

### Formal Education:

University of Florida, M.S., Systems Engineering, 1979

University of South Florida, B.A., Mathematics and Education, 1971

GE Corporate Training Center, Executive Training, 1991

[Page 92](#)

[PREV PAGE](#)

[TOP OF DOC](#)

#### Professional Experience:

Vcom3D, Inc. May 1997–present

*President and Chief Executive Officer*

Ms. Wideman is the founder of Vcom3D and originated the idea to develop 3-D virtual characters that communicate in sign language as well as voice and gestures. She has built a talented team of computer architects, software developers, educators of Deaf students, professionals who are Deaf, and digital artists to work with her to make her vision a reality.

Ms. Wideman has led product marketing, sales, and licensing for the company. She also served as Principal Investigator for the following projects:

Commercial product development for SigningAvatar™ Internet-based Characters

Internet-based Terrain Visualization Tools, Phase I and Phase II Small Business Innovative Research Projects, U.S. Department of Defense

Science Applications International Corporation (SAIC) June 1994–March 1997

*Business Area Manager, Advanced Distributed Simulation*

[Page 93](#)

[PREV PAGE](#)

[TOP OF DOC](#)

At SAIC, Ms. Wideman developed the new business strategy and roadmap, resulting in the growth of this SAIC group to a \$40M/year business.

She served as team leader for

A research team developing "intelligent decision making models and behaviors for simulation systems".

Winning proposal teams for ADST II and WARSIM U.S. Army Simulation & Training contracts.

General Electric Company (GE) October 1979–March 1994

*Manager, Simulation & Training Programs*



During her career at GE, Ms. Wideman developed real-time, interactive, visualization technology and products, contributing as an engineer, technical director, product line manager, program manager, and finally as a business leader of a \$125M/year Simulation & Training organization. She received executive training at the GE Corporate Training Center, Crotonville, NY. While in engineering, she developed algorithms, architectures, hardware, and software for image generation products and data base tools.

#### Publications & Presentations:

Wideman, C. (2001) *SigningAvatar™ Technology Offers Access to Digital Media*, Eastern Assistive Solutions with Technology Conference, Fairfax, Virginia

[Page 94](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Wideman, C. (2000) *Online Learning and Assistance*, The Visual Future of the Internet Symposium, Atlanta, Georgia

Wideman, C. (2000) *Internet-enabled, 3D, SigningAvatar™ Software Offers Accessibility for the Deaf*, 19th International Congress in Education of the Deaf, Sydney, Australia

Wideman, C. & Popson, S., (1999) *Tutor Aids Deaf Students*, Closing the Gap, Minneapolis, Minnesota, published online <http://www.seamless-solutions.com/html/presentations/ctg/index.htm>

Wideman, C. & Popson, S. (1999) *Interactive 3-D Characters Enhance Learning for Deaf and Hearing Students*, National Educational Computing Conference (NECC), Atlantic City, NJ, published online at: <http://www.seamless-solutions.com/html/presentations/Default.htm>

Wideman, C. & Sims, E. (1998) *Signing Avatars*, Proceedings of the Technology and Persons with Disabilities Conference, Los Angeles, CA.

Sims, E. & Wideman, C. (1998) *Facial Animation for Communication Enhancement*, Life-Like Computer Characters Conference, Snowbird, Utah.

Sims, E. & Wideman, C. (1998) *Emerging Desktop Visualization Technologies*, Image 98, Phoenix, Az, Tutorial Document available from Image Society ([Image@asu.edu](mailto:Image@asu.edu)).

[Page 95](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Wideman, C. (1997), *Networked PC-based Real-time 3D*, Image Society Symposium, Phoenix, AZ.

#### Synergistic Activities:

Founding member of the University of Florida, *Computer, Information Science, and Engineering Industrial Advisory Board*.

1993/4 NISA/NTSA Training 2000 Chairperson and an active member of the Interservice/Industry Training, Simulation, and Education Conference subcommittees from 1995–1998.

Former high school mathematics teacher (1972/73)

Patents:

U.S. Patent application, Method for Animating Three-Dimensional Computer-Generated Characters, submitted August 4, 2000.

Mr. **GRUCCI**. Thank you for your testimony. At this time, we will go to our first round of questions. And the Chair recognizes Mr. Moore from Kansas.

Mr. **MOORE**. Thank you, Mr. Chairman. I just got here from another committee and I would like to pass at this time and thank you and thank the Committee—I am sorry—the witnesses for appearing here today.

[Page 96](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Mr. **GRUCCI**. I just have a couple of questions. And I guess this could be directed to all of you and whoever feels more comfortable answering it, that is fine. If you all wish to answer it, that is fine too. But the threat of digital terrorism to the research and application of the projects described in your testimony, how do you—how do we sufficiently safeguard against digital sabotage? And how do we balance the need for open access, the need for safeguarding data, and, lastly, how do we train researchers in the workforce to deal with cyber security issues? Anyone that wishes to take a first stab at that is more than welcome to do so. Yes, sir.

Mr. **BRAUN**. Well, to begin with, I think that it is an illusion if people believe that the network will protect the end systems from evil. If you have a computer at home, you are responsible for the machine and you cannot expect someone else to do your homework for you. You have to apply security patches and you have to run it in ways that it is—that makes sense. And you have kind of to know a little bit what you are doing. Computers are much more complicated than telephones or so.

Relative to my project, you have to differentiate, I think, between whether you are able to address things in real time or able to assess things after the fact. On the wireless network that we have implemented, we have significant monitoring capability that we can actually trace things and perhaps find offenders who do things. And I think a lot of good could be done by properly instrumenting the network and also by users of the network being aware of those problems.

Mr. **GRUCCI**. Thank you. Anyone else wishing to comment on that?

[Page 97](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Ms. **BAJCSY**. Yes. I can comment on. There is no unique answer to this question. At the NSF, we are sponsoring substantial program in investigating new algorithms, new ways of protecting security and reliability and privacy issues. But at the same time, this is not just technological question. This is also ethical

and moral question where every citizen will have to take responsibility for a certain behavior. Because no matter how much we will make progress, unless we agree all collectively that we will be in a kind of a dictatorial, state-controlled system, which nobody will agree to, it is not possible to absolutely prevent any—all security. So there is a price, and collectively we will have to agree upon how much we are giving—willing to give up a certain amount of flexibility and freedom for the security.

Mr. **GRUCCI**. Thank you. I have two high-tech industries in my district. One of them is Computer Associates. The other one is Simple Technology. Computer Associates manufactures a lot of software programs and Simple Technologies is the discoverer, inventor, and creator of the bar-code concept and bar-code reading. And I thought that is what we were going to be doing before. When I saw the screen up behind me, I thought we were going to be taking this dialogue in bar code.

But in all seriousness, does the United States have this kind of technology available to us within side of our borders to continue to create these types of high-tech advances that we expect, or are we going to start looking outside of our borders to foreign nations to help us create this new high-tech world? Again, anyone that wishes—that wants to respond is more than welcome to. Yes, sir. Could you just speak into that mike a little closer?

[Page 98](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Mr. **BLAKE**. We have significant strength in this country from the point of view of innovating new ideas, doing relevant research, driving ahead on a number of fronts. I think the difficulty is in taking that innovation—and it is oftentimes, and more often than not, the United States that brings out new capabilities and new techniques. And as the products mature and go to very high volume, then those products go offshore.

I think the continued vigilance around being efficient manufacturers and keeping the results of a product business, a manufacturing business, engaged with the benefits of having a highly innovative research community and a highly evolved, in effect, prototyping capability, is very important. So we simply lead the world in innovation, but sometimes lose on the manufacturing side.

Mr. **GRUCCI**. Thank you.

Ms. **BAJCSY**. If I may add to it? I don't know of any other country which has such capabilities and fertile environment for innovation and exploration. And there is a tremendous amount of energy out there in our academic institutions and research institutions that just doesn't exist anywhere else. So the creativity is phenomenal and we are truly the envy of the world in this regard.

Perhaps, the concern right now is that we are—we need to continue the investment into—in order to keep this going on. And the second concern is that somehow—and this, again, seems to me is not a technological problem, but it is more cultural problem—there is true concern why American men and boys and girls are not going into engineering and science. And there is no simple answer, in my opinion. Just to say that our schools are failing, it is not good enough. There is also a certain value system that is out there that is telling the young people that it is better to go to business or law than to engineering and science.

Mr. **GRUCCI**. You mentioned a moment ago in your response, investments and keeping the investments going. How would you encourage investors to get involved in this field and continue to make the investments necessary so that these new technologies could continue to emerge and we could continue to be the leaders? What would be some suggestions that you could make?

Ms. **BAJCSY**. You mean the private investors.

Mr. **GRUCCI**. That is correct.

Ms. **BAJCSY**. Well, the government has a lot of leverage, of course, and can give various tax initiatives and other initiatives. I think if you talk to various investors in Silicon Valley and everywhere else who support—who have supported high-tech investment, the pressure these folks feel is return on the investment. The fast return on the investment is a tremendous pressure. And we have been—I have been personally watching very carefully what is happening with our research laboratories. Bell Labs, IBM—IBM, Xerox Park, and other laboratories, if not diminishing, completely closing, then certainly shrinking.

And the reason for that is that, in particular, in Bell Labs, it used to be that there was a guaranteed income, 1 percent or—from the overall Bell Telephone went to research. And that enabled them to truly look long-term problems. Now, with all the competition, there is a tremendous pressure to—research is expensive, as you well know.

Mr. **GRUCCI**. Thank you. Anyone else have a thought on that? Yes.

Ms. **WIDEMAN**. I would like to share my experience with the Small Business Innovation Research Program. We are currently being funded by the National Science Foundation and they have a part of this program where when you go into the—after Phase II you go into a Phase II–B, where they provide 50 percent matching funds to investors—outside investors, 100 percent matching—or 50 cents to a dollar type of matching.

Having just completed this round of investment, I can say that it is highly motivating to investors to take that risk with something that might be a little longer term if they find that they are having the matching funds. So that was very motivating to them and it was successful.

Mr. **GRUCCI**. Thank you. Does anyone have any questions? Okay. At this time, I think we are going to take a recess and we will wait then for the Chairman to return. And we will stand in recess.

[Recess]

Chairman **SMITH** [presiding]. The Committee on Research will reconvene at this moment. Allow me to ask you some general questions. I am going to go out of text on something that I want your opinion on and

your ideas on. Does—and that is, cyber terrorism—does our advancements in information technology make us more vulnerable to cyber terrorism, whether it is communications, whether it is the way we distribute and organize electrical dissemination in this country, or whether it is the way we communicate with each other in financial institutions? Can anybody give me any thoughts that you might have had if you have looked at it? Yes. Ms. Wideman.

[Page 101](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Ms. **WIDEMAN**. While we will always be vulnerable as long as we have global networks, there are technologies that are being developed that can help us in that regard. One of the areas that I am familiar with is the fingerprint technology that is being used and developed in the United States that allows you to use a fingerprint to screen, having access to information. I believe there is also research going on in the voice recognition like a mission impossible, but it is really happening. So while, you know, nothing is ever going to make it so that it is totally secure, there are technologies that we are developing that will help that.

Chairman **SMITH**. And, Mr. Braun, what about the wireless communication? Does the software that is developed in helping organize and transfer that kind of information—is that a vulnerable area in the whole realm of communications.

Mr. **BRAUN**. Well, as I mentioned to you before, I think it is unsafe to assume that the network is protecting you. You have to—the end systems, the machines, the computer that you have at home or at work needs to have provisions in it to protect themselves. We are vulnerable against very many things in life. We are vulnerable against—relative to what people do you or what technology does to you, in most cases, people don't even think about it. What is curious is that people think they can get away with things in cyberland that they cannot get away with anywhere else. And as Ruzena mentioned before, this is not just a technical problem. This is a social problem. This is a law problem. This is an enforcement problem.

There is also a technical component in it. But so is it in other ways. People don't lift lids from potholes all the time because they know they cannot get away with it, for example.

[Page 102](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Chairman **SMITH**. Out of sync—and anybody can keep responding to the first question also. And that is, the increased propensity that other engineers, other mathematicians, other designers, other builders, and other countries will have a greater ability to challenge our economy in the way they contribute—all the way from—I am just thinking of shopping for a new car. And all the technology in that new car and the ability of the local mechanics to simply plug in to see what is wrong with that car, it is very reasonable that mechanics in the future could be sitting in India or any other country some time to evaluate that car and, in fact, in many cases, repair the car. I mean, that is just a small example.

But what should we be looking at, what should we be doing to help assure that the—that a significant advantage accrues to the taxpayers in this country or that are paying for this research and development? Any thoughts? Any ideas?

Dr. **BAJCSY**. Mr. Chairman, I—it is trivial to say that we live in a globalized world and globalized economy and this globalization has—is taking place because of the IT and communication technology. And because of that, we do share much of it. On the other hand, I said before you came, to Mr. Grucci, that in—it is in my personal opinion that there is no other country which creates such conducive environments as the United States for creativity of new ideas, revolutionary ideas. There is a tremendous amount of energy out there among young people who feel not restrained to create new ideas. The question I think remains to be seen how can we take advantage of this creativity of these new ideas. How do we take these ideas and make for the taxpayers to use it?

Chairman **SMITH**. Okay. Specifically, maybe trying to pinpoint any suggestions that any of you might have, what areas of IT technological research should we be looking at that might be underfunded now that could use greater emphasis in helping assure that our manufacturing, for example, could stay a step ahead of our competitors in the rest of the world? Yes, Mr. Blake.

[Page 103](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Mr. **BLAKE**. I think the approach that is being taken at places like the Pittsburgh Supercomputing Center where large-scale problems are being solved uniquely with very high levels of computer capability, but the technology involved is something that can be subset and brought back to universities and smaller industrial uses. And this notion of not having the highest performance computing capability unique to one lab or unique to one use, but being able to diffuse it through the industrial base is rather important. So as you encourage the development and application of high-performance computing, which leads to new ways of designing, new ways of producing products, encourage those technologies to be subsetable and be brought out to the industrial base.

Dr. **BAJCSY**. I have been proponent of so-called cyberinfrastructure which is a high-band connectivity to high-performance computing and databases and instruments. And the reason for that is that we need to develop the software that makes these integrated systems useful so that the everyday citizen can use it. Today, IT industry is in trouble, in my opinion, because the PC market is saturated with the word processing capabilities. The next stage, the next wave, will be with the multimedia, the video and the audio. But that part of research hasn't been concluded yet. So that you, I, and my children and grandchildren can easily use it so that they don't have to have a Ph.D. in computer science.

Chairman **SMITH**. Dr. Bajcsy, and maybe Mr.—Dr. Berman, and Mr. Braun—define for me, what is open-spaced software?

Dr. **BAJCSY**. Open-spaced software is a software that it is not proprietary so that you can add to it, subtract to it, so that you can use it in a free way.

[Page 104](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Chairman **SMITH**. Dr. Berman.

Dr. **BERMAN**. Yeah. I would like to concur with Dr. Bajcsy that I believe that we need to make the real investment in infrastructure in software development, in databases, and we need to have this kind of open environment in order to make this—to really capitalize on the work that we have been doing. And I think that we also have to think about, once we develop all of these things, how we actually maintain them. And the project that I talked about has been going on for 30 years in a kind of very unstable way of funding these kinds of initiatives. And yet, they are the enabling technologies that we have to keep going. So I would like to see some way of putting more resources into this infrastructure development, as well as into applications. But I think infrastructure is absolutely key.

Chairman **SMITH**. Maybe, Mr. Blake—in the whole area of prototype development through computers versus traditional development, do you—have we done anything on failure rates in terms of the products that are ultimately developed? Is the failure rate greater or less in the computer developed prototypes?

Mr. **BLAKE**. The failure rate tends to be much, much less. As the accuracy of the model, the prototype, the virtual prototype, if you wish, of that which is being designed, becomes more and more refined, then you do have less and less surprises at the end of the manufacturing process. It is much more predictable. So the more rigorous the simulation and modeling, then the more authentic or high-fidelity the models are, meaning they track physical properties, the more likely you will have a much more predictable and more lower defect final product.

[Page 105](#)

[PREV PAGE](#)

[TOP OF DOC](#)

That is absolutely the case. And it is, in fact, very much the case that in Japan and in Europe the automobile manufacturers there are actually more aggressive in their use of computing on automotive design. And the end quality of the result is actually a result of that. A good example of—the differences would be in the United States, there is certainly a significant amount of testing going on for safety to make sure that the car is structurally sound, that it will protect occupants in a crash. Where the Japanese and the European companies will add on top of that is more rigorous analysis of noise and comfort and some of the things that are not in a crash situation, but something of appeal to a person buying a car. And it comes out of this more intense modeling and simulation.

Chairman **SMITH**. Well, I mean, maybe the meeting that I excused myself for was on the CAFÈ standards and what we do tomorrow in terms of developing—to what extent should government mandate more efficient cars and how is that going to affect the automobile industry in this country versus other countries. And so maybe more prototype.

Let me ask you a question on cost——

Mr. **BLAKE**. Uh-huh.

Chairman **SMITH** [continuing]. On the face of it. But if you were to include the cost of the software and the cost of the high-end computers, and throw that into the mix in the computer prototype development, how does the cost comparison——

Mr. **BLAKE**. The cost—although if you look at the cost of a very high-end computer, it is expensive. It is certainly not PC type of costing. There is tens of millions of dollars, in some cases. However, if the investment in a \$10 million computer defers or allows for a better decision on a hundred million or a billion dollar factory, the cost of the computing is washed out of the analysis. And that is really what is at stake—

[Page 106](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Chairman **SMITH**. But comparing it to traditional prototype development in the traditional laboratory setting.

Mr. **BLAKE**. There is almost no comparison from a time to market point of view. What—

Chairman **SMITH**. No. But in cost. If you include the cost of the software development and the cost of the high-end computing, the high-end computers, it is still—is there some analysis—is some—there have been some collection of cost figures? Or once you have a high-end computing, what—to what extent do you attribute the cost of that particular high-end computer to whatever project you are developing?

Mr. **BLAKE**. The analysis that I have seen indicate that the cost of the actual computer is relatively small compared to the savings in labor and time to produce the product. It used to take 8 years to design a new automobile, and now manufacturers are being driven to a 1 to 2-year design cycle. The cost of not having all of that investment over an 8-year period is dramatic. And the cost of making a higher degree of investment in automation that gets fed from these computer-based models, even lowers costs more and improves quality. So there is a synergistic effect once you are working from a virtual prototype than from building a sort of traditional cut-and-fit prototype going the—

Chairman **SMITH**. And so let me ask you, what—besides the auto industry, what other industries would you suggest may very well benefit from this virtual prototype supercomputing ability?

[Page 107](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Mr. **BLAKE**. Well, certainly staying on the—in the mechanical design area, the aerospace industry—really the first to fully exploit this capability. And a conversation I had 6 months ago with the airbus people, is they are designing a new 3XX jumbo—super jumbo jet. This is a double-decker jet. And the wings on that jet are an extremely, obviously important factor. It cost them \$500 million to prototype a wing. And they need to build three of them and destructively test them to validate the design.

Computers can be used to simulate and model the validation of a wing to the point where they could actually remove one degree of one whole wing test just to do test. That is a case where a \$500 million prototype and test exercise is replaced by a computer, and it is unlikely the computer will come anywhere near 1/10 of that cost. It is an example of mechanical space.

The—one of the biggest areas of impact, though, is in looking at the Solera Genomics or Geneva Proteomics type of companies where, as you know, the whole time to complete the genome, the assembly, the sequence of assembly, the genome, collapsed by years due to high-performance computing techniques.



What is happening now is companies are taking those databases, applying significant amounts of computing to search for new drug targets, to look for therapeutic proteins, to take 8 to 10 years' worth of what would traditionally be wet-lab biological work and generate results in a year or 2. That is huge——

Chairman **SMITH**. Dr.—is it Dr. Braun or Mr. Braun?

Mr. **BRAUN**. Mr. Braun.

[Page 108](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Chairman **SMITH**. We would be glad to promote you as far as this hearing goes. But let me ask you a—two questions. Why do we need the wireless computing technology that you described for emergencies or whatever when we always have the cell phones?

Mr. **BRAUN**. One of the things that I was very surprised, talking to various agencies in San Diego County, is how little of real technology they have to communicate at high-performance among each other and with—specifically with other agencies. Looking at—well, looking at movies or so, I am amazed that there is not this technology there that is portrayed everywhere, but it is really not in place.

The firefighters in San Diego, for example, were very pleased about us being able to provide Internet connectivity at megabit speed within hours after the installation. This technology is simply not there from what I have seen, and I am personally very surprised by it.

And often they need it also, in my experience, for very mundane things, not necessarily for like heads-up displace with some sophisticated technology, but often just to connect their spouses by e-mail in the evening and telling them, I am still okay, but the fire is still burning. And when they are out in the field in a rural area where you don't even have cell phone coverage, often they don't have any other means to just communicate very basic things. Let alone more technology and more advanced technology things like being able to track weather conditions in real time, like did the wind change—are they going to burn in 30 seconds because there is a wind gust?

[Page 109](#)

[PREV PAGE](#)

[TOP OF DOC](#)

And if you had technology like that, they could wear little alarm things on their belts, for example, that alerts them from sensors that could surround an incident area and say, something happened. Watch out.

Chairman **SMITH**. We have told some of you we wouldn't keep you beyond four o'clock. What I would like to do, though, is request that each of you consider being willing to respond to some written questions that we didn't get a chance to ask. And what I would like you to consider doing now is maybe just a brief, maybe 60-second, 1-minute wrap-up of any other information that you would like to have on the record as we look at how Congress and public policy develops in IT. And we will start with you Dr. Bajcsy.

Dr. **BAJCSY**. Well, I want to say that the last 2 years of investment into this ITR technology made a tremendous difference for which we are most grateful. It makes profound difference in the community on the scale of problems that we can address. But in order to truly get to the second wave of IT industry or

boom that we just got off from the first wave, we need to get the multimedia communication and connectivity and processing through. And for that, we need more investment, unfortunately.

Chairman **SMITH**. Mr. Braun.

Mr. **BRAUN**. Something that I specifically like about this project that I am running, the HPWREN project, is that it is not just focused on one area. The wireless network is almost an incidental thing. The real issue is putting together a system that benefits a broad range of research signs and educators. And this system can only be accomplished by these people really collaborating. And that, in my mind, is the biggest benefit and allows us to build these kind of systems to spend not only technical areas, but also looking into social factors, looking into economy and things like that.

[Page 110](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Chairman **SMITH**. Dr. Berman.

Dr. **BERMAN**. In my opinion, the IT advances have—are only beginning to impact biology, and the impact is going to be ever greater because of the massive amounts of data that we now have to understand from all realms of the physical and biological sciences. And so an investment in computer technologies and in algorithm development is going to make a huge difference in our understanding of basic biological systems.

Chairman **SMITH**. Mr. Blake.

Mr. **BLAKE**. I concur with Dr. Berman. Over the past 20 years, the physical sciences have driven the numerical simulation and virtual prototyping capability that we have seen with today's supercomputers as we look at the life sciences, which is where I—we clearly will see one of the biggest benefits of high-end computing that will be dealing with large amounts of data and file systems and capabilities beyond just numerically computing. And it is important to state that the sponsorship of the Federal Government on grand challenges, either through the defense side of DOE with the national labs, or NSF with places like Pittsburgh, creates an opportunity for computer companies to solve problems we would not solve in any way near the same time frame. It accelerates things by, in some cases, 5 or 10 years.

Chairman **SMITH**. Ms. Wideman.

Ms. **WIDEMAN**. Venture capital today is very focused on the next big thing. Since the dot.com crash, they are saying, okay, where is our next big thing. And it—but it must be a short-term, very high return on investment. They are not really looking at what is right for the American people, what is right for the American economy. What do we need to be healthy 5 years down the road? And I think that really depends on the Nation, as a whole, to determine those needs. And some of the areas that we have been focused on, that have very strong needs, are the information accessibility for those who have special needs. Certainly, we have implemented Section 508, but in order to actually perform on Section 508, we need new technologies which provide the interface to the people that have these needs.

[Page 111](#)

[PREV PAGE](#)

[TOP OF DOC](#)

Also, for online learning, venture capitalists are interested in developing some elements like a management system or something that might be general to everyone. But in terms of the actual basic technologies that will enable us to interact, this voice recognition and speech synthesis, both whether it is someone out in a remote area that is receiving this on the Internet, or someone that has special needs, can't hear or cannot see, these are all technologies that are really critical to their lives. Thank you.

Chairman **SMITH**. Well, I suspect that you have heard or read in the paper that last night we passed an appropriation bill out of the House for the National Science Foundation that was a 9-percent increase over last year. Last year was a 13-percent increase over the prior year. So in the last 2 years, we have increased NSF funding because of the priority that this Congress our basic research effort, by 22 percent. So it is a substantial part of the total budget as we are looking at the best way to spend money, but I am convinced that it is certainly key to our—not only our national prosperity, but our national security.

So, again, let me thank all of you for your work, your willingness today to come to this hearing and share your time. We will possibly be sending you additional questions that staff would have loved to have us ask that we didn't ask. So, again, thank you all very much. And with that, this Subcommittee is adjourned.

[Whereupon, at 4:01 p.m., the Subcommittee was adjourned.]

SPEAKER	INDEX	<u>CONTENTS</u>	<u>INSERTS</u>						
BAJCSY		<u>23</u>	<u>25</u>	<u>26</u>	<u>39</u>	<u>97</u>	<u>98</u>	<u>99</u>	<u>102</u>
		<u>103</u>							
		<u>109</u>							
BERMAN		<u>58</u>	<u>104</u>	<u>110</u>					
BLAKE		<u>72</u>	<u>73</u>	<u>74</u>	<u>98</u>	<u>103</u>	<u>104</u>	<u>105</u>	<u>106</u>
		<u>107</u>							
		<u>110</u>							
BRAUN		<u>40</u>	<u>44</u>	<u>96</u>	<u>101</u>	<u>107</u>	<u>108</u>	<u>109</u>	
DIANE JONES		<u>5</u>							
ETHERIDGE		<u>20</u>							
GRUCCI		<u>74</u>	<u>87</u>	<u>95</u>	<u>96</u>	<u>97</u>	<u>98</u>	<u>99</u>	<u>100</u>
JIM WILSON		<u>5</u>							
M		<u>92</u>	<u>93</u>	<u>94</u>					
MOORE		<u>95</u>							
NATALIE PALMER		<u>5</u>							
PETER HARSHA		<u>5</u>							
SHARON HAYS		<u>5</u>							
SMITH		<u>17</u>	<u>22</u>	<u>25</u>	<u>26</u>	<u>39</u>	<u>40</u>	<u>43</u>	<u>44</u>
		<u>57</u>							

72   73   74   100   101   102   103   104   105

106   107   108   109   110   111

SMITH OF TEXAS   39   40

T   92

T   93   94

WIDEMAN   87   100   101   110

CONTENTS   [SPEAKERS](#)   [INSERTS](#)

STATEMENT OF DR. RUZENA BAJCSY, Ph.D., ASSISTANT DIRECTOR, COMPUTER AND INFORMATION SCIENCE AND ENGINEERING, NATIONAL SCIENCE FOUNDATION; CHAIR, INTERAGENCY WORKING GROUP ON INFORMATION TECHNOLOGY R&D

[PAGE](#) \_\_\_\_\_

[23](#)

STATEMENT OF DR. HANS-WERNER BRAUN, Ph.D., RESEARCH SCIENTIST, SAN DIEGO SUPERCOMPUTING CENTER

[PAGE](#) \_\_\_\_\_

[40](#)

STATEMENT OF HELEN BERMAN, Ph.D., DIRECTOR, THE PROTEIN DATA BANK; BOARD OF GOVERNORS PROFESSOR OF CHEMISTRY, RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY

[PAGE](#) \_\_\_\_\_

[57](#)

STATEMENT OF WILLIAM BLAKE, VICE PRESIDENT FOR HIGH-PERFORMANCE TECHNICAL COMPUTING, COMPAQ COMPUTER CORPORATION

[PAGE](#) \_\_\_\_\_

[72](#)

STATEMENT OF MS. CAROL WIDEMAN, CHIEF EXECUTIVE OFFICER AND FOUNDER, VCOM3D

[PAGE](#) \_\_\_\_\_

[87](#)

INSERTS   [SPEAKERS](#)   [CONTENTS](#)

NO INSERTS IN THIS HEARING

