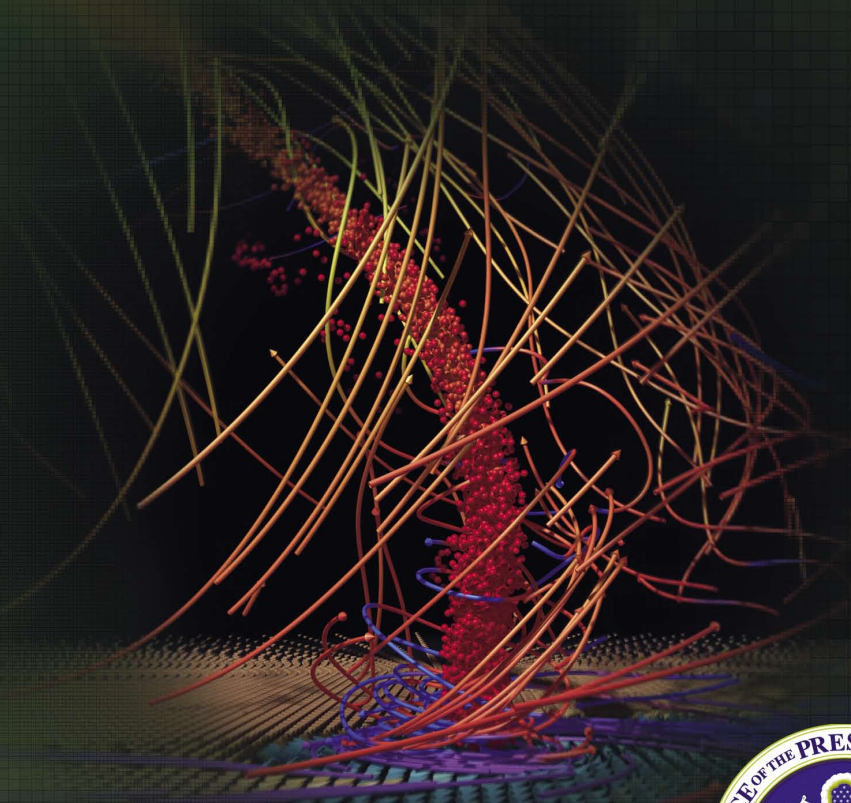
Report to the President

June 2005

Computational Science: Ensuring America's Competitiveness



President's Information Technology Advisory Committee

Cover Imagery

Tornadoes – a frightening and often devastating seasonal feature of North American weather patterns – are notoriously difficult to forecast accurately because scientists do not yet fully understand the complex interactions of forces that create and drive them. It is known that the most violent tornadoes take a particular type of rotating thunderstorm, called a supercell, to generate them. But at most 20-25 percent of supercells spawn tornadoes, and of those, under 10 percent approach the 158-318 miles-per-hour wind speeds of the most destructive F3, F4, and F5 classes of tornadoes. The destructiveness of some of these is due to the length of time they stay on the ground (~45 minutes to several hours) , producing long damage paths. Such violent long-lived tornadoes are called "long-track tornadoes."

Computational science enabled by high-end computing power now makes it possible for researchers to look inside tornadic storms and examine the intricate interplay of temperature, moisture, turbulence, air pressure, and wind in the genesis and development over time of a full-scale twister. This is a major advance toward better prediction of dangerous storms. The cover images are taken from the first scientific simulation ever done of a long-track tornado within its parent supercell storm. The simulation was developed through a collaborative effort of scientists in the Department of Atmospheric Sciences at the University of Illinois, Urbana-Champaign (UIUC) and at NOAA's National Severe Storms Laboratory, as well as a scientific visualization team from the university's National Center for Supercomputing Applications (NCSA) . The simulation was based on data collected from a 2003 F4 tornado that ripped through Manchester, South Dakota.

The simulation, which generated more than 6 terabytes of data, used a 3D stretched mesh that enabled the researchers to focus in on the forces acting within the storm where the tornado formed as the storm evolved in a 100 x 100 x 25-kilometer region over three hours' time. The front-cover image shows the data elements visualized at a stage of increasing tornadic intensity. Interactively filtered "streamtubes" colored orange when rising and blue when sinking represent the path of air through the storm. A swirling mass of red spheres in the low-pressure vortex delineates the developing tornado (the swirl becomes orange, then yellow at peak tornadic intensity) . On the ground plane, tilting cones represent wind speed and direction. Colored by temperature, they show a surface boundary where warm and cold air interact at the tornado's base. The back-cover images, top to bottom, show: the supercell's external atmospheric shape; the emerging circularity of turbulence inside the storm; the characteristic low-pressure vortex, with an orange hue indicating increasing intensity; and the tornado's eventual disintegration.

The project was supported by NSF, NCSA, NOAA, and the Intel Corporation. *Simulation Credits:* Robert Wilhelmson, Director; Matthew Gilmore; Louis Wicker; Glen Romine; Lee Cronce; and Mark Straka. *Visualization credits:* Donna Cox; Robert Patterson; Stuart Levy; Alex Betts; and Matthew Hall. For more information, see: *http://redrock.ncsa.uiuc.edu/CMG/* .

Copyright

This is a work of the U.S. government and is in the public domain. It may be freely distributed and copied, but it is requested that the National Coordination Office for Information Technology Research and Development (NCO/IT R&D) be acknowledged. Requests to use any images or figures must be made to the provider identified in the image credits, or to the NCO if no provider is cited.

REPORT TO THE PRESIDENT

Computational Science: Ensuring America's Competitiveness

President's Information Technology Advisory Committee

June 2005

President's Information Technology Advisory Committee

Co-Chairs:

Marc R. Benioff
Edward D. Lazowska

Members:

Ruzena Bajcsy
J. Carter Beese, Jr.
Pedro Celis
Patricia Thomas Evans
Manuel A. Fernandez
Luis E. Fiallo
José-Marie Griffiths
William J. Hannigan
Jonathan C. Javitt
Judith L. Klavans
F. Thomson Leighton
Harold Mortazavian
Randall D. Mott
Peter M. Neupert
Eli M. Noam
David A. Patterson
Alice G. Quintanilla
Daniel A. Reed
Eugene H. Spafford
David H. Staelin
Peter S. Tippett
Geoffrey Yang

May 27, 2005

The Honorable George W. Bush
President of the United States
The White House
Washington, D.C. 20500

Dear Mr. President:

The President's Information Technology Advisory Committee (PITAC) is pleased to submit to you the enclosed report *Computational Science: Ensuring America's Competitiveness*. Computational science – the use of advanced computing capabilities to understand and solve complex problems – has become critical to scientific leadership, economic competitiveness, and national security. The PITAC believes that computational science is one of the most important technical fields of the 21st century because it is essential to advances throughout society.

Computational science provides a unique window through which researchers can investigate problems that are otherwise impractical or impossible to address, ranging from scientific investigations of the biochemical processes of the human brain and the fundamental forces of physics shaping the universe, to analysis of the spread of infectious disease or airborne toxic agents in a terrorist attack, to supporting advanced industrial methods with significant economic benefits, such as rapidly designing more efficient airplane wings computationally rather than through expensive and time-consuming wind tunnel experiments.

However, only a small fraction of the potential of computational science is being realized, thereby compromising U.S. preeminence in science and engineering. Among the obstacles to progress are rigid disciplinary silos in academia that are mirrored in Federal research and development agency organizational structures. These silos stifle the development of multidisciplinary research and educational approaches essential to computational science. Our report recommends that both universities and Federal R&D agencies must fundamentally change these organizational structures to promote and reward collaborative research. In addition, the report calls on the National Science and Technology Council (NSTC) to commission a fast-track study by the National Academies to recommend changes and innovations in agency roles and portfolios to support advances in computational science.

c/o National Coordination Office for Information Technology Research & Development
4201 Wilson Boulevard . Suite II-405 . Arlington, VA 22230
703-292-4873 (ITRD) . FAX (703) 292-9097 . nco@nitrd.gov . www.nitrd.gov

Insufficient planning and coordination of computational science efforts across the Federal government, academia, and industry represents another obstacle to progress. Current efforts are characterized by a short-term orientation, limited strategic planning, and low levels of cooperation among the participants. To address these deficiencies, the report recommends that the NSTC commission the National Academies to convene one or more task forces to develop and maintain a multi-decade roadmap for computational science and the diverse fields that increasingly depend on it. Such a roadmap would coordinate and direct the multiple technical advances required to support computational science in order to maintain the Nation's competitive leadership in the decades ahead.

As part of this national effort, we recommend that the Federal government provide an infrastructure that includes and interconnects computational science software sustainability centers, data and software repositories, and high-end computing leadership centers with each other and with researchers. We also recommend that our computational science R&D be rebalanced to focus on improved software, systems with high sustained performance, and sensor- and data-intensive applications.

We appreciate this opportunity to provide you with our advice on computational science – an area that is central to the Nation's long-term technological leadership. We trust that the Committee's work in computational science, and our earlier reports on health care information technology and cyber security, provide useful advice on how the United States can remain a world leader in science and technology, how we can improve the effectiveness of our health care system, and how we can assure the security of our information infrastructure. These reports illustrate how critical information technology research and development is to our economic competitiveness, quality of life, and national security.

It has been an honor to serve you as PITAC Co-Chairs. We would be pleased to meet with you and members of your Administration to discuss our reports and concerns.

Sincerely,

Marc R. Benioff
PITAC Co-Chair

Edward D. Lazowska
PITAC Co-Chair

President's Information Technology Advisory Committee

CO-CHAIRS

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

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

MEMBERS

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

J. Carter Beese, Jr.
President
Riggs Capital Partners

Pedro Celis, Ph.D.
Software Architect
Microsoft Corporation

Patricia Thomas Evans
President and CEO
Global Systems Consulting Corporation

Manuel A. Fernandez
Managing Director
SI Ventures/Gartner

Luis E. Fiallo
President
Fiallo and Associates, LLC

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

William J. Hannigan
President
AT&T

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

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

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

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

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

Peter M. Neupert
Consultant

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

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

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

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

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

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

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

Geoffrey Yang
Managing Director
Redpoint Ventures

COMPUTATIONAL SCIENCE SUBCOMMITTEE

CHAIR

Daniel A. Reed

MEMBERS

Ruzena Bajcsy
Manuel A. Fernandez
José-Marie Griffiths
Randall D. Mott

CONSULTANTS

Jack Dongarra
Chris R. Johnson

About PITAC and This Report

The President's Information Technology Advisory Committee (PITAC) is appointed by the President to provide independent expert advice on maintaining America's preeminence in advanced information technologies. PITAC members are leaders in industry and academia whose reports on key issues in Federal networking and information technology research and development (R&D) help guide the Administration's efforts to accelerate the development and adoption of information technologies vital to American prosperity in the 21st century.

Authorized by Congress under the High-Performance Computing Act of 1991 (Public Law 102-194), as amended by the Next Generation Internet Act of 1998 (Public Law 105-305), and formally established and renewed through Presidential Executive Orders, PITAC is a Federally chartered advisory committee operating under the Federal Advisory Committee Act (FACA) (Public Law 92-463) and other Federal laws governing such activities.

The PITAC selected computational science as one of three topics for evaluation. The Director of the Office of Science and Technology Policy provided a formal charge (Appendix C), asking PITAC members to concentrate their efforts on the focus, balance, and effectiveness of current Federal computational science R&D activities. To conduct this examination, PITAC established the Subcommittee on Computational Science, whose work culminated in this report, *Computational Science: Ensuring America's Competitiveness.*

The PITAC found that computational science contributes to the scientific, economic, social, and national security goals of the Nation. However, much of the promise of computational science remains unrealized due to inefficiencies within the R&D infrastructure and lack of strategic planning and execution. PITAC's primary recommendations address these deficiencies, calling for a rationalization and restructuring of computational science within universities and Federal agencies, and the development and maintenance of a multi-decade roadmap for computational science R&D investments.

The report's findings and recommendations were developed by the PITAC over a year of study. The Subcommittee was briefed by computational science experts in the Federal government, academia, and industry; reviewed the current literature; and obtained public input at PITAC meetings and a town hall meeting, and through written submissions. (Appendix D summarizes the Subcommittee fact-finding process.) The Subcommittee's draft findings and recommendations were discussed and reviewed by the PITAC at its November 4, 2004 and January 12, 2005 meetings; the final findings and recommendations were approved at its April 14, 2005 meeting; and the final report was approved at its May 11, 2005 meeting.

A glossary of acronyms and abbreviations employed in the report is provided on pages 100-103.

Table of Contents

PRESIDENT'S INFORMATION TECHNOLOGY ADVISORY COMMITTEE v

ABOUT PITAC AND THIS REPORT vii

TABLE OF CONTENTS ix

EXECUTIVE SUMMARY 1

1 A WAKE-UP CALL: THE CHALLENGES TO U.S. PREEMINENCE AND COMPETITIVENESS 7

What Is Computational Science? 10

The 'Third Pillar' of 21st Century Science 12

An Unfinished Revolution 15

The PITAC's Call to Action 17

2 MEDIEVAL OR MODERN? RESEARCH AND EDUCATION STRUCTURES FOR THE 21ST CENTURY 19

Removing Organizational Silos 19

Evolving Agency Roles and Priorities 21

The Challenge of Multidisciplinary Education 23

Developing 21st Century Computational Science Leaders 24

3 MULTI-DECADE ROADMAP FOR COMPUTATIONAL SCIENCE 27

Rationale and Need 27

Computational Science Roadmap Components 29

The Computational Science Roadmap: A Schematic View 30

Roadmap Process, Outcomes, and Sustainability 33

4 SUSTAINED INFRASTRUCTURE FOR DISCOVERY AND COMPETITIVENESS 35

Software Sustainability Centers 36

National Data and Software Repositories 39

National High-End Computing Leadership Centers 41

Infrastructure, Community, and Sustainability: Staying the Course 43

5 RESEARCH AND DEVELOPMENT CHALLENGES 47

Computational Science Software 47

Programming Complexity and Ease of Use 48

Software Scalability and Reliability 50

Architecture and Hardware 51

Scientific and Social Science Algorithms and Applications 53

Scientific Algorithms and Applications 54

Social Science Applications 54

Software Integration 55

Data Management 56

CONCLUSION 58

REFERENCES 59

APPENDIX A: EXAMPLES OF COMPUTATIONAL SCIENCE AT WORK 62

Social Sciences 62

Physical Sciences 66

National Security 71

Geosciences 74

Engineering and Manufacturing 75

Biological Sciences and Medicine 81

APPENDIX B: COMPUTATIONAL SCIENCE WARNINGS – A MESSAGE RARELY HEEDED 85

APPENDIX C: CHARGE TO PITAC 94

APPENDIX D: SUBCOMMITTEE FACT-FINDING PROCESS 96

APPENDIX E: ACRONYMS 100

ACKNOWLEDGEMENTS 104

Those who cannot remember the past
are condemned to repeat it.

George Santayana

Executive Summary

Nearly half a century ago, the Soviet Union's successful launch of Sputnik – the world's first satellite – shook the political and intellectual foundations of the United States, galvanizing the Federal government to open a new era in research and education in the sciences, engineering, and technology. Today, U.S. leadership in science, engineering, and technology is again being challenged. But this time the challenge is far more diffuse, complex, and long-term than one bold technological achievement by a single U.S. competitor. In the 21st century global economy, burgeoning science and engineering capabilities of countries around the world – spurred by U.S.-pioneered computing and networking technologies – are increasingly testing the Nation's preeminence in advanced scientific research and development (R&D) and in science- and engineering-based industries.

Though the information technology-powered revolution is accelerating, this country has not yet awakened to the central role played by computational science and high-end computing in advanced scientific, social science, biomedical, and engineering research; defense and national security; and industrial innovation. Together with theory and experimentation, computational science now constitutes the "third pillar" of scientific inquiry, enabling researchers to build and test models of complex phenomena – such as multi-century climate shifts, multidimensional flight stresses on aircraft, and stellar explosions – that cannot be replicated in the laboratory, and to manage huge volumes of data rapidly and economically. Computational science's models and visualizations – of, for example, the microbiological basis of disease or the dynamics of a hurricane – are generating fresh knowledge that crosses traditional disciplinary boundaries. In industry, computational science provides a competitive edge by transforming business and engineering practices.

While it is itself a discipline, computational science serves to advance all of science. The most scientifically important and economically promising research frontiers in the 21st century will be conquered by those most skilled with advanced computing technologies and computational science applications. But despite the fundamental contributions of computational science to discovery, security, and competitiveness, inadequate and outmoded structures within the Federal government and the academy today do not effectively support this critical multidisciplinary field.

Principal Finding

Computational science is now indispensable to the solution of complex problems in every sector, from traditional science and engineering domains to such key areas as national security, public health, and economic innovation. Advances in computing and connectivity make it possible to develop computational models and capture and analyze unprecedented amounts of experimental and observational data to address problems previously deemed intractable or beyond imagination. Yet, despite the great opportunities and needs, universities and the Federal government have not effectively recognized the strategic significance of computational science in either their organizational structures or their research and educational planning. These inadequacies compromise U.S. scientific leadership, economic competitiveness, and national security.

Principal Recommendation

Universities and the Federal government's R&D agencies must make coordinated, fundamental, structural changes that affirm the integral role of computational science in addressing the 21st century's most important problems, which are predominantly multidisciplinary, multi-agency, multi-sector, and collaborative. To initiate the required transformation, the Federal government, in partnership with academia and industry, must also create and execute a multi-decade roadmap directing coordinated advances in computational science and its applications in science and engineering disciplines.

Traditional disciplinary boundaries within academia and Federal R&D agencies severely inhibit the development of effective research and education in computational science. The paucity of incentives for longer-term multidisciplinary, multi-agency, or multi-sector efforts stifles structural innovation.

To confront these issues, universities must significantly change their organizational structures to promote and reward collaborative research that invigorates and advances multidisciplinary science. They must also implement new multidisciplinary structures and organizations that provide rigorous, multifaceted educational preparation for the growing ranks of computational scientists the Nation will need to remain at the forefront of scientific discovery.

Federal R&D agencies face similar structural issues. To address them, the National Science and Technology Council (NSTC) must commission the National Academies to launch fast-track studies that recommend changes and

innovations – tied to strategic planning and collaboration – in the Federal R&D agencies' roles and portfolios to support revolutionary advances in computational science. Federal R&D agencies must be actively involved in this process, and individual agencies must implement changes and innovations in their organizational structures to accelerate the advancement of computational science.

Scientific needs stimulate exploration and creation of new computational techniques and, in turn, these techniques enable exploration of new scientific domains. The continued health of this dynamic computational science "ecosystem" demands long-term planning, participation, and collaboration by Federal R&D agencies and computational scientists in academia and industry. Instead, today's Federal investments remain short-term in scope, with limited strategic planning and little cooperation across disciplines or Federal R&D agencies.

For these reasons, the NSTC must commission the National Academies to convene one or more task forces to develop and maintain a multi-decade roadmap for computational science and the fields that require it, with a goal of assuring continuing U.S. leadership in science, engineering, the social sciences, and the humanities.

Because the Nation's research infrastructure has not kept pace with changing technologies, today's computational science ecosystem is unbalanced, with a software base that is inadequate to keep pace with and support evolving hardware and application needs. By starving research in enabling software and applications, the imbalance forces researchers to build atop inadequate and crumbling foundations rather than on a modern, high-quality software base. The result is greatly diminished productivity for both researchers and computing systems.

In concert with the roadmap, the Federal government must establish national software sustainability centers whose charge is to harden, document, support, and maintain vital computational science software whose useful lifetime may be measured in decades. Software areas and specific software artifacts must be chosen in consultation with academia and industry. Software vendors must be included in collaborative partnerships to develop and sustain the software infrastructure needed for research.

The explosive growth in the number and resolution of sensors and scientific instruments has engendered unprecedented volumes of data, presenting historic opportunities for major scientific breakthroughs in the 21st century. Given the strategic significance of this scientific trove, the Federal

government must provide long-term support for computational science community data repositories. These must include defined frameworks, metadata structures, algorithms, data sets, applications, and review and validation infrastructure. The Government must require funded researchers to deposit their data and research software in these repositories or with access providers that respect any necessary or appropriate security and/or privacy requirements.

The PITAC is also concerned about the Nation's overall computational capability and capacity. Today, high-end computing resources are not readily accessible and available to researchers with the most demanding computing requirements. High capital costs and the lack of computational science expertise preclude access to these resources. Moreover, available high-end computing resources are heavily oversubscribed.

The Government must provide long-term funding for national high-end computing centers at levels sufficient to ensure the regularly scheduled deployment and operation of the fastest and most capable high-end computing systems that address the most demanding computational problems. In addition, capacity centers are required to address the broader base of users. The Federal government must coordinate high-end computing infrastructure across R&D agencies in concert with the roadmapping activity.

The PITAC believes that supporting the U.S. computational science ecosystem is a national imperative for research and education in the 21st century. Like any complex ecosystem, the whole flourishes only when all its components thrive. Only sustained, coordinated investment in software, hardware, data, networking, and people, based on strategic planning, will enable the United States to realize the promise of computational science to revolutionize scientific discovery, increase economic competitiveness, and enhance national security.

The Federal government must implement coordinated, long-term computational science programs that include funding for interconnecting the software sustainability centers, national data and software repositories, and national high-end leadership centers with the researchers who use those resources, forming a balanced, coherent system that also includes regional and local resources. Such funding methods are customary practice in research communities that use scientific instruments such as light sources and telescopes, and increasingly in data-centered communities such as those that use biological databases.

Leading-edge computational science is possible only when supported by long-term, balanced R&D investments in software, hardware, data,

networking, and human resources. Inadequate investments in robust, easy-to-use software, an excessive focus on peak hardware performance, limited investments in architectures well matched to computational science needs, and inadequate support for data infrastructure and tools have endangered U.S. scientific leadership, economic competitiveness, and national security. The Federal government must rebalance R&D investments to:

- Create a new generation of well-engineered, scalable, easy-to-use software suitable for computational science that can reduce the complexity and time to solution for today's challenging scientific applications and can create accurate models and simulations that answer new questions
- Design, prototype, and evaluate new hardware architectures that can deliver larger fractions of peak hardware performance on key applications
- Focus on sensor- and data-intensive computational science applications in light of the explosive growth of data

The universality of computational science is its intellectual strength. It is also its political weakness. Because all research domains benefit from computational science but none is solely defined by it, the discipline has historically lacked the cohesive, well-organized community of advocates found in other disciplines. As a result, the United States risks losing its leadership and opportunities to more nimble international competitors. We are now at a pivotal point, with generation-long consequences for scientific leadership, economic competitiveness, and national security if we fail to act with vision and commitment. We must undertake a new, large-scale, long-term partnership among government, academia, and industry to ensure that the United States possesses the computational science expertise and resources to assure continuing leadership, prosperity, and security in the 21st century.

1 A Wake-Up Call: The Challenges to U.S. Preeminence and Competitiveness

The faint "beep, beep, beep" of Sputnik – the world's first satellite, launched into orbit by the former Soviet Union on October 4, 1957 – shook the political and intellectual leadership of the United States, galvanizing a flurry of private discussions and public actions that opened a new era of national attention to U.S. research and education in science, engineering, and technology. As his first step in addressing the "space race," President Eisenhower established the post of Science Advisor to the President to symbolize the great significance of the sciences for the Nation's security. Two agencies were created – the Advanced Research Projects Agency (ARPA) within the Defense Department to pursue fundamental research in advanced computing and other defense-related technologies, and the National Aeronautics and Space Administration (NASA) to spearhead space-related R&D. Grant and scholarship programs were established to encourage students to train for research and teaching positions in the sciences.

Today, U.S. leadership in science, engineering, and technology is again being challenged. But this time the challenge is far more diffuse, complex, and long-term than one bold technological achievement by a single U.S. competitor. In the 21st century global economy, burgeoning science and engineering capabilities of countries around the world – both friends and foes – are increasingly testing U.S. preeminence in advanced scientific R&D and in science- and engineering-based industries. Moreover, the rise of these global competitors is spurred by the very computing and networking technologies that were pioneered in the United States and that have been the engine of U.S. scientific discoveries, revolutionary advances in commerce and communications, and unprecedented productivity.

> Global competitors are increasingly testing U.S. preeminence in advanced R&D and in science- and engineering-based industries.

For example, vehicle crash-test simulation – a technique developed in the 1960s based on software created by NASA scientists – is now a fundamental component of automotive design and engineering by all the world's leading auto makers. In the pharmaceutical industry, computing capabilities are transforming the search for possible new drugs and therapies, dramatically increasing both productivity and competition in this key sector. In manufacturing and many other types of large-scale enterprises, specialized

software running on networked computing systems is used to manage the complex flow of information, materials, finances, and logistics that forms the enterprises' supply chains. These high-stakes supply-chain management systems are intended to increase cost-effectiveness and provide a competitive advantage. And in the financial sector, computational models have become the principal tools for both micro- and macro-level analysis and forecasting.

The global information technology-powered revolution is accelerating, but this Nation has not yet fully awakened to the implications. Consider the following new frontiers of science, engineering, and industry cited as the most economically promising and technologically important for the 21st century by various U.S. scientific and government organizations: advanced materials (including superconductors and semiconductors) , alternative energy sources, biotechnology, high-performance computing, microelectromechanical systems (MEMS) , nanotechnology, optoelectronics, sensors, and wireless communications. These diversified emerging technologies have one essential attribute in common: Breakthroughs and innovations in every single one of them will be won by those most skilled with advanced computing systems and computational science applications.

The global information technology-powered revolution is accelerating, but this Nation has not fully awakened to the implications.

In fact, the human skills and computing technologies supporting computational problem solving are now critical to achievements in all realms of scientific, social science, biomedical, and engineering research, defense and national security, and industrial innovation. As Presidential Science Advisor John H. Marburger III testified before the House Science Committee on February 16, 2005, "Research in networking and information technologies underpins advances in virtually every other area of science and technology and provides new capacity for economic productivity." [Marburger, 2005].

Now consider some indicators of the U.S. competitive situation today:

- U.S. information technology (IT) manufacturing has declined significantly since the 1970s, with the decline accelerating over the past five years [PCAST, 2004]. From 1980 to 2001, the U.S. share of global high-technology exports dropped from 31 percent to 18 percent, while the share for Asian countries rose from 7 percent to 25 percent [NSF, 2004a]. The U.S. maintained a trade surplus in high-tech products in the 1990s; since 2001, the balance has been negative [U.S. Census Bureau, 2003].

- Some of the computing system capabilities critical for U.S. national defense and national security have not improved substantially in a decade, and today's commercial high-end systems perform more poorly on some key metrics than older, custom-designed systems [DoD, 2002].
- The United States is producing a declining proportion of the world's scientists and engineers. In 2000, nearly 80 percent of the 114,000 science and engineering (S&E) doctorates awarded worldwide were from institutions outside the United States [NSF, 2004a]. Between 1994 and 2001, enrollments of U.S. citizens in U.S. graduate-level S&E programs dropped by 10 percent, while enrollments of temporary visa holders (foreign students) rose by 25 percent [NSF, 2004a].

The number of U.S. doctoral degrees awarded in computer science and computer engineering in 2002 reached the lowest level since 1989.

- In 2002, despite a welcome 5 percent upswing in U.S. students' graduate-level S&E participation, foreign-student enrollment grew by 8 percent and represented a substantial proportion of overall graduate enrollment in engineering (49 percent), computer science (48 percent), physical sciences (40 percent), and mathematical sciences (39 percent). In 2002, 58 percent of S&E postdoctoral positions at U.S. universities were held by temporary visa holders [NSF, 2004b].
- The 849 doctoral degrees in computer science and computer engineering awarded in 2002 by U.S. institutions was the lowest number since 1989, according to an annual Computing Research Association survey [NRC, 2005].
- Since 1988, Western Europe has produced more science and engineering journal articles than the United States, and the total growth in research papers is highest in East Asia (492 percent), followed by Japan (67 percent) and Europe (59 percent), compared with 13 percent for the United States. Worldwide, the share of U.S. citations in scientific papers is shrinking, from 38 percent in 1988 to 31 percent in 2001[NSF, 2004a].

In the PITAC's view, we must come to grips with both the broad science and technology challenge we face and the reality that the 21st century scientific and engineering enterprise is computational and multidisciplinary, requiring the collaborative scientific skills of diverse disciplines. This country led the world in developing the advanced information technologies that are transforming research, commerce, and communications. These capabilities place us on the threshold of revolutionary discoveries, such as in the treatment

of disease, atom-by-atom construction of materials with previously unimaginable properties, miniaturization of devices down to the quantum level, and new energy sources and fuel technologies. But we are not minding the store of U.S. intellectual resources needed to capitalize on the scientific opportunities of the new century.

A dangerous consequence of our current complacency is that, as on the eve of Sputnik's launch, we have not marshaled and focused our efforts to elevate computational science and the computing infrastructure to their appropriate status as a long-term, strategic national priority in education as well as R&D. Without such a commitment and focus, the PITAC believes, we cannot sustain U.S. scientific leadership, security, and economic prosperity in the decades ahead.

What Is Computational Science?

At one level, computational science is simply the application of computing capabilities to the solution of problems in the real world – for example, enabling biomedical researchers rapidly to identify to which protein, and where on that protein, a candidate vaccine will most effectively bind. The PITAC's definition of computational science (Sidebar 1, below, and Figure 1 on page 11) is intended, however, to underscore the reality that harnessing software, hardware, data, and connectivity to help solve complex problems necessarily draws on the multidisciplinary skills represented in the computing infrastructure as a whole.

Sidebar 1
Definition of Computational Science

As a basis for responding to the charge from the Office of Science and Technology Policy, the PITAC developed a definition of computational science. This definition recognizes the diverse components, ranging from algorithms, software, architecture, applications, and infrastructure that collectively represent computational science.

Computational science is a rapidly growing multidisciplinary field that uses advanced computing capabilities to understand and solve complex problems. Computational science fuses three distinct elements:

- Algorithms (numerical and non-numerical) and modeling and simulation software developed to solve science (e.g., biological, physical, and social), engineering, and humanities problems
- Computer and information science that develops and optimizes the advanced system hardware, software, networking, and data management components needed to solve computationally demanding problems
- The computing infrastructure that supports both the science and engineering problem solving and the developmental computer and information science

Visualization of Computational Science Definition

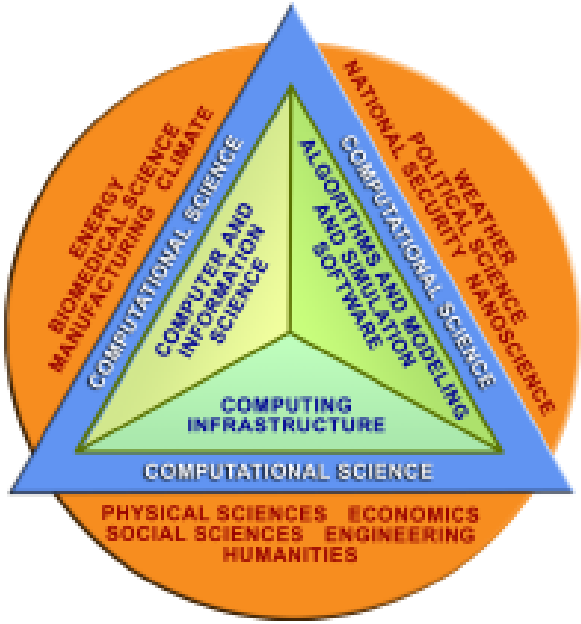


Figure 1

It takes scientific contributions across many disciplines to successfully fit software, systems, networks, and other IT components together to perform computational tasks. And it takes teams of skilled personnel representing those disciplines to manage computing system capabilities and apply them to complicated real-world challenges, much as it takes a medical team with many skills – not just a surgeon with a scalpel – to perform a complex surgical procedure. Indeed, the PITAC believes that the multidisciplinary teams required to address computational science challenges represent what will be the most common mode of science and engineering discovery throughout the 21st century.

The multidisciplinary teams required to address computational science challenges represent what will be the most common mode of 21st century science and engineering R&D.

Computational science emerged from the exigencies of World War II and the dawn of the digital computer age, when scientists trained in various disciplines – mathematics, chemistry, physics, and mechanical and electrical engineering – collaborated to build and deploy the first electronic computing machines for code-breaking and automated ballistics calculations. Today's most

advanced computing systems are fashioned from far more complex software and hardware components, and storage and communication capabilities have risen over a million-fold. These developments have qualitatively transformed not only scientific discovery but also key economic processes including industrial and pharmaceutical design and production; data-intensive analysis such as in economic forecasting, epidemiology, and weather and climate prediction; and global financial markets and systems.

The 'Third Pillar' of 21st Century Science

The first great scientific breakthrough of the new century – the decoding of the human genome announced in February 2001 – was a triumph of large-scale computational science. When the Department of Energy (DOE) and and the National Institutes of Health (NIH) launched the Human Genome Project in 1990, the most powerful computers were 100,000 times slower than today's high-end machines; private citizens using networks could send data at only 9600 baud (an outdated transmission standard; early modems transmitted at 300 baud, or a few characters per second); and many geneticists performed their calculations by hand. The challenge – determining how the genetic instructions for life are organized in the four chemical compounds that make up the biomolecule deoxyribonucleic acid (DNA) – was understood to be critical to the future of medical science, but it was expected to take decades.

The 21st century's first great scientific breakthrough was a triumph of large-scale computational science.

Ultimately, the international decoding effort, in which more than 1,000 scientists participated, became a showcase for the central role of computational science in advanced research. Distributed teams each computed pieces of possible chemical sequences and transmitted them over high-speed networks to the project's data repositories for other scientists to examine and use. Researchers devised new software that automated sequence computations and analyses. A June 2000 announcement of a "rough draft" of the genome noted that more than 60 percent of the code had been produced in the prior six months alone. Total raw sequences computed numbered more than 22 billion.

The decoding of the human genome immediately sparked a multi-billion-dollar R&D enterprise across government, academia, and industry to apply the new genetic knowledge to developing fresh understandings of biomolecular processes and inheritance factors in disease. These efforts are already generating new types of pharmaceuticals and medical interventions.

Computational science now constitutes what many call the third pillar of the scientific enterprise, a peer alongside theory and physical experimentation.[1] Indeed, as the genome decoding effort demonstrated, computational science offers powerful advantages over other research methods, enabling rapid calculations on volumes of data that no person could complete in a lifetime. The practical difference between obtaining results in hours, rather than weeks or years, is substantial – it qualitatively changes the range of studies one can conduct. For example, climate change studies, which simulate thousands of Earth years, are feasible only if the time to simulate a year of climate is a few hours. Moreover, to understand the sensitivity of climate predictions to assumptions about human impacts (e.g., generated fluorocarbon or carbon dioxide emissions) or model characteristics, one must conduct entire suites of climate simulations. This requires prodigious amounts of computing power.

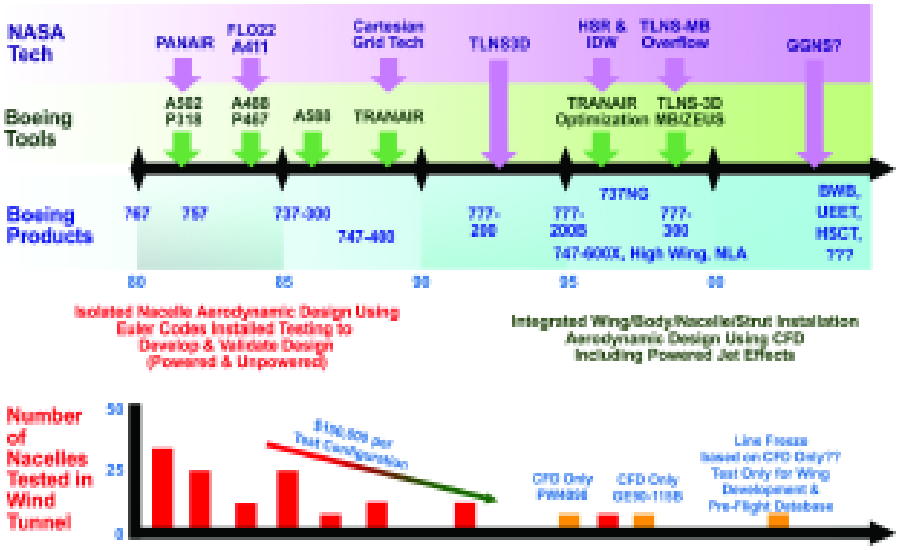
Computational science has become the third pillar of the scientific enterprise, a peer alongside theory and physical experiment.

But raw computation speeds represent only one facet of the third pillar. Computational science enables researchers and practitioners to bring to life theoretical models of phenomena too complex, costly, hazardous, vast, or small for "wet" experimentation. Computational cosmology, which tests competing theories of the universe's origins by computationally evolving cosmological models, is one such area. We cannot create physical variants of the current universe or observe its future evolution, so computational simulation is the only feasible way to conduct experiments.

To cite another example, researchers have long known that microbubbles, about 50 to 500 microns in size, can cut the drag experienced by ships (by 80 percent in some cases), reduce the amount of fuel they use, and increase their range. Microbubble effects have been studied experimentally for three decades, but the water turbulence in these physical experiments prevents precise observations and measurements of the optimum conditions for minimizing drag. Now researchers have made a major leap forward toward developing new hull technologies by creating innovative computational models that can simulate the flow and influence on hull speed of microbubbles of varying sizes. Using high-end computing systems, the researchers have been able to simulate the flow of about 20,000 microbubbles simultaneously. The next steps will involve using data from the simulations to zero in on optimal microbubble size and flow and testing the findings in physical models.

[1] The designation of computational science as the third pillar of scientific discovery has been widely cited in the scientific literature and acknowledged in Congressional testimony and Federal and private-sector reports.

Impact of Computational Fluid Dynamics on Wind Tunnel Testing for Propulsion Integration



Data courtesy of the Boeing Corporation

Figure 2

Applying NASA computational technologies, the Boeing Corporation in the 1980s developed modeling tools that enabled the company to transform the process of aircraft design from a dependence on costly physical testing of isolated structures (called nacelles) such as wings, engine housings, and pilot compartments, to fully integrated computational modeling of complete aircraft, including powered jet effects. The capability radically reduces testing costs and speeds production.

Computational science also makes it possible to examine the interplay of processes across disciplinary boundaries. For example, a model devised by a civil and environmental engineering researcher has identified the costs and benefits of various strategies for remediating groundwater contamination. Removing chemical contaminants involves many decisions about the placement of water pumps and the rate and duration of pumping. Typical plans use only rough cost estimates. Using computationally intensive genetic algorithms, the simulations demonstrated that, beyond a certain threshold, additional spending produces negligible additional reductions in groundwater contaminants. Thus, planning within the threshold's limits can rein in costs without lessening the effectiveness of remediation efforts.

Computational science makes it possible to examine the interplay of processes across disciplinary boundaries.

Understanding the environmental and biological bases of respiratory disease or biological attack, for example, requires an even more complex interdisciplinary modeling effort that couples social science and public health data and experiences with fluid dynamics models of airflow and inhalants (smoke, allergens, pathogens) , materials models of surface properties and interactions, biophysics models of cilia and their movements for ejecting foreign materials, and biological models of the genetic susceptibility to disease. The complexity of these interdisciplinary models is such that they can only be evaluated using high-performance computers. (Appendix A provides descriptions of computational science applications in many different fields.)

The universality of computational science is its intellectual strength, but it is also its political weakness.

In the marketplace, computational science provides a competitive edge by transforming business and engineering practices. Integrated modeling and simulation techniques enabled the Boeing Company to minimize wind tunnel testing as a part of its wing design process, resulting in cost savings and reduced time to market (Figure 2) . In a recent Council on Competitiveness survey of businesses [Joseph et al. , 2004], the overwhelming majority said computational science was not only beneficial but also essential to company survival.

An Unfinished Revolution

Powerful new telescopes advance astronomy, but not materials science. Powerful new particle accelerators advance high-energy physics, but not genetics. In contrast, computational science advances all of science and engineering, because all disciplines benefit from high-resolution model predictions, theoretical validations, and experimental data analysis. As with computing itself, new scientific discoveries increasingly lie at the intersections of traditional disciplines, where computational science is the research integration enabler.

The universality of computational science is its intellectual strength, but it is also its political weakness. Because all research domains benefit from it but none is solely defined by it, this quintessentially multidisciplinary field historically has lacked the cohesive, well-organized community of advocates found in other disciplines and the concomitant strategic assessment of the Nation's increasing requirements for advanced computational science. The PITAC believes that the Nation's failure to embrace computational science is symptomatic of a larger failure to recognize that many 21st-century research challenges are themselves profoundly multidisciplinary, requiring teams of highly skilled people from diverse areas of science, engineering, public policy, and the social sciences.

In consequence, despite formidable computational science successes, our R&D programs, which are predominantly Federally supported, are drifting for the most part on tradition. The norm is fragmented, discipline-based research practices that impede fully effective development and integration of computational science in advanced discovery. Moreover, today we are neither training enough computational scientists nor appropriately preparing students for the disciplinary and multidisciplinary use of leading-edge computational science techniques. Inadequate and outmoded educational structures within academia, mirrored in the Federal agencies' disciplinary silos, leave computational science students to flounder amid competing departments.

The Nation's failure to embrace computational science is symptomatic of a larger failure to recognize that many 21st-century challenges are themselves profoundly multidisciplinary.

In addition, our preoccupation with peak performance and computing hardware, vital though they are, masks the deeply troubling reality that the most serious technical problems in computational science lie in software,

Sidebar 2

Repeating History: Lessons Not Learned

During the past two decades, the national science community has produced a plethora of reports, each recommending sustained, long-term investment in the underlying technologies (algorithms, software, architectures, hardware, and networks) and applications needed to realize the benefits of computational science. These reports have stressed the now essential role that computational science plays in supporting, stimulating, catalyzing, and transforming the conduct of science and engineering.

The reports have also emphasized how computing can address applications of significantly greater complexity, scope, and scale, including problems and issues of national importance that cannot be otherwise addressed. Many of the reports generated responses, but they were often short-lived. In general, short-term investment and limited strategic planning have led to excessive focus on incremental research rather than on long-term, sustained research with lasting impact that can solve important problems. These reports and their messages are summarized in Appendix B.

A report card of national performance might record a grade of C–, with an accompanying teacher's note that says, "This student has great potential, but struggles to maintain focus and complete work on time. This student sometimes has difficulty sharing and playing well with others."

usability, and trained personnel. Heroic efforts are regularly devoted to extending legacy application codes on the latest platforms using primitive software tools and programming models. Meanwhile, the fundamental R&D necessary to create balanced hardware-software systems that are easy to use, facilitate application expression in high-level models, and deliver large fractions of their peak performance on computational science applications is perennially postponed for a more opportune time. More ominously, these difficulties are substantial intellectual hurdles that limit broad education and training.

The PITAC's Call to Action

The PITAC believes that current education and research structures and priorities must change radically if the United States is to sustain its world preeminence in science, engineering, and economic innovation. We are not alone. For two decades, organizations in government, academia, and industry have been issuing reports recommending sustained, long-term investment to realize the benefits of computational science. As Sidebar 2 notes, these calls have had only a limited impact. Instead, short-term investment and limited strategic planning have led to excessive focus on incremental research rather than on long-term, sustained research with lasting impact. Furthermore, silo mentalities have restricted the flow of ideas and solutions from one domain to another, resulting in duplication of effort and little interoperability.

Silo mentalities have restricted the flow of ideas and solutions from one domain to another, resulting in duplication of effort and little interoperability.

The PITAC's call to action begins with the following principal finding and recommendation:

Principal Finding

Computational science is now indispensable to the solution of complex problems in every sector, from traditional science and engineering domains to such key areas as national security, public health, and economic innovation. Advances in computing and connectivity make it possible to develop computational models and capture and analyze unprecedented amounts of experimental and observational data to address problems previously deemed intractable or beyond imagination. Yet despite the great opportunities and needs, universities and the Federal government have not effectively recognized the strategic significance of computational science in either their organizational structures or their research and educational planning. These

inadequacies compromise U.S. scientific leadership, economic competitiveness, and national security.

Principal Recommendation

Universities and the Federal government's R&D agencies must make coordinated, fundamental, structural changes that affirm the integral role of computational science in addressing the 21st century's most important problems, which are predominantly multidisciplinary, multi-agency, multi-sector, and collaborative. To initiate the required transformation, the Federal government, in partnership with academia and industry, must also create and execute a multi-decade roadmap directing coordinated advances in computational science and its applications in science and engineering disciplines.

We are now at a pivotal point, with generation-long consequences for scientific leadership and economic competitiveness if we fail to act with vision and commitment. As our principal finding and recommendation indicate, we must undertake a new large-scale, long-term partnership among government, academia, and industry to ensure that the United States has the computational science expertise and resources it will need to assure national security, economic success, and a rising standard of living in the 21st century. In the additional findings and recommendations in Chapters 2-5 of this report, the PITAC identifies the structural issues that must be addressed and proposes a major sustained roadmap initiative to guide the efforts of the national computational science partnership.

We are now at a pivotal point, with generation-long consequences for scientific leadership and economic competitiveness if we fail to act with vision and commitment.

2 Medieval or Modern? Research and Education Structures for the 21st Century

Finding

Traditional disciplinary boundaries within academia and Federal R&D agencies severely inhibit the development of effective research and education in computational science. The paucity of incentives for longer-term multidisciplinary, multiagency, or multisector efforts stifles structural innovation.

Recommendation

Universities must significantly change their organizational structures to promote and reward collaborative research that invigorates and advances multidisciplinary science. Universities must implement new multidisciplinary structures and organizations that provide rigorous, multifaceted educational preparation for the growing ranks of computational scientists the Nation will need to remain at the forefront of scientific discovery.

Recommendation

The National Science and Technology Council (NSTC) must commission a fast-track study by the National Academies to recommend changes and innovations – tied to strategic planning and collaboration – in the Federal R&D agencies' roles and portfolios to support revolutionary advances in computational science. Federal R&D agencies must be actively involved in this process. In addition, individual agencies must implement changes and innovations in their organizational structures to accelerate the advancement of computational science.

Removing Organizational Silos

Organizational structures in academia have antecedents reaching back to the Renaissance, with departments, schools, and colleges organized around disciplinary themes. These structures evolve so slowly that creating a new department often requires years of negotiation and resource planning, and reorganizing or creating a college occurs so rarely that each such action is national news in academic circles. The Federal R&D agencies have similar constraints on organizational change. Indeed, the current organizational structures of many Federal R&D agencies closely align with the organizational

charts for colleges of science and engineering or medical schools (Figure 3). Given the flow of people and ideas between academia and government, these similarities are hardly surprising.

The relationships among universities, agencies, and the national laboratories reinforce the organizational status quo. Universities and national laboratories provide the talent pool from which most research agency leaders are drawn. The universities and laboratories are the direct financial beneficiaries of Federally funded research, and they in turn educate and train each new generation of researchers and educators. Although this relationship has long ensured U.S. preeminence in scientific discovery and the associated research, economic, and national security benefits, its reward systems resist rapid evolution when circumstances necessitate change. The result is an architecture of organizational structures trapped in time and constrained in rigid disciplinary silos whose mutually reinforcing boundaries limit adaptation to changing research needs and competitive pressures.

The notable exception has been the rise of crosscutting centers and institutes. Most often, these entities are created in response to a funding

A Traditional University Organizational Structure

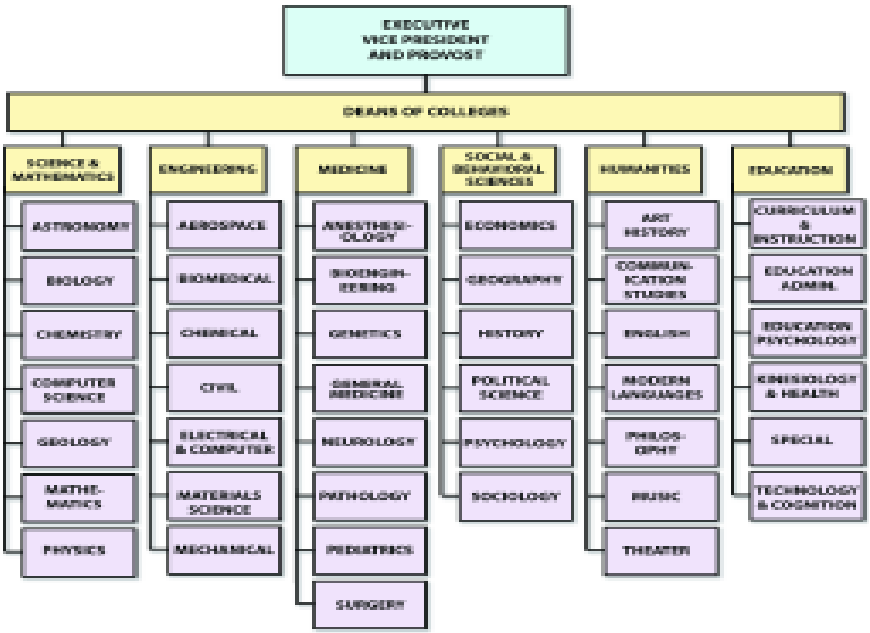


Figure 3

Traditional disciplinary boundaries within academia and Federal R&D agencies severely inhibit the development of effective research and education in computational science.

opportunity that requires a specific skill set not found solely within a particular department, or they seek to bridge the boundaries that isolate researchers, faculty, and students within departments or colleges. Because the associated Federal R&D agency programs often have sunset clauses, the entities typically are ephemeral and neither the agencies nor the universities alter their fundamental organizational structures for education and research.

Increasing international investment in science and engineering as economic drivers, together with a lack of U.S. emphasis on interdisciplinary science and engineering education and flat to declining Federal funding for long-term, basic research, have placed the historically vibrant productivity of universities, Federal R&D agencies, and national laboratories at risk. This must change. Both universities and Federal R&D agencies must escape from their disciplinary silos and rigid organizational structures if we are to realize the full potential of computational science to support our strategic national interests.

> The relationships among universities, agencies, and the national laboratories reinforce the organizational status quo.

Evolving Agency Roles and Priorities

Federal R&D agencies manage a complex portfolio of basic and applied research with widely varying time horizons. At one extreme, short-term applied research is intended to yield practical results within months. At the other, long-term basic research is driven by curiosity, without regard to expected utility but based on historical experience that basic research yields large, long-term, and unexpected benefits. A wide spectrum of basic and applied computational science research, driven by both strategic research plans and curiosity, lies between.

The missions of Federal R&D agencies range from the Defense Advanced Research Projects Agency (DARPA) focus on advancing and ensuring defense capabilities, to the NIH portfolio of basic and clinical research studies for improved health care, to the predominant DOE Office of Science and National Science Foundation (NSF) focus on long-term basic research. Historically, these agencies have each occupied unique but collaborative niches in basic and applied research planning and support.

Based on its analysis of Federal R&D agency activities, PITAC concluded that Federal support for computational science research has been overly focused on short-term, low-risk activities. In the long term, this is actually a high-risk strategy that is less likely to yield the high-payoff, strategic

innovations needed for the future. Diversifying agency research portfolios can reduce this risk. For example, a portion of each agency's research budget could be allocated to programs that exist only to foster high-risk exploration, with concurrent changes to the peer-review and funding-decision mechanisms to ensure that risk diversification actually occurs. The PITAC report *Information Technology: Investing in Our Future* [PITAC, 1999] strongly recommended an expanded, sustained program of long-term information technology research investments in the Federal R&D portfolio.

Change in a Federal R&D agency's computational science role and priorities, due to internal opportunities or external circumstances, affects allied agencies either positively or negatively. In the 1980s and 1990s, DARPA's investment in novel parallel architectures and advanced prototypes stimulated a shift from traditional vector architectures and provided an infrastructure base upon which other agencies – notably NSF, DOE, and NASA – funded research in parallel algorithms, software tools and techniques, and advanced scientific applications. DARPA's later termination of this program created an architectural research vacuum that persists today.

Substantially increased intra- and interagency coordination is required to ensure that national priorities are not harmed by such agency priority shifts. Although the Subcommittee for Networking and Information Technology R&D (NITRD) within the NSTC facilitates cross-agency coordination, large-scale changes to agency priorities are made within agencies or through the Federal budget process. These issues are discussed in Chapter 3.

Federal R&D agencies, national laboratories, and universities are subject to periodic reviews conducted by external panels of experts. At universities, these

Sidebar 3

The Limited Computational Science Talent Pool

A recent Council on Competitiveness survey of businesses revealed that the dearth of qualified computational scientists was a significant impediment to broader commercial deployment of computational science tools, techniques, and infrastructure. Researchers at national laboratories and universities have echoed this concern, noting the difficulty in finding graduate students, post-doctoral research associates, and staff members with the range of disciplinary and computational skills needed.

Of the declining number of U.S. students in science and engineering graduate study, computational scientists represent only a tiny fraction. The shortage of U.S. citizens with these skills is particularly pernicious for national laboratories, where security clearances are required for many positions.

reviews range from departmental and center reviews to university accreditation assessments. Although each of the Federal R&D agencies individually convenes a set of advisory panels and oversight groups, today there is no body that considers how Federal agency roles and priorities can best support the full ecosystem of computational science. A high-level evaluation of agency interactions, agency structures, and rewards for interagency collaboration based on emerging computational science research opportunities is urgently needed.

The Challenge of Multidisciplinary Education

Based on an intensive review of prior reports (Appendix B) and its own investigations, PITAC finds that the emerging problems of the 21st century will require insights and skills from diverse domains and often coordinated engagement by teams that collectively possess those skills. But despite growing evidence of the need for such problem-solving teams, it is often difficult to construct them (Sidebar 3). Computational scientists working on problems in a range of fields report substantial difficulty in finding students and postdoctoral research associates who can bring skills in such areas as algorithms, software, architecture, data management, visualization, performance analysis, science, engineering, and public policy.

These observations illustrate the dominance of disciplinary culture and the need to find reward metrics and mechanisms that encourage interdisciplinary collaboration and education. For example, the Biomedical Information Science and Technology Initiative (BISTI) report [NIH, 1999] noted the disparate cultures of biomedical and information technology research, with postdoctoral associates common in biomedicine but uncommon in biomedical computing (the application of information technologies in biomedical research and clinical practice).

Students benefit when their classroom research training is coupled with hands-on experiences. This suggests that new programs should provide experiential and collaborative learning environments at the graduate and undergraduate level and should tie these environments to ongoing R&D efforts, which could be supported through centers and institutes. These learning experiences should place students in real-world situations, including internships and field experiences. To devise such new program directions, we need to fund curriculum development in computational science, targeting best practices, models, and structures.

In undergraduate education, the difficulties of implementing multidisciplinary programs are particularly acute, as both students and prospective employers tend to focus on traditional single-discipline degrees. Nevertheless, undergraduates must be exposed to the capabilities and

opportunities in computational science so that they graduate with a more informed understanding of the field and more interest in pursuing graduate computational science programs or degrees. One way to begin is through individual course offerings that may eventually lead to concentrations, minors, and majors in computational science. In addition, we need to find ways to encourage faculty members to become more informed about computational science capabilities and developments in their areas of expertise.

A number of U.S. computational science education programs have emerged over the past decade in an attempt to meet these needs. A recent report on graduate computational science and engineering (CSE) education programs [SIAM, 2001] identified 28 such programs, organized in one of two general formats. The first results in a graduate degree in CSE and typically resides in an existing academic department, usually mathematics or computer science. The second results in degrees in mathematics, computer science, the sciences, or engineering but with a specialization in CSE.

The number of computational science graduates is inadequate to meet even current demand.

However, the number of graduates from computational science programs is inadequate to meet even current demand, and it is far below the number that will be needed in the future. This demand exists both in national laboratories and universities and in commercial contexts, as shown by the Council on Competitiveness survey [Joseph et al. , 2004]. It is past time for universities to take action. They must examine their educational practices and organizational structures to provide and reward interdisciplinary and collaborative research and education. New structures, programs, and institutional incentives are urgently required.

Developing 21st Century Computational Science Leaders

Addressing the interdependent, structural weaknesses in education and research will require imaginative and vigorous thinking by experienced, engaged leaders in academia and Government. But the PITAC estimates that today there are fewer than 100 senior leaders in computational science willing and able to assume national roles in government, academia, and industry. This tiny leadership talent pool signals that substantial impediments to progress and innovation may lie ahead.

As the complexity and scale of scientific infrastructure continue to rise, an increasingly sophisticated mix of skills is needed to encourage and guide the construction and operation of computational science applications, computing

infrastructure, data management and visualization tools, and collaboration environments. For example, many of NSF's Major Research Equipment and Facilities Construction (MREFC) projects – among them the Extensible Terascale Facility (ETF) program to build a comprehensive infrastructure for distributed scientific research – include construction budgets in excess of $50 million. Similarly, many Federally supported computational science applications now rival or exceed commercial software products in complexity and development time. But current graduate and postdoctoral education rarely prepares faculty for planning and managing projects of this magnitude.

The current dearth of qualified and willing leaders can be remedied only by a sustained leadership development program targeting younger researchers and exposing them to the processes and challenges of professional project planning and management, including public-service skills such as community planning and interacting with Federal agency officials, Congressional committees, and their staffs. Such skills are crucial to the success of large-scale computational science projects and infrastructure supervision and administration.

To begin to prepare such leaders, short-term management programs tailored to the culture and needs of the computational science community could be developed. Computational science graduate curricula could include courses on project management. Mentor-protégé programs could be established to foster development of promising early-career computational scientists. The PITAC offers these examples not as a prescriptive or comprehensive plan of action but as a demonstration that solutions do exist and need to be identified and implemented.

Public service can be promoted in scholarly and professional societies and the Government itself. Stakeholders should work to identify the activities most valuable and practical to implement. For example, early-career fellowship programs could be developed to cultivate national leaders in computational science. Fellows would participate in short-term (such as one semester) interagency policy development and implementation projects in Washington, D.C. Such programs would address, at least in part, a serious longstanding problem in Federal personnel (Sidebar 4, next page). The National Academies studies on organizational structures called for in this chapter and the computational science roadmap called for in Chapter 3 should also address the leadership development issue.

Addressing the interdependent structural weaknesses in education and research will require imaginative and vigorous thinking by leaders in academia and government.

Sidebar 4

The Increasing Challenge of Government Service

Concurrent with efforts to develop leaders in computational science, the Government must address the enormous challenge of luring top talent to all levels of Government service. This longstanding systemic impediment severely limits the available talent pool for most if not all Federal agencies. Each year, Federal R&D agencies must fill multiple technical positions, ranging from program officers to division directors, assistant or associate directors, and directors. For senior positions such as agency heads, prestige and potential influence on government policy are sufficient to attract and retain highly qualified applicants. At lower levels of government service, however, attracting and retaining such candidates has proven increasingly difficult. There are at least three reasons for this difficulty:

1. The rise of two career families means that accepting a position in Washington, D.C., often requires maintaining a second residence there, as family members cannot be moved without upheaval to another career. Enabling a greater number of individuals to work remotely would broaden the base of possible participants.
2. Maintaining two residences increases the financial burden of government service. Although service under the Intergovernmental Personnel Act (IPA) allows an individual to maintain the salary level earned at the home institution, the relocation offset for service away from the primary residence rarely covers the actual costs of relocation. Moreover, taking a permanent Federal position requires an academic to relinquish tenure and accept remuneration at government pay scales, which are substantially lower than those paid to senior faculty at major research universities. A more equitable housing assistance package would reduce the financial burden and increase participation.
3. Federal conflict-of-interest rules in effect levy a substantial research penalty on academics who choose IPA service. Active researchers must divest Federal funding and disassociate themselves from collaborations that might involve seeking funding from the employing agency. And it can take several years to rebuild research programs after a term of government service. Reevaluation of current conflict-of-interest rules to better distinguish between technical and actual conflicts would also increase the pool of participants.

These disincentives leave Federal R&D agencies too often unable to attract the "best and brightest" academic, national laboratory, and industry leaders to mid- and lower-level positions. Further, even when recruitment efforts are successful, promising Federal hires are often not given a clearly defined career path or challenged to assume leadership roles, and subsequently leave Government for the private sector. As a consequence, Federal programs and research initiatives do not reap the full benefits of research experience, and the community does not gain the full measure of experience in Federal planning and decision making. With many senior Federal managers now approaching retirement, and with the flow of new U.S. scientists and engineers continuing to dwindle, the Government must address this situation quickly and proactively.

3 Multi-Decade Roadmap for Computational Science

Finding

Scientific needs stimulate exploration and creation of new computational techniques and, in turn, these techniques enable exploration of new scientific domains. The continued health of this dynamic computational science "ecosystem" demands long-term planning, participation, and collaboration by Federal R&D agencies and computational scientists in academia and industry. Instead, today's Federal investments remain short-term in scope, with limited strategic planning and a paucity of cooperation across disciplines and agencies.

Recommendation

The National Science and Technology Council (NSTC) must commission the National Academies to convene, on a fast track, one or more task forces to develop and maintain a multi-decade roadmap for computational science and the fields that require it, with a goal of assuring continuing U.S. leadership in science, engineering, and the humanities. This roadmap must at a minimum address not only computing system software, hardware, data acquisition and storage, visualization, and networking, but also science, engineering, and humanities algorithms and applications. The roadmap must identify and prioritize the difficult technical problems and establish a timeline and milestones for successfully addressing them. It must identify the roles of government, academia, and industry. The roadmap must be assessed and updated every five years, and Federal R&D agencies' progress in implementing it must be assessed every two years by PITAC.

Rationale and Need

The complexity of contemporary scientific research, visible in the growing interdependencies of formerly disparate disciplines, has required new collaborative modes. Progress in some research areas has been held back or even halted by a lack of advancement or coordination in related areas. The effects in computational science are particularly dramatic. Despite two decades of efforts to highlight structural barriers limiting advances in computational science and to encourage sustained, long-term funding for the field, Federal investments remain short-term, with limited strategic planning and interagency cooperation. This has not only slowed innovation within the discipline itself but also had a negative impact on innovation within the

numerous disciplines that rely on the robustness of the computational science ecosystem.

Computational science applications, algorithms, system software, tools, and hardware, including input/output devices and networks, are core components of the overall ecosystem in which computational science is conducted. The ecosystem also encompasses a sustained research infrastructure including software repositories and data archives that researchers can exploit. Because an inadequacy in any component or an imbalance across components adversely affects the whole, the design, development, and support of computational science environments must be systemic. Failure to follow this approach inevitably results in unsatisfactory systems that do not meet the needs of application researchers.

Failure to develop and support computational science components systemically inevitably results in unsatisfactory systems that do not meet researchers' needs.

Improving computational science capabilities to face current and future challenges will require a series of complicated, interrelated, long-term projects. Taken together, these projects constitute a dynamic program that will involve a significant number of components and communities in a sustained effort to improve and enhance scientific discovery. Recent experience in other complex fields has shown that a detailed and frequently updated long-term program management plan – often called a "roadmap" – is the best way to chart and sustain coordinated innovation in such a wide-ranging effort.

The PITAC believes that the development and maintenance of a long-term roadmap for computational science is essential to its future health and advancement. The knowledge and long-term strategy derived from a roadmap will guide coordinated investments in algorithms, software, hardware, applications, and infrastructure for computational science. (Figure 4 on pages 30-31 presents a schematic view of the proposed roadmap.)

Roadmap examples are already available to the computational science community. They include SEMATECH's International Technology Roadmap for Semiconductors (ITRS) [ITRS, 2005], which regularly assesses semiconductor requirements to "ensure advancements in the performance of integrated circuits," and the recent National Institutes of Health Roadmap [NIH, 2004]. Its purpose was to "identify major opportunities and gaps in

biomedical research that no single institute at NIH could tackle alone but that the agency as a whole must address, to make the biggest impact on the progress of medical research." The agency cited the complexity of biology as "a daunting challenge" that its roadmap would need to address.

The new computational science roadmap can re-orient current support structures to address primary community goals, evolve new structures and components holistically, guide and coordinate future Federal R&D investments, minimize technological disruptions, and create a sustained infrastructure and communication system enabling researchers and skilled practitioners across the computational science spectrum to work together. Additionally, it can help address the acute shortage of educated and skilled people in computational science.

Development of a long-term roadmap for computational science is essential to its future health and advancement.

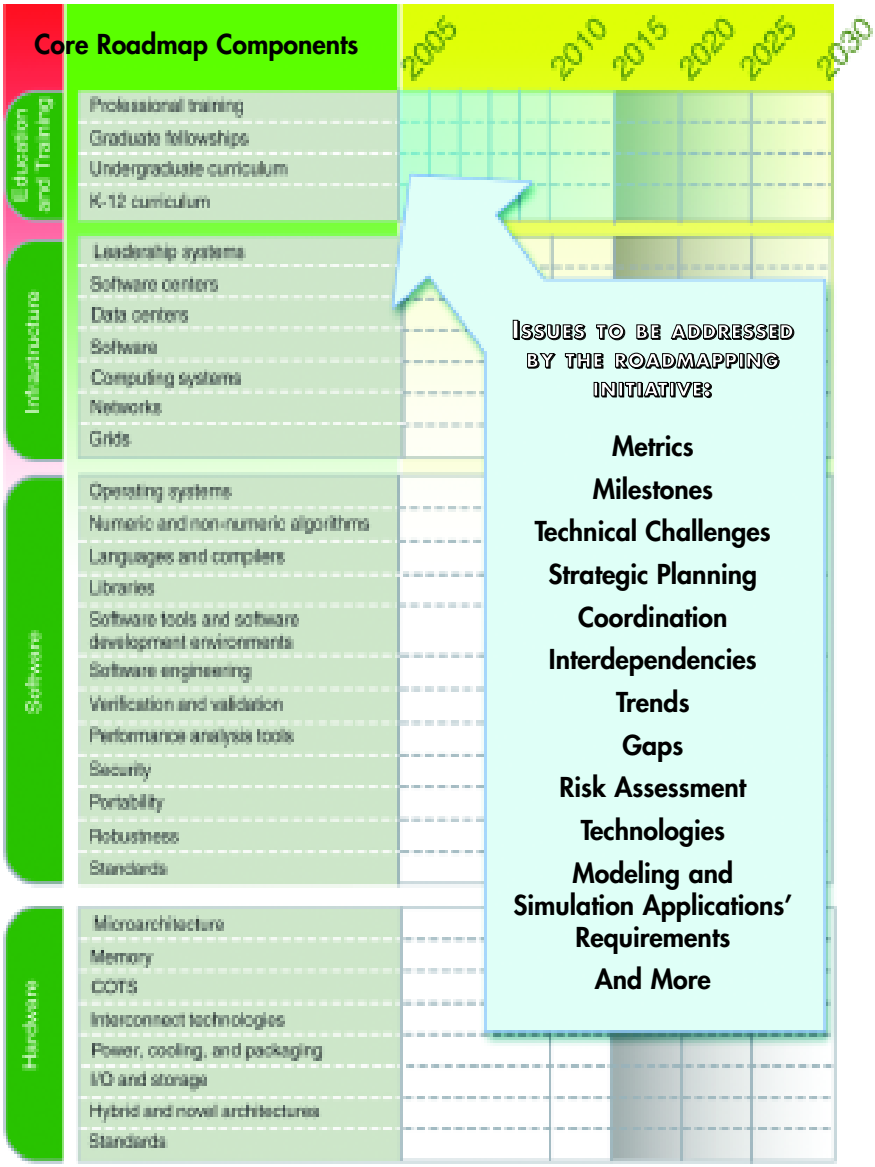
In pointing the way to future generations of computational science infrastructure, software, and technologies, the roadmap must address the multidisciplinary characteristics of the computational science community, including its complex interactions. Individual programs and solicitations must be viewed and managed within the context of the roadmap's strategic and tactical goals.

Computational Science Roadmap Components

Continued progress requires balanced investment in both computational science itself and its applications across many domains. Research in high-end architecture, systems software, programming models, algorithms, software tools and environments, data analysis and management, and mathematical methods differs from research in the use of computational science to address challenging application problems. Both kinds of research are important, but they require different expertise and generally are conducted by different people. It is a mistake to confound the two.

In addition to the lack of sustainable infrastructure, fragile, inadequate software most often limits the ability of disciplinary and interdisciplinary teams to integrate and support complex computational science R&D. As a result, software issues frequently consume the intellectual energies of students and research staff, to the detriment of research goals. Software must be a primary focus of the proposed computational science roadmap.

The Computational Science Roadmap:

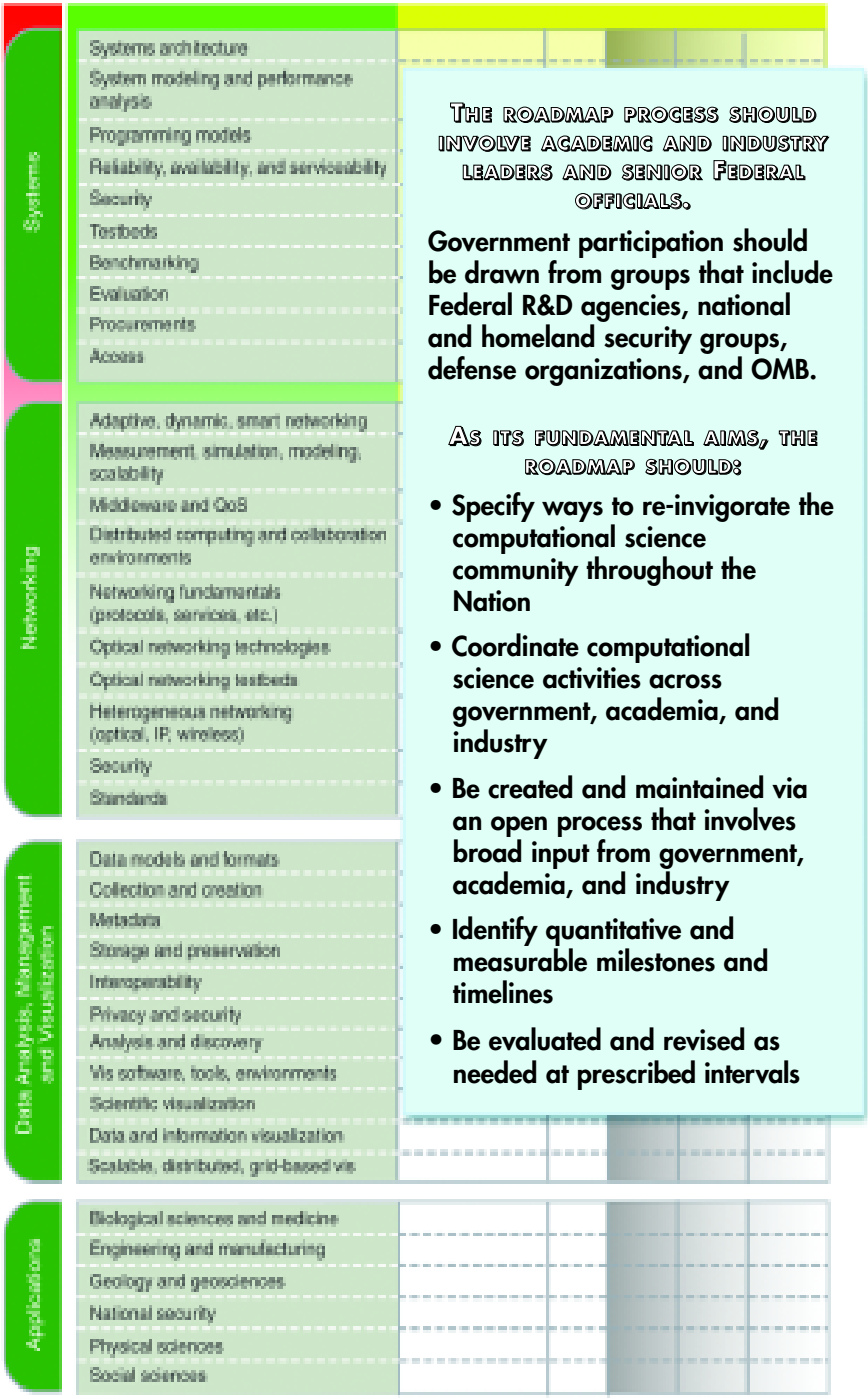


CONTINUED AT RIGHT

Figure 4

A detailed and frequently updated long-term program management plan – like SEMATECH's International Technology Roadmap for Semiconductors – is the best way to chart and sustain coordinated innovation in a wide-ranging effort. The knowledge and long-term strategy derived from the computational science roadmap will guide coordinated investments in algorithms, software, hardware, applications, and infrastructure.

A Schematic View



THE ROADMAP PROCESS SHOULD INVOLVE ACADEMIC AND INDUSTRY LEADERS AND SENIOR FEDERAL OFFICIALS.

Government participation should be drawn from groups that include Federal R&D agencies, national and homeland security groups, defense organizations, and OMB.

AS ITS FUNDAMENTAL AIMS, THE ROADMAP SHOULD:

- **Specify ways to re-invigorate the computational science community throughout the Nation**
- **Coordinate computational science activities across government, academia, and industry**
- **Be created and maintained via an open process that involves broad input from government, academia, and industry**
- **Identify quantitative and measurable milestones and timelines**
- **Be evaluated and revised as needed at prescribed intervals**

For computational science applications, the roadmapping effort will investigate a set of technological solutions (combinations of algorithms, software, and hardware) . For each application area, it will provide estimates of both the time to solution and the total cost of research, development, and ownership. As shown in Figure 4, PITAC recommends that the computational science investment priorities should include, but not be limited to, the following eight areas:

Continued progress requires balanced investment in both computational science and its applications in many domains.

1. Computational science education and training, to ensure the availability of a trained and ready workforce for research, industrial competitiveness, and national security. Sub-areas include professional training, graduate fellowships, and undergraduate and K-12 curricula.

2. Infrastructure for computational science, including high-end computing leadership centers, software sustainability centers, data and software repositories, and the middleware and networks over which users access the resources at these centers and collaborate on multidisciplinary projects.

3. The full spectrum of algorithms and software required to manage, analyze performance, and program computing systems, including numerical and non-numerical algorithms, software development environments that provide robustness and security when appropriate, and verification and validation procedures.

4. Hardware, including custom, commercial off-the-shelf (COTS) , hybrid, and novel architectures, interconnect technologies, I/O and storage, power, cooling, and packaging, to meet the growing needs of computational science applications.

5. Development of comprehensive system-wide designs using testbeds on which system modeling and performance analysis tools can be used to evaluate how effectively the interacting components perform on a given application suite. Creation of new models for system procurement that recognize the need for long-term investment and sustainability.

6. All aspects of networking including hardware technologies, middleware, protocols, and standards necessary to provide users access to computing resources, data resources, and fixed and mobile sensors with the requisite speed and security.

7. Data analysis, management, and discovery tools for heterogeneous, multimodal data, including business intelligence, scientific and information visualization, mining, and processing capabilities.

8. Applications in the biological sciences and medicine, engineering and manufacturing, geosciences, national security, physical sciences, and the social sciences.

The most critical of these topics are addressed in Chapters 4 and 5.

Roadmap Process, Outcomes, and Sustainability

Reflecting the computational science ecosystem's diverse needs and constituencies, the roadmap process should involve academic and industry leaders and senior Federal officials. Government participation should be drawn from groups that include Federal R&D agencies, national and homeland security groups, defense organizations, and the Office of Management and Budget (OMB).

Successful roadmapping generally involves planning, identifying needs, establishing process requirements and/or recommendations, and conducting periodic assessments of the roadmap itself. This roadmap should address modeling and simulation applications' requirements, interagency coordination, interdependencies among roadmap activities, trends, gaps, risk assessment of current technologies, new technologies, and more. As its fundamental aims, the roadmap should:

The roadmap process should involve academic and industry leaders and senior Federal officials.

- Specify ways to re-invigorate the computational science community throughout the Nation
- Coordinate computational science activities across government, academia, and industry
- Be created and maintained via an open process that involves broad input from government, academia, and industry
- Identify quantitative and measurable milestones and timelines
- Be evaluated and revised as needed at prescribed intervals

While planning and processes are a critical part of any roadmap, it is perhaps most important to regard it as an ongoing process. Not simply a one-

time activity, the roadmap must be a living document that is updated regularly based on objective measures of performance and evolving need.

Agency strategies for computational science should be shaped in response to the roadmap, resulting in updated strategic plans that recognize and address new roadmap priorities and funding requirements. To assist agencies in this difficult endeavor, the roadmap should specify opportunities for coordinating agency activities, successes, and challenges.

Agency strategies for computational science should be shaped in response to the roadmap.

Establishing – and following – a computational science roadmap built independently but reflecting the consensus of the R&D and associated communities will prove to be a significant step toward getting the United States "back to the future" where the Nation's technological leadership and excellence remain indisputable. The following two chapters discuss in detail the specific areas that must be addressed in order to chart a successful new course for 21st century computational science.

4 Sustained Infrastructure for Discovery and Competitiveness

The chemist Sir Humphrey Davy once shrewdly noted, "Nothing tends so much to the advancement of knowledge as the application of a new instrument. The native intellectual powers of men in different times are not so much the causes of the different success of their labors, as the peculiar nature of the means and artificial resources in their possession." [Hager, 1995]. In 2003, the National Science Board (NSB) , the policy body for NSF, made a similar point when it released its report on scientific infrastructure, defined to encompass (a) hardware (tools, equipment, instrumentation, platforms, and facilities); (b) software, libraries, databases, and data analysis systems; (c) technical support, including human experts; and (d) special environments and installations such as buildings [NSB, 2003].

Concluding that academic research infrastructure "… has not kept pace with rapidly changing technology, expanding research opportunities, and an increasing number of (facility) users," the NSB report recommended increasing the fraction of the NSF budget devoted to infrastructure support across the entire range of facility sizes. The NSB also recommended that the Federal government address the requirements of the Nation's science and engineering enterprise holistically, by developing interagency priorities and partnerships under the leadership of the Office of Science and Technology Policy (OSTP) , NSTC, and OMB. These recommendations remain on target and largely unimplemented.

Solid foundations of algorithms, software, computing system hardware, data and software repositories, and associated infrastructure are the building blocks of computational science. But our desire to support the new – exploration of newly discovered phenomena, development of new theories, and research into new ideas – has taken precedence over sustaining the infrastructure on which most scientific discoveries rest. The result has been duplication of effort, as multiple groups build and rebuild similar capabilities, to the detriment of overall scientific progress. PITAC believes we must rebalance our investments in infrastructure and research to maximize scientific productivity and intellectual progress. This chapter addresses four key components of infrastructure that warrant special attention, and Chapter 5 similarly discusses key research areas.

> We must rebalance our investments in infrastructure and research to maximize progress.

Software Sustainability Centers

Finding

Today's computational science ecosystem is unbalanced, with a software base that is inadequate to keep pace with and support evolving hardware and application needs. By starving research in both enabling software and applications, the imbalance forces researchers to build atop inadequate and crumbling foundations rather than on a modern, high-quality software base. The result is greatly diminished productivity for both researchers and computing systems.

Recommendation

The Federal government must establish national software sustainability centers whose charge is to harden, document, support, and maintain vital computational science software whose useful lifetime may be measured in decades. Software areas and specific software artifacts must be chosen in consultation with academia and industry. Software vendors must be included in collaborative partnerships to develop and sustain the software infrastructure needed for research.

Computational science software is developed and maintained by a disparate assortment of universities, national laboratories, and hardware and software vendors. Few of these groups have the human resources to support and sustain the software tools and infrastructure that enable computational science or to develop transforming technologies. Instead, academic and national laboratory researchers depend on an unpredictable stream of research grants and contracts, few of which contain explicit support for software development and maintenance.

Academic and national laboratory researchers depend on an unpredictable stream of research grants and contracts, few of which contain explicit support for software development and maintenance.

Because many of today's computational science software vendors are small companies, small changes in the software environment can drive them from the marketplace. The lack of sustainable markets built on long-term strategies and procurements means that most of these companies cannot easily recoup development costs with large sales volume. Hence, many of the products begin as derivatives of university or national laboratory software, either licensed or enhanced under an open source model.

Sidebar 5
Open Source Software Models

The rise of open source software that is developed and maintained by an international collaboration of practitioners has changed the landscape of computational science. The Linux operating system, perhaps the best-known open source software project, has become the *de facto* standard for technical computing, and a wide variety of tools have been developed upon this base or ported to it. Examples include numerical libraries such as LAPACK, message-passing libraries such as MPICH, graphics toolkits such as VTK, cluster toolkits such as ROCKS and OSCAR, and grid software such as Globus.

The rich and growing suite of open source software, together with the rise of large-scale instruments, has led to distributed, national and international research projects that require the sharing of software infrastructure across tens, hundreds, and sometimes thousands of institutions and individuals.This has necessitated a rethinking of software sharing and licensing. Negotiating a labyrinth of university licenses has proven intractable, and almost all such projects have adopted some version of an open source software model, generally a variant of the "BSD model" (derived from the original University of California at Berkeley license for UNIX), which allows reuse in new and diverse ways. Unfortunately, this often creates conflicts between the research desire to foster collaboration and sharing and the university desire to generate license revenues from research software.

In recognition of the need for sharing, DOE has begun requiring open source distribution of software developed by its academic partners. NSF, via its National Middleware Initiative (NMI) [NSF, 2005b], has funded the packaging and distribution of software for grid infrastructure deployments.

Despite this source of available software, few companies have flourished as purveyors of either software tools or applications. A series of workshops and reports examining the reasons why this market has not grown [Simmons, 1996] concluded that government support was needed to sustain software development, support, and access.

The open source model (Sidebar 5) effectively supports the rise of collaborative projects that require the free exchange of software components as part of a shared infrastructure. As Appendix A illustrates, these national and international projects are predicated on the existence of a shared base of reusable and extensible software that can interconnect scientific instruments, data archives, distributed collaborators, and scientific codes, while also enabling research in algorithms, techniques, and software tools. In this shared, open source model, development is collaborative, with contributions from a diverse set of participants supported through a variety of mechanisms.

The successful evolution and maintenance of such complex software depends on institutional memory – the continuous involvement of key

developers who understand the software's design and participate in its development and support over a period of years. Stability and continuity are essential to preserving the institutional memory but they are, unfortunately, a rarity. Research ideas can be explored by faculty and laboratory researchers with a small cadre of graduate students, but building and sustaining robust software requires experienced professionals and long-term commitments to hardening, porting, and enhancing that software infrastructure most valued by the research community.

Developing and supporting robust, user-friendly computational science software is expensive and intellectually challenging. However, effective development and support also require many activities not normally associated with academic research: software porting and testing, developing and testing intuitive user interfaces, and writing documentation and user manuals. The proposed software sustainability centers would work with academic researchers, application scientists, and vendors to evaluate, test, and extend community software. To ensure unbiased selection of the software to be supported by the centers, independent oversight bodies should be appointed, ideally with membership drawn from academia, national laboratories, and industry. Whatever funding model and structure are used, the implementation should ensure that a stable organization, with a lifetime of decades, can maintain and evolve the software.

Building and sustaining robust software requires experienced professionals and long-term commitments to hardening, porting, and enhancing that software infrastructure most valued by the research community.

At the same time, the Government should not duplicate the capabilities of successful commercial software packages. When new commercial providers emerge, the Government should purchase their products and redirect its own efforts toward developing technologies that it cannot otherwise obtain. In addition, academic researchers should leverage commercial software capabilities and best practices in the software tools they develop.

The barriers to replacement of today's low-level application programming interfaces are also high, due to the large investments in application software. Significantly enhancing our ability to program very large systems will require radical, coordinated changes to many technologies. To make these changes, the Government needs long-term, coordinated investments in a large number of interlocking technologies that create a cohesive software development and support environment.

National Data and Software Repositories

Finding

The explosive growth in the number and resolution of sensors and scientific instruments has engendered unprecedented volumes of data, presenting historic opportunities for major scientific breakthroughs in the 21st century. Computational science now encompasses modeling and simulation using data from these and other sources, requiring data management, mining, and interrogation.

Recommendation

The Federal government must provide long-term support for computational science community data repositories. These must include defined frameworks, metadata structures, algorithms, data sets, applications, and review and validation infrastructure. The Government must require funded researchers to deposit their data and research software in these repositories or with access providers that respect any necessary or appropriate security and/or privacy requirements.

The same technological advances that have produced inexpensive digital cameras and portable digital music players have enabled a new generation of high-resolution scientific instruments and sensors. Low-cost genetic sequencing, which has enabled comparative genomics across organisms, inexpensive microarrays, which can simultaneously test the differential expression of thousands of genes in a small sample, and high-resolution CCD detectors, which enable wide field surveys of the deep sky, all produce prodigious volumes of experimental data. For example, the planned Large Synoptic Survey Telescope (LSST) [LSST, 2005] will produce over 40 terabytes of data each night that must be stored, processed, and analyzed.

Large nationally or internationally distributed collaborations whose productivity depends on remote access to these often federated data require coordinated data management and long-term curation. From astronomy's International Virtual Observatory Alliance (IVOA) through the ATLAS and CMS detector groups for the Large Hadron Collider to the National Center for Biotechnology Information (NCBI) and large-scale social science data archives, long-term maintenance of distributed data, development of metadata and ontologies for interdisciplinary data sharing, and provenance validation mechanisms are all central to discovery.

Large distributed collaborations that depend on remote access to federated data require coordinated data management and long-term curation.

As with software maintenance, the support that sustains robust, user-friendly data repositories is expensive and intellectually challenging, and it requires many skills and activities not normally associated with academic research. However, without these repositories, many research activities are either impossible or the researchers involved must construct informal data archives whose long-term preservation and utility cannot be guaranteed.

The Federal government must provide long-term support for computational science community data repositories.

National data and software repositories, like software sustainability centers, will require concerted interagency development and support that must be derived from the strategic roadmap of research priorities and plans discussed in Chapter 3. These facilities are not inexpensive, but failure to support them will lead, as it has before, to wasteful research investments and lost productivity.

National High-End Computing Leadership Centers

Finding

High-end computing resources are not readily accessible and available to researchers with the most demanding computing requirements. High capital costs and the lack of computational science expertise preclude access to these resources. Moreover, available high-end computing resources are heavily oversubscribed.

Recommendation

The Government must provide long-term funding for national high-end computing centers at levels sufficient to ensure the regularly scheduled deployment and operation of the fastest and most capable high-end computing systems that address the most demanding computational problems. In addition, capacity centers are required to address the broader base of users. The Federal government must coordinate high-end computing infrastructure across R&D agencies in concert with the roadmapping activity.

Access to high-end computing systems is not merely a research or national security issue. In the Council on Competitiveness survey of business leaders [Joseph et al. , 2004], nearly 100 percent of respondents indicated that high-end computing tools are indispensable. In addition, NSF's cyberinfrastructure report [NSF, 2003], DoD's integrated high-end computing report [DoD, 2002], and DOE's SCaLeS study [DOE, 2003-2004] have all argued that today's high-end computing systems are inadequate to address 21st century research challenges and national needs.

Experts from multiple scientific disciplines and business domains have repeatedly made compelling cases for sustained performance 50 to 100 times current levels to reach new, important discovery thresholds. (Examples of current high-end computational science applications are presented in Appendix A.) The NSF cyberinfrastructure report stated, for example, that "the U.S. academic research community should have access to the most powerful computers that can be built and operated in production mode at any point in time, rather than an order of magnitude less powerful, as has often been the case in the past decade." It remains the case today.

High-end system deployments should be viewed not as an interagency competition but as a shared strategic need that requires coordinated agency responses.

The aggregate capability in open U.S. high-end computing roughly equals the scientific community's estimate of what is needed for a single, breakthrough scientific application study. Typically, though, these open systems are shared by a large number of users and the achieved application performance is often a small fraction of the peak hardware performance. This is not an agency-specific issue, but rather a shortfall in high-end computing capability that must be addressed by all agencies together to serve their communities' needs. High-end computing system deployments should be viewed not as an interagency competition but rather as a shared strategic need that requires aggressive coordinated responses from multiple agencies.

The Government must provide long-term funding for national high-end computing centers at levels sufficient to ensure deployment of the fastest and most capable systems that address the most demanding computational problems.

Today, the Nation's high-performance computing centers – notably those supported by DOE at the National Energy Research Scientific Computing Center (NERSC) and NSF at the San Diego Supercomputer Center (SDSC), the National Center for Supercomputing Applications (NCSA), and the Pittsburgh Supercomputing Center (PSC) – rely on ad hoc funding for isolated procurements that are not of leadership scale. Sustained investment and a new model of strategic procurement for these centers, as described in the following section, would help ensure that U.S. researchers and industry have access to the highest-performing computing systems and would increase their usability by amortizing software and hardware development costs across long-term contracts.

Infrastructure, Community, and Sustainability: Staying the Course

Finding

The computational science ecosystem described in this report is a national imperative for research and education in the 21st century. Like any complex ecosystem, the whole flourishes only when all its components thrive – the computational science applications, the human resources and time needed to create them, and the physical infrastructure on which they depend. Only sustained, coordinated investment in people, software, hardware, and data, based on strategic planning, will enable the United States to realize the promise of computational science to revolutionize scientific discovery, increase economic competitiveness, and enhance national security.

Recommendation

The Federal government must implement coordinated, long-term computational science programs that include funding for interconnecting the software sustainability centers, national data and software repositories, and national high-end leadership centers with the researchers who use those resources, forming a balanced, coherent system that also includes regional and local resources. Such funding methods are customary practice in research communities that use scientific instruments such as light sources and telescopes, increasingly in data-centered communities such as those that use the genome database, and in the national defense sector.

The Internet emerged as an international phenomenon and economic driver only after more than 20 years of Federally funded R&D. Similarly, developing and validating climate models that incorporate ocean, atmosphere, sea ice, and human interactions have required multiple cycles of development, computational experimentation, and analysis spanning decades. Developing leading-edge computational science applications is a complex process involving teams of people that often must be sustained for a decade or more to yield the benefits of the investment.

The HPCC Grand Challenges program [Workshop on Grand Challenges, 1993], the DOE Scientific Discovery through Advanced Computing (SciDAC) program [DOE, 2000], and others have supported teams of five to ten researchers drawn from multiple disciplines, typically computer science and a physical science domain, for three to five years. Often, the major scientific results from the collaboration have appeared long after the program ended. This suggests that the distribution of project team sizes and funding

durations most likely to maximize scientific return is not well understood. Case studies and an ethnographic assessment would help elucidate the most effective and responsible distributions of project sizes and lifetimes.

In many scientific disciplines, investment strategies take as a given the fact that large-scale scientific instruments (e.g., accelerators, telescopes, and environmental observatories) have operational lifetimes measured in decades and are expensive to relocate. Although the physical plant and ancillary support systems for computational science are much less widely recognized and understood, this infrastructure is similarly expensive to replicate. To acknowledge these costs and minimize overall program expenditures, the periodic review of infrastructure management and processes should be separated from an assessment of the infrastructure's utility and continued support. Sidebar 6 describes one emerging Federal effort to establish a comprehensive, long-term computing infrastructure for U.S. academic research.

Sidebar 6
Integrated Cyberinfrastructure

Enhanced research and learning communities are emerging to address the increasingly multidisciplinary and collaborative reach of knowledge-based activities in the United States and around the world. All disciplines, in fact, have arrived at a common inflection point, driven by the "push" of technological capacities and the "pull" of the demand to address the critical priorities for achieving revolutionary advances in science and engineering.

In the United States, NSF has adopted the term "cyberinfrastructure" to describe the complex, integrated IT tapestry of the future whose elements will include seamless networking, system software, and middleware providing the generic capabilities and specific tools for data, information, and knowledge management, processing, and transport. The NSF-commissioned report, *Revolutionizing Science and Engineering Through Cyberinfrastructure,* characterizes cyberinfrastructure as that portion of cyberspace where scientists can "build new types of scientific and engineering knowledge environments and organizations and . . . pursue research in new ways and with new efficiency."

The major components of cyberinfrastructure should include:

- High-performance, global-scale networking, whether a hybrid of traditional packet switching or a more advanced model built upon high-bandwidth optical networks
- Middleware enabling greater ease in applications building and implementation, secure communications, and collaborative research
- High-performance computation services, including data, information, and knowledge management
- Observation and measurement services
- Improved interfaces and visualization services

The U.S. has long maintained a schizophrenic approach to computational science infrastructure procurements, particularly of those specialized high-performance computing systems for which the Federal government is the primary customer. Although the Government has sought from the earliest days of computing to shape the commercial design of high-performance systems, its procurements have generally not been part of a long-term strategic plan. This is in striking contrast to the approach taken in defense procurements.

Defense procurements are long-term commitments, often for 30 or more years for multiple units, and they include ancillary support for spare parts and technical expertise. Although they involve highly competitive selection processes, this Federal policy helps ensure that multiple vendors remain viable, as even losing bidders are usually partners in the winning consortium.

High-performance computing systems share many attributes with defense hardware systems such as aircraft carriers, submarines, and fighter jets. They are built for specific technical purposes; their development involves large, non-recurring engineering costs; and they are sold in small quantities relative to the size of other commercial markets. Each procurement is essentially a stand-alone activity, and market forces are relied upon to ensure the continued viability of those companies involved in the production and maintenance of these complex systems.

Like defense systems, the Nation must regard procurement of computational science infrastructure as a long-term strategic commitment rather than a short-term tactical process.

Unlike military systems, however, the high-performance computing products developed by industry are derivatives of commercial offerings. The reason: Unlike military procurements, Federal procurements in high-performance computing systems and associated programs lack the size and long-term commitments necessary to shape corporate strategies. Thus, it is entirely too risky for industry to rely on such procurements as the basis for long-term business and development – the opposite of the situation in defense.

As a result, the dramatic growth of the U.S. computing industry, with its associated economic benefits, has shifted the balance of influence on computing-system design from the Government to the private sector. As the relative size of the high-end computing market has shrunk, we have not sustained the requisite levels of innovation and investment in high-end architecture and software needed for long-term U.S. competitiveness. It is imperative for the Nation to regard procurements of computational science

infrastructure as a long-term strategic commitment rather than a short-term tactical process. Such a shift will require deep and sustained collaboration among Federal agencies, companies, and customers to support the needed architectural and software research, develop operational prototypes, and procure and deploy multiple generations of systems.

The computational science community must confront the discipline's most intractable R&D challenges in a sustained and serious manner.

While addressing the issues of the computational science infrastructure, the community must also begin to confront the most intractable R&D challenges within the discipline itself in a sustained and serious manner. These problems, including inadequate and antiquated software, aging architecture and hardware technologies, outmoded algorithms and applications, and the overwhelming issues of data management, are explored more fully in Chapter 5.

5 Research and Development Challenges

FINDING

Leading-edge computational science is possible only when supported by long-term, balanced R&D investments in software, hardware, data, networking, and human resources. Inadequate investments in robust, easy-to-use software, an excessive focus on peak hardware performance, limited investments in architectures well matched to computational science needs, and inadequate support for data infrastructure and tools have endangered U.S. scientific leadership, economic competitiveness, and national security.

RECOMMENDATION

The Federal government must rebalance its R&D investments to: (a) create a new generation of well-engineered, scalable, easy-to-use software suitable for computational science that can reduce the complexity and time to solution for today's challenging scientific applications and can create accurate simulations that answer new questions; (b) design, prototype, and evaluate new hardware architectures that can deliver larger fractions of peak hardware performance on scientific applications; and (c) focus on sensor- and data-intensive computational science applications in light of the explosive growth of data.

The roadmap development process called for in Chapter 3 is intended to produce an R&D plan for computational science algorithms, software, architecture, hardware, data management, networking, and human resources. However, several issues are so vital to the long-term success of computational science that further explanation, as the basis for planning and scope, is required. This chapter discusses in greater detail the R&D challenges of particular concern, going beyond the findings of the High-End Computing Revitalization Task Force (HECRTF) , which captured salient technological and applications aspects [Executive Office of the President, 2004]. In addition, Appendix A details examples of diverse computational science applications and the technologies used in these domains.

Computational Science Software

As discussed in Chapter 4, the crisis in computational science software is multifaceted and remediation will be difficult. The crisis stems from years of inadequate investments, a lack of useful tools, a near-absence of widely accepted standards and best practices, a scarcity of third-party computational science software companies, and a simple lack of perseverance by the community. This

indictment is broad and deep, covering applications, programming models and tools, data analysis and visualization tools, and middleware.

Programming Complexity and Ease of Use

Over the past decade, increases in the peak performance of high-end computing systems have been due predominantly to the dramatic growth in single processor performance. Because little research was conducted in next-generation architectures, most of today's high-performance computers are based on cluster designs that interconnect large numbers of COTS computers. As of November 2004, 60 percent of the systems in the TOP500 list (the fastest 500 computers in the world based on the LINPACK linear algebra benchmark) were clusters and 95 percent of the systems used COTS processors.

Although this COTS hardware approach leverages advances in mainstream computing, with accompanying increases in peak performance and declines in financial cost, the human cost remains high. The resulting systems are difficult to program and their achieved performance is a small fraction of the theoretical peak. Today's scientific applications are generally developed with software tools from the last generation – tools that are crude when compared, for example, to those used today in the commercial sector. In some ways, programming has not changed dramatically since the 1970s.

In many environments, Fortran (50 years old) and C (35 years old) are still the main programming languages. Most low-level parallel programming is still based on MPI, a message passing model that requires applications developers to provide deep knowledge of application software behavior and its interaction with the underlying computing hardware, much like programming in assembly language. This, in turn, places a substantial intellectual burden on developers, resulting in continuing limitations on the usability of high-end computing systems and restricting effective access to a small cadre of researchers in these areas. (Sidebar 7 presents one example.)

In many environments, Fortran (50 years old) and C (35 years old) are still the main programming languages.

The problem is even more challenging for emerging areas of computational science, such as biology and the social sciences. In these domains, there is no long history of application development. Rather, researchers seek easy-to-use software that enables analysis of complex data, fusion of disparate models for interdisciplinary analysis, and visualization of complicated interactions.

Commercial desktop software has raised expectations for computational science software usability. The widespread availability of high-quality,

Sidebar 7
High-Performance Fortran (HPF): A Sustainability Lesson

High Performance Fortran (HPF) was an attempt to define a high-level data-parallel programming system based on Fortran. The effort to standardize HPF began in 1991 at the Supercomputing Conference in Albuquerque, where a group of industry leaders asked Ken Kennedy of Rice University to lead an effort to produce a common programming language for the emerging class of distributed-memory parallel computers. The proposed language would be based on some earlier commercial and research systems, including Thinking Machines' CMFortran, Fortran D (a research language defined by groups at Rice, including Kennedy, and Syracuse University, led by Geoffrey Fox), and Vienna Fortran (defined by a European group led by Hans Zima).

The standardization group, called the High Performance Fortran Forum, took a little over a year to produce a language definition that was published in January 1993 as a Rice technical report [Koelbel, et al.,1994].

The HPF project had created a great deal of excitement while it was underway and the release was initially well received in the community. However, over a period of several years, enthusiasm for the language waned in the United States, although it continues to be used in Japan.

Given that HPF embodied a set of reasonable ideas on how to extend an existing language to incorporate data parallelism, why was it not more successful? There were four main reasons: (1) inadequate compiler technology, combined with a lack of patience in the high-performance computing community; (2) insufficient support for important features that would make the language suitable for a broad range of problems; (3) the absence of an open source implementation of the HPF Library; and (4) the complex relationship between program and performance, which made performance problems difficult to identify and eliminate.

Nevertheless, HPF incorporates a number of ideas that will be a part of the next generation of high performance computing languages. In addition, a decade of R&D has overcome many of the implementation impediments. The key lesson from this experience is the importance of sustained long-term investment in technology.

inexpensive desktop software leads users to question the lack of similar computational science software, especially on high-performance systems, and to expect interoperability between desktop tools and those on high-performance systems. But developing robust software tools for a projected computational science market of 500 units is nearly as costly as developing software for the personal computer market – the former simply lacks the financial incentives.

Today, it is altogether too difficult to develop computational science software and applications. Environments and toolkits are inadequate to meet the needs of software developers in addressing increasingly complex, interdisciplinary problems. Legacy software remains a persistent problem

because the lifetime of a computational science application is significantly greater than the three- to five-year lifecycle of a computing system. In addition, since there is no consensus on software engineering best practices, many of the new computational science applications are not robust and cannot be easily extended, integrated, or ported to new hardware. The DARPA High Productivity Computing Systems (HPCS) program [DARPA, 2005] is one of the first efforts, and the only current one, seeking to measure how well our software tools are matched to problem domains. A key goal of this work is to quantify the complexity of scientific software development languages and tools, emphasizing time to solution and total development cost.

We must place a new emphasis on time to solution, the major metric of value to computational scientists.

If computing systems are to be used more widely and more easily, we must place a new emphasis on time to solution, the major metric of value to computational scientists. We must support good software engineering practices in the development of computational science software – through education, additional funding for software-oriented projects, and where appropriate, required software engineering processes for larger, multi-group projects. New programming models and languages and high-level, more expressive tools must hide architectural details and parallelism. To develop new – or even adopt more modern – advanced software will require major investments, and this expense remains a barrier, both practically and psychologically. Solving this problem will require new ideas and a long-range commitment of resources.

Software Scalability and Reliability

The complexity of parallel, networked platforms and highly parallel and distributed systems is rising dramatically. Today's 1,000-processor parallel computing systems will rapidly evolve into the 100,000-processor systems of tomorrow. Hence, perhaps the greatest challenge in computational science today is software that is scalable at all hardware levels (processor, node, and system) . In addition, to achieve the maximum benefit from parallel hardware configurations that require such underlying software, the software must provide enough concurrent operations to exploit multiple hardware levels gracefully and efficiently.

Although parallelism in computation is of the utmost importance, computational science also requires scalability in other system resources. For example, to exploit parallelism in memory architectures, software must arrange communication paths to avoid bottlenecks. Similarly, parallelism in the I/O structure allows the system to hide the long latency of disk reads and writes

and increase effective bandwidth, but only if the software can appropriately batch requests.

In distributed computing, future system software and middleware must be able to scale to hundreds of thousands of processors and enable effective fault tolerance. To achieve these goals, we must consider both network behavior and I/O interfaces that are designed as integral parts of a complete system.

Architecture and Hardware

In the past decade, the Federal government's strategy for technical computing has been predicated on acquiring COTS products. Although this has yielded systems with impressive theoretical peak performance, the fraction of peak that can be sustained for scientific workloads is much lower than that for commercial ones. For commercial workloads, caches – small, high-speed memories attached to the processor – can hold the key data for rapid access. In contrast, many computational science applications have irregular patterns of access to a large percentage of a system's memory. Sidebar 8 shows that capability has actually declined for some critical national applications.

Sidebar 8
Limitations of COTS Architectures

In October 2000, the Defense Science Board issued a report by its Task Force on DoD Supercomputing Needs, which analyzed the capabilities of current computer systems for critical national problems, including national security and signals intelligence analysis [DoD, 2000]. One metric of system capability is billions of updates per second (GUPS), which measures the ability to address large amounts of memory in an irregular way. As the table below shows, today's COTS systems perform more poorly than older, custom-designed high-performance computing systems, notably vector systems with high-bandwidth memory access.

Architecture (Year)	*GUPS (4 GB Memory)*
Cray Y-MP (1988)	0.16
Cray C90 (1991)	0.96
Cray T90 (1995)	3.2
Cray SV1 (1999)	0.7
Cray T3E (1996)	2.2
Symmetric multiprocessors (2000)	0.35-1.00
COTS clusters (2000)	0.35-1.00

The rapid growth of the Internet and commercial computing applications has diverted attention away from industry development of computing components suited to computational science and government needs. The technical computing market is too small to garner much industry interest. High-end computing procurements are estimated at $1 billion per year, compared with a server market of more than $50 billion [Kaufmann, 2003]. To support the demands of scientific workloads, new high-end computing designs are needed – both fully custom high-end designs and more appropriate designs based on commodity components.

Unfortunately, the research pipeline in computer architecture has almost emptied. NSF awards for high-performance computer architecture research have decreased by 75 percent, published papers have decreased by 50 percent, and no funding is available for significant demonstration systems. The human pipeline is also empty. For the U.S. to maintain a leadership role in computational science, we must ensure the involvement and viability of domestic suppliers of components, systems, and expertise. To meet current and future needs, the U.S. government must take primary responsibility for accelerating advances in computer architectures and ensuring that there are multiple strong domestic suppliers of both hardware and software for computational science problems. As noted in Chapter 4, this R&D must be either subsidized by the Federal government or supported by means of stable, long-term procurement contracts.

The Government must launch a next-generation algorithms, software, and hardware program whose goal is to build advanced prototypes of novel computing systems.

The PITAC believes that the Government must launch a next-generation algorithms, software, and hardware program whose goal is to build advanced prototypes of novel computing systems. Much as DARPA funded creation of ARPANet, ILLIAC IV, and other systems in the 1970s, 1980s, and 1990s, these prototyping projects would have lifetimes of sufficient length and budgets of sufficient scope to develop, test, and assess the capabilities of alternative designs. These "expeditions to the 21st century" were recommended in the 1999 PITAC report as a means to create systems better matched to the needs of computational science applications [PITAC, 1999].

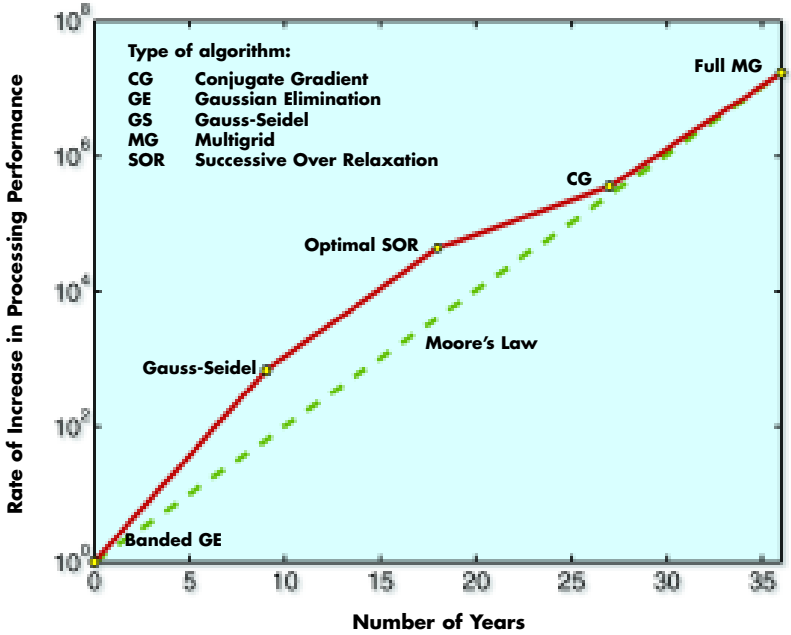
In the 1990s, the Government supported the development of several new parallel computing systems. In retrospect, it is clear that we did not learn the critical lesson of vector computing, namely the need for long-term, sustained, and balanced investment in both hardware and software. We underinvested in

software and expected innovative research approaches to yield robust, mature systems in only two to three years. One need only look at the history of any large-scale software system to recognize the importance of an iterated cycle of development, deployment, and feedback in producing an effective, widely used product. Effective computational science architectures will not be inexpensive. They will require sustained investment, long-term research, and the opportunity to incorporate lessons learned from previous versions.

Scientific and Social Science Algorithms and Applications

Historically, computational science has largely been associated with the physical sciences and engineering. However, with the growth of quantitative biological models and data, biomedicine and biology have emerged as beneficiaries of but also dependent on new computational science algorithms, tools, and techniques. Equally important, the social sciences and humanities are now major consumers of computing technology, with a set of data-rich problems distinctly different from those found in the physical sciences. All domains would benefit from improved numerical and non-numerical

Improvements in Algorithms Relative to Moore's Law



The relative gains in some algorithms for the solution of an electrostatic potential equation on a uniform cubic grid compared to improvements in the hardware (Moore's Law).

Figure 5

algorithms, data management and mining technologies, and easier-to-use software suites. (Appendix A cites examples of such problems.)

Scientific Algorithms and Applications

Although dramatic increases in processor performance are well known, improved algorithms and libraries have contributed as much to increases in computational simulation capability as have improvements in hardware. Figure 5 on page 53 shows the performance gained from improved algorithms for solving linear systems arising from the discretization of partial differential equations. These gains either track or exceed those from hardware performance improvements from Moore's Law.

Computational science applications software must continually be infused with the latest algorithmic advances. In turn, these applications must actively drive research in algorithms. This interplay was highlighted by the 2003 activities of the HECRTF, which solicited input from leading scientists in a variety of physical science and engineering disciplines [CRA, 2003]. The scientists were asked to identify the important computational capabilities needed to achieve their research goals. They said that it will take a combination of new theory, new design tools, and high-end computing for large-scale simulation to achieve fundamental understanding of the emergence of new behaviors and processes in nanomaterials, nanostructures, nanodevices, and nanosystems. Similarly, it will take ensembles of ultra-high-resolution simulations on high-end systems to improve our ability to provide accurate projections of regional climate. The scientists also pointed out that the intelligence community's ability to safeguard the Nation hinges to a substantial degree on high-end computing capabilities with diverse specialized computational applications.

Improved algorithms and libraries have contributed as much to increases in capability as have improvements in hardware.

Social Science Applications

To date, relatively few computational efforts have focused on the social dynamics and organizational, policy, management, and administration decision making in the purview of the social sciences and their application to solving complex societal problems. However, expanding methods for collecting and analyzing data have enabled the social and behavioral sciences to record more and more information about human social interactions, individual psychology, and human biology. Rich data sources include national censuses, map-making, psychophysical comparison, survey research, field archaeology, national income accounts, audio and video recording, functional magnetic resonance imaging

(fMRI) , genetic sampling, and geographic information systems. Now, using analytical techniques in computation, including statistical methods, spatial analysis, archaeometry, content analysis, linguistic annotation, and genetic analysis, researchers can work with the data to understand the complex interactions of psychology and biology.

A recent NSF workshop [NSF, 2005a] noted that continued advances in social and behavioral science methods and computational infrastructure will make it possible to:

- Develop data-intensive models sophisticated enough to accurately model lifetime decision-making by individuals with respect to such matters as work, marriage, children, savings, and retirement
- Code the verbal and non-verbal cues in large numbers of videotaped physician-patient interactions and analyze their relationship to the resulting medical diagnoses
- Perceive changes in metropolitan areas by coding and analyzing land-use, environmental, social-interaction, institutional, and other data over time
- Map the sequence of biochemical interactions through which the human brain makes decisions by analyzing MRI data for many individuals
- Develop and analyze databases of tens of thousands of legislative votes, speeches, and actions to better understand the functioning of government
- Understand the development and functioning of social networks on the Web by modeling key usage characteristics over time
- Develop better institutional and technical methods to reduce malevolent behavior on the Web by understanding not only the Web's technical vulnerabilities but also the realistic and feasible threats from human agents

Developing the algorithms and applications that can provide these capabilities, as well as establishing the necessary infrastructure, will require ongoing collaborations among social and computer scientists and engineers.

Software Integration

Too often, researchers spend much more time coupling disparate application programs and software systems than they do conducting research. The limited interoperability of the tools and their complexity have become major hindrances to further progress. Sources of this complexity include the number of equations and variables required to encapsulate realistic function, the size of

the resulting systems and data sets, and the diverse range of computational resources required to support major advances [Bramley, et al. , 2000].

Today, a typical computational researcher must use software, libraries, databases, and data analysis systems from a variety of sources. Most of these tools are incompatible, most likely written in different computer languages, for different operating systems, using different file formats. The need to integrate algorithms and application software is especially acute when researchers seek to create models that span spatial or temporal scales or cross physical systems.

No single researcher has the skills required to master all the computational and application domain knowledge needed to gather data from databases or experimental devices, create geometric and mathematical models, create new algorithms, implement the algorithms efficiently on modern computers, and visualize and analyze the results. To model such complex systems faithfully requires a multidisciplinary team of specialists, each with complementary expertise and an appreciation of the interdisciplinary aspects of the system, and each supported by a software infrastructure that can leverage specific expertise from multiple domains and integrate the results into a complete application software system.

We must continue to develop and improve the mathematical, non-numeric, and computer science algorithms that are essential to the success of future computational science applications. Computational researchers also need enabling, scalable, interoperable application software to conduct computational examinations of their ideas and data. To be successful, application software must provide infrastructure for vertical integration of computational knowledge, including knowledge of the relevant discipline(s); the best computational techniques, algorithms, and data structures; associated programming techniques; user interface and human-computer interface design principles; applicable visualization and imaging techniques; and methods for mapping the computations to various computer architectures.

Data Management

Today, most data and documents are born digital, rather than being converted from analog sources. Multi-megapixel images are now commonplace, whether from consumer cameras or instrument detectors, and our collective store of digital data is expanding at an estimated rate of 30 percent per year [Lyman, 2003]. Examples of this explosive data growth abound. In 2007, the new ATLAS and CMS detectors for the Large Hadron Collider (LHC) will produce tens of petabytes of raw and processed detector data each year. In the biomedical domain, brain data captured with high-

resolution instruments can easily exceed several petabytes. The social sciences are experiencing a similar data explosion.

These enormous repositories of digital information require a new generation of more powerful analysis tools. What was appropriate for a modest volume of manually collected data is wholly inadequate for a multiple-petabyte archive. Large-scale data sets cannot be analyzed and understood in a reasonable time without computational models, data and text mining, visualizations, and other knowledge discovery tools. Moreover, extraction of knowledge across heterogeneous or federated sources requires contextual knowledge, typically provided through metadata. For example, knowledge to be derived from data captured through an instrument requires some knowledge of the instrument's characteristics, the conditions in which it was used, and the calibration record of the instrument. Metadata are necessary to determine the accuracy and provenance (heredity) of the individual datasets as well as the validity of combining data across sets.

Large-scale data sets cannot be effectively analyzed without computational models, visualizations, and other knowledge discovery tools.

Computational science researchers often gather multichannel, multimodal, and sensor data from real-time collection instruments, access large distributed databases, and rely on sophisticated simulation and visualization systems for exploring large-scale, complex, multidimensional systems. Managing such large-scale computations requires powerful, sometimes distributed, computing resources and efficient, scalable, and transparent software that frees the user to engage the complexity of the problem rather than of the tools themselves. Such computational application software does not currently exist.

Data-intensive computational science, based on the emergence of ubiquitous sensors and high-resolution detectors, is a new opportunity to couple observation-driven computation and analysis, particularly in response to transient phenomena (e.g., earthquakes or unexpected stellar events). Moreover, the explosive growth in the resolution of sensors and scientific instruments – a consequence of increased computing capability – is creating unprecedented volumes of experimental data. Such devices will soon routinely produce petabytes of data.

A consequence of the explosive growth of experimental data is the need to increase investment and focus on sensor- and data-intensive computational science. We must act now to develop the requisite data-mining, visualization, and information-extraction tools to gain knowledge from these data collections.

Conclusion

Unlike the space race that captured the national imagination nearly five decades ago, our diminishing leadership role in computational science is a quiet crisis. While computational science is the key field contributing to rapid advances in the physical and social sciences and in industry, its largely behind-the-scenes role is unknown to the millions of citizens who regularly enjoy its benefits through improvements to our national security, energy management and usage, weather forecasting, transportation infrastructure, health care, product safety, financial systems, and in countless other ways large and small. But the near-invisibility of computational science does not signify its lack of importance – merely our own lack of understanding.

Although the PITAC did not plan the convergence, the same themes emerged in its two previous studies, *Cyber Security: A Crisis of Prioritization* and *Revolutionizing Health Care Through Information Technology*. The diverse technical skills and technologies underlying software, computing systems, and networks themselves constitute a critical U.S. infrastructure that we underappreciate and undervalue at our peril. Computational science is a foundation of that infrastructure.

Given all that depends on the field's vitality, it is imperative that the leaders in academia and the Federal government who are responsible for assuring the continued health of computational science spearhead the design and implementation of new multidisciplinary research and education structures that will assure the United States the advanced capabilities to address the 21st century's most important problems. In addition, the Federal government, in partnership with academia and industry, must commission – and execute – a multi-decade computational science roadmap that will direct coordinated advances in computational science and its underlying technologies, paving the way to greater breakthroughs in the many disciplines that will require these capabilities in the years ahead.

By following the computational science roadmap and moving decisively forward to build a sustained software/data/high-end computing infrastructure and support R&D investments in new generations of well engineered and easy-to-use software for scalable and reliable hardware architectures, the Federal government – together with its partners – can help elevate computational science to the status it has already earned as a strategic, long-term national priority.

References

Bramley, R., B. Char, D. Gannon, T. Hewett, C. Johnson, and J. Rice, "Enabling Technologies for Computational Science: Frameworks, Middleware and Environments." In *Proceedings of the Workshop on Scientific Knowledge, Information, and Computing,* E. Houstis, J. Rice, E. Gallopoulos, and R. Bramley, eds., Kluwer Academic Publishers, Boston, 2000, pp. 19-32.

Computing Research Association (CRA), *Workshop on the Roadmap for the Revitalization of High-End Computing,* 2003.

Council on Competitiveness, Final Report from the High Performance Users Conference: *Supercharging U. S. Innovation & Competitiveness,* Washington, D.C., July 2004.

Davies, C. T. H., et al., "High-Precision Lattice QCD Confronts Experiment," *Physical Review Letters*, 92, 022001, January 16, 2004.

Defense Advanced Research Projects Agency (DARPA), High Productivity Computing Systems (HPCS) Program, http://www.high productivity.org, 2005.

Department of Defense (DoD), *Report on High Performance Computing for the National Security Community*, http://WWW.HPCMO.HPC.mil/Htdocs/DOCUMENTS/04172003_hpc_report_unclass.pdf, July 2002.

Department of Defense, Defense Science Board, *Report of the Defense Science Board Task Force on DoD Supercomputing Needs*, http://www.acq.osd.mil/dsb/reports/dodsupercomp.pdf, October 2000.

Department of Energy (DOE), Office of Science, *A Science-based Case for Large-Scale Simulation*, Vol. 2, http://www.pnl.gov/scales, September 2004.

Department of Energy, Office of Science, *A Science-based Case for Large-Scale Simulation*, Vol. 1, http://www.pnl.gov/scales, July 2003.

Department of Energy, Office of Science, "Scientific Discovery through Advanced Computing," March 2000.

Executive Office of the President, Office of Science and Technology Policy, *Federal Plan for High-End Computing: Report of the High-End Computing Revitalization Task Force (HECRTF)*, May 2004.

Hager, Thomas, *Force of Nature: The Life of Linus Pauling,* Simon & Schuster, N.Y., 1995.

Interagency Panel on Large Scale Computing in Science and Engineering, sponsored by the Department of Defense and the National Science Foundation, in cooperation with the Department of Energy and the National Aeronautics and

Space Administration, *Report of the Panel on Large Scale Computing in Science and Engineering*, 1982.

International Technology Roadmap for Semiconductors (ITRS) , http://public.itrs.net, 2005.

Joseph, E., A. Snell, and C. G. Willard, Council on Competitiveness Study of U.S. HPC Users, White Paper, http://www.compete.org/pdf/HPC_Users_Survey.pdf, July 2004.

Kaufmann, N. J., C. G. Willard, E. Joseph, and D. S. Goldfarb, "Worldwide High Performance Systems Technical Computing Census," IDC Report No. 62303, June 2003.

Koelbel, C. H., D. B. Loveman, and R. S. Schreiber, *The High Performance Fortran Handbook*, MIT Press, Cambridge, Mass. ,1994.

Large Synoptic Survey Telescope (LSST) , http://www.lsst.org, 2005.

Lyman, P. and H. R. Varian, "How Much Information," http://www.sims.berkeley.edu/how-much-info-2003, 2003.

Marburger, John H. III, "Statement on the Fiscal Year 2006 Federal R&D Budget," Committee on Science, United States House of Representatives, February 2005.

National Institutes of Health (NIH) , National Institutes of Health Roadmap: Accelerating Medical Discovery to Improve Health, http://nihroadmap.nih.gov, 2004.

National Institutes of Health, Working Group on Biomedical Computing Advisory Committee to the Director, "The Biomedical Information Science and Technology Initiative," June 1999.

National Research Council (NRC) , Computer Science and Telecommunications Board, *Getting Up To Speed: The Future of Supercomputing*, National Academies Press, Washington, D.C., 2005.

National Research Council, Computer Science and Telecommunications Board, *Embedded Everywhere: A Research Agenda for Networked Systems of Embedded Computers*, National Academy Press, Washington, D.C. , 2001.

National Research Council, Computer Science and Telecommunications Board, *Making IT Better: Expanding Information Technology Research to Meet Society's Needs*, National Academy Press, Washington, D.C. , 2000.

National Science Board (NSB) , *Science and Engineering Infrastructure Report for the 21st Century: The Role of the National Science Foundation*, February 2003.

National Science Foundation (NSF) , *Final Report: NSF SBE-CISE Workshop on Cyberinfrastructure and the Social Sciences*, http://vis.sdsc.edu/sbe, May 2005a.

National Science Foundation, National Middleware Initiative (NMI) , http://www.nsf-middleware.org, 2005b.

National Science Foundation, *Science and Engineering Indicators 2004*, May 2004a.

National Science Foundation, Science Resources Statistics, *InfoBrief,* June 2004b.

National Science Foundation, *Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure*, January 2003.

National Science Foundation, *Report of the Task Force on the Future of the NSF Supercomputer Centers Program*, September 1995.

National Science Foundation, NSF Blue Ribbon Panel on High Performance Computing, *From Desktop to Teraflop: Exploiting the U.S. Lead in High Performance Computing*, August 1993.

President's Council of Advisors on Science and Technology, *Report to the President: Sustaining the Nation's Innovation Ecosystems, Information Technology Manufacturing and Competitiveness*, January 2004.

President's Information Technology Advisory Committee (PITAC) , *Report to the President: Information Technology Research: Investing in Our Future*, February 1999.

Society for Industrial and Applied Mathematics (SIAM) Working Group on Computational Science and Engineering (CSE) Education, "Graduate Education in Computational Science and Engineering," *SIAM Review*, Vol. 43, No. 1, pp. 163-177, 2001.

Simmons, M., A. Hayes, J. Brown, and D. Reed, eds., *Debugging and Performance Tuning for Parallel Computing Systems*, Wiley-IEEE Computer Society Press, July 1996.

University of Pittsburgh, *Knowledge Lost in Information: Report of the NSF Workshop on Research and Directions for Digital Libraries*, June 2003.

U.S. Census Bureau Foreign Trade Statistics, U.S. International Trade in Goods and Services, 2003.

Workshop and Conference on Grand Challenges Applications and Software Technology, Pittsburgh, Pennsylvania, May 1993.

APPENDIX A

Examples of Computational Science at Work

Computational science enables important discoveries across the entire range of social and physical sciences. It serves, for example, as the basis for design optimizations in engineering and manufacturing and provides tools for understanding biological processes and biomedical solutions. The vignettes below, though by no means exhaustive, illustrate the breadth of computational science applications as well as the opportunities that the Nation can realize by providing broader support.

SOCIAL SCIENCES

Monitoring the U.S. Economy

Though invisible to most citizens, computational science plays a central day-to-day role in the deliberations and decisions of the Federal Reserve Bank's Board of Governors, the group of top regional Reserve Bank officers – currently chaired by Alan Greenspan – whose task is to guide U.S. monetary policy. Wielding substantial influence over the direction of the economy, the Federal Reserve Board was an early adopter of computational science techniques and has used macroeconomic modeling and simulation for more than three decades to analyze national and international economic processes and evaluate the possible impacts of shifts in monetary policy.

With advances in macroeconomic theory, the mathematics underlying computational economics, the power of computing systems, and mass storage capacity enabling preservation and use of large quantities of historic data, the Board's first-generation computer models eventually became outmoded despite constant incremental improvements. In the mid-1990s, Federal Reserve researchers unveiled a new set of models that incorporate significant dynamic attributes that were not possible in the older models – in particular, adaptive specifications for the role of expectations in economic activity and dynamic adjustments to equilibrium conditions. The new U.S. model, FRB/US, and a second version called FRB/WORLD – which links FRB/US to an international model of 11 other countries and regions – together contain 250 behavioral equations. Forty of the equations describe the U.S. economy. The large size and disaggregation of the models enable researchers to execute a wide range of types of simulations and provide estimates of outcomes for a large set of variables.

With FRB/US, for example, the Board's staff can gauge the likely consequences of specific events through computational "what-if" exercises. By

setting the model's equations to represent alternative assumptions about such variables as fiscal policy, business output, cost of capital, household income, energy prices, and interest rates, researchers can run simulations that forecast outcomes over time of the interactions among the variables, and they can examine the impacts of economic shocks such as a sudden stock market drop or a sharp rise in inflation. In the same way, the model can be used to predict the likely implications for economic performance of a given change in monetary policy. In one frequently cited study using FRB/US, Federal Reserve researchers examined the problems that could result from a monetary policy setting a lower boundary of zero on nominal interest rates, and they proposed a policy modification that would prevent economic instabilities in such a low-interest-rate climate.

For more information, see:
http://www.federalreserve.gov/pubs/feds/1997/199729/199729abs.html and *http://ideas.repec.org/p/sce/scecf9/843.html.*

Cyberinfrastructure and the Social Sciences

Cyberinfrastructure is defined as the coordinated aggregate of software, hardware, and other information technologies, as well as the human expertise, required to support current and future discoveries in science and engineering. Less explored, however, is the potential impact of the cyberinfrastructure in disciplines such as the humanities and the social sciences.

In a recent NSF-supported workshop on "Cyberinfrastructure and the Social Sciences," participants reached several important conclusions that could lead to more robust cooperation and collaboration between computational scientists and social scientists. Particularly striking is the potential for social scientists to collaborate with computational scientists to collect better data through experiments and simulations on the Internet. Social scientists could also conduct experiments of unprecedented scale and intensity using distributed networks and powerful tools. Such collaboration would prove highly beneficial today, as social and behavioral scientists face the possibility of becoming overwhelmed by the massive amount of data available and the challenges of comprehending and safeguarding it.

In turn, social scientists could assist computational scientists in achieving a better understanding how computational science exists in the social ecosystem. Organizational researchers and political scientists can help develop appropriate management, decision-making, and governance structures for Web-enabled research communities and the cyberinfrastructure providers that support them, while behavioral scientists can help develop better modes of human-computer interaction. Sociologists can analyze the implications for knowledge

production of social networks developed on the Web. Psychologists and linguists can collaborate with computer scientists to develop computer programs that readily understand, employ, and translate natural languages.

By increasing their understanding of large-scale social changes, social science and computational science researchers can significantly assist the Nation in maximizing the societal benefits from the evolving cyberinfrastructure.

For more information, see: *http://vis.sdsc.edu/sbe/reports/SBE-CISE-FINAL.pdf.*

Agent-based Computational Economics

Agent-based computational economics (ACE) is the computational study of economies modeled as dynamic systems of interacting agents. Here "agent" refers broadly to a bundle of data and behavioral methods representing an entity in a computationally constructed world. Agents can include individuals (such as consumers and producers), social groupings (families, firms, communities, government agencies), institutions (markets, regulatory systems), biological entities (crops, livestock, forests), and physical entities (infrastructure, weather, and geographical regions). Thus, agents can range from active data-gathering decision makers with sophisticated learning capabilities to passive world features with no cognitive function. Moreover, agents can be composed of other agents, permitting hierarchical constructions.

Current ACE research divides roughly into four strands differentiated by objective. One primary objective is empirical understanding. Why have particular macro regularities evolved and persisted, despite the absence of top-down planning and control? Examples of such regularities include trade networks, socially accepted monies, market protocols, business cycles, and the common adoption of technological innovations. ACE researchers seek causal explanations grounded in the repeated interactions of agents operating in realistically rendered worlds.

A second primary objective is normative understanding. How can agent-based models be used as laboratories for the discovery of good economic designs? ACE researchers pursuing this objective are interested in evaluating whether designs proposed for economic policies, institutions, or processes will result in socially desirable system performance over time. A third primary objective is qualitative insight and theory generation: How can the full potentiality of economic systems be better understood? A final object is methodological advancement: How can ACE researchers best be provided with the methods and tools they need to undertake the rigorous study of economic systems through controlled computational experiments?

Researchers with the non-profit Electrical Power Research Institute, for example, developed an elaborate model of what they termed the U.S. "electric enterprise." The model simulates the evolution of the power industry using autonomous adaptive agents to represent both the possible industrial components and the corporate entities that own these components. The model includes an open-access transmission application and real-time pricing. The goals of the effort were to provide high-fidelity simulations offering insight into the operation of the deregulated power industry; suggest how intelligent software agents might be used in the management of complex distributed systems and for transactions in the electric marketplace; and illuminate how such agents might contribute to a self-optimizing and self-healing electric power grid.

For more information, see: *http://www.econ.iastate.edu/tesfatsi/ace.htm* and *http://www.econ.iastate.edu/tesfatsi/SEPIA.EPRI.pdf.*

Political and Social Science Archives

The growing interdependence of society's most challenging economic, political, and technical issues makes social science data and methodologies increasingly significant in the public policy arena. But in the debates surrounding policy decision making, the validity of data can itself become an issue. Within the social science community, this problem is well recognized and it is addressed by organizations such as the Inter-university Consortium for Political and Social Research (ICPSR) . Established in 1962, ICPSR maintains and provides access to a vast archive of original-source social science data for research and instruction and offers training in quantitative methods to facilitate effective data use. A unit within the Institute for Social Research at the University of Michigan, ICPSR is a membership-based organization with more than 500 member colleges and universities around the world.

The ICPSR data holdings contain some 6,000 studies and 450,000 files covering a wide range of social science areas such as population, economics, education, health, aging, social and political behavior, social and political attitudes, history, crime, and substance abuse. While the archive includes several time series and other types of aggregate data, most holdings consist of raw data derived from surveys, censuses, and administrative records. The data security and preservation unit of ICPSR is charged with ensuring that ICPSR data are secure at all times and not vulnerable to intrusion or violation. It also protects and preserves ICPSR's data resources by securing back-up copies of data and documentation that are stored off-site and migrating them to new storage media as changes in technology warrant.

For more information, see: *http://www.icpsr.umich.edu/*.

PHYSICAL SCIENCES

Quantum Chromodynamics: Predicting Particle Masses

High-energy physicists have arrived at a picture of the microscopic physical universe called "The Standard Model," which unifies the nuclear, electromagnetic, and weak forces and enumerates the fundamental building blocks of the universe, quarks and leptons. However, the model has serious flaws – it does not account for gravity, does not explain or predict the masses of the various particles, and requires a number of parameters to be measured and inserted into the theory.

Quantum chromodynamics (QCD) is the theory of how the nuclear force binds quarks together to form a class of particles call hadrons (that include protons and neutrons). For 30 years, researchers in lattice QCD have been trying to use the basic QCD equations to calculate the properties of hadrons, especially their masses, using numerical lattice gauge theory calculations in order to verify the standard model. Unfortunately, limited by the speed of available computers, they have had to simplify their simulations to get results in a reasonable amount of time, and those results typically have had an error rate of around 15 percent when compared with experimental data.

Now, with significantly faster computers, improved algorithms that employ fewer simplifications of physical processes, and better-performing codes, four QCD collaborations involving 26 researchers have reported calculations of nine different hadron masses, covering the entire range of the hadron spectrum, with an error rate of 3 percent or less. This work [Davies et al., 2004] marks the first time that lattice QCD calculations have achieved results of this precision for such diverse physical quantities using the same QCD parameters.

QCD theory and computation are now poised to fulfill their role as equal partners with experiment. A significant fraction of the $750 million per year that the United States spends on experimental high-energy physics is devoted to the study of the weak decays of strongly interacting particles. To capitalize fully on this investment, the lattice calculations must keep pace with the experimental measurements.

For more information, see: *http://www.usqcd.org*.

High-Temperature Superconductor Models

Experimental high-temperature superconductors (HTSC), such as cuprate superconductors, can transport electrical current without significant resistance at unusually high temperatures. The perfection and deployment of such novel

ceramic materials could have a significant economic impact, allowing, for example, a few superconducting cables to channel electricity to entire cities or enabling a new generation of powerful, light-weight motors.

Despite years of active research, however, understanding superconductivity in cuprate HTSC remains one of the most important unsolved problems in materials science. In the superconducting state of a material, electrons pair to form so-called Cooper-pairs, allowing them to condense into a coherent macroscopic quantum state in which they conduct electricity without resistance. Although conventional superconductors are well understood, the pairing mechanism in HTSC is of an entirely different nature. Models describing itinerant correlated electrons – in particular, the two-dimensional Hubbard model – are believed to capture the essential physics of the copper dioxide (CuO_2) planes of HTSC. But despite intensive studies, this model remains unsolved.

A recent concurrence of new algorithmic developments and significant improvements in computational capability has enabled massively parallel computations for the two-dimensional Hubbard model and opened a clear path to solving the quantum many-body problem for HTSC. The solution of this model in the thermodynamic limit requires an approximation scheme. Simulations of small, four-atom clusters have shown that the model reproduces the antiferromagnetic and superconducting phases as well as the exotic normal-state behavior observed in the cuprates. However, the scale of the computation increases dramatically with larger cluster sizes, necessitating high-performance computing resources.

For more information, see:
http://nccs.gov/DOE/mics2004/Cuprates.Maier.doc.

Fusion Plasmas and Energy Sources

Our ever-increasing dependence on foreign petroleum resources has sparked renewed interest in fusion as a long-term energy source. ITER (Latin word for "the way") , the proposed international fusion testbed, is being designed to test new ideas and serve as a precursor to realistic designs. Central to eventual success is developing an infrastructure that can contain a stable plasma at temperatures high enough to sustain nuclear fusion. But determining what is happening inside a fusion plasma is very difficult experimentally. A conventional probe inserted into the hot plasma is likely to sputter and contaminate the plasma, leading to a loss of heat. Experimentalists must use non-perturbative diagnostics – such as laser scattering, and measurements with probes and magnetic loops around the edge of the plasma – to deduce the plasma conditions and the magnetic field structures inside the plasma.

An important aid to the experiments is work undertaken with computational scientists to create detailed simulations of fusion plasmas. Researchers at Lawrence Livermore National Laboratory, in collaboration with others at the University of Wisconsin-Madison, have developed simulations using the NIMROD code on the National Energy Research Scientific Computing Center's (NERSC's) supercomputer that accurately reproduce experimental results. With recent changes to their code, the collaborators have created simulations with temperature histories – measured in milliseconds – that are closer to the temperature histories observed in experiments. This follows the group's prior success in simulating the magnetics of experiments.

Although the simulations cover only four milliseconds in physical time, they involve more than 100,000 time steps. As a result, the group ran each of the simulations in 50 to 80 shifts of 10 to 12 hours each, consuming more than 30,000 processor hours in each complete simulation, and multiple simulations were needed.

For more information, see:
http://www.nersc.gov/news/nerscnews/NERSCNews_2004_12.pdf.

Designing Compact Particle Accelerators

For a quarter of a century, physicists have been trying to push charged particles to high energies with devices called laser wake field accelerators. In theory, particles accelerated by the electric fields of laser-driven waves of plasma could reach, in fewer than 100 meters, the high energies attained by miles long machines using conventional radiofrequency acceleration. Stanford University's linear accelerator, for example, is two miles long and can accelerate electrons to 50 GeV (50 billion electron volts) . Laser wake field technology offers the possibility of a compact, high-energy accelerator for probing the subatomic world, for studying new materials and technologies, and for medical applications.

Researchers at Lawrence Berkeley National Laboratory have taken a giant step toward realizing the promise of laser wake field acceleration by guiding and controlling extremely intense laser beams over greater distances than ever before to produce high-quality, energetic electron beams. By tailoring the plasma channel conditions and laser parameters, researchers are first able to achieve clean guiding of laser beams of unprecedented high intensity while suppressing electron capture. This paves the way for using laser-powered plasma channels as ultra-high-gradient accelerating structures. Next, by using even higher peak powers, plasma waves are excited that are capable of picking up background plasma electrons, rapidly accelerating them in the wake's electric field, then finally subsiding just as the surfing electrons reach the dephasing length, when they are on the verge of outrunning the wake.

These experimental results were validated using the VORPAL plasma simulation code at NERSC. The model allowed the researchers to see the details of the experiment's evolution, including the laser pulse breakup and the injection of particles into the laser plasma accelerator, a prerequisite for optimizing the process.

For more information, see: *http://www.nersc.gov/news/nerscnews/NERSCNews_2004_10.pdf*.

Discovering Brown Dwarves via Data Mining

An innovative approach to finding undiscovered objects buried in immense astronomical databases has produced an early and unexpected payoff: the discovery of a new occurrence of a hard-to-find star known as a brown dwarf. Scientists creating the National Virtual Observatory (NVO) , an online portal for astronomical research unifying dozens of large astronomical databases, confirmed the existence of the new brown dwarf in 2003. The star emerged from a computerized search of information on millions of astronomical objects in two separate astronomical databases.

The new discovery came from one of three scientific prototypes that NVO scientists presented at the January 2003 meeting of the American Astronomical Society. NVO partners at the California Institute of Technology's Infrared Processing and Analysis Center (IPAC) implemented the software for the prototype that found the new brown dwarf.

A search for this type of celestial object formerly required weeks or months of close human attention. But the new NVO-based search discovered the star in approximately two minutes. NVO researchers emphasized that a single new brown dwarf, added to a list of approximately 200 known brown dwarves, is not as scientifically significant as the rapidity of the new discovery and the tantalizing hint it offers for the potential of NVO.

The new star's discovery was unexpected. Researchers had simply hoped to demonstrate the software's feasibility and to confirm existing science, not make new findings. But the very first time the NVO devices were powered up, they immediately yielded the new discovery from data that had been publicly available for at least 18 months. That is precisely the type of result scientists hope will begin to cascade from the NVO in a few more years: revelations hidden in data already gathered by observatories, probes, and surveys that remain undiscovered because new technology is pouring fresh data so rapidly into a variety of different databases.

For more information see: *http://www.us-vo.org*.

Dark Matter, Dark Energy, and the Structure of the Universe

About five years ago, cosmologists discovered that the universe is expanding at an accelerating pace. This finding was contrary to the behavior of matter in Einstein's well-tested theory of general relativity, which predicted that the universe's expansion would slow with time. The finding forced cosmologists to contemplate the possibility that, besides dark matter, the universe also contains "dark energy" that experiences gravity as a repulsing force and thus speeds expansion. The cosmological constant is one type of dark energy model, originally considered by Einstein, in which the cosmic repulsion is built into the fabric of space-time.

A team at the University of Illinois has conducted large-scale cosmological computational simulations that show the distribution of cold dark matter in a model of cosmic structure formation incorporating the effects of a cosmological constant (Lambda) on the expansion of the universe. The simulation contained 17 million dark matter particles in a cubic model universe that is 300 million light-years on a side. It relied on an expanded version of the adaptive mesh refinement (AMR) code FLASH, developed by a team of researchers at the ASCI Center for Astrophysical Thermonuclear Flashes at the University of Chicago. Though FLASH was originally intended to simulate supernova explosions, the Illinois team led an effort to enhance it with self-gravity, expansion, and the ability to track particles. These modifications have extended FLASH's capabilities to cosmological simulation.

For additional information, see:
http://www.ncsa.uiuc.edu/News/Access/Stories/LambdaCDM.

Supernova Modeling

Four hundred years after Galileo's observation of the massive exploding star now known as SN1604, the mechanism for explosions of core collapse supernovae (stars at least 10 times as massive as our sun) remains unknown. Today, scientists in many disciplines are working with computational scientists to perform one-, two-, and three-dimensional simulations that may lead to a greater understanding of this phenomenon, adding to our understanding of the nature of the universe.

Over the past decade, the development of multidimensional supernova models has allowed scientists to explore the roles that convection, rotation, and magnetic fields might have in the occurrence of supernovas. Important research in this area is currently being conducted under the TeraScale Supernova Initiative (TSI), a national, multi-institution, multidisciplinary collaboration of astrophysicists, nuclear physicists, applied mathematicians,

and computer scientists. TSI currently involves 34 U.S. researchers from 11 institutions and a total of 89 researchers from 28 institutions worldwide.

TSI's principal goals are to understand the mechanism(s) responsible for the explosions of core collapse supernovae and all the phenomena associated with these stellar explosions. Such associated phenomena include a supernova's contribution to the synthesis of the chemical elements in the Periodic Table; the emission of an unfathomable flux of nearly massless, radiation-like particles known as neutrinos; the emission of gravitational waves (ripples in space predicted by Einstein's theory of gravity); and in some cases the emission of intense bursts of gamma radiation.

For additional information, see: *http://www.phy.ornl.gov/tsi*.

NATIONAL SECURITY

Signals Intelligence

While human intelligence (HUMINT) and signals intelligence (SIGINT) capabilities are both acknowledged pillars of the Nation's overall intelligence effort, the technological problems involved in collecting and processing data in the latter arena have consistently proved daunting. Even before 9/11, the demand for significant computational power by DoD, intelligence community agencies, and related organizations was difficult to address. But after the 2001 attacks, this demand grew substantially. To enhance the security of the United States and its allies, including anticipating the actions of terrorists and rogue states, R&D in supercomputing and advanced computational science has assumed a pivotal role in the intelligence community as we attempt to stay at least one step ahead of our enemies.

SIGINT takes aim at the capabilities and electronic communications of hostile foreign powers, organizations, or individuals. Like HUMINT, this intelligence also can play a part in counterintelligence, helping buttress the Nation's active defense against rogue nations, terrorists, or criminal elements.

The area of SIGINT processing employs supercomputing and parallel computing technologies to transform a veritable worldwide tsunami of intercepted communications signals of varying quality into useful, actionable information on our adversaries' intentions. The process of intercepting, sifting, analyzing, and storing this almost incomprehensible amount of data, however, is overwhelming, involving technical challenges such as overcoming an adversary's sophisticated cryptographic systems or rapidly reconstructing messages when confronted with incomplete or corrupted data in a foreign alphabet or language.

The key computational elements involved in solving signals intelligence problems differ considerably from those used in other types of scientific problems. In addition, the massive scale of the intelligence community's knowledge discovery effort, particularly at the National Security Agency, is significantly larger than that of the most substantial commercial "data mining" operations. The requirement for continual advances in computational science capabilities for SIGINT makes computational science R&D a high priority for the intelligence community's role in the war against terrorism.

For more information see: *http://www.nsa.gov/sigint/*.

Modeling Real-Time Complex Systems in the Human Environment

Modeling and simulation techniques are increasingly being applied to complex, large-scale systems that have an impact on people or are affected by people in real time. The ability to simulate, for example, the spread of a disease epidemic over time or the daily traffic patterns across a metropolitan transportation system is providing public health officials and emergency-response coordinators with a powerful new planning tool that provides visual representations of the interactions of complex data. Seeing the "big picture" of what might transpire during a crisis helps planners anticipate and address issues in advance, such as which hospitals and how many hospital beds would be needed at what points during the spread of an epidemic.

Because wildfires are a series of small, intense physical phenomena affected by terrain and atmospheric conditions, their spread could not be reliably predicted before the availability of supercomputers and high-resolution modeling techniques. Ecologists and fire behavior specialists at Los Alamos National Laboratory (LANL) have developed a real-time wildfire modeling application to assist in fighting wildfires as they occur. The forested areas of northern New Mexico are prone to catastrophic wildfires, particularly in recent years as a regional drought continues. In 2000, the 43,000-acre Cerro Grande Fire burned a significant fraction of LANL's lands as well as the nearby town. The cost in physical damage and lost work time approached $1 billion. To assist in preventing such catastrophic losses from future fires, laboratory scientists have adapted topographic, vegetation, and weather data layers to work with the Fire Area Simulator (FARSITE) model to predict fire behavior on a real-time basis during a wildfire emergency and to develop fire-fighting plans.

For more information, see:
http://www.lanl.gov/news/index.php?fuseaction=home.story&story_id=2032,
http://www.esh.lanl.gov/~esh20/projects.shtml, and
http://www.esh.lanl.gov/~esh20/pdfs/Cerro_Bx_Narr.pdf.

Dynamic Modeling of the Spread of Infectious Disease

The impact of infectious diseases in humans and animals is enormous, in terms of both suffering and social and economic consequences. Studying the spread of diseases, in both space and time, provides a better understanding of transmission mechanisms and those features most influential in their spread, allows predictions to be made, and helps determine and evaluate control strategies. The emergence of diseases such as Lyme disease, HIV/AIDS, hantavirus, West Nile virus, SARS, and the newest avian flu has raised the stature and visibility of epidemiological modeling as a vital tool in public health planning and policy making.

In recent years, epidemiologists have developed agent-based computational models for simulating the spread of infectious disease through a population. These models are based on understanding the details of disease transmission as well as the dynamics of the community, using mathematics and computational science to integrate this knowledge in simulation programs. Such programs can provide scenarios to help planners envision the results of such strategies as vaccination and quarantine in the face of a pandemic.

Modeling software has progressed to the point that it must be deployed on high-performance computers to achieve useful sensitivity analysis and parameter definition, explore various intervention strategies to alter the course of pandemic disease, and become part of an emergency response to pandemics, either naturally occurring or caused by bioterrorism. A major reason for the need for supercomputing power is that the models and the phenomena being modeled are inherently probabilistic. In computational science terms, this means that particular scenarios must be simulated over and over again – with variables modified to reflect differing probabilities – in order to generate ensembles of results from which the likelihood of particular outcomes can be inferred. The most intensive current work, aimed at response to avian flu, is extendable to other infectious diseases.

For more information, see: *http://jasss.soc.surrey.ac.uk/5/3/5.html*.

GEOSCIENCES

Predicting Severe Storms

Severe storms spawn about 800 tornadoes a year in the United States, mostly in the Great Plains states. The toll in property and economic losses runs to billions of dollars, in addition to an annual average of 1,500 injuries and 80 deaths. Today, weather forecasters can frequently identify storms with tornadic potential. But with current technology, it is seldom possible to air public warnings of potential tornadoes more than half an hour before a twister might strike, and such warnings are still imprecise about timing and location. Largely as a result of this imprecision and lack of timeliness, three of four tornado warnings still prove to be false alarms.

To pave the way for a more advanced and comprehensive approach to storm data-gathering, researchers at the University of Oklahoma recently used the Pittsburgh Supercomputing Center's terascale system to conduct the largest tornado simulation ever performed. The simulation required an area 50 kilometers on each side and an altitude of 16 kilometers. Using 24 hours of computing time with 2,048 processors, the simulated storm yielded 20 terabytes of data.

This simulation successfully reproduced a 1977 storm and the high-intensity tornado it spawned. The results – which captured the tornado's vortex structure, with a wind speed of 260 miles per hour – represented the first simulation of an entire thunderstorm to realistically replicate the complete evolution of a tornado. Simulations like this are an important step in developing scanning algorithms for a new form of low-altitude radar that will be mounted on cell-phone towers. These new radar installations will be used to gather comprehensive forecast data from the cyclonic storms that spawn tornadoes. Scheduled to begin deployment in 2006, these devices and the information that they will provide are expected to reduce the incidence of false tornado alarms from the current 75 percent of warnings to 25 percent – a significant improvement that will add an extra measure of safety for individuals and structures in the paths of these dangerously unpredictable storms.

For more information see:
http://www.psc.edu/science/2004/droegemeier/retwistered_twister.html.

California Earthquake Modeling and Data Analysis

California's southern San Andreas Fault region has not experienced a major earthquake since 1690. It is estimated that the accumulated stress could eventually lead to a catastrophic magnitude 7.7 event in this area. Researchers

are continually seeking ways to secure structures and saves lives in the event of such a disaster, wherever it might occur.

Recently, earthquake scientists produced the largest and most detailed computational simulation yet of a major earthquake. Their primary goal was to explore the response of Southern California's deep, sediment-filled basins to a significant temblor. Researchers modeled a volume 600 kilometers long by 300 kilometers wide and 80 kilometers deep, spanning all major population centers in Southern California.

Dividing the volume into a grid of 1.8 billion cubes, 200 meters on a side, their simulation project, dubbed TeraShake, generated an unprecedented 47 terabytes of data. Two complementary simulations were run for the same 230-kilometer stretch of the fault. A key finding was that the direction of the rupture dramatically focused the energy of the quake. When the fault ruptured from north to south, the energy was focused in the Imperial Valley region in the south, whereas in the northward-running rupture the shaking was stronger and longer in the San Bernardino and Los Angeles basins.

In addition to advancing basic earthquake science, such detailed simulations can lead to new designs by architects and structural engineers for more earthquake-resistant structures, limiting potential human and economic losses even in the event that a major disaster strikes.

For more information, see: *http://www.scec.org/cme*.

ENGINEERING AND MANUFACTURING

Efficient Highway Engineering

The Federal Highway Administration estimates that a staggering $94 billion will be spent on transportation infrastructure every year for the next 20 years. The average large-scale construction project consists of 700 separate activities, each involving a number of variables.

The duration of a highway construction project and the quality and the durability of the product are major considerations for Federal, state, and local transportation officials, as important as the cost of each project. Not surprisingly, state and Federal transportation departments want to ensure that such significant infrastructure investments are indeed worthwhile. The old rule of thumb, "Faster, cheaper, better – pick any two" still seems to be in play today. But how does one reach a logical, comfortable tradeoff among

conflicting objectives in a major construction project? And is it actually possible to objectify quality?

A team at the University of Illinois at Urbana-Champaign has developed a multi-objective genetic algorithm that can weigh more than two factors in determining the combinations of duration, cost, and quality to produce the best possible outcome in a given situation. The model allows an engineer or construction manager to generate a large number of possible construction resource utilization plans that provide a range of tradeoffs among project duration, cost, and quality factors. The options help rapidly eliminate the vast majority of sub-optimal plans from the outset. The model also permits the project planner to assign a quality level to specific resource combinations, based on extensive data from the Illinois Department of Highways. Decision makers would ultimately be provided with a range of optimal tradeoffs that could be used to determine the best possible combination of resources for a specific project.

Older methods for generating such models on personal computers are available, but can consume a month or more of valuable time to produce results. The ability to evaluate these models on parallel systems can reduce elapsed time to a day or less, making this form of evaluation practical for rapid development of project management schedules.

For more information, see: *http://access.ncsa.uiuc.edu/Stories/construction/*.

Converting Biomass to Ethanol for Renewable Energy

The National Renewable Energy Laboratory (NREL) is striving to develop new technologies and processes that enable efficient large-scale conversion of biomass to ethanol to provide a clean-burning and renewable fuel source. Such a breakthrough could reduce dependence on fossil fuels and increasingly expensive imported oil. A major bottleneck to making this process economically viable, however, is the slow breakdown of cellulose by the enzyme cellulase. Scientists hope to understand this key process at the molecular level so they can target further research toward speeding it up.

To explore the intricate molecular dynamics involved in the breakdown of cellulose, researchers have employed CHARMM, a versatile community code for simulating biological reactions. But the size of new simulations needed is so large – more than 1 million atoms – and the simulation times are so long – more than 5,000 time steps for the 10-nanosecond simulations – that they exceed CHARMM's current capabilities.

To make simulating the cellulase reaction feasible, researchers at the San Diego Supercomputer Center (SDSC) , NREL, Cornell University, the Scripps

Research Institute, and the Colorado School of Mines are working to enhance CHARMM so that the simulations can scale up to millions of atoms and run on hundreds of processors on today's largest supercomputers. The research is enabling the largest simulations ever of an important scientific problem that will yield economic and environmental benefits. In addition, improvements to the CHARMM code will be available for the scientific community to use on a wide range of challenging problems.

For more information, see: *http://www.nrel.gov/biomass/*.

Seismic Modeling and Oil Reservoir Simulations

Old-time oil prospectors once relied on hunches as much as anything else to discover promising new sites for wells. Today, oil companies demand the latest technologies to analyze geological features and minimize risk.

Using the NSF's TeraGrid resources, a multidisciplinary research team is currently at work creating software tools that could significantly improve energy companies' oil reservoir management techniques. Using these tools, a hypothetical reservoir is subdivided into a mesh of blocks. Wells, pumps, and other equipment are associated with individual blocks, and an approximate model of each block's fluid dynamics is created. Equipment is moved around within the blocks in order to compare different configurations and determine the most cost-effective one. Since this process could yield billions of possible configurations, a dynamic, data-driven optimization system helps narrow the field of choices.

Middleware tools manage data generated from a rough sampling of the search space and identify good starting points to conduct more comprehensive searches. Dynamic steering and collaboration tools allow on-the-fly searches within these subsections. Sophisticated optimization algorithms guide searches by comparing configurations in the subsections. Seismic models reveal likely geological conditions, based on simulated soundings. These conditions, in turn, help fine-tune the reservoir models, making them as realistic as possible.

In one NSF TeraGrid study, a set of about 25,000 reservoir optimization runs were completed in less than a week, translating into 200 to 400 runs at any given time. More than eight terabytes of seismic simulation data are now being integrated into the reservoir models. Research like this will become increasingly valuable to 21st century energy prospectors attempting to search out ever more scarce resources with less time, manpower, and cost.

For more information, see: *http://access.ncsa.uiuc.edu/Stories/oil*.

Cooling Turbine Blades for Efficient Propulsion and Power

High-efficiency turbines used in propulsion and power generation are operated at near stoichiometric temperatures – i.e., near the point where the fuel is burned completely. Consequently, the gases exiting the combustor into the first stage of the turbine are at temperatures a few hundred degrees Centigrade higher than the melting point of the turbine components. A few tens of degrees increase in surface temperatures can cut blade life in half. So cooling these components is critical to turbine durability and safety.

Turbine vanes and blades are cooled by circulating compressor bypass air through internal passages in the blade (internal cooling). To enhance internal heat transfer, these passages are configured with turbulence promoting augmentors in the form of ribs, pin fins, and impingement cooling. But the turbulent flow is difficult to predict accurately by standard prediction techniques. New computation models have successfully simulated turbulent flow and heat transfer for these complex systems, enabling reliable prediction of design characteristics.

For additional information, see: *http://access.ncsa.uiuc.edu/Stories/blades*.

Microbubbles and Drag Reduction for Ships

Researchers have long known that microbubbles, roughly 50 to 500 microns in size, can cut the drag experienced by ships by 80 percent in some cases, reducing fuel use and increasing range. For 30 years, microbubble systems have been studied experimentally. Pistons push air through porous plates that represent a ship's hull and into tanks of moving water. Researchers have moved the locations of the plates and increased or decreased the number and size of the bubbles. They have seen a wide range of changes in drag, but they have not been able to determine the characteristics of an optimal microbubble system – where to insert bubbles, how many to insert, and how big to make them.

Microbubbles foil traditional methods of measuring the flow details in an experimental tank because optical systems cannot see through the turbulence created by the bubbles. To get around that problem, a group at Brown University created novel first-principles computational models of microbubbles in action. The presence of the bubbles and their influence on the flow are represented by a force-coupling method that tracks the flow and influence of the bubbles without requiring models of the bubbles' surface physics. Bubbles are represented by spherical "force envelopes" instead of solid spheres. By using high-performance computing systems, the Brown team improved the state of the art by a factor of 40, moving from models that track 500 microbubbles to ones that track about 20,000.

The Brown computational model has been distributed to universities, national laboratories, and industry for diverse applications such as combustion, flow-structure interactions, and supersonic flows. This work is part of DARPA's Friction Drag Reduction program, which combines the efforts of 14 research teams around the country. The teams are looking for ways to reduce drag by creating models and experiments at a variety of scales – from computational models that follow the behavior of individual bubbles to mockups that are about 3 meters by 13 meters and run in the world's largest recirculating water tunnel.

For more information, see: *http://access.ncsa.uiuc.edu/Stories/microbubbles*.

Tailoring Semiconducting Polymers for Optoelectronics

Semiconductors and other inorganic crystals serve as the basis for electronics and other technologies. But aside from small changes that can be caused by doping them with impurities, their chemical properties remain fairly inflexible. Soft materials such as polymers, on the other hand, have almost unlimited possibilities because the chemical repeat groups can be modified to suit a particular application. However, commonly used techniques for producing the needed types of soft materials structures such as thin-film or self-assembly processes suffer from substrate and other molecular interactions that may dominate or obscure the underlying polymer physics.

By combining experimental observations and developments with extensive computational chemistry studies, researchers have developed a fundamentally new processing technique for generating optoelectronic materials that is largely controlled by the choice of the solvent involved. By achieving uniform orientation perpendicular to the substrate with enhanced luminescence lifetimes and photostability under ambient conditions, these researchers have opened the door to major developments in molecular photonics, display technology, and bio-imaging, as well as new possibilities for optical coupling to molecular nanostructures and for novel nanoscale optoelectronics devices.

For more information, see:
http://nccs.gov/DOE/mics2004/Sumpter.NanoHighlight.doc.

High-Performance Computing for the National Airspace System

The task of achieving efficient air traffic control services will benefit from the development of high performance computational systems. In the tactical control of air traffic, plans call for increased automation to detect conflicts and provide resolutions to controllers in the en route domain (between airport terminals) . In today's airspace, aircraft are required to fly over radio beacons first designed in the 1930s along marked "airways," rather than flying directly

from point to point. This causes the typical aircraft to fly a route that is 10 percent or more longer than the direct path between its origin and its destination. The basis for this antiquated approach is the need for human controllers to visualize the flight paths of all aircraft in their sectors and order course adjustments manually to maintain adequate separation.

The only solution to this problem lies in the use of high performance computers to anticipate conflicts and issue routing changes to aircraft in real time. An "integrated resolution" algorithm could, for example, balance possible conflicts between two or more aircraft; calculate the extent of rerouting around severe weather; and evaluate the impact of traffic flow imperatives such as meeting specified terminal arrival metering times.

The air traffic control system also needs sophisticated traffic flow management (TFM) , the strategic control of aircraft in order to minimize delays, wasted fuel, and needless cost. TFM is the process of planning and coordinating day-of actions in anticipation of flow-constraining conditions such as thunderstorms, communications outages, or flight demand that exceeds airport capacity. Future TFM systems will acknowledge the uncertain nature of the system and employ probabilistic problem-solving techniques. These advanced capabilities will rely on computational science to assist in the estimation of probabilities in real time and to suggest small changes in the system to maintain a desired level of performance.

The Traffic Flow Management-Modernization (TFM-M) Program of the Federal Aviation Administration (FAA) is addressing the need for an improved infrastructure to support the strategic planning and management of air traffic demand and ensure smooth, efficient traffic flow. Hardware modernization was completed at the end of 2004 and efforts are now focused on reengineering and rearchitecting applications software to achieve a modern, standards-based, open system. Efforts also continue to achieve a robust, scalable, standards-compliant TFM infrastructure and enhance availability, performance, security, expandability, maintainability, and human computer interaction. FAA and the National Oceanic and Atmospheric Administration are collaborating in this research to test and demonstrate the use of innovative science, technology, and computer communication interfaces in developing new weather products for decision makers.

For more information, see: *http://www.faa.gov/aua/aua700/default.shtml* and *http://www-sdd.fsl.noaa.gov/FIR_01_02/FIR_01_02_AD.html#D1*.

BIOLOGICAL SCIENCES AND MEDICINE

Identifying Brain Disorders via Shared Infrastructure

Researchers participating in NIH's Biomedical Informatics Research Network (BIRN) are collaborating in basic medical research that can lead to improved clinical tools. BIRN is a consortium of 15 universities and 22 research groups that participate in testbed projects on brain imaging of human neurological disorders. Through large-scale analyses of patient data acquired and pooled across collaborating sites, the scientists are investigating how to identify and use specific structural differences in patients' brains to help clinicians distinguish diagnostic categories such as Alzheimer's disease. Such research could lead to earlier and more accurate diagnosis of serious brain disorders.

As one component of this large research program, researchers at the Center for Imaging Science (CIS) at Johns Hopkins University and other BIRN researchers collaborated on a processing pipeline for seamless analysis of shape data for brain structures. Computational anatomy tools were integrated in the testbed to perform semi-automated statistical analysis of shapes of anatomical structures. The CIS Large Deformation Diffeomorphic Metric Mapping (LDDMM) tool was used to study hippocampal data from three categories of subjects: Alzheimer's, semantic dementia, and control subjects. The data involved 45 subjects scanned using high-resolution structural magnetic resonance imaging (MRI) at one BIRN site. The data sets were then accessed, aligned, and processed using LDDMM.

LDDMM computes a mathematical description of the shapes that are similar and different by computing metric distances in the space of anatomical images, which allows direct comparison and quantitative characterization of differences in brain structure shapes.

For more information, see: *http://www.nbirn.net/ and http://cis.jhu.edu*.

Decoding the Communication of Bees

Biologists are pursuing research to understand why some bee species have evolved the capability for abstract language to describe their surroundings. Relying on digital video to record bee communication, the researchers have discovered that some bees use sounds to encode information about food location. This ability can prevent other bee species from intercepting the information. Such eavesdropping may have helped drive the development of

sophisticated bee languages as anti-espionage techniques to transmit food source information to nest mates inside the hive.

Using digital video requires storing and accessing massive amounts of information. For each bee species, scientists record 1.2 terabytes of digital video annually. Researchers expect the archive to grow to 30 terabytes or more. Networking infrastructure provides widely separated collaborating labs in Mexico, Brazil, Panama, and San Diego with efficient distributed access to the data, allowing scientists to analyze millions of video frames of bee behavior. Such research may help explain why certain species continue to thrive as a result of sophisticated evolutionary adaptations.

For more information, see: *http://www-biology.ucsd.edu/faculty/nieh.html*.

Modeling Protein Motors

The protein adenosine triphosphate synthase, or ATPase, is the power plant of metabolism, producing ATP, the basic fuel of life and the chemical energy that fuels muscle contraction, transmission of nerve messages, and many other functions. The 1997 Nobel Prize in Chemistry recognized Paul Boyer and John Walker for their work in assembling a detailed picture of ATPase and its operation. Subsequent research has added to the picture, but many challenging questions remain.

Examining the crucial details of how bonds break and reform during a chemical reaction requires the use of quantum theory. A team at the University of Illinois used a method called QM/MM (quantum mechanics/molecular mechanics) , which made it possible to simulate the molecular mechanics of the unit that houses the ATPase's active site, while employing quantum theory selectively like a zoom lens to focus on the active site itself where "combustion" occurs. This model consumed over 12,000 hours of computation time.

Among several new findings, the simulations reveal that one of the amino acids of ATPase appears to coordinate the timing among the protein's three active sites, where ATP is produced. This amino acid – referred to as the arginine finger – operates somewhat like a spark plug, shifting position depending on whether ATP or the reaction products are in the active site. This finding may be a key to resolving the story of how this protein does its vital job, potentially leading to future medical breakthroughs.

For more information see:
http://www.psc.edu/science/2004/schulten/protein_motors_incorporated.html.

Protein Dynamics and Function

Computational methods have long been used to extend the reach of experimental biology by means of data analysis and interpretation. However, the real power of computational science in this area is in biomolecular simulations that explore areas of research that are impossible via experimentation.

One area where biomolecular simulations are starting to make an impact is in how biologists think about the function of proteins. Previously, protein complexes were viewed as static entities, with biological function understood in terms of direct interactions among components. Based on computational simulations, proteins are now viewed as efficient molecular machines that are dynamically active in ways closely associated with their structure and function. This emerging view has broad implications for protein engineering and improved drug design.

Using biomolecular simulations and advanced visualization techniques, a network of protein vibrations in the enzyme cyclophilin A has been identified. The discovery of this network is based on investigation of protein dynamics at picosecond to microsecond-millisecond time scales. This network plays a vital role in the function of this protein as an enzyme. Cyclophilin A is involved in many biological reactions, including protein folding and intracellular protein transport, and is required for the infectious activity of the human immunodeficiency virus (HIV-1) .

Currently, researchers are attempting to make software improvements that will more fully exploit the power of next-generation supercomputers to better understand protein dynamics. Such improvements can be achieved through the parallelization and optimization of molecular dynamics (MD) code for supercomputers. Parallelization of MD codes is of wide interest to the biological community. With current computational resources, MD modeling falls short of simulating biologically relevant time scales by several orders of magnitude. The ratio of desired and simulated time scales is somewhere between 100,000 and 1,000,000. In addition, today's biological systems of interest consist of millions of atoms, which will require substantially more computing power for extended periods of time.

For more information, see:
http://nccs.gov/DOE/mics2004/Agarwal.VibrationsHighlight.doc.

Computational Science and Medical Care

A major national initiative is currently underway to computerize the nation's health care infrastructure. Current estimates suggest that as much of 25 percent of the cost of today's health care delivery is associated with the cost of the paper-bound systems through which health care is provided. Moreover, there is substantial evidence that one in seven hospitalizations occurs because critical patient information was not transmitted from one caregiver to another. Similarly, it is well established that one in seven diagnostic tests is performed simply because the results of the last test are not available at the time of care and that one in five paper-based physician orders is carried out incorrectly.

The solutions to problems like these lie in the nationwide adoption of electronic health records, computerized order entry and execution, and computer-aided decision support – all within a context of secure, interoperable health information exchange. It is envisioned that the universal adoption of computerized health care records and systems will vastly improve the efficiency of medical care. Such gains have already been demonstrated by the Veterans Administration, which is now able to care for twice as many patients as it did a decade ago on a budget that has increased by only 33 percent. The PITAC's findings and recommendations on the R&D necessary to realize the promise of IT to improve health care are presented in its June 2004 report, *Revolutionizing Health Care Through Information Technology.*

For more information, see:
http://www.nitrd.gov/pitac/reports/20040721_hit_report.pdf and
http://www.os.dhhs.gov/healthit/.

APPENDIX B

Computational Science Warnings – A Message Rarely Heeded

During the past two decades, the national science community has produced a number of reports, each recommending sustained, long-term investment in the underlying technologies and applications needed to realize the full benefits of computational science. Instead, short-term investment and limited strategic planning have led to an excessive focus on incremental research rather than long-term research with lasting impact. The recommendations and warnings of these reports often triggered short-term responses. But their admonitions to ensure long-term, strategic investment have rarely been heeded, to the detriment of U.S. competitiveness.

Twenty Years of Recommendations

Each of these reports stressed the catalytic role that computational science plays in supporting, stimulating, and transforming the conduct of science, engineering, and business. The reports also emphasized how computing can address problems of significantly greater complexity, scope, and scale than was previously possible, including issues of national importance that cannot be otherwise addressed. U.S. leadership in computational science, the reports concluded, can and should yield a wide range of ongoing benefits for innovation, competitiveness, and quality of life.

The reports identified a range of barriers and concerns that must be overcome if these benefits are to be fully realized. First, they argued that the Federal government must take primary responsibility, in partnership with industry and academia, for achieving and retaining international leadership in computational science via sustained, long-term investment. Second, they emphasized that computational science now encompasses a broad range of components, including hardware, software, networks, data and databases, middleware and metadata, people, and organizations, and that significant development is needed in each area.

Organizations and their support mechanisms will need to change, the reports agreed, as multidisciplinary teams and distributed and federated approaches become the norm. The reports also argued that innovative incentive, reward, and recognition systems must be put in place to draw new people into emerging areas of computational science specialization.

Many themes recurred throughout the reports. They can be summarized as follows:

- **Opportunity:** the enormous opportunities to advance scientific discovery, enhance economic competitiveness, and help ensure national security
- **Sustainability:** the importance of long-term, sustained investment at adequate levels to reap the rewards of computational science
- **Leading-Edge Capability:** the need for deployment of leading-edge computing systems and networks for scientific discovery
- **Data Management:** the emergence of instruments and the data they capture as part of a larger computational environment, with large-scale data archives for community use
- **Education:** the importance of a trained and well-educated workforce with state-of-the-art computational science skills
- **Software:** the need for easy-to-use, effective software and tools for computational science discovery
- **Research Investment:** the need for continued investment in computer and computational science research
- **Cyberinfrastructure:** the emerging opportunity to interconnect instruments, computing systems, data archives, and individuals in an international cyberinfrastructure
- **Coordination:** the importance of coordinated planning and implementation across Federal R&D agencies

Following are brief synopses of the major reports the PITAC reviewed.

PITAC: Information Technology Research

The PITAC examined contemporary Federal IT R&D activities in its 1999 report entitled *Information Technology Research: Investing in Our Future.* The PITAC concluded that Federal IT R&D investment was inadequate and too heavily focused on near-term problems. The Committee recommended a strategic initiative in long-term IT R&D, highlighting five priorities for the overall research agenda: (1) software; (2) scalable information infrastructure; (3) high-end computing; (4) socioeconomic impacts; and (5) management and implementation of Federal IT research.

Department of Energy: SCaLeS

In *A Science-Based Case for Large-Scale Simulation,* commissioned by DOE's Office of Science, the research community stated that computational simulation has attained peer status with theory and experiment in many areas of science. The two-part report, released in 2003 and 2004, noted that there were both responsibilities and opportunities to initiate a vigorous research effort that could bring the power of advanced simulation to many scientific frontiers, while simultaneously leapfrogging theoretical and experimental progress in addressing such questions as the fundamental structure of matter, production of heavy elements in supernovae, and the functions of enzymes.

The report called for new, sustained, and balanced funding for: (1) scientific applications; (2) algorithm research and development; (3) computing system software infrastructure; (4) network infrastructure for access and resource sharing, including software to support collaboration among distributed teams of scientists; (5) computational facilities supporting both capability computing for "heroic simulations" that cannot be performed any other way and capacity computing for "production simulations" that contribute to a steady stream of new knowledge; (6) innovative computer architecture research for the facilities of the future; and (7) recruiting and training a new generation of multidisciplinary computational scientists.

Council on Competitiveness: Supercharging Innovation

A 2004 report from the Council on Competitiveness entitled *Supercharging U. S. Innovation & Competitiveness* stressed the importance of high-performance computing as a business tool for innovation and transformation, but observed that it was currently underutilized. The report noted several barriers to high-performance computing in the private sector, including: (1) a business culture that views high-performance computing as a cost of doing business rather than an investment that produces returns; (2) the lack of personnel capable of using high-performance computing productively or fully exploiting its potential for innovation; and (3) difficulty in using current high-performance computing hardware, software, and models.

The report noted that opportunities for boosting innovation and competitiveness through high-performance computing included creating new government-industry-university partnerships, developing next-generation computational simulations, and improving correspondence between the computational knowledge and skills required by businesses and those taught by universities.

Department of Defense: HPC For National Security

Until the mid-1990s, national security interests drove the supercomputing industry and its advances. As the non-defense industrial, scientific, and academic markets for high-end computing grew, and as foreign competition emerged for market share and technology leadership, both government and industry focused on developing and manufacturing supercomputers based on commodity components. Although this significantly increased the affordability of solving many important national security problems, other critical application areas remain unaddressed by the commercial sector.

DoD's 2002 report *High-Performance Computing for the National Security Community* outlined a plan to rebuild and sustain a strong industrial base in high-end computing, including applied research, advanced development, and engineering and prototype development. The plan also called for establishing high-end computing laboratories to test system software on dedicated, large-scale platforms; supporting the development of software tools and algorithms; developing and advancing benchmarking and modeling and simulation for system architectures; and conducting detailed technical requirements analyses.

National Academies: Future of Supercomputing

Getting up to Speed: The Future of Supercomputing, a 2005 report by the National Academies, examined U.S. needs for supercomputing and recommended a long-term strategy for Federal government support of high-performance computing R&D. The report recognized the central contribution of supercomputing to the economic competitiveness of many industries (e.g., automotive, aerospace, health care, and pharmaceutical) but raised concerns about the rate of progress in other areas of science and engineering. This study was part of a broader initiative by the U.S. to assess its current and future supercomputing capabilities. The assessment was spurred in part by the introduction of Japan's Earth Simulator, which could process data at three times the speed of the fastest U.S. supercomputer available at the time.

The report recommended that investment decisions regarding supercomputing research and development should not be based on whether the U.S. possesses the world's fastest supercomputer. Instead, the Government should make long-term plans to secure U.S. leadership in the hardware, software, and other technologies that are essential to national defense and scientific research. The report concluded that the demands for supercomputing to strengthen U.S. defense and national security cannot be satisfied with current policies and levels of spending. It called on the Federal government to provide stable, long-term funding and support multiple supercomputing hardware and software companies to give scientists and

policymakers better tools for problem solving in such areas as intelligence, nuclear stockpile stewardship, and climate change.

National Institutes of Health: BISTI

NIH's Biomedical Information Science and Technology Initiative (BISTI) report cited the tremendous progress in computation and the scope of its impact on biomedicine in the latter half of the 20th century, and it described the challenges and opportunities presented to NIH by the convergence of computing and biomedicine. The report highlighted the transition of biology from a bench-based science to a computation-based science, from individual researchers to interdisciplinary teams, and from a focus on the application of digital technologies to the development of computational methods that are changing the way biomedical research is pursued.

The report recommended creating National Programs of Excellence in Biomedical Computing to conduct research into all facets of biomedical computation and play a major role in the education of biomedical computation researchers. It also called for establishing a new program directed toward the principles and practice of data and information storage, curation, analysis, and retrieval (ISCAR) . Other recommendations included providing adequate resources and incentives for those working on the tools of biomedical computing and supporting a scalable and balanced national computing infrastructure to address a dynamic range of computational needs and accompanying support requirements. In response to these recommendations, NIH Director Elias Zerhouni convened a series of meetings to chart a roadmap for medical research in the 21st century.

Interagency: High-End Computing Revitalization Task Force

The 2004 HECRTF report, *Federal Plan for High-End Computing,* addresses three components of a plan for high-end computing: (1) an interagency research and development roadmap for high-end core technologies, (2) a Federal high-end computing capacity and accessibility improvement plan, and (3) recommendations relating to Federal procurement of high-end computing systems. Based on independent review and planning efforts by DoD, DOE, and NSF, the report notes that the strategy of pursuing high-end computing capability based on COTS components is insufficient for applications of national importance.

The report recommends: (1) a coordinated, sustained research, development, testing, and evaluation program over 10 to 15 years to overcome major technology barriers limiting effective use of high-end computers, including detailed roadmaps for hardware, software, and systems; (2) providing high-end computing across the full scope of Federal missions,

including both production and "leadership-class" systems offering leading-edge capability for high-priority research and guiding the next generation of production systems; and (3) improved efficiency in Federal procurement processes for high-end computing through benchmarking, development of total-cost-of-ownership models, and shared procurement across agencies. The HECRTF assumes that agency investments in the broader computing environment – including networking, applications software development, computational science education, general computing and storage systems, and visualization – will be at the levels required to support high-end computing as an effective tool in national defense, national security, and scientific research missions.

National Science Foundation: Cyberinfrastructure

NSF's *Revolutionizing Science and Engineering Through Cyberinfrastructure* report (the Atkins report) found that today's computing, information, and communication technologies now make possible development of a comprehensive cyberinfrastructure to support a new era of research whose complexity, scope, and scale would once have been beyond imagination. The 2003 report's key recommendation urges the foundation to establish and lead a large-scale, interagency, and internationally coordinated Advanced Cyberinfrastructure Program (ACP) to create, deploy, and apply that infrastructure to radically empower all scientific and engineering research and allied education.

This report proposes a large, long-term, and concerted effort, not merely a linear extension of current investment levels and resources. The report also envisions the education and involvement of more broadly trained personnel with blended expertise in a disciplinary science or engineering as well as the skill sets encompassed by computational science, such as mathematical and computational modeling, numerical methods, visualization, and socio-technical understanding about working in new grid or collaboratory organizations.

National Academies: Making IT Better

The 2000 National Academies report, *Making IT Better,* found that the United States – indeed much of the world – is in the midst of a transformation wrought by information technology (IT). Fueled by continuing advances in computing and networking capabilities, IT has moved from the laboratories and back rooms of large organizations and now touches people everywhere. The indicators are almost pedestrian: computing and communications devices have entered the mass market and the language of the Internet has become part of the business and popular vernacular.

The report observed that the critical role of the first half of the R&D process is often overlooked, namely the research that uncovers underlying principles, fundamental knowledge, and key concepts that fuel the development of numerous products, processes, and services. Research has been an important enabler of IT innovations – from the graphical user interface to the Internet itself – and it will continue to enable the more capable systems of the future, the forms of which have yet to be determined. When undertaken in the university environment in particular, it also serves as a key educational tool as well, helping build a broader and more knowledgeable IT workforce.

The future of IT and of the society it increasingly powers depends on continued investments in research, the report concludes. New technologies based on quantum physics, molecular chemistry, and biological processes are being examined as replacements for or complements to the silicon-based chips that perform basic computing functions. Research is needed to enable progress along all these fronts and to ensure that IT systems can operate dependably and reliably, meeting the needs of society and complementing the capabilities of their users.

But key questions remain to be answered, according to the report: Can the Nation's research establishment generate the advances that will enable tomorrow's IT systems? Are the right kinds of research being conducted? Is there sufficient funding for the needed research? Are the existing structures for funding and conducting research appropriate to the challenges IT researchers must address?

National Academies: Embedded Infrastructure

The 2001 National Academies report, *Embedded Everywhere,* found that IT is on the verge of another revolution. Driven by the increasing capabilities and declining costs of computing and communications devices, IT is being embedded in a growing range of physical devices linked together through networks and will become ever-more pervasive as the component technologies become smaller, faster, and cheaper. These changes are sometimes obvious – in pagers and Internet-enabled cell phones, for example. But often IT is buried inside larger (or smaller) systems in ways that are not easily visible to end users. These networked systems of embedded computers have the potential to change the way people interact with their environment by linking together a range of devices and sensors that will allow information to be collected, shared, and processed in unprecedented ways.

The range of applications continues to expand with continued research and development. Examples include instrumentation ranging from in situ

environmental monitoring to battlespace surveillance. Embedded networks will be employed in defense-related and civilian personal monitoring strategies combining information from sensors on and within a person with information from laboratory tests and other sources. These networks will dramatically affect scientific data collection capabilities, ranging from new techniques for precision agriculture and biotechnological research to detailed environmental and pollution monitoring.

National Science Foundation: Digital Libraries

Knowledge Lost in Information, an NSF workshop report published in 2003 by the University of Pittsburgh, found that digital libraries are transforming research, scholarship, and education at all levels. Vast quantities of information are being collected and stored online and organized to be accessible to everyone. Substantial improvements in scholarly productivity are already apparent. Digital resources have demonstrated the potential to advance scholarly productivity, most likely doubling research output in many fields within the next decade. These resources will become primary resources for education, with the potential for making the kinds of significant advances in lifelong learning that have been sought for many years. This report details the nature of the Federal investment required to sustain the pace of progress.

Digital library programs have engaged international partners, with several U.S. projects coordinated with counterpart projects in the United Kingdom and Germany, as well as with broader international projects involving the European Union and Asian countries. Moreover, the kinds of information created and examined have moved well beyond text and book-like objects to include scans of fossils, images of dolphin fins, cuneiform tablets, and videos of human motion, potentially enabling more sophisticated analysis in domains that range from archaeology and paleontology to physiology, while exploring the engineering issues that are exposed in the course of such investigations.

Legacy Reports and Implications

The 2005 National Academies study, *Getting up to Speed: The Future of Supercomputing,* contains a cogent summary of early assessments of the importance of computational science and high-end computing. In 1982, the *Report of the Panel on Large Scale Computing in Science and Engineering* (the Lax report) made four recommendations: (1) increase access for the science and engineering research community to regularly upgraded supercomputing facilities via high-bandwidth networks; (2) increase research in computational mathematics, software, and algorithms necessary for effective and efficient use of supercomputing systems; (3) train people in scientific computing; and (4) invest in the R&D basic to the design and implementation of new

supercomputing systems of substantially increased capability and capacity, beyond that likely to arise from computational requirements alone.

A 1993 successor report, *From Desktop to Teraflop: Exploiting the U.S. Lead in High Performance Computing* (the Branscomb report), recommended significant expansion in NSF investments, including accelerating progress in high-performance computing through computer and computational science research.

In 1995, NSF formed a task force to advise it on the review and management of its supercomputer centers program. The chief finding of the *Report of the Task Force on the Future of the NSF Supercomputer Centers Program* (the Hayes report) was that the supercomputing centers funded by NSF had enabled important research in computational science and engineering and had also changed the way that computational science and engineering contribute to advances in fundamental research across many areas. The recommendation of the task force was to continue to maintain a strong advanced scientific computing centers program.

APPENDIX C

Charge to PITAC

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

June 9, 2004

Mr. Marc R. Benioff
Chairman and CEO Salesforce.com
Suite 300
The Landmark@One Market
San Francisco, CA 94105

Dear Mr. Benioff:

Again, I want to thank you for your service as co-chair of the President's Information Technology Advisory Committee (PITAC) and your excellent leadership at the Apri1 13, 2004 PITAC meeting. This letter outlines my expectations regarding PITAC's plans to address issues related to computational science. I look forward to PITAC's engagement in this issue.

The importance of computational science as a complement to experiment and theory is increasing, with applications that are relevant to numerous Federal agency missions. The Federal government has funded much of the development of computational science and is a major beneficiary of its use, making it an appropriate area for PITAC to consider. I would like PITAC to address the following questions in the context of the Networking and Information Technology Research and Development (NITRD) program, as well as other relevant Federally funded research and development:

1. How well is the Federal government targeting the right research areas to support and enhance the value of computational science? Are agencies' current priorities appropriate?

2. How well is current Federal funding for computational science appropriately balanced between short term, low risk research and longer term, higher risk research? Within these research arenas, which areas have the greatest promise of contributing to breakthroughs in scientific research and inquiry?

3. How well is current Federal funding balanced between fundamental advances in the underlying techniques of computational science versus the application of computational science to scientific and

engineering domains? Which areas have the greatest promise of contributing to breakthroughs in scientific research and inquiry?

4. How well are computational science training and research integrated with the scientific disciplines that are heavily dependent upon them to enhance scientific discovery? How should the integration of research and training among computer science, mathematical science, and the biological and physical sciences best be achieved to assure the effective use of computational science methods and tools?

5. How effectively do Federal agencies coordinate their support for computational science and its applications in order to maintain a balanced and comprehensive research and training portfolio?

6. How well have Federal investments in computational science kept up with changes in the underlying computing environments and the ways in which research is conducted? Examples of these changes might include changes in computer architecture, the advent of distributed computing, the linking of data with simulation, and remote access to experimental facilities.

7. What barriers hinder realizing the highest potential of computational science and how might these be eliminated or mitigated?

Based on the findings of PITAC with regard to these questions, I request that PITAC present any recommendations you deem appropriate that would assist us in strengthening the NITRD program or other computational science research programs of the Federal government.

In addressing this charge, I ask that you consider the appropriate roles of the Federal government in computational science research versus those of industry or other private sector entities.

I request that PITAC deliver its response to this charge by February 1, 2005.

Sincerely

John H. Marburger, III
Director

Letter also sent to: Edward D. Lazowska, Ph.D.

APPENDIX D

Subcommittee Fact-Finding Process

The Computational Science Subcommittee studied and deliberated on an array of relevant reports and trade publications. The Subcommittee also held a series of meetings during which Federal government leaders and experts from academia and industry were invited to provide input. The meetings held were as follows:

- June 17, 2004 PITAC meeting
- September 16, 2004 Computational Science Subcommittee meeting
- October 19, 2004 Computational Science Subcommittee meeting
- November 4, 2004 PITAC meeting
- November 10, Computational Science Subcommittee Birds of a Feather Town Hall meeting at the Supercomputing (SC) 2004 conference
- January 12, 2005 PITAC meeting
- April 14, 2005 PITAC meeting
- May 11, 2005 PITAC meeting

June 17, 2004 PITAC Meeting (Arlington, Virginia)

Formal presentations were given by:

- Eric Jakobsson, Ph.D., Director, Center for Bioinformatics and Computational Biology, National Institute of General Medicine, National Institutes of Health
- Michael Strayer, Ph.D., Director, Scientific Discovery through Advanced Computation, Office of Science, Department of Energy
- Arden L. Bement, Jr., Ph.D., Director, National Science Foundation
- Ken Kennedy, Ph.D., John and Ann Doerr University Professor, Department of Computer Science, Rice University

To view or hear these presentations, or to read the meeting minutes, please visit: *http://www.nitrd.gov/pitac/meetings/2004/index.html.*

September 16, 2004 Subcommittee Meeting (Chicago, Illinois)

Formal presentations were given by the following experts:

- James Crowley, Ph.D., Executive Director, Society for Industrial and Applied Mathematics
- Robert Lucas, Ph.D., Director, Computational Science Division, Information Sciences Institute, University of Southern California
- Phillip Colella, Ph.D., Leader, Applied Numerical Algorithms Group, Lawrence Berkeley National Laboratory

- Edward Seidel, Ph.D., Director, Center for Computation and Technology, Louisiana State University
- Charbel Farhat, Ph.D., Professor, Department of Mechanical Engineering and Institute for Computational and Mathematical Engineering, Stanford University
- Kelvin Droegemeier, Ph.D., Director, Center for Analysis and Prediction of Storms; Regents' Professor, School of Meteorology, College of Geoscience, University of Oklahoma
- Michael Vannier, Ph.D., Professor of Radiology, University of Chicago
- Jonathan C. Silverstein, M.D., M.S., FACS, Assistant Professor of Surgery, University of Chicago
- John Reynders, Ph.D., Information Officer, Lilly Research Labs
- Vernon Burton, Ph.D., Associate Director, Humanities and Social Sciences, National Center for Supercomputing Applications, University of Illinois, Urbana-Champaign
- Daniel E. Atkins, Ph.D., Professor, School of Information; Executive Director, Alliance for Community Technology, University of Michigan
- Jack Dongarra, Ph.D., University Distinguished Professor, Innovative Computing Laboratory; Computer Science Department, University of Tennessee

October 19, 2004 Subcommittee Meeting (Arlington, Virginia)

Formal presentations were given by:

- Alvin W. Trivelpiece, Ph.D., Director, Oak Ridge National Laboratory (Retired)
- André van Tilborg, Ph.D., Director, Information Systems, Deputy Under Secretary of Defense (Science and Technology), DoD
- Walt Brooks, Ph.D., Chief, Advanced Supercomputing Division, National Aeronautics and Space Administration
- Timothy L. Killeen, Ph.D., Director, National Center for Atmospheric Research
- Chris R. Johnson, Ph.D., Director, Scientific Computing and Imaging Institute, University of Utah
- Michael J. Holland, Ph.D., Senior Policy Analyst, Office of Science and Technology Policy

November 4, 2004 PITAC Meeting (Arlington, Virginia)

This meeting was held by WebEx/teleconferencing at which Subcommittee Chair Daniel A. Reed provided an update on the Subcommittee's activities. PITAC members discussed these activities and solicited comments from the public. Dr. Reed's presentation can be found at: *http://www.nitrd.gov/pitac/meetings/2004/20041104/agenda.html.*

November 10, 2004 Subcommittee Meeting (Pittsburgh, Pennsylvania)

The Subcommittee held a Birds of a Feather (BOF) Town Hall meeting at the SC 2004 conference. The purpose of the meeting was to solicit input from the SC 2004 community as part of gathering broader input from the public. Subcommittee Chair Reed provided a presentation and a list of questions to focus on particular areas of interest. Chair Reed's presentation and list of questions can be found at:

http://www.nitrd.gov/pitac/meetings/2004/20041110/reed.pdf and *http://www.nitrd.gov/pitac/meetings/2004/20041110/bof_pitac.pdf.*

January 12, 2005 PITAC Meeting (Arlington, Virginia)

At this meeting Chair Reed gave an update on the Subcommittee, and formal presentations on computational science in education programs were given by:

- Linda Petzold, Ph.D., Professor and Chair, Department of Computer Science; Professor, Department of Mechanical and Environmental Engineering; and Director, Computational Science and Engineering Program, University of California, Santa Barbara
- J. Tinsley Oden, Ph.D., Associate Vice President for Research, Director, Institute for Computational Engineering and Sciences, Cockrell Family Regents' Chair #2 in Engineering, University of Texas

PITAC members discussed the Subcommittee's preliminary draft findings and recommendations. Chair Reed's presentation from the meeting can be found at:

http://www.nitrd.gov/pitac/meetings/2005/20050112/agenda.html.

April 14, 2005 PITAC Meeting (Washington, D.C.)

Computational Science Subcommittee Chair Reed presented the draft report and solicited discussion by the PITAC and comments from the public. The PITAC approved the report's findings and recommendations and asked the Subcommittee to revise the text in response to the comments from PITAC members and the public. To view these presentations, please visit:

http://www.itrd.gov/pitac/meetings/2005/20050414/agenda.html.

May 11, 2005 PITAC Meeting (Arlington, Virginia)

At this meeting, held by WebEx/teleconferencing, Computational Science Subcommittee Chair Reed outlined the editorial revisions the Subcommittee had made to the report, highlighting the substantive rewrites of several sections of the document responding to comments at the April 14 meeting. In

discussion, PITAC members praised the revisions as significant improvements to the overall quality of the report. The report was then approved by a unanimous vote.

Agency Information

A number of agencies provided written information about their computational science R&D investments in response to a formal request from PITAC. Senior officials from several agencies made presentations to the Subcommittee to provide further insights into agency policies and practice with regard to computational science.

APPENDIX E

Acronyms

ACE
Agent-based computational economics

ACP
Advanced Cyberinfrastructure Program

AMR
Adaptive mesh refinement

ARPA
Advanced Research Projects Agency

ARPANet
Advanced Research Projects Agency Network

ASCI
DOE/National Nuclear Security Administration's Accelerated Strategic Computing Initiative

ATLAS
A ToroidaLHC ApparatuS

ATP
Adenosine triphosphate

BIRN
Biomedical Informatics Research Network

BISTI
Biomedical Information Science and Technology Initiative

BOF
Birds of a feather

BSD
Berkeley Software Distribution

CCD
Charge Coupled Device

CHARMM
Chemistry at Harvard Molecular Mechanics

CIS
Center for Imaging Science

CMS
Compact Muon Solenoid

COTS
Commercial-off-the-shelf

CRA
Computing Research Association

CSE
Computational science and engineering

DARPA
Defense Advanced Research Projects Agency

DNA
Deoxyribonucleic acid

DoD
Department of Defense

DOE
Department of Energy

ETF
Extensible Terascale Facility

FAA
Federal Aviation Administration

FACA
Federal Advisory Committee Act

FARSITE
Fire Area Simulator

FLASH
State-of-the-art simulator code for solving nuclear astrophysical problems related to exploding stars

fMRI
Functional magnetic resonance imaging

FORTRAN
Formula Translation (programming language)

FRB
Federal Reserve Bank

GeV
Giga-electron-Volt (one billion electron-volts)

GUPS
Giga updates per second

HECRTF
High-End Computing Revitalization Task Force

HIV/AIDS
Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome

HPC
High-Performance Computing

HPCC
High-Performance Computing and Communications

HPCS
DARPA's High Productivity Computing Systems Program

HPF
High-Performance FORTRAN

HTSC
High-temperature superconductors

HUMINT
Human intelligence

ICPSR
Inter-university Consortium for Political and Social Research

ILLIAC IV
Illinois Integrator and Automatic Computer

IPA
Intergovernmental Personnel Act

IPAC
Infrared Processing and Analysis Center

ISCAR
Information storage, curation, analysis, and retrieval

IT
Information technology

ITER
International Thermonuclear Experimental Reactor

ITRS
International Technology Roadmap for Semiconductors

IT R&D
Information Technology Research and Development

IVOA
International Virtual Observatory Alliance

I/O
Input/output

LANL
Los Alamos National Laboratory

LAPACK
Linear Algebra PACKage

LDDMM
Large Deformation Diffeomorphic Metric Mapping

LHC
Large Hadron Collider

LINPACK
LINear algebra software PACKage

LSST
Large Synoptic Survey Telescope

MD
Molecular dynamics

MEMS
Microelectromechanical systems

MPI
Message Passing Interface

MPICH
Argonne National Laboratory MPI implementation

MREFC
Major Research Equipment and Facilities Construction, an NSF budget line

MRI
Magnetic resonance imaging

NASA
National Aeronautics and Space Administration

NCBI
National Center for Biotechnology Information

NCO
National Coordination Office

NCSA
National Center for Supercomputing Applications

NERSC
National Energy Research Scientific Computing Center

NIH
National Institutes of Health

NIMROD
Non-ideal MHD with Rotation Open Discussion

NITRD
Networking and Information Technology Research and Development Program

NMI
National Middleware Initiative

NOAA
National Oceanic and Atmospheric Administration

NRC
National Research Council

NREL
National Renewable Energy Laboratory

NSA
National Security Agency

NSB
National Science Board

NSF
National Science Foundation

NSTC
National Science and Technology Council

NVO
National Virtual Observatory

OMB
Office of Management and Budget

OSCAR
Open Source Clustering Application Resource, a Linux cluster distribution

OSTP
Office of Science and Technology Policy

PCAST
President's Council of Advisors on Science and Technology

PITAC
President's Information Technology Advisory Committee

PSC
Pittsburgh Supercomputing Center

QCD
Quantum chromodynamics

QM/MM
Quantum mechanics/molecular mechanics

R&D
Research and development

ROCKS
Linux cluster distribution

S&E
Science and engineering

SARS
Severe Acute Respiratory Syndrome

SCaLeS
Science-based Case for Large-scale Simulation

SciDAC
Scientific Discovery Through Advanced Computing

SDSC
San Diego Supercomputer Center

SEMATECH
Semiconductor Manufacturing Technology

SGI
Silicon Graphics Incorporated, now SGI

SIAM
Society for Industrial and Applied Mathematics

SIGINT
Signals intelligence

TCO
Total cost of ownership

TFM
Traffic flow management

TFM-M
Traffic Flow Management-Modernization program

TSI
Terascale Supernova Initiative

UC
University of California

UNICOS
UNIX operating system for Cray computers

VORPAL
A parallel, object-oriented hybrid (fluid and particle-in-cell) code for modeling systems of electromagnetic fields, charged particles, and/or neutral gases

VTK
Visualization Toolkit

XML
Extensible Markup Language

Acknowledgements

The PITAC co-chairs gratefully acknowledge the members of the Committee's Subcommittee on Computational Science for their contributions. Subcommittee Chair Daniel A. Reed deserves special mention for his strong leadership and commitment, which were vital to the development of this report. Dr. Reed also developed the analytical framework for the report and drafted significant portions of the text. The PITAC also appreciates the valuable contributions of consultants Jack Dongarra and Chris R. Johnson to the Subcommittee's work.

Many individuals and organizations generously provided input to PITAC during the data-collection process for this report. A number of experts made presentations to the Subcommittee or to the Committee (Appendix D). Federal R&D agencies responded to a PITAC request for information. Several individuals and organizations provided written submissions to PITAC. Others provided input during the public comment periods at PITAC meetings. Collectively, these inputs helped the Committee produce a report that reflects the broad range of relevant perspectives. PITAC is grateful to these individuals and organizations for their efforts.

The PITAC thanks the staff members of the National Coordination Office for Information Technology Research and Development for their contributions in supporting and documenting meetings; drafting sections of the report; critiquing, editing, and proofreading the numerous drafts; and contributing to the substantive dialogue that led to the final report. Staff members Alan S. Inouye, William "Buff" Miner, Martha K. Matzke, and Terry L. Ponick provided primary support in the report's development, under the guidance and oversight of David B. Nelson and Sally E. Howe. Staff members Nekeia J. Bell, Vivian Black, Stephen Clapham, William C. Harrison, Jr., Virginia Moore, Alan Tellington, and Diane Theiss provided technical and administrative support for the Subcommittee's work.

The Committee appreciates the thoughtful advice provided by Sharon L. Hays and Charles H. Romine of the Office of Science and Technology Policy. As was true in PITAC's prior reports on health care and cyber security, their inputs stimulated our thinking and led to a better report.

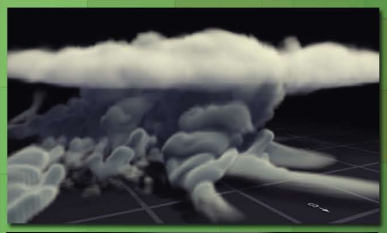
Finally, the PITAC acknowledges Scientific Designer-Illustrator James J. Caras of NSF for his work on the cover and figures of this report and NCSA Visualization Research Artist Robert Patterson, who rendered the cover images and supplied technical information on the tornado simulation.

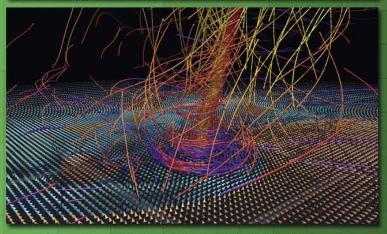
Ordering Copies of PITAC Reports

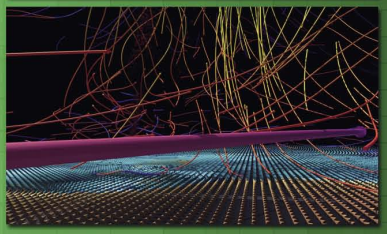
This report is published by the National Coordination Office for Information Technology Research and Development. To request additional copies or copies of other PITAC reports, please contact:

National Coordination Office
for Information Technology Research and Development
4201 Wilson Blvd., Suite II-405
Arlington, Virginia 22230
(703) 292-4873
Fax: (703) 292-9097
Email: nco@nitrd.gov

PITAC documents are also available on the NCO Web site:
http://www.nitrd.gov







National Coordination Office
for Information Technology
Research and Development
Suite II-405
4201 Wilson Boulevard
Arlington, Virginia 22230
(703) 292-4873
Email address: nco@nitrd.gov
Web address: http://www.nitrd.gov/pitac