SEPTEMBER 2003

Advancing High End Computing:
Linking to National Goals

Innovation Series

Juan D. Rogers
Associate Professor of Public Policy and
Director, Research Value Mapping Program
School of Public Policy
Georgia Institute of Technology

Barry Bozeman
Regents' Professor of Public Policy
School of Public Policy
Georgia Institute of Technology

IBM Center for
**The Business
of Government**

I N N O V A T I O N   S E R I E S

Advancing High End Computing: Linking to National Goals

Juan D. Rogers
Associate Professor of Public Policy and
Director, Research Value Mapping Program
School of Public Policy
Georgia Institute of Technology

Barry Bozeman
Regents' Professor of Public Policy
School of Public Policy
Georgia Institute of Technology

September 2003

TABLE OF CONTENTS

FOREWORD

September 2003

On behalf of the IBM Center for The Business of Government, we are pleased to present this report, "Advancing High End Computing: Linking to National Goals," by Juan D. Rogers and Barry Bozeman.

Over the past 18 months, there has been considerable discussion about the United States strategy for advancing high end computing (HEC) to achieve national goals. Several public and private sector sponsored studies are now under way to examine the federal government's current strategy in high end computing. These reports will make recommendations on how future advances in HEC will meet this nation's scientific and business goals and challenges.

The IBM Center for The Business of Government commissioned this study to further explore the issues surrounding United States strategy in high end computing. The goal of this report is to add to the base of knowledge on this important issue. Professors Rogers and Bozeman describe the questions surrounding HEC in order to establish a common ground for continuing dialogue about the future of HEC. The report addresses the need for strong partnerships between the government, universities, and the business community to ensure long-term, significant, and thoughtful advancement in high end computing.

The report addresses the critical importance of high end computing to science, engineering, and the overall research and development system of the nation, as well as the role of policy makers in ensuring HEC's continued advancement. Professors Rogers and Bozeman discuss the importance of high end computing as a tool for achieving national goals and the application needs of the scientific, research, and business community.

This report comes at a critical time in the history of high end computing. Today, increased application capability, combined with manageable costs, will be critical to a high end computing strategy in government, universities, and the private sector. It is also our belief that long-term success in HEC will require sustainable strategies, products, and strong, enduring partnerships. It is our hope that this report will assist the federal government in developing new ways to encourage such strategies, products, and partnerships.

We trust that this report will be helpful to public and private sector leaders as they work to revitalize this nation's high end computing strategy, and that the findings and recommendations contained in this report will further the national debate and cement the partnerships critical to advancing high end computing.

Paul Lawrence
Partner-in-Charge
IBM Center for The Business of Government
paul.lawrence@us.ibm.com

Anne Altman
Managing Director
IBM U.S. Federal
Altmana@us.ibm.com

EXECUTIVE SUMMARY

Government policy on High End Computing (HEC) in the United States is at a crossroads. After pioneering the field of HEC and dominating it for several decades, recent developments on several fronts are challenging the nation's community of scientists, policy makers, and industry to devise a path forward that continues to address the country's scientific, economic, and strategic needs at a reasonable cost. Over the last few years, government agencies and task forces have dedicated a significant effort to articulating those needs for different constituencies, producing reports and recommendations for action in this area. The aim of this report is twofold. First, it provides a summarized mapping of the key issues for HEC policy. Second, it offers findings and recommendations relevant to HEC policy. After consulting the documentation available in the public domain and interviewing individuals in various government agencies, academia, and industry, we found that the key issues in HEC policy are the following:

- **HEC has a special role in science and engineering because of the critical importance of simulations and other computational approaches.** The growth and increasing complexity of cutting-edge science, including defense and security applications, demands an integrated environment of computational capabilities. Therefore, HEC deserves special attention by policy makers as the new realities demand new policy approaches that align with the evolution of the research and development (R&D) system of the nation.
- **The evolutionary paths of both HEC technology and most fields of science and engineering are deeply intertwined.** Therefore, HEC policy must allow for the generation of sufficient diversity in HEC innovations so that technology choices are not arbitrarily curtailed prior to a fair assessment of their value for the overall environment of science and engineering.
- **The agenda for HEC policy cannot be set according to strict ideological commitments either to the market or to the mission of government agencies.** The perceived problems with the current situation are, in a way, due to too much of a good thing. With the mainstreaming of supercomputing during the 1990s, little policy attention was paid to HEC because it seemed market forces alone would address needs in the area. There are now many unmet needs, and HEC policy must design an institutional arrangement that allows each sector to do what it knows best to enhance opportunities in the innovation system.
- **The leadership of the United States in HEC is understandably a driver of the policy discussion.** However, HEC policy must not be dominated by a short-term crisis perception. Rather, the field needs a longer-term pragmatic leadership vision that is derived from the political articulation of national goals and focuses on the abilities required to achieve desired results.

To address these issues we offer the following recommendations:

Recommendation 1

Make HEC policy an explicit and integrated component of the national goals articulated in contemporary political deliberations.

What is needed is a balanced HEC environment that contributes to all the relevant national goals in a coordinated and consistent fashion. It is not possible to succeed by isolating an area and picking an HEC component or set of projects to address it. National goals—such as national security, for example—are tied to other national goals in one way or another and so are the various components of HEC. Without the political leadership to articulate such a vision, it is not possible to design an HEC policy that will succeed in the medium to long term.

Recommendation 2

Create a high-level coordinating entity for HEC that has enough power to overcome the zero-sum game dynamics that plague policy in this area.

The decentralization of American R&D across various federal agencies that compete for resources has worked well in many ways for the U.S. innovation system. It also has serious disadvantages reflected in the case of HEC. On the one hand, it tends to force stereotypical forms of division of labor, such as charging one agency with the development of new hardware architectures and another with application software, with no provision for their coordination or compatibility. On the other, it may lead to unnecessary duplication. The coordinating role has mostly been played by ad hoc panels and task forces with some success. There are numerous panels and councils that have coordinating roles but depend largely on the special abilities of individuals to bridge the gaps in institutional design. The HEC policy agenda, to the extent it can be said to exist, is fragmented and has no high-level institutional champion. HEC needs its own "national initiative," its own high-level coordinating body, a sustained policy and coordination effort, and the continuing attention of the Office of Science and Technology Policy (OSTP) and high-level representatives from industry and from the many federal agencies having a crucial stake in the development of HEC.

Recommendation 3

Implement an HEC policy that addresses the incentives of researchers in the different sectors—government, academia, and industry—to explore alternatives in hardware and software and avoid either premature "lock-in" by suboptimal technologies or premature abandonment of immature technologies that may still hold promise.

On the surface of things, it may seem that, in the end, the solution to all science and technology policy issues is to give everybody more money. In this view, policy "issues" arise in science and technology only because resources are finite and we must decide what to fund. However, that is not the argument behind this recommendation. What matters is strategic support to insure that more choices will be available down the road.

Recommendation 4

After creating the institutional and policy mechanisms for a National Supercomputing Initiative, build into the policy the time frame factor in this field and provide for a periodic, objective assessment of progress in achieving national policy goals.

HEC is a moving target. Today's high end is tomorrow's mainstream, and newly implemented architectures and software applications remain at the true high end for periods that last only a few years. The predicament the field finds itself in today is in part due to lack of attention to its changing dynamics. The new challenges that will emerge in the medium to long term should not have to be confronted without the benefit of what has been learned today. Therefore, the policy should anticipate some of the possible "forks in the road" that will demand an update of HEC policy. We recommend the appointment of a National Research Council (NRC) panel to provide a yearly assessment of the progress in achieving goals set forth in a National Supercomputing Initiative, complete with recommendations for changing policy directions that fail to live up to promise, including recommendations for changes in federal funding for HEC.

Introduction

Common sense indicates that technologies are essentially means to achieve desired ends. We use transportation technologies, for example, to move things we cannot using human strength alone or to go longer distances and achieve greater speeds than humans can on their own. This is generally the way policy understands technologies, too. Technology policy commits some of the nation's resources to acquire or develop technologies deemed necessary to achieve national goals in defense, health, or economic development. In other words, technology policy seeks to answer the question: "What technological means do we need to achieve agreed upon national goals?"

This simple, common-sense way of understanding technology and technology policy is no longer appropriate for the complex role technologies play in contemporary society. There are two main reasons for this. First, national goals are generally broad and the object of constant political discussion. Therefore, there is no simple logical way of determining what technologies are better suited to achieve them. For example, the goal of maintaining the health of the population can be served by developing drugs for treating known diseases or by developing new fuels that do not contain poisonous chemicals. A choice between the two seems misplaced, and the second can hardly be considered just a tool to achieve a health goal. Second, most contemporary technologies are no longer isolated devices that are applied to a single function. They exist in the context of systems in which our social life is embedded. Therefore, technological choices are not equivalent to selecting a single tool for a job but instead involve an aggregation of choices related to all the activities associated with the system of technology we are referring to. For example, in choosing to use cars as a means of transportation, we also choose certain possibilities for urban development that include freeways and suburbs, and networks of gas stations and so on. As a result, technology policy requires, first, an explicit political articulation of the national goals that are relevant at the moment. And, second, it involves developing the context and rules for an aggregation of technological choices rather than selecting a single technological tool.

High End Computing constitutes an excellent example of these two dimensions of technology policy. It is tied to critical national goals in national security, economic competitiveness, and scientific advancement that are broad and the subject of much political deliberation. And it is a complex set of technologies that cannot be identified with a single type of device.

Historically, U.S. science and technology policy has connected national goals to science and technology choices according to three different paradigms: (1) market failure, (2) government mission, and (3) cooperative technology.[1] The first is based on the notion that markets constitute the most efficient way to distribute goods and services. Therefore, government intervention in science and technology is warranted only when the markets do not provide proper incentives for innovation by private enterprises. The second gives government a role in scientific and technological development to perform clearly defined government functions such as defense and public health. The third emerged

Definitions of "High End Computing" (HEC) and "Supercomputer"

There are certain sorts of problems, such as nuclear weapon simulations, climate modeling, cryptography, and other cutting-edge science applications, that still impose such enormous demands on computing resources that the arrangements of hardware and software set up to tackle them are set apart from the more widely available computers and programs, even if they share some features or components. The achievement of the highest performance in computing is a function of many dimensions of the hardware/software arrangement including vast amounts of very fast access memory, high-speed communication paths internal to the computer and between it and its peripherals, special architecture solutions using multiple processors, and software that is designed to use all of these special features in an optimal way for the problems it must solve. This emphasis on achieving the highest possible performance of a computing arrangement for the solution of certain types of problems is what defines High End Computing (HEC).

HEC is a moving target. The highest performance achieved today may be commonplace in tomorrow's retail machines. But at any point in time, there will be a set of problems that demand faster solutions than those available, and arrangements of hardware and software must be conceived to achieve them. The pursuit of the highest possible computing performance for any set of problems of interest is what defines High End Computing.

Another way to focus the discussion is to refer to "supercomputers." These are the fastest available machines according to certain benchmarks accepted by the community. In general, these measures are based on some type of mathematical computation that is sufficiently general to indicate the speed with which the machine would calculate solutions to most problems of interest. These machines are termed "supercomputers" because they are several orders of magnitude faster than the computers available to most ordinary users. When the fastest machines are custom built at very high expense, ranging from the tens to hundreds of millions of dollars, the distinction between supercomputers and all other computers is relatively easy.

During the 1990s, one remarkable development was the emergence of commercial off-the-shelf (COTS) based machines that made very high-speed computers available at a fraction of the cost using components also produced for servers and other computers in the commercial market. This development blurred the boundary between supercomputers and other machines. Furthermore, one approach to the achievement of high end computing is to share many ordinary machines on a network or grid and develop programs that allocate portions of the problem to each one and then aggregate their contributions to obtain the final result. Many important scientific problems are amenable to such approaches, but many more are not. The difficulty in making these distinctions suggests that the focus on individual fast machines—i.e., supercomputers—is not the only aspect of interest in this area.

during a period of perceived crisis for American industry in the face of foreign competition, especially from Japan. As Japanese industry succeeded in capturing high-technology markets with government support of technology development, it seemed the market-failure rationale was not sufficient to answer the challenge, and cooperation between sectors and among rival firms was encouraged. HEC is one of the main high-technology areas targeted by Japanese policy. In response, most U.S. policy in HEC from the early 1980s to the mid-1990s has followed the cooperative technology paradigm. The aim of maintaining or enhancing the competitiveness of American industry in HEC during that period was achieved. Today, U.S. industry controls roughly two-thirds of a market for supercomputing that has grown significantly to include many industrial applications as well as cutting-edge science.

It may, however, be too much of a good thing. With this success, the last several years have seen U.S. policy revert to a market-failure approach. Furthermore, most recent HEC policy has assumed the markets are working well and has ignored many critical dimensions of the HEC field. The lack of attention to HEC is perceived by many to threaten the leadership position that the nation has enjoyed since the first supercomputers were invented.

Our report focuses on the issues and controversies presently surrounding U.S. policy options for HEC. We employ a variety of sources for our analysis, including published articles and reports, historical accounts of the development of HEC, and interviews we conducted with HEC experts in industry, government, and academia. We do not identify these experts, except by the most general descriptions. In some cases the respondents asked explicitly that we not identify them, and in other cases that we not attribute statements to them. We were also quite impressed by the difficulty of obtaining interviews. In several other projects[2] we have conducted similar interviews among scientists, engineers, and public- and private-sector policy makers, and we have had little or no difficulty obtaining interviews, and among those interviewed we have had very few ask that their comments remain anonymous or on background. This experience was different. The fact that so many of our potential interviewees were not eager to participate seemed to us a signal that we should protect the anonymity of those who did, including those who did not explicitly ask to remain anonymous.

Another limitation of this study is that we cannot assess the merits of arguments about the specific scientific fields in which HEC is applied nor can we add any information to financial statements and other plans circulated in the government and other HEC-interested communities. We focus on the multiple paths of policy deliberation that seem to be in place and how they weigh on the reality of policy making in the current institutional context. We also look at the conceptual models and their implications found in the deliberations and how they impact the chances of a coherent policy in this area. Ultimately, our study is a description—and in some instances an interpretation—of what various communities related to or interested in HEC say they need.

This report will take most of the technical considerations of machines, software, and computational science research areas presented and described in other reports and published documents at face value. The authors claim no specialized technical expertise in computing or the computational research areas to assess claims internal to those fields. However, the implications of such claims for policy and the presentation of arguments that lead to competing claims at the policy level are the rightful and main concern of this report. Specialized technical concerns are frequently translated into non-specialized language when the policy implications are derived. Therefore, it is possible to assess most of the arguments for compatible or competing demands on policy making without distorting their internal content.

Acronyms

ACP	Advanced Cyberinfrastructure Program
COTS	Commercial off-the-shelf
DARPA	Defense Advanced Research Projects Agency
HEC	High End Computing
HPCC	High Performance Computing and Communication
IHEC	Integrated High End Computing
NNI	National Nanotechnology Initiative
NRC	National Research Council
NSF	National Science Foundation
NSC	National Security Community
OSTP	Office of Science and Technology Policy
R&D	Research and Development
SAPCWG	Systems, Architecture, Programmability, and Components Working Group

Following this introduction, our report examines the following issues in HEC:

- Does HEC deserve greater attention by policy makers? On what grounds?
- What are the key features of HEC that should be addressed in designing policy?
- What are the roles of the market and government in setting the agenda for HEC technology policy?
- What does U.S. leadership in HEC mean and is it in crisis?

In the concluding section we summarize our findings, present recommendations for addressing these issues, and analyze the perception of crisis in HEC development and policy. We frame the discussion of issues and policy options using a simple model of the equipment intensiveness of research.

Key Issues Facing HEC

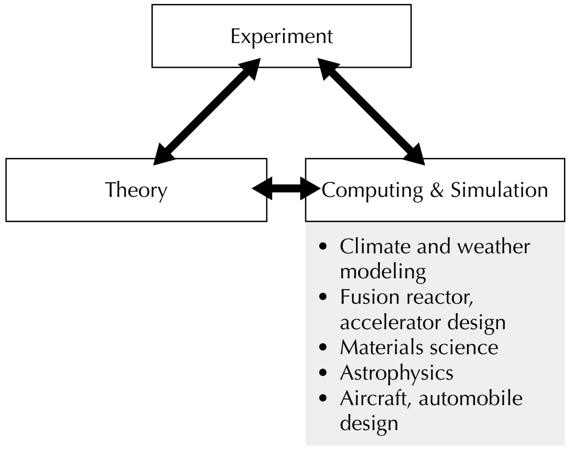
The Importance of HEC to U.S. Research and Development Policy

The simple answer to the question of why HEC deserves greater attention by policy makers is that HEC now plays a central role in contemporary science and engineering. Computers are often thought of as support machines that any user will always want more of. Having the newest gadget or the most recent version of a machine may seem to be something scientists and technologists will always wish for and cannot be expected to be objective about. Therefore, in times of tight budgets, the arguments for increased spending will receive more careful scrutiny, and policy makers will focus on tough questions about the real need for spending more money on new computers.

However, this obvious sense in which computers are tools in science may obscure the fact that HEC has acquired an intrinsic role in the knowledge creation process that is without precedent in research history. In the last 20 years or so, increased computational capabilities have made possible realistic simulations of physical processes, such as nuclear explosions and atmospheric phenomena, and of complex technological systems, such as airplanes, cars, and nuclear reactors. The list of areas of scientific research and technological development in which these simulations are critical is indeed lengthy, and most government agency reports on HEC contain detailed descriptions of them.[3] These simulations provide a new way to grasp reality by making manipulation of physical systems possible either more cheaply, such as in the case of testing a car before it is built, or where it was earlier impossible, such as simulating a storm that cannot be reproduced in a lab or simulating the full set of phenomena unleashed by a nuclear explosion. The importance of these simulations for scientific research has elevated computer simulation to the status of theory and experiment and has been termed the "third mode" of science and engineering.[4]

Using the language of philosophy of science, we can say that simulations have become established as a scientific activity with an *epistemic status* of their own. In other words, HEC is not merely very fast number crunching. It allows scientists and engineers to interact with nature or large technological systems in such a direct way that the validity of their conclusions is as good as a laboratory experiment.

Figure 1: Computer Simulation—The Third Mode of Science and Engineering



Scientific advancements in the last decade have placed demands on simulation capabilities that result in more complicated relationships between hardware and software. It is no longer possible to think of HEC as isolated machines and programs cobbled together to solve individual problems. Even though it is still possible to buy an individual supercomputer and load programs on it to do certain things, the multiplication and interdependence of scientific activities relying on simulations requires an integrated approach to the computational resources the scientific community relies upon.

From the point of view of the health of the nation's system of research and development (R&D), addressing HEC needs in piecemeal fashion, one machine and one program at a time for each individual project or center, is no longer adequate. Therefore, the division of labor in science that gave different government agencies implicit leading roles in different areas of research seems no longer to be an adequate basis for developing the computational needs for the scientific activities under the aegis of each. With such division of labor, under the best of circumstances (i.e., an abundance of resources) we would have parallel paths of development in computational resources. Under more realistic or even difficult circumstances, with budget reductions, a lower-cost common denominator solution to the problem of computational resources might be the result. With a more complicated horizon of scientific advancement, the need for better adjusted computational resources for the problems at hand suggests a computational environment that relates multiple developments in science to a variety of alternatives in computational capability. The institutional correlate of this situation is a true coordination of efforts in all aspects of the resulting policy, from program design and funding categories to procedural issues such as procurement.

A recent detailed presentation of this argument is contained in a blue ribbon report for the National Science Foundation (NSF) that states:

> Digital computation, data, information, and networks are now being used to replace and extend traditional efforts in science and engineering research, indeed to create new disciplines. The classic two approaches to scientific research, theoretical/analytical and experimental/observational, have been extended to *in silico* simulation to explore a larger number of possibilities at new levels of temporal and spatial fidelity. Advanced networking enables people, tools, and information to be linked in ways that reduce barriers of location, time, institution, and discipline. In numerous fields, new distributed-knowledge environments are becoming essential, not optional, for moving to the next frontier of research.[5]

Science and engineering have advanced to the point that many computational resources and approaches are developing in fields across the disciplinary spectrum with no coordination or standardization. If this process is allowed to develop much further, many incompatible systems will be in use in the scientific community, and many artificial boundaries between fields will arise as researchers' skills are tailored to them. Many areas of research will suffer as opportunities for interdisciplinary research are lost to the lack of awareness of each other's results, lack of compatible information gathering and storage formats, and waste of time and financial resources on software that serves very narrow purposes. It is not a matter of whether HEC resources will be applied across the spectrum of scientific disciplines, but how that will happen. The academic community, and NSF in particular, is calling for coordination in building a "cyberinfrastructure" so that the net result is a unified system enhancing opportunities for scientific communication and sharing rather than a fragmented one with multiple semi-autonomous fields that fail to stimulate and enrich each other.

In sum, the answer to the question that constitutes the first issue raised for HEC in this report is that due to the special role that simulations and other computational approaches now play in science and engineering, HEC deserves special attention by policy makers because the new realities demand new policy approaches that fit the evolution of the R&D system of the nation. As further discussed in the sections that follow, HEC cannot be approached as the mere supply of localized resources to projects and centers. Rather, an integrated and coordinated policy is necessary in HEC to allow for both the continued advancement in a growing number of fields of science and engineering and the exploration of new HEC developments on most fronts.

The Evolution of HEC Systems and Applications

Just as science and engineering have developed recently and grown ever more critically dependent on computational modeling and simulation, HEC machines and applications have had a progression of their own over the last three decades. Until the early 1990s, all HEC machines were custom-built individual computers using the most advanced architecture solutions and specialized components. With the amazing advance of microprocessor technology, a new approach to HEC machines emerged that connected commercially available microprocessors in parallel, achieving HEC performance at a fraction of the cost. Even newer approaches to computing such as teragrids, based on networks and distributed computing, and quantum computing, taking advantage of properties of materials at the particle level, are on the horizon now. However, the two paths of development—that of HEC and of science and engineering fields themselves—are intertwined and growing more integrated into each other than ever. So the assessment of HEC capability for a broad set of needs is what is really important. At the same time, it is complicated by the fact that it has become a multidimensional problem.

It is clearly impossible to accurately map the future of every aspect of the leading edge of science and technology for making exact, forward-looking policy decisions. However, we can draw on an analogy with evolutionary thinking in the biological sciences to understand how wise policy decisions can be made. It is widely accepted that technological development often follows evolutionary paths, analogous to a species in an ecological niche.[6] Therefore, the generation of new and improved technologies is analogous to the emergence of new species better adapted to different ecological conditions. For this to happen, just as in the biological realm, the environment must provide the resources to insure a continuous ability to generate mutations from which the evolutionary process may select.

This reasoning shows the importance of various types of investment in the technological development process. Once the environment is sufficiently complex that a single device can no longer be expected to satisfy "survivability" conditions, it is necessary to insure that enough activity in the field is ongoing to allow for a broadly based selection process to take place. In other words, from the point of view of the overall socio-technical system, the investment in HEC technological development should not just be "intra-species," that is, in existing types of machines. There should also be investment in the overall mutation and selection process from which multiple technological species may evolve. Ultimately, this is what the user community is calling for. Avoid settling for a single species that has already adapted to a particular ecological niche. Address the overall ecology of HEC and sustain the mutation and selection process that will keep the entire environment healthy. The National Security Community (NSC) report states this quite clearly:

> The Systems, Architecture, Programmability, and Components Working Group (SAPCWG) also recommended increasing the number and level of advanced development efforts. This would facilitate the transition and combination of promising applied research ideas from the laboratory into subsystem prototypes and concept test beds and promote healthy competition amongst good ideas. Additionally, the team recommended increasing the number of engineering and operational prototype efforts supported by the national security community from one to at least two. The SAPCWG reasoned the broad diversity of the community's HEC needs could not be adequately served by a single development.[7]

There is a strong argument from HEC developers and researchers that exploration of new architecture alternatives and opportunities should be encouraged, though it is now discouraged in several ways. One HEC researcher we interviewed stated:

> *We have an uncoordinated and anemic effort to develop new architectures. From a computer science standpoint, there is no serious effort to discuss architectures.... We cannot depend on efforts by vendors. Even IBM cannot explore architectures across the board. The government needs a policy supporting architecture research.*[8]

The clear implication of this reasoning is that the critical level of decision making for future technological development in HEC is at the policy level because it is the only perspective that can take the overall HEC environment into its purview and provide for the health of its evolutionary processes. For example, if the vision for this overall environment were that only HEC technologies that are commercially viable today must survive, then the implicit policy decision is simple: Let businesses make all the choices to suit their bottom line. Or, if the vision for this environment were to insure defense superiority, then let the defense leaders specify and decide what HEC technologies to pursue. However, we have established from the beginning that due to the role of modeling and simulation in science and engineering today, all aspects of national life critically dependent on cutting-edge knowledge are, by extension, critically dependent on the health of the HEC sector. And, given that it is not possible to predict exactly what advances in science and engineering will move beyond the extremely short term, policy decisions are not direct technology choices but support for a complex environment that selects HEC advancements as the particular ecology of each field requires.

Complementarities and Tensions between Hardware and Software

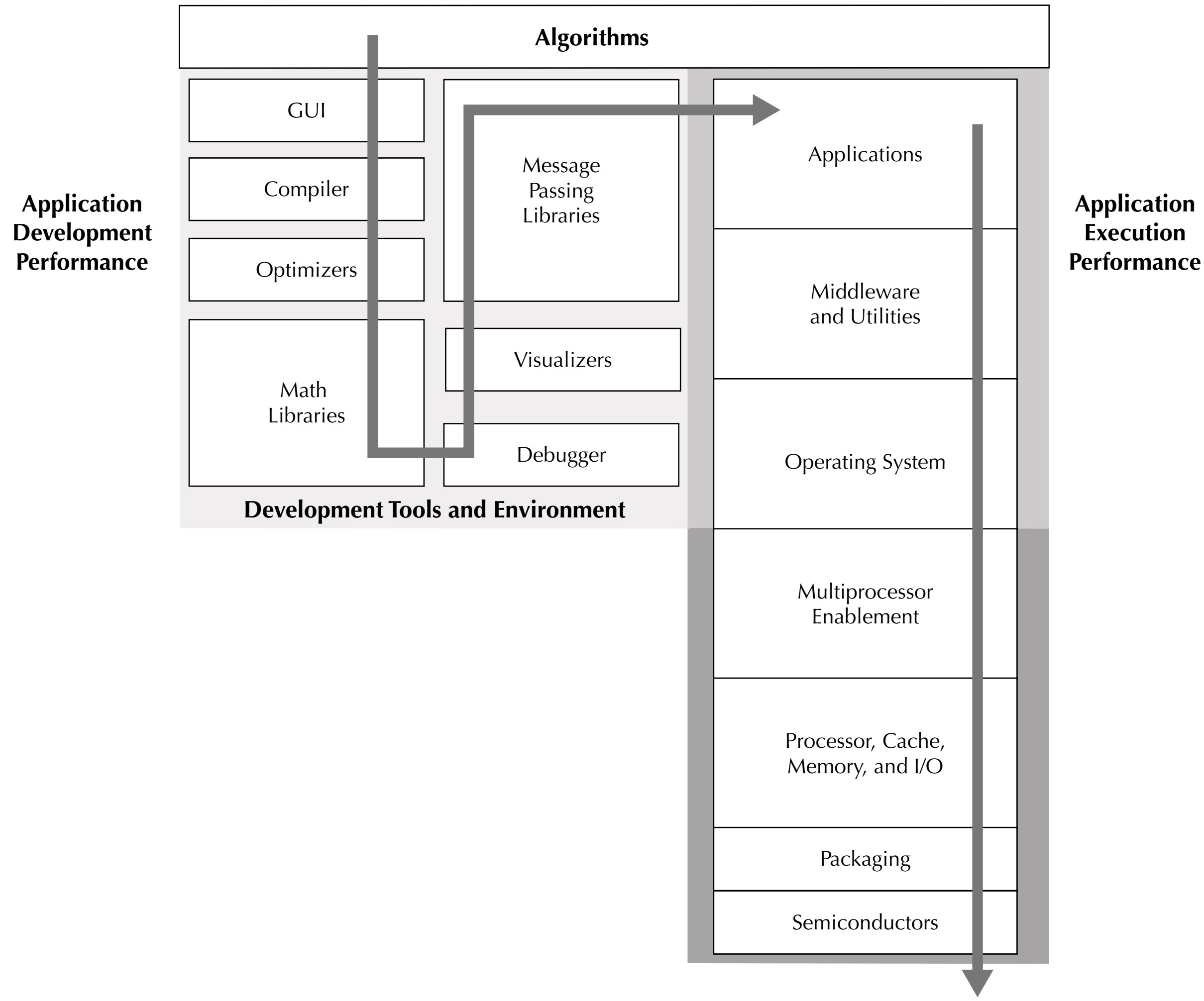
Within the larger environment that sustains HEC development, the relationship between hardware and software is critical. It is known across the field of computing that the boundary between hardware and software is very diffuse and changing. In the specific case of HEC, some of the arguments have become stereotyped by affirming too strict a divide between the two. Some claim that too much attention is being paid to hardware at the expense of software for useful applications, which are what matters in the end. Others claim that most of the problems faced by software developers for HEC applications have feasible solutions if architecture opportunities that are at hand were allowed to proceed.

In some cases, technologies are (fairly fixed) means toward achieving higher policy goals. However, in cases like HEC, the changing technological environment must itself be shaped by carefully designed policies. Often politics and policy making requires that issues be simplified, categorized, and condensed to meet the small window of political attention that one issue has in the context of many other policy issues at hand. However, the importance of hardware, software, and other elements in designing and building HEC systems must be viewed in light of the complexities. The critical importance of all elements must be understood for true HEC advances to be made. If this is not accomplished, the gradual and varying boundary between hardware, software, and other elements will become much sharper and fixed than it really is. Figure 2 describes the elements necessary to advance high end computing.

The structure of practitioner communities compounds the problem. Even though many HEC researchers understand the interdependencies, they are still specialized enough to conceive of their needs in competition with each other. There is still much more identity formation around "hardware" on the one hand, and "software" on the other, to make community-wide problems like HEC policy concerns difficult to grapple with. *"People are applications oriented, software oriented, not comfortable with hardware,"* one HEC researcher said. Another said, *"A huge unidentified cost is changing architecture on people, so it introduces a huge lag of about four years. During that time they do technology adaptation, not science.… Sustaining large projects with large amounts of code on several architectures is not a trivial task."*

As a result, with budgetary constraints, the short-term goals of making available HEC machines that are more usable, reducing the time-to-solution in critical problems in academic science, commercial product development, and defense and security applications, would focus on software improvements keeping hardware relatively constant. Then medium- and longer-term goals aiming at new architectures and hardware that change the conditions for software development down the road will be competing for resources. This tension will arise because the categories toward which the policy discussion converges, given the factors mentioned above, generally make software, hardware, and other elements distinct alternatives rather than an essentially unified set of problems.

Figure 2: Components of High End Computing



In sum, a critical policy issue for the HEC environment is to insure that hardware, software, and other elements can co-evolve without arbitrary constraints resulting from zero-sum games in the budgeting process.

Government's Role in Setting the HEC Policy Agenda

The Role of Government in Addressing Market Failure

Most of the arguments for government policy in HEC are a very direct application of the market-failure rationale. As this sort of argument goes, the incentives for commercial development of what is needed to satisfy a social need may not be well structured and, therefore, we would expect private industry to underinvest. For example, the policy argument for the cyberinfrastructure is as follows:

> The Advanced Cyberinfrastructure Program (ACP) requires government investment in research and development of cyberinfrastructure technologies (principally software) for several reasons. First, the marketplace underinvests in long time-horizon research. The cyberinfrastructure and application technologies are within the domain of NSF responsibility for government-funded research, and the ACP will maintain U.S. leadership in these technologies through research, experimentation, and commercialization. Second, infrastructure and applications suffer from a chicken-and-egg conundrum that infrastructure requires a diversity of successful applications for its commercial viability, while commercial applications target only widely deployed infrastructure. This ACP will follow the successful model of the Internet, with targeted and coordinated government investment in infrastructure and applications, experimentation and refinement in actual uses, and coordinated commercialization of both elements together. Third, while we expect many if not most of the technologies developed in this ACP to be of broad applicability, science and engineering research has special needs in functionality, performance, and scale that are unlikely to be fully served by commercial firms, at least not without government assistance.[9]

Some of our interviewees suggested that recent developments in the HEC field can in part be attributed to the role of government and industry in the development of the Internet to the exclusion of other technological needs such as HEC. The U.S. government was a significant facilitator of the creation of a market for high technology and all sorts of services delivered via the Internet during the 1990s. This was the period during which the gap in the development path of HEC computers and the needs of the user community diverged and significant underinvestment is said to have occurred. People's attention was so taken by the sensational economic and social phenomena associated with the Internet in the 1990s that it was hard to communicate that HEC had significant strategic importance.

One role the government has played in HEC in the past was merely to provide specifications and issue procurement requests, leaving the rest of the evolution of the market to its own forces and the ability of industry suppliers. All voices speaking about these matters seem to agree that now government must do more than that. It must get involved in carefully designed partnerships with industry developers and suppliers as well as academic researchers to support an HEC environment on many fronts. One of the objectives of the Integrated High End Computing (IHEC) program proposed by the NSC report is to:

> Assure a healthy domestic high end computing research and development environment and production capability focused on national security.[10]

The cyberinfrastructure proposal by NSF contains similar statements. An HEC researcher said, *"The key is well-placed money in a well-balanced environment."*

Having said this, however, the reports and commentary that have been circulating during the last year or so, triggered in part by the Japanese Earth Simulator, point to a shift in the outlook for policy making in this area. Prior policy was based on the idea of "grand challenges" in science and technology. It was predicated on a science market rationale that then led to agency division of labor for funding supercomputer hardware and software projects to solve the grand-challenge scientific problems in the agency's area. The particular scientific problem focused the effort for a while and gave it its own time frame based on scientific competition for being first to publish in that area. Over the period dominated by this policy rationale, efforts were clustered around these problems both over time and in content.[11]

The grand-challenge approach to HEC policy contributed to many successes both in science and in the emergence of new HEC technologies. Therefore, the contemporary predicament is related to the failure of prior policies not because of their conceptual design but because of the actual allocation of resources that was made in its implementation. One HEC researcher, referring to the High Performance Computing and Communication (HPCC) program of the 1990s, said:

> *People cared about HPCC but it ended up "repainting" the money. Agencies said, "You are taking our money and giving it back to us." DARPA [Defense Advanced Research Projects Agency] said it was already doing what HPCC required them to do. NASA was doing part of the software side of it. So HPCC failed because there was no new money.*

The dominant request today is to fill in the spaces in time and content that the grand-challenge approach left empty. This approach leads to an effort that is integrated across application areas and sustained in time over the long term. If the prior approach already consisted of several moving targets, the new one has carried this logic to its finality. All dimensions of leadership and all possible needs must be continuously pursued, redefining policy goals with the needs and applications of the next stage determined by the integrated, high-level mechanism that must be put in place.

Assessment of this new situation calls for restoring the original intent and level of funding in HEC.[12] This means that the problems of computational research must be looked at again as a full set of problems deserving new and fundamental approaches. The new approach must break out of a seemingly self-reinforcing logic given by the evolution of the markets and technology during the past decade.

There are areas of application that are very important but are not large enough to command the resources to pursue their own path in computational development. The government has a clear role in developing those islands of demand into a workable market. According to one of our interviewees, the medical establishment, for example, has many needs and opportunities for the use of computational science but lacks the coordination and the culture to overcome its commitment to established laboratory approaches and therefore does not take advantage of simulations.

HEC User Community Perspectives

Academia

Over the last decade, one of the objectives of earlier R&D policies in HEC has been achieved, namely, the broadening of the market for supercomputing to non-defense applications.[13] The expansion of this market has been enormously beneficial for U.S. industry, both for vendors and users of supercomputers, by bringing the price of high-performance machines down dramatically and enhancing numerous industrial application areas.

However, the niche areas related to defense from which the machines emerged and were maintained for many years are claimed to have suffered by losing visibility, and some of its needs are now going unmet.[14]

On the other hand, the academic community has always felt it lagged behind the defense community in HEC. For a brief period the NSF supercomputer centers of the 1980s put the academic community on a par with their defense colleagues. Two decades later, a renewed focus on HEC is required on basic research on supercomputer architectures. Most efforts by academics are going into much more futuristic areas such as quantum computing. As a member of HEC policy circles put it:

> *Another part of possible policy is that the academic community has pretty much abandoned research into HEC in the near term. There is quantum computing, "wet" computing, demonstrating computing techniques using DNA. Most of the academic community has moved into those areas and abandoned the near term.*

Federal Government

While some participants in HEC deliberations point to some sort of market-failure argument to justify government intervention in HEC, all appeared committed to a healthy market for American industry. Beyond the differences in the specific market failure each part of the community may be interested in, from large machines with custom architecture to various software components, every proposal plans to draw on private industry for those products it already has available and expects the end result of government intervention to lead to a new healthy market for private industry once the incentive for underinvestment is overcome. In other words, there are no serious proposals in HEC for long-term government subsidies that do not take great care in assessing the impact on the information technology markets for American industry. The long-term government commitment to HEC that most proposals contain is believed to contribute continuously to the health and capabilities of the American HEC industry. The NSF report perspective is typical:

> The vision of an ACP cannot be achieved by procuring existing commercial technologies alone. Of course, to the extent that commercial technologies and services are available off the shelf, they should be incorporated. But information technology is hardly mature; in fact, it is always evolving toward greater capabilities. Its applications are even less mature, and there are many opportunities to mold it to better meet the needs of end users.… Thus, research in new information technologies and applications utilizing those technologies often have important commercial spin-offs. This situation is illustrated by supercomputing, first applied to scientific and military applications and later to many commercial purposes.[15]

While articulating the HEC needs of the national security community, the NSC report states the market failure argument in similar terms:

> … The computational needs of technologies critical to national security are likely to continue to grow at a faster rate than is the development of new computational capabilities. And, given the current focus of the HEC industry, without government intervention this trend is expected to accelerate.
>
> However, the SAPCWG also found that this unwanted trend can be reversed. The creative talent and skills needed to revitalize the high-end supercomputer industry are still resident in the U.S.A. New, more powerful, and needed capabilities can be developed with prompt action. Hence, the SAPCWG also made recommendations as to how the government should engage academia and industry to deliver significantly increased HEC capabilities. The team proposed a spiral model starting with applied technology research, proceeding through advanced development and prototyping phases, leading to vendor products that support national security requirements.[16]

Sometimes the promise of commercial spin-offs and the potential for large profits to industry in the future is waved in front of policy makers without a solid basis. For a long time, it was assumed that any investment in basic research led inexorably to commercial products and profits down the road.

There is a healthy skepticism about this today since it has been hard to prove the connection between many investments in research and specific commercial payoffs. However, the antecedents of the Internet and today's supercomputer market are more favorable in that regard. In these cases, it is not limited to the profits derived from specific commercial products but to the emergence of complete new markets that did not exist before. There is obviously no guarantee that it will happen again with new investments in HEC technology, but the relationships between government, universities, and industry to exploit opportunities as they arise have become much more widespread and effective than in the past.

Competing Interests in Setting HEC Policy

The relationship between the market for HEC and the high-end needs of various sectors of the user community has some tensions. For instance, according to the NSC report, commercial vendors must make design trade-offs to supply the server market. This sometimes diminishes industry's ability to supply the higher-end needs of the national security community.[17] This would also be true, by extension of the same argument, for other high-end users of the broader scientific community, as we know from talking to academic scientists and researchers.

Such tensions, compounded by a sense of urgency brought about by the success of the Japanese Earth Simulator, have sometimes made for some heated discussions and comments sharply derogatory toward American commercial vendors by prominent computer specialists, who suggest that American companies will be hard-pressed to develop machines that match the performance of the Earth Simulator.[18] Others similarly deride the commercial off-the-shelf (COTS) approach by pointing to their connection with web servers:

> *It's just not going to work any longer to leverage web servers for scientific research.*[19]

The questions that some researchers and agency program managers raise have to do with the details of how the current markets are working. Many industry players in the supercomputer market have significant commitments to massive parallel architectures based on COTS components. Therefore, the assertion that there is a need for custom-designed alternative architecture machines is the subject of some disagreement. The objective for industry, understandably, would be to question the real demand and cost-effectiveness of non-COTS machines. Researchers and agency managers with projects that explore non-COTS alternatives emphasize the inability of the current industry suppliers to solve their applications problems. In many ways these arguments have been played out in the popular press over the past year.[20] An agency manager we interviewed exemplified these exchanges:

> *No amount of money in IBM's hands or Sun Microsystem's and Hewlett-Packard's will give them the memory bandwidth we need.… If you have applications that aren't communications intensive, okay; otherwise, it isn't going to work using bolted together commodity network components.*

Historically, dialogue and consensus building have diminished these types of debates. In fact, one HEC researcher observed that industry questioned the COTS approach before it became a market success. The feasibility of the COTS approach had to be demonstrated through research and partnership building by government, universities, and industry.

Another issue is the role of business models in advancing HEC. In a recent document circulated by IBM, their HEC strategy is to develop systems that satisfy the needs of customers in government, academia, and private industry. In contrast, many vendors who have dedicated most of their efforts to provide machines for one customer—i.e., specialized needs of government—have gone out of business over the last decade or so. In this sense, most industry leaders don't believe the Japanese Earth Simulator has much market significance. By extension, efforts at developing custom design architectures are "single trick ponies" doomed to fail in the market.

On the other hand, the NSC report posits that the Japanese firm NEC has received a significant boost in the HEC market as a result of the Earth Simulator and that it is well positioned to market smaller scale versions of the systems (the NEC

SX-6) with non-recurring development costs having been covered by the Earth Simulator project.[21]

In sum, it is not surprising that this discussion has taken the shape it has given the key goals and incentives for each set of participants. At the same time, they all recognize they need each other to continue their work. What is certain is that U.S. government, industry, and the user community will have to work in partnership to advance HEC. Industry needs the partnerships with government and academia, in spite of the pressures from them to explore things that tend to break their business model. Government and academia have no other way to satisfy their computational needs without the products and services industry is able to develop and support. An HEC researcher we interviewed stated this well:

> *We say we need disruptive technologies: It breaks their market model. They say, like [President] Reagan, "we want the status quo to stay the way it is." But we cannot do it without their help.… The role of government is to make partnerships and invest in the weak spots while the giants continue to work at their strengths.*

As we requested interviews from various members of the HEC community, we were surprised to find that several potential interviewees were very keen to understand our relationship with the commercial vendors and whether their identity would be revealed and comments made known to them. Many did not want to jeopardize their relationship with commercial vendors. This seems to indicate that members of the non-commercial HEC community feel the influence of commercial vendors very strongly and, to a certain extent, sacrifice academic freedom to maintain the relationship with them. We believe this is an example where a clear policy framework and a strong government/university/industry partnership is extremely important to insure that no technological exploration by any partner is closed down prematurely or not considered due to the perception that it may alienate another partner in an asymmetric relationship.

The conceptual discussion over the best policy on HEC is affected by the fact that most participants anticipate the end game of the policy deliberations around budget allocations. The interplay of various emphases that stretches or shrinks the inclusiveness of perspectives with respect to each other, in an effort to maintain a positive-sum game to build in broad-based support for HEC, changes sharply when strict budget limitations revert the entire discussion back to a zero-sum game. It may then be true that if cyberinfrastructure dominates the policy discussion at that stage, there will be money available for high-speed networks and teragrids but none for new high-end architecture research. Or the proponents of particular HEC projects may succeed in getting their bits and pieces of the overall landscape funded, but no specific support is awarded for coordinating or integrating the efforts. In other words, the interpretation frames for various policy documents and statements must include the fact that participants will offer their views with an eye to how their main concern will be positioned for the final budgetary decisions. (For a more detailed discussion of cyberinfrastructure, see the Appendix.)

From this discussion it is clear that U.S. policy in HEC cannot be limited to addressing isolated technological choices. It must create an institutional framework that avoids the detrimental effect of zero-sum games such as the one described. In other words, HEC technology policy has to adopt a vision that allows technological choices to occur in the most favorable context rather than force isolated technology choices to compete for pre-established budget lines.

U.S. Leadership in HEC

Defining Leadership in Terms of Policy Goals

The notion of leadership is a strong driver of the policy deliberations on HEC. From the discussion presented in this report, it is not surprising. The idea of an overall environment necessary for the generation of diversity in HEC technological development and selection of specific machines and applications begs the question of what specific purpose the United States has in supporting such a thing. The answer that most participants in discussions about HEC policy offer is articulated in terms of U.S. leadership in both HEC specifically and various fields of science and engineering generally.

However, the diversity of roles played by HEC in science and engineering and the complexity of HEC technologies themselves have, in turn, complicated the concept of leadership that inspires U.S. policy in this area. On the surface of things, leadership is identified with a competition in which the leader is also, obviously, the winner. Coverage by the popular press of supercomputing discussions has picked up on this idea and has helped to divulge the notion that there is a competition that the United States is now losing to Japan by virtue of the successful implementation of the Earth Simulator. This serves as a clear indicator of the power a concept such as leadership can have to frame the discussion and policy alternatives available to the country. There is a simple equation that policy makers must confront: Loss of leadership—meaning the United States is losing an important competition—creates a crisis that must be resolved.

The definition of the competition, in this view, consists of building the machine capable of performing a certain benchmarking test at the fastest rate. Therefore, in order to recover its leadership, the United States must respond by building a machine that can beat the Earth Simulator at performing similar benchmark tests. This notion has the virtue of being easy to understand and straightforward in its adjudication. At the same time, its very simplicity is deceiving and has at least two consequences that are very detrimental to the policy deliberation that must take place. First, by contrast, the discussion of more thoughtful notions of leadership sounds like a defensive argument to rationalize a losing situation. Second, HEC policy is trivialized by reducing it to the decision on whether or not to build a machine to beat the Earth Simulator at the benchmarking game. Rather, HEC policy must be considered in the context of the key U.S. policy goals in national security, commerce, education, and research, for the achievement of which High End Computing is critical.

As HEC is embedded in the pursuit of critical national goals, leadership as it relates to HEC is tied into the position the United States is trying to secure. For example, given recent events, U.S. leadership in HEC should have a prominent national security perspective in order to insure the nation's security in the face of external threats. Leadership should also have a commercial meaning related to the ability of U.S. industries to dominate various markets, especially those in information technology that are directly connected to HEC. It should also have an academic meaning related to the ability of the American scientific community to consistently produce the best new scientific results in most disciplines of research, especially those that are based on computational research. In the end, because of the ubiquity of HEC in key government, industry, and scientific endeavors, U.S. leadership in HEC must be judged by the nation's ability to succeed on all these fronts.

National security concerns yield the most clear-cut, straightforward sense of a competitive environment in which leadership must be pursued. Considering all possible threats from foreign nations or terrorist organizations, the government must be prepared with a high degree of certainty to defeat them all. The heightened sense of threat brought about by the terrorist attacks of 2001 has driven home to the American public, as well as government officials at all levels, the point that information gathering and processing is absolutely crucial to this sense of leadership. And in this form of competition a lot is at stake, arguably the survival of the country. However, in this case it is clear that leadership as status does not mean much. It must be the practical ability to defeat those who pose a threat. Admittedly, translating this into the specific HEC resources needed to achieve the desired degree of security is a little more difficult.

Similarly, in other areas in which HEC contributes to national priorities, status matters less than results. In the commercial realm, leaders are those whom the market rewards financially. This is a point that industry providers of COTS supercomputers are keen to emphasize. American supercomputer vendors control at least two-thirds of the world's market for these machines and provide support and reliable HEC services over time. However, market share is not all that market leadership is about. The ability to introduce innovations to achieve or retain market share projects the leadership position into the future. As so many IT firms have learned from direct experience, it is not wise to be smug about a position in the market; it can change very quickly. So even though the market may be the ultimate adjudicator of leadership, the results that matter are obtained from the ability to introduce new products that succeed in meeting future customer demand. Rather than the isolated datum of market share today, *"being the first one to market with new ideas,"* as one of our interviewees put it, or *"taking the high ground so everything else falls into place"*—meaning the commitment to develop the crucial innovation that enables others—as another interviewee said, may be better indicators of leadership in the commercial sense. High payoff for investments in innovation as a sustained result is, then, the key to leadership in the commercial sense.

Science is also a competitive enterprise. Winning scientific races may get researchers the title of leaders in their field. In this regard, scientific leadership might have more of a status or reputation connotation. However, one of the simplest and most important results in the sociological study of science is that scientific productivity is self-reinforcing. Therefore, early results by a researcher or a team hold much promise of more and greater things to come. Again, the point is that leadership in science has a fundamentally pragmatic meaning related to getting things done. With the increased significance of computational resources, high-end computational resources in the hands of productive researchers will yield exponential increases in scientific results. Even though this may be enough of a reason to relate HEC to scientific leadership, the most important factor may still be the indirect effect it has on future scientific potential due to the recruiting effect of leadership. Scientific leaders attract the most capable students, the researchers of the not too distant future. Therefore, the enabling ability of leadership may have its most critical meaning in the development of human capital for all dimensions of HEC into the future. Commitment to HEC is a strong indicator of the nation's commitment to such a broad spectrum of scientific endeavors that the catalytic effect of HEC as it relates to scientific leadership may be the most significant of all.

It seems leadership is truly at stake as it relates to HEC in the pursuit of several key U.S. policy goals in national security, commerce, and academics. But it is not a simple notion of leadership in the sense of winning a race. Rather it is leadership as the ability to achieve important results in several important areas. Therefore, sustaining all these types of leadership at once requires a multidimensional program with efforts on many fronts at the same time, with tightly coupled coordination across institutional boundaries. Institutions that have long-standing traditions of autonomy in defining their mission and the requirements to fulfill them must become much more permeable to a sort of coordination that cannot be conceptualized as a mere agreement on division of labor.

Loss of Leadership and the Crisis Approach to Technology Policy

Among stakeholders in HEC, the polarizing term is "crisis": Is there a crisis? If so, what is the crisis? What can or should be done about the crisis? While some may wish to dismiss the alleged HEC crisis and its attendant controversies as "mere rhetoric," doing so shows a misunderstanding of policy making in the United States, especially policy making pertaining to science and technology. In the first place, rhetoric matters. In policy making, one who wins the rhetorical battle often enhances the chance of winning budgetary battles. In the second place, the term "crisis" has a power in science and technology policy exceeding its power in most other policy domains. Post-war science policy has been rife with crises, real or imagined, and in many instances the correspondence between the rhetoric and the reality correlated little with the resultant policy outcomes. The Sputnik crisis, for example, resulted in the transformation of U.S. higher education, despite the fact that the satellite program at that time had very limited relevance to national security and, moreover, that the U.S. nuclear capability eclipsed that of the Soviet Union. Perhaps more relevant, the fact that the engineering education "crisis" in the 1970s was apparently fabricated[22] seems to have had little bearing on the policy responses. The energy "crisis" during the Carter administration today seems less a matter of U.S. inability to produce synthetic fuels (the chief response) or to conserve fuel than an artifact of international petroleum politics.

Why does "crisis" rhetoric have such power in U.S. science and technology policy? One reason is the notion of science and technology progress and that a failure to "keep up" may have more dire consequences than in policy domains where infusion of funds can have swift and predictable results. The second reason that crisis rhetoric has such power in science and technology policy is that it is so often and so easily tied up with issues of national leadership and all that goes with it: national prestige, national security, and market leadership.

Status Leadership Versus Pragmatic Leadership

Leadership as status: perceived to be the winner of a race

versus

Pragmatic, results-oriented leadership: ability to secure desired results into the future

- National security leadership: ability to defeat internal and external threats
- Economic leadership: ability to sustain high-payoff innovation investments
- Academic leadership: ability to sustain continuous generation of new scientific results

Although in some cases the crisis card is not so easily played (witness the superconducting supercollider, an instance where the "crisis" was not easily communicated), the HEC case, by its nature, is well positioned for exploiting crisis rhetoric. Among the reasons: (1) HEC is demonstrably integral to national security and defense, (2) many quite distinct scientific fields share a critical dependence on HEC, (3) HEC development at least has the potential to develop in great technological leaps so that the fear of being left behind is not unrealistic. None of this implies, of course, that the crisis is real or, even if it is, that any particular development or policy path is superior. Therefore, the importance we attach to HEC policy in the United States today and the approach to addressing its key issues are not based on a crisis characterization.

Findings and Recommendations

Findings

Policies in HEC have implications for national security, economic competitiveness, and academic preeminence. Much of the United States' advantage in all three areas is based on superior technological ability. Therefore, the voices of concern in this area cannot be dismissed as mere special pleading for the advantage of certain groups. There is a genuine argument for the centrality of these issues to some of the basic thrusts of American national security, economic, and academic research policies. The question is to what extent HEC deserves attention in the national policy scene vis-à-vis other important factors contributing to the same larger objectives.

Finding 1

HEC entails a complex set of problems, and policy in this area cannot be reduced to a single approach, idea, machine, or project.

Finding 2

HEC policy constitutes an institutional design challenge given that the array of needs and programs to address them requires organizational efforts that established agency practices are not well suited for.

The first two findings summarize the main discussion in the body of this report. On the one hand, there are many specific projects and initiatives proposed or requested by various stakeholders in HEC policy. Each one addresses particular problems of one sector or organization. However, no single project taken individually will enhance the overall position of the United States in the policy areas for which HEC is critical. What is needed is a balanced approach that seriously addresses all issues such as the need for research in new architectures as well as the problem-driven software issues related to usability of systems.

On the other hand, we must be aware of the limitations of instrumental approaches to technology policy in areas such as this one that require complex technological systems, true environments of computing activity that cannot be paired one-to-one with narrowly defined objectives. The achievement of broad national goals with a multidimensional approach to HEC is mostly an institutional design challenge rather than an exercise in picking tools or machine designs. Our recommendations suggest ways to address this challenge.

Finding 3

The alleged crisis in HEC policy is contingent upon perspectives of stakeholders given by the role computational technology plays in their research.

From the first two findings, a third finding is derived that merits a more extended discussion at this stage. What do these findings imply for the alleged HEC crisis? The most important finding is that crisis is contingent. It is easy to see why some HEC stakeholders view HEC development in crisis terms and others do not. Among those who view HEC development and policy as a crisis, there are different crises from different perspectives. Among other variants, there is the "we are going to be left behind crisis" and its two chief sub-categories: "hence, we lose prestige" and "hence, our range of HEC-dependent scientific and technical capabilities is reduced." Another crisis variant is the "politics of opportunism" in which there is little concern about particular substantive issues and much more about using any available device to enhance funding. It is this variant that is present in nearly any aspect of science and technology policy but especially in large-scale technology development policy making. It is important because it affects the politics of agenda setting and priority making in both Congress and the science bureaucracies, but there is little hope of generating much light from the heat.

If we consider the simple model depicted in Figure 3, a depiction of the equipment intensiveness of research, we can see at least one reason why different stakeholders have different views of crisis and different interpretations of HEC policy-making realities.

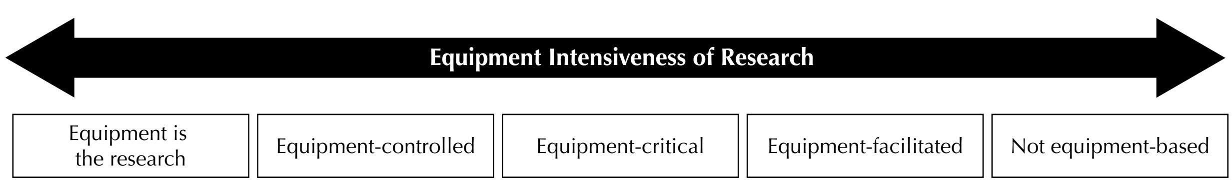
The model in Figure 3 can be used to help understand the role of equipment in the issues and controversies of several types in most fields of science and technology. But if we limit ourselves to computers—and, specifically, HEC computers—we can use these categories to help sort out the policy stakes. It is perhaps not too much of an oversimplification to suggest that among researchers the perception of the HEC crisis depends on one's position on the equipment-intensiveness arrow. To be sure, several other factors explain views about HEC, but the extent to which researchers depend critically upon computer technology development is among the crucial determinants.

For those whose research is the development of HEC equipment *(equipment is the research),* the stakes are clear cut. If advances in HEC state-of-the-art prove not to be a priority, then their research livelihood is at stake. In some ways this set of interests is less interesting because it is so easily understood. It is the three middle categories where policy options and positions are more complex.

The *equipment-facilitated* group is certainly the largest and most diverse. For this group, HEC has the ability to enhance their work, but there are many specialists not dependent upon HEC and there are alternative means of accomplishing work. Thus, users of some HEC simulation applications would fall in this group if, in fact, simulation was just one of many tools to accomplishing research objectives. For most of the members of this group, HEC policy and state-of-the-art is of only modest importance, and they likely track issues only a little more closely than the well-educated member of the general population. For many in this group there are significant applications in HEC, but the applications are easily accommodated by the current state-of-the-art. It is difficult to get these research stakeholders motivated about HEC policies, because for them these issues are viewed, at least for the present, as largely peripheral. Few in this group would likely perceive an HEC crisis, at least insofar as it pertained to their ability to do their work.

The members of the *equipment-critical* group have a quite different attitude about HEC developments. Within this group, HEC plays a fundamental role in research for many, some members of the field depend upon HEC, and the development of the state-of-the-art matters. Examples would be many users in such fields as structural genomics and, perhaps, the computational study of macromolecular interactions. Certain research pathways are dependent upon HEC, and further developments of HEC may well open new research directions. For this group, it is not difficult to get motivated about HEC policies, but views are diverse because of diverse needs. For some users in this category, grid computing is sufficient for any current or currently contemplated research problem, but for others more dedicated, specialized computers are needed. Within this group, some percentage face a crisis in their ability to do research, but the configuration of technology and access to existing technology are

Figure 3: Equipment Intensiveness of Research



often more the issue than pressing the state-of-the-art. Members of this group may perceive a wide variety of "small crises," but probably not *the* crisis. They are concerned about relatively small steps in specific software and hardware development and applications. This is a large group and perhaps the fulcrum one must look to for mobilizing support among researchers for HEC.

For the *equipment-controlled* group, HEC development and policy is especially salient. These are fields that cannot exist in the current state without HEC, and most were in a sense invented by the opportunities accruing from HEC technology. Applications in fields as diverse as astronomy and astrophysics and advanced medical imaging cannot easily advance in the absence of adequate HEC. This group is perhaps the smallest in number but also in some respects the most cohesive inasmuch as they agree about the need to pursue aggressively advances in new architecture and, if possible, diverse high-end architectures rather than settling on a single design. For this group, the "crisis" is about their ability to actually conduct research that moves the frontiers of their field.

While researchers' dependence upon HEC tells us much about the salience, character, and chief issues in HEC policy controversies, one must bear in mind that researchers comprise just one stakeholder set. Public policy makers, HEC application users, and manufacturers have quite different needs and perspectives. Nevertheless, we feel that an understanding of the diversity of research and application needs and priorities serves to underscore the two fundamental points above—that no single path is likely to prove optimal and certainly no one approach will accommodate the very different priorities and "crises." Second, the institutional design challenge for HEC development must take into account not only multiple paths but diverse and sometimes conflicting needs. Some stakeholders and some applications need more computing power, and larger grids will serve those needs. For others, the machines must advance; it is not more computing capacity but the very computing capability that is the key.

What this suggests, then, is a latticework of approaches, with different investments and different investors. There is a significant private role, a significant public role, and a significant public-private partnership role. But absent the leadership to establish the latticework and manage its parts (at least to the extent of rationalizing roles), many of the scientific, technical, and commercial opportunities presented by HEC may be lost.

Recommendations

Each of the four main issues we addressed in this report leads to possible policy actions. We built the case from the bottom up, that is, from the grounds for giving HEC more attention in policy making to the way the political articulation of national goals establishes an implied sense of leadership. Now we offer recommendations in logical order, or top down, from the highest, most encompassing category to the lowest.

Recommendation 1

Make HEC policy an explicit and integrated component of the national goals articulated in contemporary political deliberations.

In recent years, the absence of an overall policy framework has been a significant gap and a failure of political leadership. While there is not a need for an "industrial policy" guiding HEC development, a policy framework is required to coordinate activities of disparate developers of software and hardware for highly diverse applications, many of which are vital niche applications rather than abundant market opportunities. Absent strong industrial developers working in a competitive market, the United States will fall behind in HEC. But industry cannot provide the overall policy framework within which the market will operate.

Several reports by different sectors of the HEC community have been written recently showing directly how specific needs in their area will only be satisfied with new HEC resources. However, as we have argued in this report, what is needed is a balanced HEC environment that contributes to all the relevant national goals in a coordinated and consistent fashion. It is not possible to succeed by isolating an area, such as national security, for example, and picking an HEC component or set of projects to address it. As we now know, national security, to continue with the same example, is tied to most other national goals in one way or another and so are the various components of HEC. Without the political leadership to articulate such a vision, it is not possible to design an HEC policy that will succeed beyond the short term.

The development of the Internet as a broad technological infrastructure serving commercial, academic, and defense goals of the nation owes its success in large measure to the attention it received from political leaders. They were able to envision the relationship between some of their key political objectives and the system that was proposed. HEC may suffer from the perception that its main uses are still confined to esoteric fields of activity. However, we argue in this report that the importance of HEC has brought it to center stage for many key activities in the knowledge-based society of today. Following the example of the inclusion of the Internet in the political agenda of the country, we anticipate that political leaders will have to go through a learning process to develop their own understanding of the contemporary challenges and the role of HEC. It is the responsibility of political leadership to articulate these new roles of HEC for the political agenda of today. Without this commitment, the nation will be making suboptimal use of its resources and investments in this key area of contemporary science and technology.

Recommendation 2

Create a high-level coordinating entity for HEC that has enough power to overcome the zero-sum game dynamics that plague policy in this area.

The decentralization of American R&D across various federal agencies that compete for resources has worked well in many ways for the U.S. innovation system. However, it has serious disadvantages. On the one hand, it tends to force stereotypical forms of division of labor, such as charging one agency with the development of new hardware architectures and another with application software, with no provision for their coordination or compatibility. On the other, it may lead to unnecessary duplication. The coordinating role has mostly been played by ad hoc panels and task forces with some success. In the current situation there are numerous panels and councils that have coordinating roles but depend largely on the special abilities of individuals to bridge the gaps in institutional design.

The creation of a high-level coordinating entity certainly is not unprecedented. An excellent case in point is the National Nanotechnology Initiative (NNI), an effort that has many similarities to HEC policy agendas.[23] The NNI is a multi-agency, multi-sector effort to galvanize research to exploit the potential for nanoscale research. The parallels with supercomputing are clear: Both sets of issues relate to infrastructure and expensive equipment, both are highly interdisciplinary and affect the course of many fields of science, and both have important commercial implications. The key difference is that the HEC policy agenda, to the extent it can be said to exist, is fragmented and has no high-level institutional champion. HEC needs its own "national initiative," its own high-level coordinating body, a sustained policy and coordination effort, and the continuing attention of the OSTP and high-level representatives from industry and from the many federal agencies having a crucial stake in the development of HEC.

The fragmentation of initiatives in HEC also conspires against the intent of our first recommendation. It is very difficult for political leaders to go beyond a piecemeal instrumental view of specific projects and devices if HEC policy is a mere aggregation of agency-based initiatives. No overall vision, with its corresponding coherent policy and strong political support, is possible under those conditions. In this sense, oversight of HEC policy by political leaders will be significantly enhanced by following this recommendation. Implementation of the policy by a high-level coordinating entity will respond directly to the political agenda that grounds HEC policy design.

Recommendation 3

Implement an HEC policy that addresses the incentives of researchers in the different sectors—government, academia, and industry—to explore alternatives in hardware and software and avoid either premature "lock-in" by suboptimal technologies or premature abandonment of immature technologies that still hold promise.

On the surface of things, it may seem that, in the end, the solution to all science and technology policy issues is to give everybody more money. In this view, policy "issues" arise in science and technology only because resources are finite and we must decide what to fund. However, that is not the argument behind this recommendation. What matters is strategic support in order to insure that more choices will be available down the road. When one approach is allowed to dominate prematurely, many decisions follow, such as the areas targeted by leading researchers to explore new promising results or the career decisions of graduate students who will look elsewhere for exciting fields to work in. When the limitations of a poorly selected dominant approach become visible, money alone cannot fix the problem, because the community has lost expertise and does not have options available to address the new set of problems.

This aspect of science and technology policy may appear to be very ambiguous and open ended. When the consideration of optimal use of scarce resources is factored in, it is difficult to offer criteria to decide how far down this road to go. It could be a recipe for wasteful spending on dead-end projects. However, there are several reasons why this need not be the case. First, there are strong competitive forces in the research community to sustain a healthy technology selection process. Researchers do not have a significant incentive to pursue dead-end technologies once a significant group in the community loses faith in its possibilities. If anything, the pressures on researchers to show results in these areas or risk losing support are much greater than advisable, and many opportunities have been lost because a first round of disappointing results was enough to terminate initiatives. Some of our interviewees pointed out that the success of the Japanese Earth Simulator was partly due to their researchers' picking up both hardware and software problems where American researchers had abandoned too soon.

Second, the academic research community, especially in high-technology areas, has developed very close ties with private industry, which has been a source of inspiration, with new problems and ideas for academics to work on. These ties have also added a new level of accountability since the promise of academic research in the marketplace has been implicitly adopted as an evaluation criterion for their work. It has been widely reported that one side effect of this phenomenon is the introduction of a sense of urgency to produce results and a short-term vision in much academic research, which could undermine other aspects of its specific mission. In sum, the research system has a very strong selection process, but its generation of diversity is beginning to suffer, and government policy should insure that there is a large enough set of ideas and possibilities to select from in the future.

Recommendation 4

After creating the institutional and policy mechanisms for a National Supercomputing Initiative, build into the policy the time frame factor in this field and provide for a periodic, objective assessment of progress in achieving national policy goals.

As we have pointed out, HEC is a moving target. Today's high end is tomorrow's mainstream, and newly implemented architectures and software applications remain at the true high end for periods that last only a few years. The predicament the field finds itself in today is in part due to lack of attention to its changing dynamics. The new challenges that will emerge in the medium to long term should not have to be confronted without the benefit of what has been learned today. Therefore, the policy should anticipate some of the possible "forks in the road" that will demand an update of HEC policy. The policy should contain mechanisms to generate the relevant information for the next round of HEC policy making.

We recommend the appointment of a National Research Council (NRC) panel to provide a yearly assessment of the progress in achieving goals set forth in a National Supercomputing Initiative, complete with recommendations for changing policy directions that fail to live up to promise, including corresponding changes in federal funding for HEC.

This recommendation is based on two key priorities for HEC policy, namely, the time dependence of its content and independent assessment of its merit. The NRC provides a forum with an established reputation serving the science and technology policy needs of the country. This meeting point between scientific expertise and political leadership is crucial to the ongoing learning process that this field demands today. The nature of learning in this context must not be construed in a single direction—science or technology informing politics—but as a two-way street in which the achievement of national goals is informed by the state-of-the-art, and the opportunities for research into new scientific and technological opportunities are informed by the contemporary articulation of those national goals. The time dimension inherent in changes in HEC developments and capabilities and the objectivity required for properly grounded policy choices would be addressed in this arrangement.

Appendix: The Implementation of a Cyberinfrastructure

A significant perspective on the overall HEC environment described in this report is the proposal of a "cyberinfrastructure" by NSF.[24] The basis for this proposal is that science and engineering today require a common set of sophisticated information technology resources that should sustain digital knowledge environments rather than think of information technology as bounded tools for individual projects.[25] The exact shape of the common information technology resources that would constitute the infrastructure is still part of the research agenda. Because its shape is still undefined, this emphasis on cyberinfrastructure that comes from looking at work in science and engineering as a whole worries some people in the HEC community because it has the potential for drawing attention away from some specific problems. Some important applications may require a pointed effort to overcome current limitations that a generalized discussion of infrastructure will not capture. For example, many HEC researchers believe there is a frequent conflation of the cyberinfrastructure theme with the teragrid approach to HEC, which will certainly be useful for many applications but not as useful for many others. If the strategy that is finally enacted committed a disproportionate amount of resources to this approach, the overall balance that is needed to pursue most of the critical needs in computational science would not be achieved. One researcher in charge of a large supercomputing facility voiced this concern:

> *There are a number of agencies that are saying "grid" is the answer. That scares me. It may be useful for some applications, but it is hard to foresee how it will assist in modeling in computational science. You can't do computing like that over a distributed set of centers. There are a number of agencies strongly investing in the grid that will not work at the highest end.*

Another HEC researcher said:

> *A distributed environment is of extreme value and needs to be pursued. But one thread of this argument is that we can meet our computational needs by using machines distributed across the country. But we are trading in microseconds to get milliseconds and most machines are already oversubscribed. You can't get the*

Infrastructure and Cyberinfrastructure

The term *infrastructure* has been used since the 1920s to refer collectively to the roads, power grids, telephone systems, bridges, rail lines, and similar public works that are required for an industrial economy to function. Although good infrastructure is often taken for granted and noticed only when it stops functioning, it is among the most complex and expensive thing that society creates. The newer term *cyberinfrastructure* refers to infrastructure based upon distributed computer, information, and communication technology. If *infrastructure* is required for an *industrial* economy, then we could say that *cyberinfrastructure* is required for a *knowledge* economy. (From NSF 2003, p. 5.)

Definition of "TeraGrid"

TeraGrid is a multiyear effort to build and deploy the world's largest, fastest, distributed infrastructure for open scientific research. When completed, the TeraGrid will include 20 teraflops of computing power distributed at five sites, facilities capable of managing and storing nearly 1 petabyte of data, high-resolution visualization environments, and toolkits for grid computing. These components will be tightly integrated and connected through a network that will operate at 40 gigabits per second—the fastest research network on the planet.

The TeraGrid project was first launched by the National Science Foundation in August 2001 with $53 million in funding to four sites: the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Urbana–Champaign; the San Diego Supercomputer Center (SDSC) at the University of California, San Diego; Argonne National Laboratory in Argonne, Illinois; and Center for Advanced Computing Research (CACR) at the California Institute of Technology in Pasadena. In October 2002, the Pittsburgh Supercomputer Center (PSC) was added to the TeraGrid partnership when NSF announced $35 million in supplementary funding. Primary corporate partners are IBM, Intel Corporation, and Qwest Communications. Other partners are Myricom, Sun Microsystems, and Oracle Corporation. (From www.teragrid.org)

> *whole machine now for your application, so why are we going to get time on oversubscribed machines then?…*
>
> *It doesn't do anything about the high end. The belief is that productivity is ultimately a software problem. But the truth is, in my opinion, that we are programming the wrong machines. Now you have to delineate the parallelism with the physical environment and allocation, but the hardware should do this … new classes of architecture have to come in.*

The main point here is to highlight the fact that the argument for a "cyberinfrastructure" focusing on the overall computational environment for science must be neither reduced to one particular computational approach, such as "teragrids," for example, nor made into an either/or argument between the emphasis on communication and sharing that is implicit in descriptions of the environment created by cyberinfrastructure and the specific high-end solutions needed to achieve levels of performance critical for certain key applications not available with current machines. A researcher pleading the cyberinfrastructure case said:

> *Supercomputers are the big mountains, but there is a big landscape around it which is also really important for the high end. The big machines have their role and should not be forgotten, but it is a much more complex landscape using them in conjunction with other resources. Big data, linking supercomputers with environments, remote instruments, and so on, [are] all part of the changing modern landscape of science. The big instrument is not the only thing, and software is absolutely key in this because it must be possible to use them in an integrated way: Not all machines are yours if you are a user, not all have the same performance, the need for low latency access to your data—these are all very different problems. You must worry about the linkability of all of them as in transportation: planes, taxis, buses, etc. You want to get from here to there without having to worry about the details of how the individual pieces work. If you think about the supercomputer, it is the Concorde, but not the end-to-end problem.…*
>
> *Other people not in the academic community are putting it as an "either/or" thing. It is false because cyberinfrastructure is not laptops for everybody. It is not mutually exclusive with building an Earth simulator. But we have to decide as a nation how to spend the money. You can spend the money on networks and smaller-level platforms or giant computational platforms, and it does become a weighting of the budget in a zero-sum game and it affects the costing and how you do this job.*

The way the NSF report presents the cyberinfrastructure initiative may be the source of this concern. It does not include the "base technologies" in the infrastructure frame it is proposing:

> [Figure A.1] illustrates the types of facilities and services to be provided in an integrated way by a cyberinfrastructure layer (shaded). This layer is built upon base technology for computation, storage, and communication. Cyberinfrastructure should be produced and managed in a way that enables research communities/projects to tailor efficient and effective application-specific, *but interoperable,* knowledge environments for research and education. Interoperability is important for facilitating multidisciplinary projects as the evolution of discovery dictates. The Panel has learned that new types of scientific organizations and supporting environments (*"laboratories without walls"*) are essential to the aspirations of growing numbers of research communities/projects and that thus they have begun creating such environments under various names including *collaboratory*, *co-laboratory, grid community, e-science community,* and *virtual community.*[26]

The proposed cyberinfrastructure is a middle layer between base technologies and the discipline- and application-specific development. Much of the concern of HEC over architectures and performance of the higher-end machines would belong in the base, and much of the effort in software to increase scientific productivity in computational science areas would belong on the top layer. This depiction of cyberinfrastructure, building on the bottom one to support the other one at the top, important as it may be, could be interpreted as not contributing directly to either. In other words—and using the evolutionary thinking analogy introduced earlier—tying the cyberinfrastructure to one arbitrarily selected approach creates a distorting effect on the overall environment for the development of HEC technologies. The design of HEC policy must insure that this sort of arbitrary selection does not occur to avoid closing down other technological paths too soon.

Figure A.1: Integrated Cyberinfrastructure Services to Enable New Knowledge Environments for Research and Education

<table>
<tr><td colspan="5">Community-Specific Knowledge Environments for Research and Education
(collaboratory, co-laboratory, grid community, e-science community, virtual community)

Customization for discipline- and project-specific applications</td></tr>
<tr><td>High performance computation services</td><td>Data, information, knowledge management services</td><td>Observation, measurement, fabrication services</td><td>Interfaces, visualization services</td><td>Collaboration services</td></tr>
<tr><td colspan="5">Networking, Operating Systems, Middleware</td></tr>
<tr><td colspan="5">Base Technology: computation, storage, communication</td></tr>
</table>

☐ *cyberinfrastructure: hardware, software, services, personnel, organizations*

Source: "Revolutionizing Science and Engineering through Cyberinfrastructre: Report of the National Science Foundation," January 2003, p. 13.

Endnotes

1. Michael Crow and Barry Bozeman. *Limited By Design: R&D Laboratories in the U.S. National Innovation System*. Columbia University Press, 1998, pp. 26ff.

2. Barry Bozeman and Juan D. Rogers, "Strategic Management of Government-Sponsored R&D Portfolios: Project Outputs and 'Scientific and Technical Human Capital,'" *Environment and Planning C.: Government and Policy* Vol. 19, pp. 413-442, 2001; Juan D. Rogers and Barry Bozeman, "Knowledge Value Alliances: An Alternative Method to R&D Project Evaluation," *Science, Technology and Human Values* Vol. 26, No. 1, pp. 23-55, Winter 2001.

3. See "Report on High Performance Computing for the National Security Community," July 2002 (cited as NSC Report 2002). The report was prepared by members of the National Security Agency, the Defense Advanced Research Projects Agency, the Department of Defense's High Performance Computing Modernization Program, the National Reconnaissance Office, National Nuclear Security Administration of the Department of Energy, and the National Aeronautics and Space Administration. The authors referred to the composition of this team as a representation of the 'National Security Community' and we have adopted the same name for this group in this paper. See also "Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation" January 2003 (cited as NSF 2003). Other agency and blue ribbon panel reports are currently in progress. We have had access to only those mentioned above.

4. U.S. Congress. House. *Federal Supercomputer Programs and Policies*. Committee on Science and Technology. 99th Congress. No. 44. Hearing June 10, Washington, D.C.: GPO, 1985.

5. NSF 2003, p. 4.

6. R. Nelson and S. Winter, *An Evolutionary Theory of Economic Change*, Harvard University Press, 1982; P. David, "Clio and the Economics of QWERTY," *American Economic Review,* 75(2), May 1985, pp. 332-37.

7. NSC Report 2002, p. 12.

8. Quotes from interviews conducted by the authors are italicized throughout the report.

9. NSF 2003, p. 15.

10. NSC Report 2002, p. 13.

11. National Research Council, Computer Science and Technology Board, *Evolving the High Performance Computing and Communications Initiative to Support the Nation's Information Infrastructure*, National Academic Press, 1995.

12. NSC Report 2002, p. 11.

13. U.S Congress, Office of Technology Assessment, *High Performance Computing and Networking for Science—Background Paper,* OTA-BP-CIT-59 (Washington, D.C.: U.S. Government Printing Office, September 1989; National Research Council, Computer Science and Technology Board, *Keeping the U.S. Computer Industry Competitive: Defining the Agenda*, National Academy Press, 1990.

14. NSC Report 2002, p. 1.

15. NSF 2003, p. 14.

16. NSC Report 2002, p. 12.

17. NSC Report 2002, p. 8.

18. "Supercomputing Resurrected" by Claire Tristram, *Technology Review,* February 2003, p. 54.

19. Horst Simon, Department of Energy, quoted in "World Beater: Computing power to humble all rivals and forecast Earth's future," by Kenneth Terrell, in *U.S. News & World Report,* February 10, 2003, p. 83.

20. See the articles referenced above in notes 18 and 19. As recently as June 1, 2003, *The New York Times* carried an article by John Markoff about Gordon Bell and Jim Gray's argument that government should not fund

supercomputer development. Rather, they maintain, it should fund large-scale data-storage systems for the new data-intensive scientific applications.

21. The NSC report states, "IDC's current market forecast for technical capability computing predicts that NEC will have a strong growth in revenue with sales of the SX-6. NEC will have a strong presence in Europe and Japan, and since Cray has signed a distribution agreement to sell NEC SX-6 systems within North America, there may also be a renewed presence in North America." (NSC Report 2002, p. 10).

22. Daniel Greenberg, *Science, Money, and Politics: Political Triumph and Ethical Erosion,* University of Chicago Press, 2001.

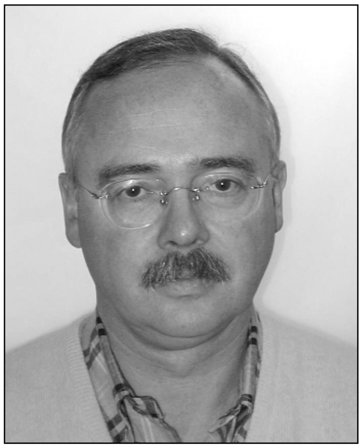
23. It emerged from a scientific working group, the Interagency Working Group on Nanotechnology, formed in 1996, which took its case to the Clinton administration and the White House Office of Science and Technology Policy, laying the groundwork for what became, in January 2000, a White House-backed policy initiative, the NNI, including a substantial budget submission to the U.S. Congress. Since that time the NNI has been active in galvanizing the research and policy agenda for nanoscale research, coordinating activity (along with OSTP and the Office of Management and Budget), meeting monthly with agency and industry representatives, conducting an annual workshop, and sponsoring and implementing regional workshops throughout the country. The result has been a sustained focus on the nanotechnology agenda, substantial funding, and support at the highest level for public-private partnerships.

24. NSF 2003, p. ES 1.

25. Use of the term "environment" must not confuse the implementation of common information technology resources for science and engineering with the analogy of evolutionary thinking. In the latter case, the environment is where technological "species" are selected following their evolutionary paths. Actually, the shape of the first meaning of "environment" would be determined in the selection process of the second.

26. Text and figure from NSF 2003, p. 13.

ABOUT THE AUTHORS



Juan D. Rogers is Associate Professor of Public Policy and Director of the Research Value Mapping Program at the School of Public Policy, Georgia Institute of Technology. He teaches courses on science and technology policy, knowledge management, logic of policy inquiry, and bureaucracy and policy implementation.

His current research interests include modeling the R&D process, assessment of R&D impacts—especially in the formation of human capital—technology transfer, R&D policy and evaluation, the interaction of social and technical factors in the development of information technology, and information technology policy. Recent publications include "A Churn Model of Knowledge Value: Internet Researchers as a Knowledge Value Collective," in *Research Policy,* Vol. 31, 2002 (with Barry Bozeman); "Knowledge Value Alliances: An Alternative Method to R&D Project Evaluation," in *Science, Technology and Human Values,* Vol. 26, No. 1, Winter 2001 (with Barry Bozeman); "Software's 'Functional Coding' and Personnel Mobility in Technology Transfer: Linkage Fields between Industry and Publicly Funded Research," in *International Journal of Technology Management,* Vol. 22, No. 8, 2001; and "Science and the Politics of Internetworking: NSFNET in Internet History," in *The Information Society,* Vol. 14, No. 3, July-September 1998.

Professor Rogers received his Ph.D. in science and technology studies from Virginia Polytechnic Institute and State University and an EE from the University of Buenos Aires, Argentina.

Barry Bozeman is Regents' Professor of Public Policy, Georgia Institute of Technology. He previously served as director of the School of Public Policy and was founding director of the Research Value Mapping Program. Before joining Georgia Tech in 1994, Bozeman was Professor of Public Administration and Affiliate Professor of Engineering at Syracuse University's Maxwell School of Citizenship and Public Affairs and the L. C. Smith College of Engineering.

Bozeman's research interests have focused on public management and science and technology policy. His two most recent books are *Bureaucracy and Red Tape* (Prentice-Hall, 2000) and *Limited by Design: U.S. R&D Laboratories in the U.S. National Innovation System* (Columbia University Press, 1998), written with Michael Crow. He is also the author of a previously published IBM Center report: "Government Management of Information Mega-Technology: Lessons from the Internal Revenue Service's Tax Systems Modernization" (March 2002).

Professor Bozeman has served as an advisor to a number of government agencies and worked briefly at the National Science Foundation's Division of Information Science and Technology and the Japanese government's National Institute for Science and Technology Policy. He received his Ph.D. in political science from the Ohio State University in 1973.

KEY CONTACT INFORMATION

To contact the authors:

Juan D. Rogers
Associate Professor of Public Policy
Research Value Mapping Program
School of Public Policy
201 D.M. Smith Building
Georgia Institute of Technology
Atlanta, GA 30332-0345
(404) 894-6697

e-mail: juan.rogers@pubpolicy.gatech.edu

Barry Bozeman
Regents' Professor of Public Policy
School of Public Policy
Georgia Institute of Technology
D. M. Smith Building
Atlanta, GA 30332
(404) 894-0093

e-mail: barry.bozeman@pubpolicy.gatech.edu

REPORTS

E-Government

Supercharging the Employment Agency: An Investigation of the Use of Information and Communication Technology to Improve the Service of State Employment Agencies (December 2000)

Anthony M. Townsend

Assessing a State's Readiness for Global Electronic Commerce: Lessons from the Ohio Experience (January 2001)

J. Pari Sabety
Steven I. Gordon

Privacy Strategies for Electronic Government (January 2001)

Janine S. Hiller
France Bélanger

Commerce Comes to Government on the Desktop: E-Commerce Applications in the Public Sector (February 2001)

Genie N. L. Stowers

The Use of the Internet in Government Service Delivery (February 2001)

Steven Cohen
William Eimicke

State Web Portals: Delivering and Financing E-Service (January 2002)

Diana Burley Gant
Jon P. Gant
Craig L. Johnson

Internet Voting: Bringing Elections to the Desktop (February 2002)

Robert S. Done

Leveraging Technology in the Service of Diplomacy: Innovation in the Department of State (March 2002)

Barry Fulton

Federal Intranet Work Sites: An Interim Assessment (June 2002)

Julianne G. Mahler
Priscilla M. Regan

The State of Federal Websites: The Pursuit of Excellence (August 2002)

Genie N. L. Stowers

State Government E-Procurement in the Information Age: Issues, Practices, and Trends (September 2002)

M. Jae Moon

Preparing for Wireless and Mobile Technologies in Government (October 2002)

Ai-Mei Chang
P. K. Kannan

Public-Sector Information Security: A Call to Action for Public-Sector CIOs (October 2002, 2nd ed.)

Don Heiman

The Auction Model: How the Public Sector Can Leverage the Power of E-Commerce Through Dynamic Pricing (November 2002, 2nd ed.)

David C. Wyld

The Promise of E-Learning in Africa: The Potential for Public-Private Partnerships (January 2003)

Norman LaRocque
Michael Latham

Digitally Integrating the Government Supply Chain: E-Procurement, E-Finance, and E-Logistics (February 2003)

Jacques S. Gansler
William Lucyshyn
Kimberly M. Ross

Using Technology to Increase Citizen Participation in Government: The Use of Models and Simulation (April 2003)

John O'Looney

Services Acquisition for America's Navy: Charting a New Course for SeaPort (June 2003)

David C. Wyld

Financial Management

Credit Scoring and Loan Scoring: Tools for Improved Management of Federal Credit Programs (July 1999)

Thomas H. Stanton

Using Activity-Based Costing to Manage More Effectively (January 2000)

Michael H. Granof
David E. Platt
Igor Vaysman

Audited Financial Statements: Getting and Sustaining "Clean" Opinions (July 2001)

Douglas A. Brook

An Introduction to Financial Risk Management in Government (August 2001)

Richard J. Buttimer, Jr.

Understanding Federal Asset Management: An Agenda for Reform (July 2003)

Thomas H. Stanton

Efficiency Counts: Developing the Capacity to Manage Costs at Air Force Materiel Command (August 2003)

Michael Barzelay
Fred Thompson

Human Capital Management

Profiles in Excellence: Conversations with the Best of America's Career Executive Service (November 1999)

Mark W. Huddleston

Reflections on Mobility: Case Studies of Six Federal Executives (May 2000)

Michael D. Serlin

Managing Telecommuting in the Federal Government: An Interim Report (June 2000)

Gina Vega
Louis Brennan

Using Virtual Teams to Manage Complex Projects: A Case Study of the Radioactive Waste Management Project (August 2000)

Samuel M. DeMarie

A Learning-Based Approach to Leading Change (December 2000)

Barry Sugarman

Labor-Management Partnerships: A New Approach to Collaborative Management (July 2001)

Barry Rubin
Richard Rubin

Winning the Best and Brightest: Increasing the Attraction of Public Service (July 2001)

Carol Chetkovich

Organizations Growing Leaders: Best Practices and Principles in the Public Service (December 2001)

Ray Blunt

A Weapon in the War for Talent: Using Special Authorities to Recruit Crucial Personnel (December 2001)

Hal G. Rainey

A Changing Workforce: Understanding Diversity Programs in the Federal Government (December 2001)

Katherine C. Naff
J. Edward Kellough

Life after Civil Service Reform: The Texas, Georgia, and Florida Experiences (October 2002)

Jonathan Walters

Leaders Growing Leaders: Preparing the Next Generation of Public Service Executives (November 2002, 3rd ed.)

Ray Blunt

The Defense Leadership and Management Program: Taking Career Development Seriously (December 2002)

Joseph A. Ferrara
Mark C. Rom

The Influence of Organizational Commitment on Officer Retention: A 12-Year Study of U.S. Army Officers (December 2002)

Stephanie C. Payne
Ann H. Huffman
Trueman R. Tremble, Jr.

Human Capital Reform: 21st Century Requirements for the United States Agency for International Development (March 2003)

Anthony C. E. Quainton
Amanda M. Fulmer

Modernizing Human Resource Management in the Federal Government: The IRS Model (April 2003)

James R. Thompson
Hal G. Rainey

Managing for Performance and Results

Corporate Strategic Planning in Government: Lessons from the United States Air Force (November 2000)

Colin Campbell

Using Evaluation to Support Performance Management: A Guide for Federal Executives (January 2001)

Kathryn Newcomer
Mary Ann Scheirer

Managing for Outcomes: Milestone Contracting in Oklahoma (January 2001)

Peter Frumkin

The Challenge of Developing Cross-Agency Measures: A Case Study of the Office of National Drug Control Policy (August 2001)

Patrick J. Murphy
John Carnevale

The Potential of the Government Performance and Results Act as a Tool to Manage Third-Party Government (August 2001)

David G. Frederickson

Using Performance Data for Accountability: The New York City Police Department's CompStat Model of Police Management (August 2001)

Paul E. O'Connell

Moving Toward More Capable Government: A Guide to Organizational Design (June 2002)

Thomas H. Stanton

Performance Management: A "Start Where You Are, Use What You Have" Guide (October 2002)

Chris Wye

The Baltimore CitiStat Program: Performance and Accountability (May 2003)

Lenneal J. Henderson

How Federal Programs Use Outcome Information: Opportunities for Federal Managers (May 2003)

Harry P. Hatry
Elaine Morley
Shelli B. Rossman
Joseph S. Wholey

Market-Based Government

Determining a Level Playing Field for Public-Private Competition (November 1999)

Lawrence L. Martin

Implementing State Contracts for Social Services: An Assessment of the Kansas Experience (May 2000)

Jocelyn M. Johnston
Barbara S. Romzek

A Vision of the Government as a World-Class Buyer: Major Procurement Issues for the Coming Decade (January 2002)

Jacques S. Gansler

Contracting for the 21st Century: A Partnership Model (January 2002)

Wendell C. Lawther

Franchise Funds in the Federal Government: Ending the Monopoly in Service Provision (February 2002)

John J. Callahan

Making Performance-Based Contracting Perform: What the Federal Government Can Learn from State and Local Governments (November 2002, 2nd ed.)

Lawrence L. Martin

Moving to Public-Private Partnerships: Learning from Experience around the World (February 2003)

Trefor P. Williams

IT Outsourcing: A Primer for Public Managers (February 2003)

Yu-Che Chen
James Perry

The Procurement Partnership Model: Moving to a Team-Based Approach (February 2003)

Kathryn G. Denhardt

Moving Toward Market-Based Government: The Changing Role of Government as the Provider (June 2003)

Jacques S. Gansler

Innovation, Collaboration, and Transformation

Innovation

Managing Workfare: The Case of the Work Experience Program in the New York City Parks Department (June 1999)

Steven Cohen

New Tools for Improving Government Regulation: An Assessment of Emissions Trading and Other Market-Based Regulatory Tools (October 1999)

Gary C. Bryner

Religious Organizations, Anti-Poverty Relief, and Charitable Choice: A Feasibility Study of Faith-Based Welfare Reform in Mississippi (November 1999)

John P. Bartkowski
Helen A. Regis

Business Improvement Districts and Innovative Service Delivery (November 1999)

Jerry Mitchell

An Assessment of Brownfield Redevelopment Policies: The Michigan Experience (November 1999)

Richard C. Hula

San Diego County's Innovation Program: Using Competition and a Whole Lot More to Improve Public Services (January 2000)

William B. Eimicke

Innovation in the Administration of Public Airports (March 2000)

Scott E. Tarry

Entrepreneurial Government: Bureaucrats as Businesspeople (May 2000)

Anne Laurent

Rethinking U.S. Environmental Protection Policy: Management Challenges for a New Administration (November 2000)

Dennis A. Rondinelli

The Challenge of Innovating in Government (February 2001)

Sandford Borins

Understanding Innovation: What Inspires It? What Makes It Successful? (December 2001)

Jonathan Walters

Government Management of Information Mega-Technology: Lessons from the Internal Revenue Service's Tax Systems Modernization (March 2002)

Barry Bozeman

Advancing High End Computing: Linking to National Goals (September 2003)

Juan D. Rogers
Barry Bozeman

Networks, Collaboration, and Partnerships

Leveraging Networks to Meet National Goals: FEMA and the Safe Construction Networks (March 2002)

William L. Waugh, Jr.

21st-Century Government and the Challenge of Homeland Defense (June 2002)

Elaine C. Kamarck

Assessing Partnerships: New Forms of Collaboration (March 2003)

Robert Klitgaard
Gregory F. Treverton

Leveraging Networks: A Guide for Public Managers Working across Organizations (March 2003)

Robert Agranoff

Extraordinary Results on National Goals: Networks and Partnerships in the Bureau of Primary Health Care's 100%/0 Campaign (March 2003)

John Scanlon

Public-Private Strategic Partnerships: The U.S. Postal Service-Federal Express Alliance (May 2003)

Oded Shenkar

The Challenge of Coordinating "Big Science" (July 2003)

W. Henry Lambright

Transforming Organizations

The Importance of Leadership: The Role of School Principals (September 1999)

Paul Teske
Mark Schneider

Leadership for Change: Case Studies in American Local Government (September 1999)

Robert B. Denhardt
Janet Vinzant Denhardt

Managing Decentralized Departments: The Case of the U.S. Department of Health and Human Services (October 1999)

Beryl A. Radin

Transforming Government: The Renewal and Revitalization of the Federal Emergency Management Agency (April 2000)

R. Steven Daniels
Carolyn L. Clark-Daniels

Transforming Government: Creating the New Defense Procurement System (April 2000)

Kimberly A. Harokopus

Trans-Atlantic Experiences in Health Reform: The United Kingdom's National Health Service and the United States Veterans Health Administration (May 2000)

Marilyn A. DeLuca

Transforming Government: The Revitalization of the Veterans Health Administration (June 2000)

Gary J. Young

The Challenge of Managing Across Boundaries: The Case of the Office of the Secretary in the U.S. Department of Health and Human Services (November 2000)

Beryl A. Radin

Creating a Culture of Innovation: 10 Lessons from America's Best Run City (January 2001)

Janet Vinzant Denhardt
Robert B. Denhardt

Transforming Government: Dan Goldin and the Remaking of NASA (March 2001)

W. Henry Lambright

Managing Across Boundaries: A Case Study of Dr. Helene Gayle and the AIDS Epidemic (January 2002)

Norma M. Riccucci

Managing "Big Science": A Case Study of the Human Genome Project (March 2002)

W. Henry Lambright

The Power of Frontline Workers in Transforming Government: The Upstate New York Veterans Healthcare Network (April 2003)

Timothy J. Hoff

Making Public Sector Mergers Work: Lessons Learned (August 2003)

Peter Frumkin

Efficiency Counts: Developing the Capacity to Manage Costs at Air Force Materiel Command (August 2003)

Michael Barzelay
Fred Thompson

SPECIAL REPORTS

Government in the 21st Century

David M. Walker

Results of the Government Leadership Survey: A 1999 Survey of Federal Executives (June 1999)

Mark A. Abramson
Steven A. Clyburn
Elizabeth Mercier

Creating a Government for the 21st Century (March 2000)

Stephen Goldsmith

The President's Management Council: An Important Management Innovation (December 2000)

Margaret L. Yao

Toward a 21st Century Public Service: Reports from Four Forums (January 2001)

Mark A. Abramson, Editor

Becoming an Effective Political Executive: 7 Lessons from Experienced Appointees (January 2001)

Judith E. Michaels

The Changing Role of Government: Implications for Managing in a New World (December 2001)

David Halberstam

BOOKS*

E-Government 2001
(Rowman & Littlefield Publishers, Inc., 2001)

Mark A. Abramson and
Grady E. Means, editors

E-Government 2003
(Rowman & Littlefield Publishers, Inc., 2002)

Mark A. Abramson and
Therese L. Morin, editors

Human Capital 2002
(Rowman & Littlefield Publishers, Inc., 2002)

Mark A. Abramson and
Nicole Willenz Gardner, editors

Innovation
(Rowman & Littlefield Publishers, Inc., 2002)

Mark A. Abramson and
Ian Littman, editors

Leaders
(Rowman & Littlefield Publishers, Inc., 2002)

Mark A. Abramson and
Kevin M. Bacon, editors

Managing for Results 2002
(Rowman & Littlefield Publishers, Inc., 2001)

Mark A. Abramson and
John Kamensky, editors

New Ways of Doing Business
(Rowman & Littlefield Publishers, Inc., 2003)

Mark A. Abramson and
Ann M. Kieffaber, editors

Memos to the President: Management Advice from the Nation's Top Public Administrators (Rowman & Littlefield Publishers, Inc., 2001)

Mark A. Abramson, editor

The Procurement Revolution
(Rowman & Littlefield Publishers, Inc., 2003)

Mark A. Abramson and
Roland S. Harris III, editors

Transforming Organizations
(Rowman & Littlefield Publishers, Inc., 2001)

Mark A. Abramson and
Paul R. Lawrence, editors

* Available at bookstores, online booksellers, and from the publisher (www.rowmanlittlefield.com or 800-462-6420).

About IBM Business Consulting Services
With consultants and professional staff in more than 160 countries globally, IBM Business Consulting Services is the world's largest consulting services organization. IBM Business Consulting Services provides clients with business process and industry expertise, a deep understanding of technology solutions that address specific industry issues, and the ability to design, build and run those solutions in a way that delivers bottom-line business value. For more information visit www.ibm.com/bcs.

About the IBM Center for The Business of Government
Through research stipends and events, the IBM Center for The Business of Government stimulates research and facilitates discussion on new approaches to improving the effectiveness of government at the federal, state, local, and international levels.

The Center is one of the ways that IBM seeks to advance knowledge on how to improve public sector effectiveness. The IBM Center focuses on the future of the operation and management of the public sector.

For additional information, contact:
Mark A. Abramson
Executive Director
IBM Center for The Business of Government
1616 North Fort Myer Drive
Arlington, VA 22209
(703) 741-1077, fax: (703) 741-1076

e-mail: businessofgovernment@us.ibm.com
website: www.businessofgovernment.org

IBM Center for
The Business
of Government

1616 North Fort Myer Drive
Arlington, VA 22209-3195

PRST STD
US Postage
P A I D
Permit 1112
Merrifield, VA