Revolutionizing Science and Engineering Through Cyberinfrastructure:
Report of the National Science Foundation
Blue Ribbon Advisory Panel on Cyberinfrastructure

Executive Summary

This is the final report of a **Blue Ribbon Advisory Panel on Cyberinfrastructure**, a panel of experts formed and charged by the National Science Foundation (NSF) Assistant Director for the Computer and Information Science and Engineering (CISE) Directorate to evaluate current major investments in cyberinfrastructure and its use, to recommend new areas of emphasis relevant to cyberinfrastructure, and to propose an implementation plan for pursuing them. We carried out this charge through individual interactions with researchers, surveys, testimony, review of prior relevant reports, requests for comments, participation in workshops, and extensive deliberation.

The Panel's overarching finding is that a new age has dawned in scientific and engineering research, pushed by continuing progress in computing, information, and communication technology, and pulled by the expanding complexity, scope, and scale of today's challenges. The capacity of this technology has crossed thresholds that now make possible a comprehensive "cyberinfrastructure" on which to build new types of scientific and engineering knowledge environments and organizations and to pursue research in new ways and with increased efficacy.

Such environments and organizations, enabled by cyberinfrastructure, are increasingly required to address national and global priorities, such as understanding global climate change, protecting our natural environment, applying genomics-proteomics to human health, maintaining national security, mastering the world of nanotechnology, and predicting and protecting against natural and human disasters, as well as to address some of our most fundamental intellectual questions such as the formation of the universe and the fundamental character of matter.

The Panel's overarching recommendation is that the National Science Foundation should establish and lead a large-scale, interagency, and internationally coordinated Advanced Cyberinfrastructure Program (ACP) to create, deploy, and apply cyberinfrastructure in ways that radically empower all scientific and engineering research and allied education. We estimate that sustained new NSF funding of $1 billion per year is needed to achieve critical mass and to leverage the coordinated co-investment from other federal agencies, universities, industry, and

international sources necessary to empower a revolution. The cost of not acting quickly or at a subcritical level could be high, both in opportunities lost and in increased fragmentation and balkanization of the research communities.

The amounts of calculation and the quantities of information that can be stored, transmitted, and used are exploding at a stunning, almost disruptive rate. Vast improvements in raw computing power, storage capacity, algorithms, and networking capabilities have led to fundamental scientific discoveries inspired by a new generation of computational models that approach scientific and engineering problems from a broader and deeper systems perspective. Scientists in many disciplines have begun revolutionizing their fields by using computers, digital data, and networks to extend and even replace traditional techniques. Online digital instruments and wide-area arrays of sensors are providing more comprehensive, immediate, and higher-resolution measurement of physical phenomena. Powerful "data mining" techniques operating across huge sets of multidimensional data open new approaches to discovery. Global networks can link all these together and support more interactivity and broader collaboration.

A central goal of ACP is to define and build cyberinfrastructure that facilitates the development of new applications, allows applications to interoperate across institutions and disciplines, insures that data and software acquired at great expense are preserved and easily available, and empowers enhanced collaboration over distance, time and disciplines. The individual disciplines must take the lead in defining specialized software and hardware environments for their fields based on common cyberinfrastructure, but in a way that encourages them to give back results for the general good of the research enterprise.

The emerging vision is to use cyberinfrastructure to build more ubiquitous, comprehensive digital environments that become interactive and functionally complete for research communities in terms of people, data, information, tools, and instruments and that operate at unprecedented levels of computational, storage, and data transfer capacity. Increasingly, new types of scientific organizations and support environments for science are essential, not optional, to the aspirations of research communities and to broadening participation in those communities. They can serve individuals, teams, and organizations in ways that revolutionize what they can do, how they do it, and who participates. This vision also has profound broader implications for education, commerce, and social good.

Our findings are supported by substantial grass roots activity in research communities. The Internet, World Wide Web, and supercomputing have already provided new tools for science, but glimpses of much more powerful and comprehensive environments for discovery and learning can be seen in a landscape of projects

focusing on creating advanced cyberinfrastructure and/or using it to create new knowledge environments for specific fields of science. Included in this landscape are the NSF Partnerships for Advanced Computing Infrastructure (PACI), the Pittsburgh Terascale Computing System (TCS), and the Distributed Terascale Facility (DTF) that "grids" together resources at all these centers plus others. Also included are a series of NSF networking, digital library, scientific database, advanced interface, and middleware research initiatives. Through the NSF Information Technology Research (ITR) initiative and other NSF programs, projects have emerged from many disciplines involving computer science and engineering researchers working to develop and use cyberinfrastructure in specific projects.

Testimony from research communities indicate that many contemporary projects require effective federation of both distributed resources (data and facilities) and distributed, multidisciplinary expertise, and that cyberinfrastructure is a key to making this possible. There is no standard term for such environments enabled by cyberinfrastructure; some of the names in use are collaboratory, co-laboratory, grid community/network, virtual science community, and e-science community. A few examples are the Network for Earthquake Engineering Simulations (NEES), the Space Physics and Aeronomy Research Collaboratory (SPARC), the National Ecological Observatory Network (NEON), the Grid Physics Network (GriPhyN), the International Virtual Data Grid Laboratory (iVDGL), and the High Energy Physics Collaboratory for the ATLAS project. Research mission agencies are also initiating similar projects, for example, the NIH Biomedical Informatics Research Network (BIRN), the Department of Energy (DOE) National Collaboratories Program, and the DOE program in Scientific Discovery through Advanced Computing (SciDAC). Relevant international activities include the UK E-science program, parts of the European Union 6th Framework Project, and the Japanese Earth Simulator. Because of the extent of cyberinfrastructure investment under way in other countries and the intrinsic global nature of science, **an effective response to our primary finding should be interagency and international in scope.**

Achieving this vision will **challenge our fundamental understanding of computer and information science and engineering as well as parts of social science, and it will motivate and drive basic research** in these areas. We envision radical improvements in cyberinfrastructure and its impact on all science and engineering over time, as work ripens at the intersection of fundamental social and technical research about cyberinfrastructure and its application to advance discovery and learning.

This vision of science and engineering research involves **significant educational dimensions**. The research community needs more broadly trained personnel with blended expertise in disciplinary science or engineering, mathematical and computational modeling,

numerical methods, visualization, and the sociotechnical understanding about working in new grid or collaboratory organizations. Grid and collaboratory environments built on cyberinfrastructure can enable people to work routinely with colleagues at distant institutions, even ones that are not traditionally considered research universities, and with junior scientists and students as genuine peers, despite differences in age, experience, race, or physical ability. These new environments can contribute to science and engineering education by providing interesting resources, exciting experiences, and expert mentoring to students, faculty, and teachers anywhere there is access to the Web. The new tools, resources, human capacity building, and organizational structures emerging from these activities will also eventually have even broader beneficial impact on the future of education at all levels and likely on all types of educational institutions.

A vast opportunity exists for creating new research environments based upon cyberinfrastructure, **but there are also significant risks and costs if we do not act quickly and at a sufficient level of investment.** The dangers, all increasing with the passage of time, include adoption of incompatible data formats in different fields; permanent loss of observational data due to lack of well-curated, long-term archives; increased technological ("not invented here") balkanizations rather than interoperability among disciplines; wasteful redundant system-building activities among science fields or between science fields and industry; lack of synergy among information technology research, the IT industry, and domain science users resulting in under- or overestimating technological futures; lost opportunity from not driving basic computer science research with advanced applications; loss of leadership to other countries and a falloff of research and economic vigor; lack of understanding of social/cultural barriers to new ways of doing research; inadequate supporting or supported educational activities; and an inadequate, piecemeal cyberinfrastructure program.

We propose a **large, long-term, and concerted new effort, not just a linear extension** of the current investment level and resources. NSF must recognize that the scope of shared cyberinfrastructure is far broader than in the past – it includes computing cycles, higher capacity networking, massive storage, and managed information. NSF must ensure that the exponentially growing amounts of data are collected, curated, managed, and stored for broad, long-term access by scientists everywhere. The new effort must create and continually renovate a new "high end," so that selected research projects can use centralized resources 100-1000 times faster and bigger than are available locally. But even this is not sufficient. There must be high-level leadership on shared standards, middleware, and building advanced scientific tools that enable scientists to follow new paths, try new techniques, build better models, and test them in new ways and that facilitates innovative interdisciplinary activities. The program must also have a **component to empower more people and more disciplines** to benefit from the use of cyberinfrastructure. It must especially encourage science

and engineering communities to exploit the new opportunities that cyberinfrastructure brings for including people who, because of physical capabilities, location, or history, have been excluded from the frontiers of scientific and engineering research and education.

NSF's prior investments provide a sound foundation for the ACP. In particular, the two NSF Partnerships for Advanced Computational Infrastructure (PACI) established in 1997 have been pioneers in activities closely related to the ACP. They have provided high-end computing cycles; developed software tools for helping people to use architecturally diverse machines; supported education, outreach, and training with a special focus on underrepresented groups; and nurtured specific testbed projects for science-driven collaboratories or grids. Much of the experience and expertise represented in the PACIs is highly relevant to the ACP, and we believe that subject to appropriate review they should be competitive for expanded missions and continuing or expanding resources within ACP.

NSF has both a unique breadth of scientific scope and a mandate for the health of the scientific research enterprise in the U.S., and therefore **NSF should lead the ACP for the federal government.** We estimate that sustained new funding for NSF of $1 Billion per year is required to achieve the critical mass necessary for revolutionary changes, reusable assets and experiences, and to be a true partner with other agencies. Only then will it be able to leverage the coordinated co-investment from other federal agencies, universities, industry, and international sources required to empower a revolution. An NSF-led ACP can be catalytic and provide over-the-horizon views for other agencies, research labs, and education at large.

We estimate that the **new funding will be distributed into four coordinated areas:** fundamental research to create advanced cyberinfrastructure ($60M); research on the application of cyberinfrastructure to specific fields of science and engineering research ($100M); acquisition and development of production quality software for cyberinfrastructure and supported applications ($200M); provisioning and operations (including computational centers, data repositories, digital libraries, networking, and application support) ($660M). These are recurring annual figures.

The opportunity is enormous, but also enormously complex and must be approached in a long-term, comprehensive way with great attention to a management structure that can identify and act on the common interests of a large and varied set of stakeholders. Some of the most critical challenges to this ambitious program are to 1) build real synergy between computer and information science research and development, and its use in science and engineering research and education; 2) capture the cyberinfrastructure commonalities across science and engineering disciplines; 3) use cyberinfrastructure to empower and enable, not impede, collaboration across science and

engineering disciplines; 4) exploit technologies being developed commercially and apply them to research applications, as well as feed back new approaches from the scientific realm into the larger world; and 5) engage social scientists to work constructively with other scientists and technologists.

We recommend that the **organization of the ACP be overlaid in a matrix fashion** on the existing organizational structures of NSF with the addition of a single new coordinating ACP Office (ACPO). Achieving sufficient coordination within the proposed matrix management structure will be formidable; the roles of the ACPO are to provide overall vision and guidance and to exercise budgetary planning and responsibility. Wherever it is administratively placed within NSF, the ACPO must have significant autonomy. Its leader must have fundamental responsibility for achieving the goals of the ACP, with sufficient credibility, power, resources, and authority to succeed in working with all NSF directorates and other domestic and international agencies. Domain science and engineering directorates must take the lead in revolutionizing their respective fields through new research organization and processes, supported by new applications of information technology. CISE must be deeply involved as a technology user and as a technology leader for the overall program. It should also benefit from advanced scientific applications informing and validating its own research.

The ACP requires an organization for internal NSF coordination, as well as a central point of coordination in its external implementation. Several development centers should be devoted to activities at the core of the program. These core activities include the planning, acquisition, integration, and support of the major software platforms and components at the foundation of cyberinfrastructure, as well as the management of consistency and sharing across the program.

Human resources are critical to making cyberinfrastructure and applications work, keeping them working, and providing user support. In the interest of funding more grants, NSF has arguably under-supported the recurring costs of permanent staff, preferring to focus resources on direct research costs and "hard" or "tangible" assets. In the ACP, human resources are the primary requirement in both development and operations, and success is clearly dependent on adequate funding both in centers and in the end-user research groups.

To be successful, the ACP will require committed champions and leaders from the research community, long-term focus and commitment, innovative organizational structures, and a sustained high level of support and commitment from the upper levels of NSF, other federal agencies, and Congress.

This Panel believes that the National Science Foundation has a once-in-a-generation opportunity to lead the revolution in science and engineering through coordinated development and expansive use of cyberinfrastructure.