Cyber-Enabled Scientific Discovery

To cite this article: Tony Chan and Leland Jameson 2007 J. Phys.: Conf. Ser. 78 011003

View the article online for updates and enhancements.

Related content
- The science that never was
  Michelle Jeandron
- VisIVOWeb: A WWW Environment for Large-Scale Astrophysical Visualization
  A. Costa, U. Becciani, P. Massimino et al.
- Numerical simulation of Flow Pressure Drop and Friction Factor of Water in 2D channel
  H Aya Baquero and L H Camargo Casallas
Cyber-Enabled Scientific Discovery

Tony Chan
Assistant Director
Directorate for Mathematical and Physical Sciences
National Science Foundation
(on leave from UCLA)
tfchan@nsf.gov

Leland Jameson
Program Director
Computational Mathematics
National Science Foundation
ljameson@nsf.gov

ABSTRACT:

It is often said that numerical simulation is third in the group of three ways to explore modern science: theory, experiment and simulation. Carefully executed modern numerical simulations can, however, be considered at least as relevant as experiment and theory. In comparison to physical experimentation, with numerical simulation one has the numerically simulated values of every field variable at every grid point in space and time. In comparison to theory, with numerical simulation one can explore sets of very complex nonlinear equations, such as the Einstein equations, that are very difficult to investigate theoretically. Cyber-enabled scientific discovery is not just about numerical simulation but about every possible issue related to scientific discovery by utilizing cyberinfrastructure such as the analysis and storage of large data sets, the creation of tools that can be used by broad classes of researchers and, above all, the education and training of a cyber-literate workforce. The National Science Foundation shares many of these objectives for cyber-enabled scientific discovery with the Office of Science of the Department of Energy, as exemplified in the SciDAC program, but has its own charter as illustrated herein.

A Brief Introduction to the NSF

The National Science Foundation (NSF) is an independent federal agency created by Congress in 1950 "to promote the progress of science; to advance the national health, prosperity, and welfare; to secure the national defense."

We fulfill our mission chiefly by issuing limited-term grants -- currently about 10,000 new awards per year, with an average duration of three years -- to fund specific research proposals that have been judged the most promising by a rigorous and objective merit-review system. Most of these awards go to individuals or small groups of investigators. Others provide funding for research centers, instruments and facilities that allow scientists, engineers and students to work at the outermost frontiers of knowledge.

NSF’s goals -- discovery, learning, research infrastructure and stewardship -- provide an integrated strategy to advance the frontiers of knowledge, cultivate a world-class, broadly inclusive science and engineering workforce and expand the scientific literacy of all citizens. It builds the nation's research capability through investments in advanced instrumentation and facilities, and support excellence in science and engineering research and education through a capable and responsive organization. We like to say that NSF is "where discoveries begin."

Many of the discoveries and technological advances have been truly revolutionary. In the past few decades, NSF-funded researchers have won more than 170 Nobel Prizes as well as other honors too numerous to list. These pioneers have included the scientists or teams that discovered many of the fundamental particles of matter, analyzed the cosmic microwaves left over from the earliest epoch of the universe, developed carbon-14 dating of ancient artifacts, decoded the genetics of viruses, and created an entirely new state of matter called a Bose-Einstein condensate.
NSF also funds equipment that is needed by scientists and engineers but is often too expensive for any one group or researcher to afford. Examples of such major research equipment include giant optical and radio telescopes, Antarctic research sites, high-end computer facilities and ultra-high-speed connections, ships for ocean research, sensitive detectors of very subtle physical phenomena and gravitational wave observatories.

Another essential element in NSF’s mission is support for science and engineering education, from pre-K through graduate school and beyond. The research we fund is thoroughly integrated with education to help ensure that there will always be plenty of skilled people available to work in new and emerging scientific, engineering and technological fields, and plenty of capable teachers to educate the next generation.

No single factor is more important to the intellectual and economic progress of society, and to the enhanced well-being of its citizens, than the continuous acquisition of new knowledge. NSF is proud to be a major part of that process.

As described in our strategic plan, NSF is the only federal agency whose mission includes support for all fields of fundamental science and engineering, except for medical sciences. NSF is tasked with keeping the United States at the leading edge of discovery in a wide range of scientific areas, from astronomy to geology to zoology. So, in addition to funding research in the traditional academic areas, the agency also supports "high risk, high pay off" ideas, novel collaborations and numerous projects that may seem like science fiction today, but which the public will take for granted tomorrow. And in every case, we ensure that research is fully integrated with education so that today's revolutionary work will also be training tomorrow's top scientists and engineers.

Unlike many other federal agencies, NSF does not hire researchers or directly operate our own laboratories or similar facilities. Instead, we support scientists, engineers and educators directly through their own home institutions (typically universities and colleges). Similarly, we fund facilities and equipment such as telescopes, through cooperative agreements with research consortia that have competed successfully for limited-term management contracts.

The NSF's job is to discern where the frontiers are, identify the leading U.S. pioneers in these fields and provide money and equipment to help them continue. The results can be transformative. For example, years before most people had heard of "nanotechnology," NSF was supporting scientists and engineers who were learning how to detect, record and manipulate activity at the scale of individual atoms—the nanoscale. Today, scientists are adept at moving atoms around to create devices and materials with properties that are often more useful than those found in nature. Dozens of companies are gearing up to produce nanoscale products. NSF is funding the research projects, state-of-the-art facilities and educational opportunities that will teach new skills to the science and engineering students who will make up the nanotechnology workforce of tomorrow. At the same time, we are looking for the next frontier. As noted above, the NSF currently has a total budget of roughly six billion dollars and as one can see from the chart below the NSF, as well as other federal agencies, has benefited from a steady growth in its budget for the last 30 years:

![Figure 1: Budget growth of Federal funding agencies since 1975](image-url)
The NSF is comprised of various directorates and offices. The following table shows the actual budgets for each NSF entity for the last two years and the budget requests for the current fiscal year as well as FY08:

|       | FY 2005 Actuals | FY 2006 Actuals | Change from 05 to 06 | FY 2007 Request | Change from 06 to 07 | FY 2007 Request | Change from 07 to 08 |
|-------|-----------------|-----------------|---------------------|-----------------|---------------------|-----------------|---------------------|
| BIO   | 576.78          | $580.90         | 0.7%                | $607.85         | 4.6%                | $633.00         | 4.1%                |
| CISE  | 490.20          | 496.35          | 1.3%                | 526.69          | 6.1%                | 574.00          | 9.0%                |
| ENG   | 557.09          | 585.46          | 5.1%                | 628.55          | 7.4%                | 683.30          | 8.7%                |
| GEO   | 697.17          | 703.95          | 1.0%                | 744.85          | 5.8%                | 792.00          | 6.3%                |
| MPS   | 1,069.36        | 1,086.61        | 1.6%                | 1,150.30        | 5.9%                | 1,253.00        | 8.9%                |
| SBE   | 196.80          | 201.23          | 2.3%                | 213.76          | 6.2%                | 222.00          | 3.9%                |
| OCI   | 123.40          | 127.14          | 3.0%                | 182.42          | 43.5%               | 200.00          | 9.6%                |
| OISE  | 43.38           | 42.61           | -1.8%               | 40.61           | -4.7%               | 45.00           | 10.8%               |
| OPP   | 348.53          | 390.54          | 12.1%               | 438.10          | 12.2%               | 464.90          | 6.1%                |
| OIA   | 130.92          | 233.30          | 78.2%               | 231.37          | -0.8%               | 263.00          | 13.7%               |
| USARC | 1.19            | 1.17            | -1.7%               | 1.45            | 23.9%               | 1.49            | 2.8%                |
| NSF R&RA | 4234.82   | 4449.25         | 5.1%                | 4765.95         | 7.1%                | 5131.69         | 7.7%                |

Figure 2: Budget growth by directorate and office within the NSF.

The American Competitiveness Initiative

The National Science Foundation, as well as other Federal agencies, will benefit from and work towards achieving the goals outlined in the American Competitive Initiative (ACI). Keeping our competitive edge in the world economy requires focused policies that lay the groundwork for continued leadership in innovation, exploration, and ingenuity. America's economic strength and global leadership depend in large measure on our Nation's ability to generate and harness the latest in scientific and technological developments and to apply these developments to real world applications. These applications are fueled by: scientific research, which produces new ideas and new tools that can become the foundation for tomorrow’s products, services, and ways of doing business; a strong education system that equips our workforce with the skills necessary to transform those ideas into goods and services that improve our lives and provide our Nation with the researchers of the future; and an environment that encourages entrepreneurship, risk taking, and innovative thinking. By giving citizens the tools necessary to realize their greatest potential, the American Competitiveness Initiative (ACI) will help ensure future generations have an even brighter future.

"The role of government is not to create wealth; the role of our government is to create an environment in which the entrepreneur can flourish, in which minds can expand, in which technologies can reach new frontiers." – President George W. Bush, May 2001
Sustained scientific advancement and innovation are key to maintaining our competitive edge, and are supported by a pattern of related investments and policies. Periodically the NSF clarifies research investment priorities in the form of a strategic plan that is intended to address the needs of the nation and to respond to such Presidential directives as the ACI. The following is an outline of the strategic plan.

The NSF Strategic Plan, 2006-2011

The National Science Foundation on June 25, 2007 released its strategic plan to guide the agency's priorities and investments for the next five years.

The plan, titled Investing in America's Future, reflects months of discussion with the research and education community. It is the first to be published under NSF's current leadership, and will influence the agency's FY 2008 budget request.

"Strategic planning is important for every federal agency, because it helps ensure that public funds are well spent," said NSF Deputy Director Kathie L. Olsen. "Taxpayers can be confident that we have thought carefully about giving them their money's worth. This strategic plan keeps us focused on specific goals and outcomes for the next five years.

The new plan identifies four (outcome-based) goals: discovery, learning, research infrastructure and stewardship. Each is linked to specific investment priorities, and emphasizes actual results, or outcomes. See, http://www.nsf.gov/pubs/2006/nsf0648/nsf0648.jsp for the complete document. A prominent part of the strategic plan is focused on cyber-infrastructure in support of scientific discovery.

Cyber-infrastructure Vision for the 21st Century Discovery

Historically scientific discovery has been achieved through a combination of theory, experiment, and careful observation. During the 20th century scientific computing has become an additional mechanism through which one can explore numerically various theoretical ideas and compare such numerical experimentation with physical experiments. As computational platforms have grown in capability numerical experiments have progressed in size and gone from 1-dimensional simulations 30 years ago to the current ability to simulate 3-dimensional physical processes. At times such simulations have led to a very deep understanding of basic physical processes that could not be understood from theory and physical experiments alone. In the 21st Century it is expected that numerical simulation will become an increasing important method of scientific exploration. In order to help define the NSF's path forward in the world of simulation and cyberinfrastructure the following vision document has been produced: http://www.nsf.gov/pubs/2007/nsf0728/nsf0728_1.pdf The following statement from the director is extracted from the Vision document:

"At the heart of the cyberinfrastructure vision is the development of a cultural community that supports peer-to-peer collaboration and new modes of education based upon broad and open access to leadership computing; data and information resources; online instruments and observatories; and visualization and collaboration services.

Cyberinfrastructure enables distributed knowledge communities that collaborate and communicate across disciplines, distances and cultures. These research and education communities extend beyond traditional brick-and-mortar facilities, becoming virtual organizations that transcend geographic and institutional boundaries. This vision is new, exciting and bold."

Dr. A. Bement, NSF Director
The Office of Cyber-Infrastructure

In 2005 a new Office of Cyber-Infrastructure (OCI) was created in the Office of the Director of the NSF. The Director of OCI is Dan Atkins and the Deputy Office Director is Jose Muñoz. The Office of Cyberinfrastructure coordinates and supports the acquisition, development and provision of state-of-the-art cyberinfrastructure resources, tools and services essential to the conduct of 21st century science and engineering research and education.

OCI supports cyberinfrastructure resources, tools and related services such as supercomputers, high-capacity mass-storage systems, system software suites and programming environments, scalable interactive visualization tools, productivity software libraries and tools, large-scale data repositories and digitized scientific data management systems, networks of various reach and granularity and an array of software tools and services that hide the complexities and heterogeneity of contemporary cyberinfrastructure while seeking to provide ubiquitous access and enhanced usability.

OCI supports the preparation and training of current and future generations of researchers and educators to use cyberinfrastructure to further their research and education goals, while also supporting the scientific and engineering professionals who create and maintain these IT-based resources and systems and who provide essential customer services to the national science and engineering user community.

The following graphic illustrates the interdependency of three critical parts of OCI:

**Office of Cyberinfrastructure (since 2005)**

*Mission: provide leadership to nurture applied research and development, and to provision advanced shared and connecting cyberinfrastructure that supports and bridges science & engineering, research & learning.*

**Transformative Application** - to enhance discovery & learning

**Provisioning** - Creation, deployment and operation of advanced CI

**R&D** to enhance technical and *social* effectiveness of future CI environments

*Borromean Ring: The three rings taken together are inseparable, but remove any one ring and the other two fall apart. See [www.liv.ac.uk/~sprmd2/rings/](http://www.liv.ac.uk/~sprmd2/rings/)*

*Figure 3: The OCI combines Transformative Application, Provisioning and R&D in an interdependent way just as the three rings of a Borromean Ring depend on each other.*

Scientific computing can lead to breakthroughs in many disciplines such as physics, chemistry, biology, etc. The Office of Cyberinfrastructure has an approach to hardware acquisition that has three distinct categories: Track 1 that is capability class hardware, Track 2 that is for capacity class hardware, and Track 3 that is for campus level class hardware.
Figure 4: Scientific discovery and the three hardware acquisition tracts of the Office of Cyberinfrastructure

The Office of CyberInfrastructure has numerous solicitations that are intended to help accomplish its goals. A few examples are the following:

- PetaApps: $21.5M (11 – 22 awards)
- Software Development for CyberInfrastructure
  FY07 (closed): HPC, Middleware, Digital Data
- High-End Computing University Research Activity (NSF/ DARPA/ DOE)
  FY06 $14.5M, 19 awards, 62 props (I/O, file, storage)
  FY08 $8M. Focus: Programming Models
- Community Based Data Interoperability Networks
  FY08: ~10 $250K/yr awards
- Engineering Virtual Organizations
  ~10-15 awards $100-200K/yr

Investments in Cyber-Enabled Scientific Discovery throughout the NSF

Throughout the various directorates one finds a variety of investments that are in support of cyber-enabled science. Such investments include investments in algorithms, virtual observatories, workforce training, storage and analysis of large data sets, etc. The following examples illustrate the breadth and importance of some of these investments.
Algorithms

First of all, advances and increased capability of scientific computing require that not only improvements in computing platforms but constant and continued investments in algorithms. In fact, investments in broad classes of algorithms can often outpace Moore’s law in terms of effective computational speedup as can be seen in the following figure from the ScaLeS report, Volume 2:

![Figure 5: Investments in Algorithms can outpace investments in hardware for certain applications.](image)

The following two quotes further summarize the importance of algorithms:

For the solution of important differential equations “... the progress made through better methods from 1945 to 1978 exceeds the progress made through faster computers...” by 40 times [Rice].

Factoring a large integer using modern mathematical techniques would be 120,000 times faster than using techniques from the 1970's. So if it took 1 day to do the problem now, it would have taken over 300 years [Odlyzko].

It is often said that numerical simulation is the “third way” to explore modern science: theory, experiment and simulation. Carefully executed modern numerical simulations should, however, be considered at least as relevant as experiment and theory. There are published examples of simulations that have uncovered man-made artifacts from physical experiments. In one example in a wind-tunnel experiment of flow over a cylinder acoustic waves were reflecting off the end of the wind tunnel and propagating upstream and corrupting the experiment. At first it was thought that the so-called secondary instability seen downstream from the cylinder was physical. However, a carefully executed simulation using spectral methods showed that the secondary instability was caused by the wind tunnel itself. When a longer wind tunnel was used the secondary instability was no longer observed, [Arbabanel]. This example is not unique. Furthermore, numerical simulations provide the numerical value of every field variable at every grid point. Physical experiments or observations, on the other hand, offer field variable values only where observations are taken such as the point values along a curve indicating sea-surface height as observed from a satellite such as the TOPEX/Poseidon satellite.
As with physical experiments, numerical experimentation must be carefully conducted from beginning to end. A variety of artifacts or errors can be introduced or can accumulate as a simulation progresses. The two figures below represent the mixing zones for two dimensional simulations of identical Richtmyer-Meshkov Instability (RMI) simulations. The 9th order Weighted Essentially Non-Oscillatory (WENO), see [Liu], has far less numerical dissipation than the 3rd order WENO simulation and captures much more small-scale structure than the 3rd order simulation. Note that these two schemes are of the same type and only differ in the order of accuracy. Both simulations are advanced in time with the identical 4th order Runge-Kutta time integration scheme yet the two results are visually very different and if one quantitatively estimates the mixing in the two simulations one finds far more mixing with the 9th order simulation than the 3rd simulation. The importance of this example is that it shows the impact that the numerics alone can have on the result of a simulation. If one does exactly the same simulation using an Alternating Lagrangian-Eulerian (ALE) scheme one will find that the result is very different from the above two WENO simulations illustrating yet again the importance of having a deep understanding of numerical errors or artifacts. The importance of these simulations is that they are used for calculating the shock-induced mixing zone of a RMI. The results will be used to enhance the understanding of the onset of fusion at the National Ignition Facility and for a better understanding of Supernova Hydrodynamics.

Figure 6: Mixing zone, 3rd order WENO.                  Figure 7: Mixing zone, 9th order WENO.

The above calculations of the RMI come from the field of physics which is probably the field that has the longest history of numerical simulation.

Biology and Peta-Scale Computing

Biology is being transformed into a data-rich science. Whether comparing genomes to improve annotation, or predicting protein structures through homology modeling and molecular simulations, biologists are dealing with a torrent of data. Add to that emerging capabilities to simulate multi-protein machines, regulatory pathways, and groups of cells and organisms working in concert, and you encounter high-performance computing problems. Biology requires the integration of data over many scales of space and time, e.g., regulatory changes in a protein conformation influence pathways and lead to macroscopic changes of a cell interacting with the environment. That means that many lower level activities cannot simply be replaced by heuristic assumptions and parameters, but need to simulated at the appropriate level. While many of these calculations can be embarrassingly parallel, holding the biological structures usually exceeds memory capacity at the processor level. Flop rate is usually not such a big issue and communication between processors can be minimal as long as the individual problem fits on one processor. The following figure illustrates the complex process of going from DNA all the way to an organism:
Figure 8: Biology requires the integration of data over many scales of space and time, e.g., regulatory changes in a protein conformation influence pathways and lead to macroscopic changes of a cell interacting with the environment.

The Challenge of Large Data Sets

Many fields are inundated with very large data sets that can be either the product of a very large numerical simulation or perhaps from very high resolution observations. Astronomy is one such field that must meet the challenge of the huge volumes of data from observations arriving from numerous sources. The National Virtual Observatory Sky Statistics Survey allows astronomers to get a fast inventory of astronomical objects from various catalogs. For example, the Large Synoptic Survey Telescope captures very large amounts of data daily. The nightly data volumes generated by LSST will be an order of magnitude larger than those estimated for Pan-STARRS 4 and 2 orders of magnitude larger than the Sloan Digital Sky Survey (SDSS). The data archive will grow at a rate of roughly 7 PB/yr. This requires scalability and reliability in LSST data management systems well beyond pre-cursors.

Estimated Nightly Data Volumes

Figure 9: Astronomy is faced with huge amounts of data as illustrated here.
Education and Workforce Development

The future prosperity and security of our nation depends on a well-trained and internationally engaged scientific workforce. Workforce development begins in K-12 and continues through the end of Postdoctoral programs. Every directorate of the NSF plays a critical role in preparing our nation for its future scientific workforce needs. This training comes in many forms from large formal programs to each program officer choosing to fund as many undergraduates, graduate students, and postdocs as budgets allow.

Figure 10: The Materials Computation center at the University of Illinois fosters world-class research and education in computational materials science (Johnson and Martin).

Another example of workforce development is the JASON project. In order to bring the thrill of discovery to millions of students worldwide a year-round scientific expedition designed to excite and engage students in science and technology and to motivate and provide professional development for teachers, see http://www.jason.org/jason_home/home_old.htm.
Figure 11: The JASON Project uses cyberinfrastructure to connect middle school students with great explorers and great events to inspire and motivate them to learn science.

National Defense

As with many US government agencies, part of our mission is to secure the national defense. Throughout the NSF there are numerous programs that support fundamental science that can be utilized in a variety of ways including helping to defend the nation from numerous threats. One such activity is called the Approached to Combat Terrorism which was a collaboration between the Mathematical and Physical Sciences Directorate and the Intelligence community. Numerous topics were emphasized in this collaboration such as fundamental research in novel power sources, image and data analysis, etc. The top figure on the right shows a path that could be a road. The bottom image on the right show the result after modern image analysis techniques were applied to the top image and the gaps in the road that were covered by trees were filled in with mathematically sophisticated image processing methods [Bertozzi].

Figure 12: In the ACT program the NSF worked with the IC for the purpose of investing is fundamental science that could used to help protect the nation such as the imaging techniques in the right two frames.
The NSF Supports Cyber-Enabled Science at Many Levels

The NSF supports fundamental research in various parts of the foundation that leads to the next generation of technology that leads to the next generation of microprocessors, memory chips, etc. Much of the research leads to the next generation of microprocessors involves numerical simulation using the current generation of microprocessors.

Figure 13: Advances in chemistry and materials science lead to the next generation of microprocessors, chemistry PI Ralph Nuzzo, UIUC.

Building on advances in microprocessors, the NSF supports research in novel computer architectures, operating systems, etc. Yet other parts of the foundation support research in fundamental research that could lead to revolutionary changes in computation and simulation such as quantum computing.

Figure 14: Atoms strongly coupled to a cavity are a resource for scalable quantum logic. Figure from Nature, October 2006.
Summary:

As one can see, the NSF is involved in cyber-enabled discovery from almost every point of view from investing in futuristic quantum computers to providing infrastructure to help to stimulate and educate the next generation of scientists. The NSF portfolio, as illustrated herein, is complementary to that of DOE’s in many respects, while both envision an increasingly fruitful role for simulation in scientific discovery.

References

J. Rice, "Numerical Methods, Software, and Analysis", 1992, Purdue University

A. Odlyzko, “The future of integer factorization”, 1995, AT&T Bell Laboratories

S. Abarbanel, W. S. Don, D. Gottlieb, D. Rudy & J. Townsend, Secondary frequencies in the wake of a circular cylinder with vortex shedding, Journal of Fluid Mechanics, 225, pp. 557-574, 1991

X-D Liu, S. Osher, and T. Chan, Weighted Essentially Non-oscillatory Schemes, J. Comput. Phys., Vol. 115, (1994)

A. Bertozzi, Identification and Classification in Aerial Images, Report on NSF award 0442037.

Cyber-enabled science related reports

The following list of reports is only a subset of all the NSF reports related to cyber-enabled science.

Building a Cyberinfrastructure for the Biological Sciences; workshop held July 14-15, 2003
http://research.calit2.net/cibio/archived/CIBIO_FINAL.pdf
http://research.calit2.net/cibio/report.htm

CHE Cyber Chemistry Workshop; workshop held October 3-5, 2004
http://bioeng.berkeley.edu/faculty/cyber_workshop

Computation as a Tool for Discovery in Physics; report by the Steering Committee on Computational Physics
http://www.nsf.gov/pubs/2002/nsf02176/start.htm

Cyberinfrastructure for the Atmospheric Sciences in the 21st Century; workshop held June 2004
http://netstats.ucar.edu/cyrdas/report/cyrdas_report_final.pdf

Cyberinfrastructure for Engineering Design workshop report held February 28 - March 1, 2005
http://www.mne.psu.edu/simpson/NSF/EXCITED/

Cyberinfrastructure for Engineering Research and Education; workshop held June 5 - 6, 2003
http://www.nsf.gov/eng/general/Workshop/cyberinfrastructure/index.jsp

Cyberinfrastructure for Environmental Research and Education (2003); workshop held October 30 - November 1, 2002
http://www.ncar.ucar.edu/cyber/cyberreport.pdf

Geoinformatics: Building Cyberinfrastructure for the Earth Sciences (2004); workshop held May 14 - 15, 2003; Kansas Geological Survey Report 2004-48
http://www.geoinformatics.info/

Geoscience Education and Cyberinfrastructure, Digital Library for Earth System Education, (2004); workshop held April 19-20, 2004
http://www.dlese.org/documents/reports/GeoEd-CI.pdf
Identifying Major Scientific Challenges in the Mathematical and Physical Sciences and their CyberInfrastructure Needs, workshop held April 21, 2004
http://www.nsf.gov/attachments/100811/public/CyberscienceFinal4.pdf

Materials Research Cyberscience enabled by Cyberinfrastructure; workshop held June 17 - 19, 2004
http://www.nsf.gov/mps/dmr/csci.pdf

Multiscale Mathematics Initiative: A Roadmap; workshops held May 3-5, July 20-22, September 21-23, 2004
http://www.sc.doe.gov/ascr/mics/amr/Multiscale%20Math%20Workshop%20Roadmap%20Report%20edition.pdf

Planning for Cyberinfrastructure Software (2005); workshop held October 5 - 6, 2004
http://www.nsf.gov/od/oci/ci_workshop/index.jsp

Preparing for the Revolution: Information Technology and the Future of the Research University (2002); NRC Policy and Global Affairs, 80 pages
http://www.nap.edu/catalog/10545.html

Revolutionizing Science and Engineering Through Cyberinfrastructure: report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure; Daniel E. Atkins (Chair), January 2003
http://www.nsf.gov/od/oci/reports/atkins.pdf

Workshop on the Challenges of Scientific Workflows, May 1-2, 2006, Arlington, VA
http://vtpc.isi.edu/wiki/index.php/Main_Page

Workshop on Cyber-Based Combustion Science, April 19-20, 2006, Arlington, VA
http://www.nsf-combustion.umd.edu/

Identifying Major Scientific Challenges in the Mathematical and Physical Sciences and Their CyberInfrastructure Needs, held on April 21, 2004;
http://www.nsf.gov/attachments/100811/public/CyberscienceFinal4.pdf

A Science-Based Case for Large-Scale Simulation; workshop held June 24-25, 2003;
http://www.pnl.gov/scales/docs/volume1_72dpi.pdf
http://www.pnl.gov/scales/docs/SCaLeS_v2_draft_toc.pdf

Computation as a Tool for Discovery in Physics,” October 2002,
http://www.nsf.gov/pubs/2002/nsf02176/start.htm

Cyber Chemistry Workshop; workshop held October 3-5, 2004;
http://bioeng.berkeley.edu/faculty/cyber_workshop

Materials Research Cyberscience enabled by Cyberinfrastructure; workshop held November 2004;
http://www.nsf.gov/mps/dmr/csci.pdf