Future Perspectives from the Advanced Scientific Computing Advisory Committee

Highlights of the ASCAC March 2006 Meeting

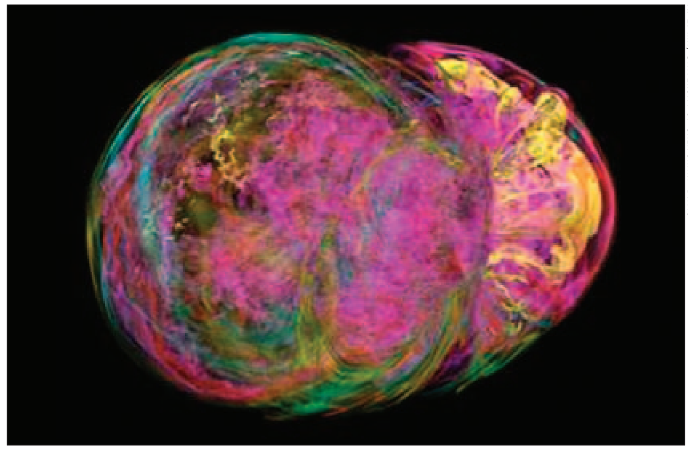
Dr. Jill Dahlburg

Dr. Raymond Orbach, Department of Energy Office of Science (DOE SC) Director, termed the 2006 State of the Union address as "historic" because President Bush announced the American Competitiveness Initiative that is intended to boost critical basic research in the physical sciences by doubling funding to the DOE SC, the National Science Foundation, and the National Institute of Standards and Technology core over the next decade. During his Advanced Scientific Computing Advisory Committee (ASCAC) meeting overview presentation in March 2006, he called this commitment a testimony to ASCAC leadership. Ultra-scale scientific computing, the centerpiece initiative of the Office of Advanced Scientific Computing Research (ASCR), is the second priority in the DOE SC twenty-year plan, "Facilities for the Future of Science: a Twenty-Year Outlook." Dr. Orbach noted that ASCR will receive a 35.7% funding increase this year for research towards this initiative; this is the largest percentage rise within the DOE SC.

In his overview of ASCR following Dr. Orbach's opening talk, Dr. Michael Strayer (DOE SC ASCR Associate Director) expressed that his aim for ASCR with SciDAC is that they will form the nexus of a global village for high-end computing (HEC), which will enable whole new levels of scientific discovery to be realized. The goal is to sponsor the best-in-class HEC research and capabilities—including facilities, simulation, middleware and collaboratory tools, applied mathematics, and computational and computer science—to advance forefront DOE SC science and technology, with an emphasis on the petascale.

Both successes and challenges abound. Dr. Strayer described the recent discovery of a new instability important to core collapse supernovae (figure 1) in two and three dimensions. Briefing about progress toward developing BlueGene/P and /Q, Dr. Rick Stevens (Argonne National Laboratory; University of Chicago) noted that significant innovations are needed in order to achieve sustained petascale computing in the next decades (see table). Networking researchers are preparing for the tremendous quantities of future HEC data: ESNet has stably connected SC assets to scientists worldwide since 1987, and ASCR-developed middleware recently enabled a 10-day flow of 500 TBytes from CERN to its associated sites. The ASCR plans to participate in the new DOE Global Nuclear Energy Partnership by providing 100-Tflop computing resources for simulation of reactor designs, reprocessing, fuel fabrication, and waste disposal.

Dr. Orbach issued two new charges to the ASCAC during the March meeting. ASCAC members Gordon Bell (Microsoft®) and Dr. James Hack (National Center for Atmospheric Research) are spearheading a committee response to the first charge, on the topic of ASCR facilities science-based performance metrics. They expect to provide an interim report at the August 8–9, 2006 ASCAC meeting, and a final report at the November 8–9, 2006 ASCAC meeting. ASCAC members Dr. Stephen Wolff (Cisco®) and Dr. Ellen Stechel (Sandia National Laboratories) are leading development of the response to the charge on the roles and efficacy of networking within the DOE SC. The interim report for this second charge is to be presented at the November 2006 ASCAC meeting, with the final report due in the autumn of 2007.



K. L. Ma, University of California–Davis

Figure 1. *The supernova shock wave instability in three dimensions.*

Dr. Jill Dahlburg, ASCAC Chair; Naval Research Laboratory, Washington DC, on behalf of the ASCAC

Acknowledgements
These highlights were largely summarized from the March 2006 ASCAC Meeting Minutes (Fred O'Hara, recording secretary)

Further Reading
All talks refer to those presented at ASCAC Meeting, March 15, 2006; http://www.er.doe.gov/ascr/March06presentationpage.htm

Future Scaling Without Innovation

Scaling current peak performance numbers for various architectures and allowing system peak doubling every 18 months

	Projected Year	BlueGene/L	Earth Simulator	Mare Nostrum
250 TF	2005	1.0 MWatt	100 MWatt*	5 MWatt
1 PF	2008	2.5 MWatt	200 MWatt*	15 MWatt
10PF	2013	25 MWatt*	2,000 MWatt*	150 MWatt*
100 PF	2020	250 MWatt*	20,000 MWatt*	1,500 MWatt*

**Trouble Ahead.*

High Productivity Computing Systems

DARPA HPCS Program

Dr. Frederick Johnson

The key to understanding the DARPA (Defense Advanced Research Projects Agency) HPCS (High Productivity Computing Systems) program is that the 'P' in HPCS stands for productivity and not petascale or performance. While execution performance is an essential component of system productivity, the HPCS program also targets major improvements in ease-of-use and development time activities to enhance productivity for users of HPCS systems.

In the latter part of the 1990s almost all high-performance computing (HPC) systems were based on clusters of commodity nodes, and the programming model for these systems was MPI coupled with a standard computing language such as C, C++, or Fortran. Many mission-critical applications of the federal government were difficult to implement on these machines, and after many detailed studies, DARPA was charged to develop the HPCS program. The program began in 2002 with an overall goal of providing a new generation of economically viable, high productivity computing systems for the national security, mission agency, and industrial user communities by 2010. Specifically, HPCS systems are expected to impact HPC in many areas, including:

- **Performance** (time-to-solution): speed up critical applications by factors of 10 to 40;

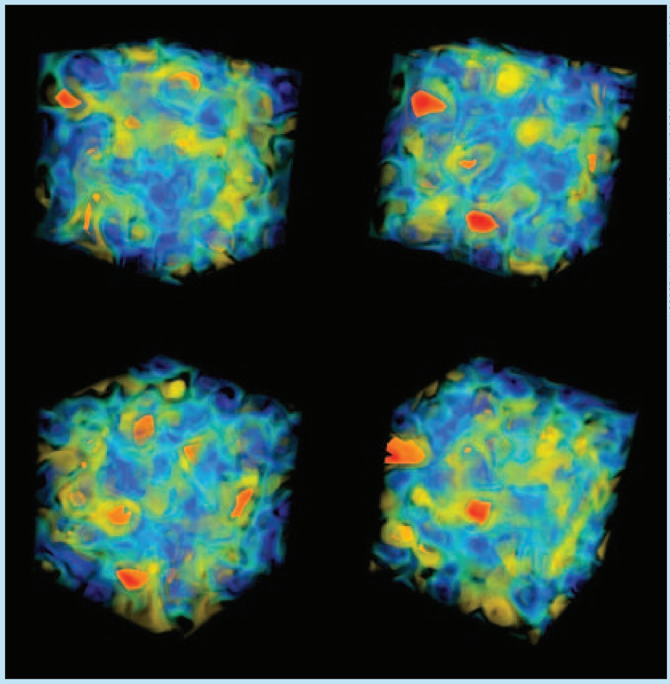
Nobel Laureate Frank Wilczek on:

Quantum Chromodynamics

Dr. Frank Wilczek shared the 2004 Nobel Prize with Dr. David J. Gross and Dr. H. David Politzer for "the discovery of asymptotic freedom in the theory of the strong interaction," or the interpretation of the high energy, short distance interactions of quarks and gluons inside protons, neutrons, and other particles.

Quantum chromodynamics (QCD) is our theory of the strong interaction, the most powerful of the four fundamental interactions in nature (the others are gravity, electromagnetism, and the weak nuclear interaction). It has been used with great success to describe how quarks and gluons are produced in accelerators, and to calculate the mass of the proton. That result, one of the greatest in the history of science, was achieved through efforts of the lattice gauge theory community. It gives a new and fundamental account, based on equations of high symmetry, of how most of the mass of matter arises from pure energy.

QCD is potentially our most perfect theory. Uniquely among the fundamental theories of physics, it gives us precise equations that can be extrapolated to arbitrarily small distances or arbitrarily high energies without breaking down. As befits the grandeur and generality of these equations, they are difficult to solve. To build upon our success, and do justice to QCD's potential to describe extreme conditions in the early universe, the deep interior of neutron stars, supernovae explosions, and more—not to mention more conventional nuclear and accelerator physics—we must continue to harness the full power of modern computing, and keep pushing it further.



VISUALIZATION: M. McGUIGAN AND S. TOMOV DATA: T. IZUBUCHI

Figure 2. *Shown above are four three-dimensional slices of the topological charge density of a lattice-QCD gauge configuration with red indicating high values of the topological charge. It was generated at the RIKEN BNL Research Center which operates two QCDOC supercomputers each with 12,288 nodes (10 teraflops).*

Dr. Frank Wilczek, 2004 Nobel Laureate, Herman Feshbach Professor of Physics, Center for Theoretical Physics, MIT

- **Programmability** (idea-to-first-solution): reduce cost and time for developing application solutions;
- **Portability:** insulate application software from system specifics;
- **Robustness:** protect applications from hardware faults and system software errors.

While we have many metrics for evaluating execution performance, there are few metrics for development time productivity. However, an expert in the intelligence community has said that HPCS systems would be successful in his area if they increased the number of people able to effectively use the systems by a factor of ten. This is one practical measure of development productivity.

The HPCS program has three phases. Phase one was a one-year concept study phase, in which five supercomputer vendors Cray, HP, IBM, SGI, and SUN participated. Phase two is a three-year research and development period, in which Cray, IBM, and SUN participate. Phase two will conclude this summer, and phase three is about to commence. All three phase two vendors have submitted proposals, which are now under evaluation.

Currently, few details concerning the HPCS architectural designs are available publicly, but more and more details will emerge as phase three progresses. This information will allow application developers and system software experts to prepare for these systems.

The system vendor activities represent the principal focus of the HPCS program. However, the system has several other components as well:

- **Productivity Assessment**—The productivity assessment team conducts research in development and execution time productivity metrics and benchmarks. Part of these activities is the HPC Challenge benchmark suite and result tables, which augment the Top500 Linpack results with several metrics that are more meaningful to the HPC community.
- **High Performance Language Systems**—Each of the phase two vendors is developing and studying new languages for simplifying the programming of petascale parallel systems. Cray's language is Chapel, IBM's language is X10, and SUN's language is Fortress. Additional information about these languages is readily available on the Internet.
- **Application Development**—Key mission partners will undertake petascale application development activities to accelerate the availability of petascale applications for HPCS systems. It is likely that certain SciDAC applications will have the opportunity to participate in this process as these ongoing plans mature.

The DOE SC and National Nuclear Security Administration (NNSA) have been active in supporting the HPCS program from its inception. Representatives from these two organizations have participated in proposal evaluations for all three phases, and in all of the quarterly progress reviews with each vendor—over 50 reviews. Much more importantly, SC and NNSA are HPCS mission-partners, having provided financial backing, with plans to continue funding. Mission-partner application needs have been incorporated into the requirement specifications for the phase three systems. SC has funded research activities in development and execution time productivity in phase two, and that will continue and possibly expand this effort in phase three.

With strong participation from a wide spectrum of government agencies, including SC, NNSA, NSA, NASA, NSF, NRO and the DOD Modernization Program, the HPCS program exhibits innovation in interagency coordination, as well as HPC hardware and software. It is very likely that the HPCS systems that emerge in 2010 will significantly impact agency mission-critical applications.

Dr. Frederick C. Johnson, Senior Technical Manager for Computer Science, Office of Advanced Scientific Computing Research

Further Reading

http://www.darpa.mil/ipto/programs/hpcs/index.htm
http://www.highproductivity.org/
http://icl.cs.utk.edu/hpcc/

Funding

SciDAC-2: Reaching Out to New Communities

Dr. Walter M. Polansky

As most of the projects funded under the initial SciDAC program officially end and the resulting success stories underscore the value of multidisciplinary collaborations, SciDAC-1 might appear to be a very hard act to follow. But with the announcement of SciDAC-2 projects, it is clear that the Department of Energy (DOE) and the nation's scientific computing community are up to the task.

To start with, the response to the SciDAC-2 call for proposals was overwhelming. In all, we received over 350 letters of intent, resulting in 240 full proposals. After a month of internal review, those proposals were scrutinized for three weeks in intensive peer review panels. Based on these peer review recommendations 33 of the most promising projects were selected to be supported in the second round of SciDAC funding.

Among the reasons for the strong response to SciDAC-2 was the addition of new scientific application areas. These new disciplines and grand challenges, along with the core areas of research from SciDAC-1, offer the promise of an even broader range of scientific discoveries. The research areas to be supported under SciDAC-2 are summarized below:

- **Physics**—The DOE is the federal government's primary source of support for physics research. Therefore, physics is the largest scientific application area in SciDAC-2. These efforts include projects in astrophysics that are striving to shed light on the dark matter and dark energy that make up 95% of our universe, as well as research aimed at learning more about supernovae and their role in the creation of the chemical elements. Other efforts focus on nuclear structure, lattice quantum chromodynamics, turbulence, and preparing to manage and analyze the massive amounts of data that are expected from large physics experiments such as the Large Hadron Collider in Europe. Some of these efforts are also partially supported by the National Science Foundation (NSF) and/or the National Nuclear Security Administration (NNSA).
- **Climate Modeling and Simulation**—The DOE's charge to elucidate the impacts of energy production and use on the environment continues to lead the evolution of climate modeling and simulation research in conjunction with the Climate Change Prediction Program (CCPP). SciDAC efforts will advance the development of future climate models based on theoretical foundations and improved computational methods that dramatically increase both the accuracy and throughput of computer model-based predictions of future climate system response to the increased atmospheric concentrations of greenhouse gases.
- **Computational Biology**—The Department's missions in energy and the environment include life sciences research on microbes and microbial communities that have the potential to generate hydrogen or ethanol, to sequester carbon dioxide, or help with environmental remediation. This field is new to the SciDAC portfolio and will focus on developing new methods for modeling complex biological systems, including molecular complexes, metabolic and signaling pathways, individual cells, and ultimately interacting organisms and ecosystems. Such systems act on time scales ranging from microseconds to thousands of years. Additionally, the systems must be coupled to huge databases created by an ever-increasing number of high-throughput experiments.
- **Fusion Energy Science**—SciDAC efforts will continue to develop and improve the simulation and modeling of fusion systems essential for achieving the predictive scientific understanding needed to make fusion energy practical. Current large-scale simulations in fusion

Conference

SCIDAC 2006 Conference Summary

Dr. William Tang

The second annual Scientific Discovery through Advanced Computing (SciDAC) conference was held June 25–29, 2006 at the new Hyatt Regency Hotel in Denver, Colorado. This conference showcased outstanding SciDAC sponsored computational science results achieved during the past year, emphasizing science at terascale. The program also featured exciting developments in computational science from national and international contributors outside of the SciDAC program, to help foster communication between SciDAC computational scientists and those funded by other organizations. The conference emphasized the cooperative efforts of domain scientists, applied mathematicians, and computer scientists to effectively take advantage of terascale computers to accelerate progress in scientific discovery in a variety of fields, as well as the development of new applications for petascale computing in the near future.

Application domains within the SciDAC 2006 conference agenda encompassed a broad range of science including:

- The DOE core mission of energy research involving combustion, including studies relevant to current issues of fuel efficiency and pollution as well as magnetic fusion investigations impacting prospects for future energy sources.
- Fundamental explorations into the building blocks of matter, ranging from quantum chromodynamics—the basic theory describing the composition of all matter—to the design of modern high-energy accelerators.
- The formidable challenges of predicting the behavior of molecules in quantum chemistry, and investigating the complex biomolecules that form the basis of all biological systems.
- Studies of the dynamics of exploding stars for insights into the nature of the universe.
- Integrated climate modeling to enable realistic analysis of Earth's changing climate.

KERR PHOTOGRAPHY

Figure 3. *A presentation captures the attention of SciDAC 2006 participants.*

plasma science include integrated modeling of electromagnetic wave interactions with plasmas, research on understanding the plasma edge, and modeling plasma turbulence and macroscopic stability.

- **Groundwater Reactive Transport Modeling and Simulation**—The DOE's efforts to contain and remediate contaminated sites challenge the state of the art in many areas. Scientifically rigorous models of subsurface reactive transport that accurately simulate contaminant mobility across multiple length scales remain elusive. New SciDAC efforts in this area aim to provide more advanced models for better following the movement of underground contaminates. This will benefit environmental cleanup efforts at DOE facilities and improve the monitoring of contaminants in groundwater around existing and future radionuclide waste disposal and storage sites with the aim of ultimately preventing environmental hazards. These efforts will also assist the Department's research on deep geologic carbon sequestration.

The response to the SciDAC-2 call for proposals was overwhelming. The addition of new scientific application areas, along with the core areas of research from SciDAC-1, offer the promise of an even broader range of scientific discoveries.

- **Materials Science & Chemistry**—Ongoing SciDAC efforts in materials science and chemistry will be supplemented with efforts, in partnership with, and focused on the needs of the NNSA, that are focused on the needs of that program. These include quantum simulations of materials and nanostructures, simulations of stress corrosion cracking, and multiscale simulations of strongly correlated materials. New efforts will be coordinated with current projects to improve understanding and attain accurate modeling of material properties, reactions and interactions, on length scales that are extended by ten orders of magnitude or more.

In support of these advanced scientific applications, SciDAC-2 will also include multidisciplinary teams to create computational technologies that overcome some of the mathematics, computer science, and networking challenges of petascale computing. These teams bring together experts in the various scientific disciplines, computer scientists, and applied mathematicians to focus on the immediate needs of the applications and to anticipate future challenges. Under SciDAC-2, teams will be supported in the following three different organizational structures.

Centers for Enabling Technologies (CETs) will address the need for new algorithms that scale to parallel systems having hundreds of thousands of processors. CETs will also work to attain methodologies that can achieve portability and interoperability of complex high performance scientific software packages, and to develop operating systems and runtime tools that support application execution performance and system management. Finally CETs will develop effective tools for remote access, feature identification, data management, and visualization of petabyte-scale scientific datasets. These CETs will broadly serve the scientific application teams and the SciDAC community.

SciDAC Institutes are university-led centers of excellence that will complement the efforts of the CETs, as well as focus on outreach to new communities and on educating the next generation of computational scientists in a range of scientific domains.

Science Application Partnerships (SAPs, or Partnerships) also complement the CETs but are targeted efforts attached to a specific scientific application.

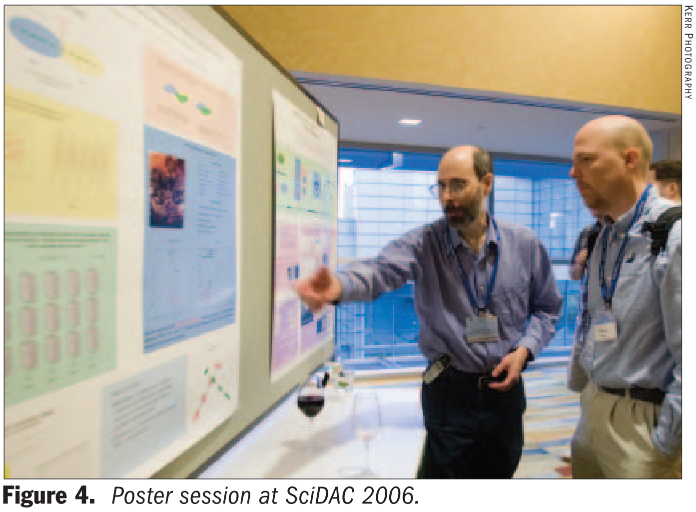
All SciDAC teams will develop project specific websites. More information on SciDAC, including a listing of new and continuing projects can be found at www.scidac.org.

Together the SciDAC-2 teams will strive to match or exceed the advances that were realized through SciDAC-1, while broadening both the community of practitioners and the areas of scientific application.

Dr. Walter M. Polansky, Senior Technical Advisor for Project Management at the Office of Advanced Scientific Computing Research, Office of Science, DOE

Because of the huge temporal and spatial scales involved, advanced computation is often the only method by which significant progress can be made when dealing with these complex, multicomponent physical, chemical, and biological systems. Working with domain scientists, applied mathematicians and computer scientists have continued to develop improved discretizations of the underlying equations and their complementary algorithms in order to enable faster and more detailed solutions on modern parallel computing platforms, particularly as they evolve from the terascale toward the petascale regime. Moreover, the associated tremendous growth of data generated from the terabyte to the petabyte range demands not only advanced data analysis and visualization methods to extract useful information but also the development of efficient workflow strategies which can deal with the data input/output, management, movement, and storage challenges. If scientific discovery is expected to keep pace with the continuing progression from tera- to petascale platforms, the vital alliance between domain scientists, applied mathematicians, and computer scientists will be even more crucial. During the SciDAC 2006 Conference, some of the future challenges and opportunities in interdisciplinary computational science were discussed in the Advanced Architectures Panel and by Dr. Victor Reis, Senior Advisor to the Secretary of Energy, who gave a featured presentation on "Simulation, Computation, and the Global Nuclear Energy Partnership."

Overall, the conference provided an excellent opportunity to highlight the rising importance of computational science in the scientific enterprise and to motivate future investment in this area. As Dr. Michael Strayer, SciDAC Program Director, has noted: "While SciDAC may have started out as a specific program, Scientific Discovery through Advanced Computing has become a powerful concept for addressing some of the biggest challenges facing our nation and our world." Looking forward to next year, SciDAC 2007 will be held June 24–28 at the Westin Copley Plaza in Boston, Massachusetts, and will be chaired by Dr. David Keyes of Columbia University.



KERR PHOTOGRAPHY

Figure 4. *Poster session at SciDAC 2006.*

Dr. William Tang, Chief Scientist at the Princeton Plasma Physics Laboratory; Associate Director of the Princeton Institute of Computational Science and Engineering at Princeton University; Chair of the SciDAC 2006 Conference

Further Reading

http://www.scidac.org/Conference2006