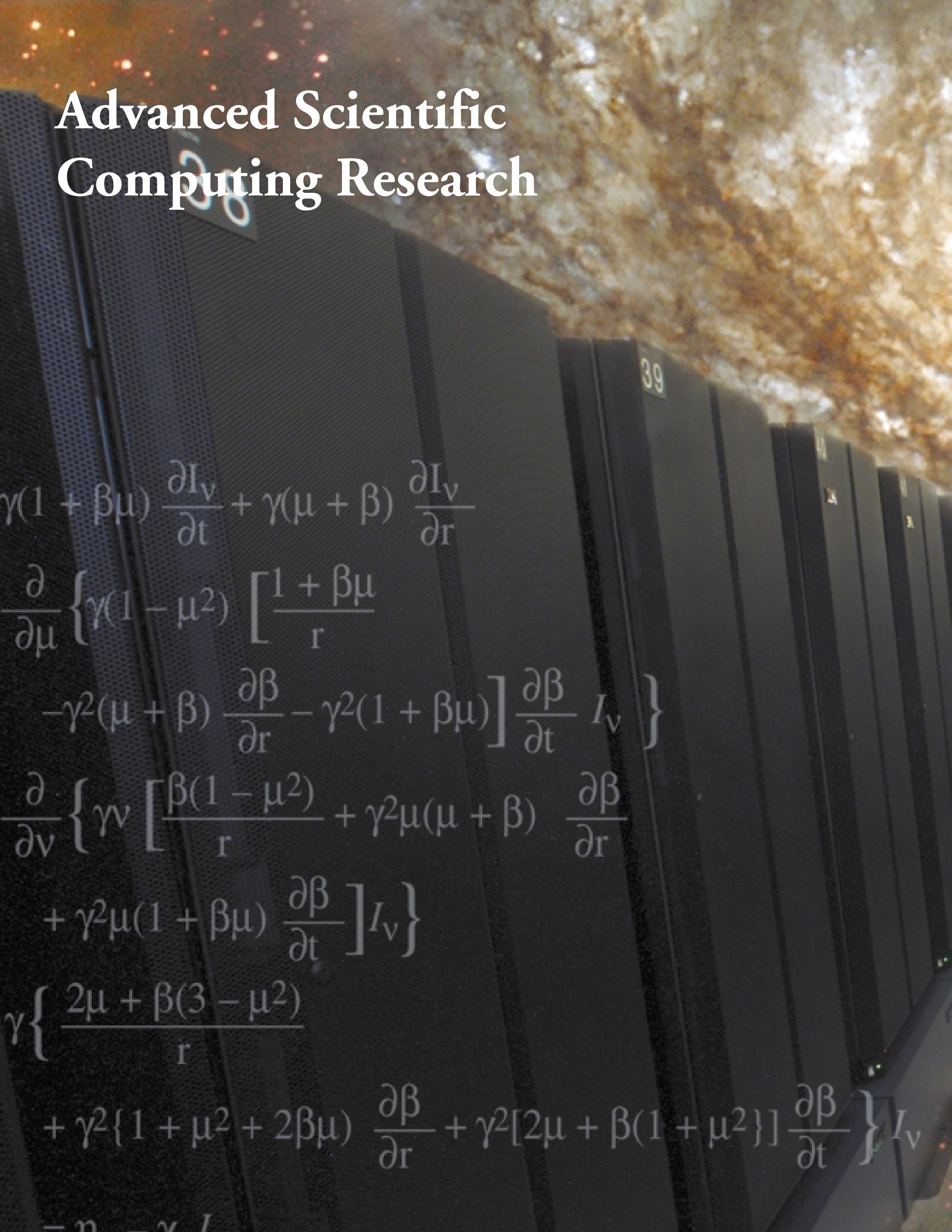
Advanced Scientific Computing Research



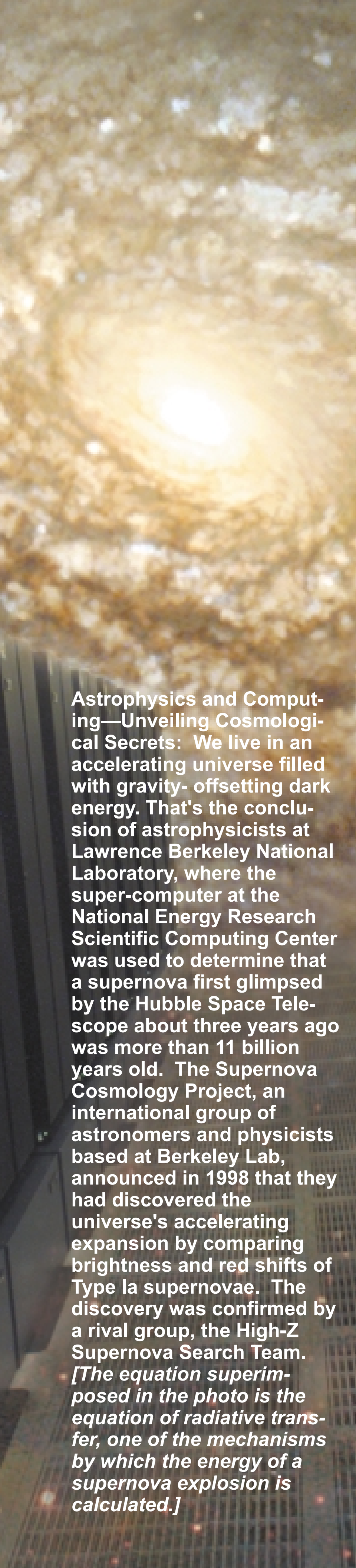
6 Deliver Computing for the Frontiers of Science

Deliver forefront computational and networking capabilities to scientists nationwide that enable them to extend the frontiers of science, answering critical questions that range from the function of living cells to the power of fusion energy.

Computer-based simulation enables us to predict the behavior of complex systems that are beyond the reach of our most powerful experimental probes or our most sophisticated theories. Computational modeling has greatly advanced our understanding of fundamental processes of Nature, such as fluid flow and turbulence or molecular structure and reactivity. Through modeling and simulation, we will be able to explore the interior of stars and learn how protein machines work inside living cells. We can design novel catalysts and high-efficiency engines. Computational science is increasingly central to progress at the frontiers of almost every scientific discipline and to our most challenging feats of engineering.

The science of the future demands that we advance beyond our current computational abilities. Accordingly, we must address the following challenges:

- What new mathematics are required to effectively model systems such as the Earth's climate or the behavior of living cells that involve processes taking place on vastly different time and/or length scales?
- Which computational architectures and platforms will deliver the most benefit for the science of today and the science of the future?
- What advances in computer science and algorithms are needed to increase the efficiency with which supercomputers solve problems for the Office of Science?
- What operating systems, data management, analysis, model development, and other tools are required to make effective use of future-generation supercomputers?
- Is it possible to overcome the geographical distances that often hinder science by making all scientific resources readily available to scientists, regardless of whether they are at a university, national laboratory, or industrial setting?



Astrophysics and Computing—Unveiling Cosmological Secrets: We live in an accelerating universe filled with gravity- offsetting dark energy. That's the conclusion of astrophysicists at Lawrence Berkeley National Laboratory, where the super-computer at the National Energy Research Scientific Computing Center was used to determine that a supernova first glimpsed by the Hubble Space Telescope about three years ago was more than 11 billion years old. The Supernova Cosmology Project, an international group of astronomers and physicists based at Berkeley Lab, announced in 1998 that they had discovered the universe's accelerating expansion by comparing brightness and red shifts of Type Ia supernovae. The discovery was confirmed by a rival group, the High-Z Supernova Search Team. *[The equation superimposed in the photo is the equation of radiative transfer, one of the mechanisms by which the energy of a supernova explosion is calculated.]*

LBNL

LANL

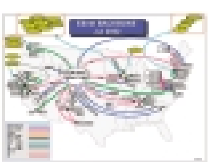
Pioneering computers—then and now: The first "high-speed" computer, MANIAC (Mathematical Analyzer, Numerical Integrator And Computer), was developed as part of the nuclear weapons program at Los Alamos National Laboratory in 1952. MANIAC was only available to the foremost scientists around the country to help solve the critical scientific problems of that era. It occupied a large room and was the first computer programmed to play chess, possessing enough memory to store up to 5000 words. It is hardly a comparison to the laptops of today, each with gigabytes of memory and available now to children in grade schools. Many technologies from Office of Science programs have contributed to the present generation of computers, and the Office of Science operates the premier supercomputer available for civilian research and development within DOE at NERSC.

The Office of Science will deliver models, tools, and computing platforms to dramatically increase the effective computational capability available for scientific discovery in fusion, nanoscience, high-energy and nuclear physics, climate and environmental science, and biology. We will develop new mathematics and computational methods for modeling complex systems; work with the scientific community and vendors to develop computing architectures tailored to simulation and modeling; develop improved networking resources; and support interdisciplinary teams of scientists, mathematicians, and computer scientists to build sophisticated computational models that fully exploit these capabilities. Our role complements and builds on the National Nuclear Security Administration's Accelerated Strategic Computing Initiative, delivering forefront modeling capabilities for stockpile stewardship, the basic computer science and mathematics research programs conducted by the National Science Foundation, and mission-focused programs of other agencies.

As an integral part of this Strategic Plan, and in *Facilities for the Future of Science: A Twenty-Year Outlook*, we have identified the need for three future facilities to realize our Advanced Scientific Computing Research vision and to meet the science challenges described in the following pages. All three of the facilities are near-term priorities: the **UltraScale Scientific Computing Capability (USSCC)**, the **Energy Sciences Network (ESnet) Upgrade**, and the **National Energy Research Scientific Computing Center (NERSC) Upgrade**. The USSCC, located at multiple sites, will increase by a factor of 100 the computing capability available to support open (as opposed to classified) scientific research—reducing from years to days the time required to simulate complex systems, such as the chemistry of a combustion engine, or

Our History of Discovery…Select Examples

1970s
Established the first national unclassified computer center, the Controlled Thermonuclear Research Center, the forerunner to today's National Energy Research Scientific Computing Center.



1980s
Built ESnet to link research facilities and supercomputers to users and the emerging Internet.

1991
Pioneered the transition to massively parallel supercomputing, enabling 1000 or more processors to work together.

1970 *1980* *1990*

1970s
Determined the emergence of chaotic behavior in systems thought to be stable.

1990s
Installed the first supercomputer available to the civilian research community that broke the peak performance barrier of 1 teraflop computing speed.

weather and climate—and providing much finer resolution. The ESnet upgrade will enhance the network services available to support Office of Science researchers and laboratories and maintain their access to all major DOE research facilities and computing resources, as well as fast interconnections to more than 100 other networks. The NERSC upgrade will ensure that DOE's premier scientific computing facility for unclassified research continues to provide high-performance computing resources to support the requirements of scientific discovery. All three facilities are included in our Advanced Scientific Computing Research Strategic Timeline at the end of this chapter and in the facilities chart in Chapter 7 (page 93), and they are discussed in detail in the *Twenty-Year Outlook*.

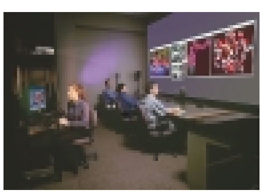
Our Strategies

6.1 Advance scientific discovery through research in the computer science and applied mathematics required to enable prediction and understanding of complex systems.

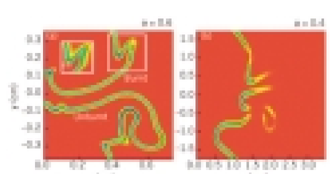
New computational methods are needed to make possible the simulation of the most complex physical and biological systems and to gain efficiency on multiprocessor terascale computers. Effective application of supercomputers requires sophisticated, scalable, operating systems; large-scale data management tools; and other computer science tools. We will support individual investigators and teams to develop new methods and tools, and encourage their transition to advanced computational science applications.

Our strategy includes the following emphases:

- Develop new and improved mathematical methods for addressing the challenges of multi-scale problems.
- Create methods and capabilities to address large-scale data management.
- Develop and apply middleware tools that enable researchers to focus on science while obtaining effective computational performance.



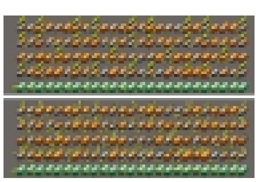
1992
Launched the first Internet videoconference.



1998
Simplified the development of scientific simulations in complex geometrics such as diesel engines.

1998
Wrote the first application code that surpassed one teraflop.

2000



1995
Developed a "spin dynamics" computational method to accurately model magnetic materials.

1998
Announced the discovery from the Supernova Cosmology Project that the universe is expanding.

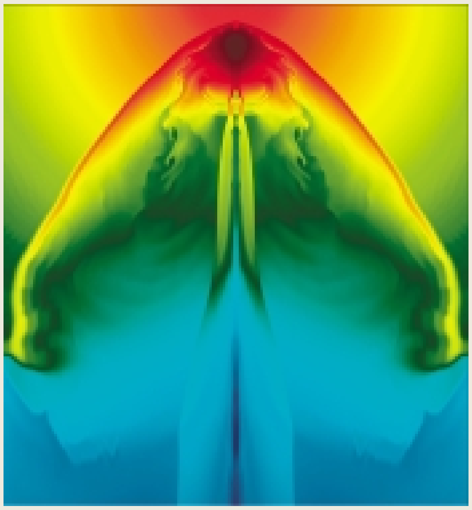
2001
Launched Scientific Discovery through Advanced Computing (SciDAC), a program that accelerates advances in computing and information technologies as tools for scientific discovery.

Scientific Discovery Through Advanced Computing (SciDAC)

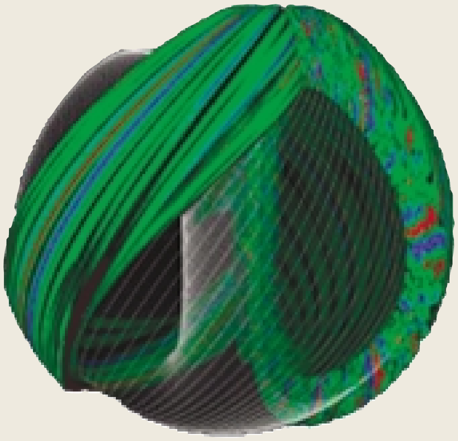
SciDAC is a research program with the goal to achieve breakthrough scientific advances through computer simulation. SciDAC has established a new model for collaboration among the scientific disciplines, computer scientists, and mathematicians. The SciDAC program is creating a new generation of scientific simulation codes and developing collaboratory software to enable geographically separated scientists to use scientific instruments and computers remotely, enabling distant colleagues to share data and function together as a team.

Current projects involve collaborations among 13 DOE laboratories and more than 50 colleges and universities in a broad spectrum of projects such as:

- Climate simulation and prediction
- Quantum chemistry and fluid dynamics
- Plasma systems to advance fusion energy science
- High energy and nuclear physics
- Software infrastructure
- Applied Mathematics Integrated Software Infrastructure Centers
- Computer Science Integrated Software Infrastructure Centers
- National collaboratory, middleware, and network research.



Model of a Supernova Blast



Plasma Microturbulence Simulation

6.2 Extend the frontiers of scientific simulation through a new generation of computational models that fully exploit the power of advanced computers and collaboratory software that makes scientific resources available to scientists anywhere, anytime.

Scientific discovery in many areas requires computational models that incorporate more complete and realistic descriptions of the phenomena being modeled than are possible today.

Our strategy includes the following emphases:

- Create, in partnerships across the Office of Science, new generations of models for fusion science, biology, nanoscience, physics, chemistry, climate, and related fields that provide high-fidelity descriptions of the underlying science.
- Incorporate the new models into scientific simulation software that achieves substantially greater performance from terascale supercomputers than we can achieve today.
- Build on the successes of the SciDAC program.

6.3 Bring dramatic advances to scientific computing challenges by supporting the development, evaluation, and application of supercomputing architectures tailored to science.

Major improvements in scientific simulation and analysis can be obtained through advances in the design of supercomputer architectures. Most of today's supercomputers were designed for commercial applications. However, computational science places stringent requirements on supercomputer designs that are often quite different from what arise in commercial applications. To meet the need for effective computing performance in the 100-teraflop range and beyond, we will support the evaluation, installation, and application of new very high-end computing architectures for computational science.

Our strategy includes the following emphases:

- Develop partnerships with U.S. industry in the near term to adapt current and next-generation products to more fully meet the needs of visionary computational science.
- Develop partnerships with the Department of Defense, the Defense Advanced Research Projects Agency (DARPA), and other Federal agencies to evaluate long-term architecture developments at the scale needed for Office of Science computation.
- Advance the focused research and development of systems software for radical increases in performance, reliability, manageability, and ease of use.

6.4 Provide computing resources at the petascale and beyond, network infrastructure, and tools to enable computational science and scientific collaboration.

Work at the forefront of science can require the dedicated availability of the most advanced supercomputers for extended periods of time. Furthermore, it is likely that at least a few different supercomputer designs will offer significant advantages for different classes of problems.

Our strategy includes the following emphases:

- Provide sustained, high-bandwidth access to the highest possible performance computers for the most demanding applications at the scientific frontiers.
- Upgrade the network and data management infrastructure supporting these resources to enable computational scientists to manage the extraordinarily large volumes of data often generated by large-scale scientific computing and modern experiment.
- Create supporting resources, grid nodes, and tools that enable teams of scientists to collaborate effectively at a distance.

Computing test beds: Advanced Computing Research test beds evaluate new computing hardware and software, such as Oak Ridge National Laboratory's IBM Power4 Cheetah (pictured left) and Cray Xl, and Argonne National Laboratory's IBM/Intel/ Cluster.

ORNL

"It is unworthy of excellent men to lose hours like slaves in the labor of calculation which could be relegated to anyone else if machines were used."

—Gottfried Wilhelm von Leibnitz (1646-1716), German philosopher and mathematician

Our Timeline and Indicators of Success

Our commitment to the future and to the realization of **Goal 6: Deliver Computing for the Frontiers of Science** is not only reflected in our strategies, but also in our Key Indicators of Success, below, and our Strategic Timeline for Advanced Scientific Computing Research (ASCR), at the end of this chapter.

The ASCR Strategic Timeline charts a collection of important, illustrative milestones, representing planned progress within each strategy. These milestones, while subject to the rapid pace of change and uncertainties that belie all science programs, reflect our latest perspectives on the future—what we hope to accomplish and when we hope to accomplish it—over the next 20 years and beyond. Following the science milestones, toward the bottom of the timeline, we have identified the required major new facilities. These facilities, described in greater detail in the DOE Office of Science companion report, *Facilities for the Future of Science: A Twenty-Year Outlook*, reflect time-sequencing that is based on the general priority of the facility, as well as critical-path relationships to research and corresponding science milestones.

Additionally, the Office of Science has identified Key Indicators of Success, designed to gauge our overall progress toward achieving Goal 6. These select indicators, identified below, are representative long-term measures against which progress can be evaluated over time. The specific features and parameters of these indicators, as well as definitions of success, can be found on the web at **www.science.doe.gov/measures**.

Key Indicators of Success:

- Progress toward developing the mathematics, algorithms, and software that enable effective scientifically critical models of complex systems, including highly nonlinear or uncertain phenomena, or processes that interact on vastly different scales or contain both discrete and continuous elements.
- Progress toward developing, through the Genomics: GTL partnership with the Biological and Environmental Research program, the computational science capability to model a complete microbe and a simple microbial community.

Strategic Timeline for Advanced Scientific Computing Research

Strategic Timeline—Advanced

2003 | 2005 | 2007 | 200

The Science

Computer Science and Applied Mathematics Research

- Complete ASCR roadmap that defines national approach to the challenges of mathematics for complex systems (2003)
- Deliver operating systems for scientific computers that incorporate fault tolerance (2005)
- Deliver algorithms that scale to tens of thousands of processors for key mathematical libraries (2007)

Extending Science through Computation and Collaboration

- Simulate gyrokinetic transport of fusion plasma without detailed electron dynamics (2004)
- Enable secure, remote operation of fusion facilities (2004)
- Complete computational model of gene regulation (2005)
- Calculate enhanced optical properties at the nanoscale (2007)
- Simulate the catalyst action in automobile exhaust (2007)
- Perform full three-dimensional supernova simulation (2008)
- Simulate soot formation in diesel engines (2008)
- Enable real-time collaborative remote teams at the Spallation Neutron Source (2008)

Supercomputing Architectures for Science

- Complete evaluation of Cray X1 (2003)
- Complete evaluation of first computer with more than 50,000 processors (2005)
- Initiate evaluation of systems from DARPA High Productivity Computing Systems program (2007)

Computational and Network Infrastructure and Tools

- Deliver computing facilities for open science with a 50-fold increase in capability (2005)
- Complete expansion of ESnet to deliver core bandwidth of 10 gigabytes per second (2005)
- Deliver computing facilities for open science with 100-fold increase in capability relative to 2004 (2007)
- Increase ESnet core capability by 400% (2008)

Future Facilities**

UltraScale Scientific Computing Capability (USSCC): The USSCC, located at multiple sites, will increase by a factor of 100 the computing capability available to support open (as opposed to classified) scientific research—reducing from years to days the time required to simulate complex systems.

***These strategic milestones are illustrative and depend on funds made available through the Federal budget process.**

****For more detail on these facilities and the overall prioritization process, see the companion document, *Facilities for the Future of Science: A Twenty-Year Outlook.***

Scientific Computing Research*

09	2011	2013	2015

- Complete programming model that enables scientists to use 100,000 processors (2009)
- Deliver mathematics of complex systems that enables accurate linkage of multiple time and length scales (2011)
- Deliver mathematics of complex systems that enables simulations of microbes (2013)
- Revolutionize computing in U.S. industry through research results from applied mathematics and computer science (2015)

- Achieve seamless integration of astrophysics simulation and data (2009)
- Perform climate simulations that incorporate biological carbon sequestration (2011)
- Deliver virtual catalogue that enables access to all climate data worldwide (2012)
- Complete simulation of tokamak disruptions that enable design of active control system to avoid disruptions (2013)
- Enable computational design of microbe for energy production (2014)
- Compete first integrated burning plasma simulation (2014)
- Deliver hundreds of petabytes per year of data to scientists, routinely (2015)

- Complete tests of computer systems that lead to the first system with sustained application performance over 10 petaflops (2011)

- Achieve computational capability for open science that reaches one petaflop (2012)
- Expand ESnet core capability to exceed 100 gigabytes per second (2013)

Energy Sciences Network (ESnet) Upgrade: The ESnet upgrade will enhance the network services available to support Office of Science researchers and laboratories and maintain their access to all major DOE research facilities and computing resources.

National Energy Research Scientific Computing Center (NERSC) Upgrade: This upgrade will ensure that NERSC continues to provide high-performance computing resources to support the requirements of scientific discovery.