Minutes for the
Advanced Scientific Computing Advisory Committee Meeting
May 2-3, 2001, Crowne Plaza Hotel, Washington, D.C.

ASCAC members present:
- John W. D. Connolly, Vice Chair
- Jill P. Dahlburg (Wednesday only)
- Roscoe C. Giles
- Helene E. Kulsrud
- William A. Lester, Jr.
- Gregory J. McRae
- Juan C. Meza
- Ellen B. Stechel
- Warren Washington
- Stephen Wolff
- Margaret H. Wright, Chair

ASCAC members absent:
- Karen R. Sollins

Also participating:
- Melea Baker, Office of Advanced Scientific Computing Research, USDOE
- Michael Colvin, Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory
- James Decker, Acting Director, Office of Science, USDOE
- Daniel Drell, Office of Biological and Environmental Research, USDOE
- Steven Eckstrand, Acting Assistant Director, Office of Science, USDOE
- Mike Holland, Office of Management and Budget
- Fred Johnson, Office of Science, USDOE
- Michael Knotek, Office of the Under Secretary, USDOE
- James Leighton, ESnet Project Manager, Lawrence Berkeley National Laboratory
- Frederick O'Hara, ASCAC Recording Secretary
- Edward Oliver, Office of Advanced Scientific Computing Research, USDOE
- Walter Polansky, Acting Director, Mathematical, Information, and Computational Sciences Division, Office of Science, USDOE
- Horst Simon, Director, National Energy Research Scientific Computing Center, Lawrence Berkeley National Laboratory
- Rick Stevens, Director, Mathematics and Computer Science Division, Argonne National Laboratory
- Andreene Witt, Oak Ridge Institute of Science and Education
- Thomas Zacharia, Director, Center for Computational Sciences, Oak Ridge National Laboratory

About 30 others also were in attendance.

Wednesday, May 2, 2001

The Chair, Margaret Wright, called the meeting to order at 9:47 a.m. and reported that James Decker has asked the Committee to report by January on the status of facilities and biotechnology. She noted that Edward Oliver is ASCAC's Designated Federal Officer (DFO), and she introduced **James Decker**, Acting Director of the Office of Science (SC), who spoke about the budget of the Department of Energy (DOE), especially that of SC.

This transition year is more difficult than others. Science in general is up by 6%, but most of that results from a 15% increase in the budget for the National Institutes for Health (NIH). DOE's budget is down several hundred millions of dollars, resulting from the one-time expenditure for the Los Alamos National Laboratory (LANL) fires last year. The Department is awaiting results of the Cheney report to see what the policy directions might be for energy. The report will have a large impact on the shape of the budget.

The FY02 budget for SC is essentially the same as the FY01 appropriation. Highlights include a $13 million increase for the Spallation Neutron Source (SNS); essentially the same level of funding as last year for Scientific Discovery Through Advanced Computing (SciDAC); a $4 million increase for nanoscience; an additional $20 million for a new initiative, Genomes to Life; an increase to $721 million for high-energy physics; a flat budget for fusion; and an increase of about a million dollars for science education.

The SNS is a big project; it is on budget and schedule. Construction has started and is expected to be completed in June 2006. Instruments are being procured.

Key facilities are being upgraded: instruments at the Intense Pulsed Neutron Source (IPNS), computers at the Environmental Molecular Sciences Laboratory (EMSL), detectors and increased operating time at Fermilab, increased operating time at the Stanford Linear Accelerator Center (SLAC), and terascale computing capabilities at the National Energy Research Scientific Computing Center (NERSC).

The Nanoscale Science Engineering and Technology initiative was launched last year with a request for proposals (RFP). $30 million was made available to national laboratories and universities; 745 preproposals resulted in 497 proposals from universities and 46 proposals from laboratories. As part of the initiative, DOE is planning to establish one or more Nanoscale Science Research Centers. And $3 million was allocated to support increased facility operation and greater researcher access to facilities for nanoscience research.

The Genomes to Life initiative is a partnership between the Office of Biological and Environmental Research (BER) and the Office of Advanced Scientific Computing Research (OASCR). Its goals are to identify and characterize molecular machines of life, characterize gene regulatory networking, characterize the functional repertoire of microbial communities, and develop the computational methods and capabilities needed to advance the understanding of complex biological systems and to predict behavior. Interests in finding biological solutions include human susceptibility to disease, bioremediation, carbon sequestration, etc.

This is an exciting time in physics. The physics community hopes to observe the Higgs boson at Fermilab. At SLAC, preliminary observations are indicating that neutrinos have mass, raising the question of whether they are part of the dark matter, so they are investigating CP (charge conjunction-parity) violation. With the G Minus 2 Experiment, Brookhaven National Laboratory (BNL) is measuring the magnetic properties of the muon. And at the Relativistic Heavy-Ion Collider, they may be producing a "new" state of matter, the quark-gluon plasma.

The Fusion Program is producing a lot of interesting science with computational models. They are looking at heavy-ion accelerators.

Science education is a small program and is in addition to the fellowships provided by the research programs. It includes the Undergraduate Research Experience at national laboratories. The National Science Foundation (NSF) will add funding for expansion of this program.

Decker issued a new charge to ASCAC: Look across all the computational capabilities in SC. DOE would like an outside look at this program because it holds such importance for the future. The chair would come from this Committee and members from a variety of other advisory committees.

Washington asked if he would like to see a consolidation of funding for OASCR or to continue to have it spread out. Decker replied that having it spread out is better from a budgetary strategy point of view.

Connolly asked if DOE wanted a full-scale review of facilities. Decker replied, not just facilities, but the whole program. The Department needs an outside look at the whole activity. Facilities would certainly be a component. Wright pointed out that Jill Dahlburg and her Subcommittee are already looking at facilities.

McRae asked what problems the Division had. Decker said that staffing is an issue, but that the Division had made some progress. Oliver commented that, at the previous meeting of ASCAC, the Division had five vacancies. After 6 months, it still has five vacancies. Four individuals have been identified, but DOE has yet to close the deal with any of them. Decker commented that the Department has problems living within the civil service program: salaries are not competitive, the process of hiring someone is time-consuming, and the system allows few flexibilities (e.g., filling critical positions and paying retention bonuses).

Kulsrud asked what year's budget would be affected by the report Decker is currently requesting. Decker said, FY04. A preliminary report might help formulate the FY03 budget.

Wright asked if there was any chance that Congress will change the SC budget. Decker noted that there is a lot of sympathy for the SC budget on the Hill. There may be additions made during the budget-hearings process. Members feel there are a lot of problems in the budget [especially Environmental Management (EM)] and they want to fix those.

Wright asked if there was any progress on filling the position of Director of the Office of Science. Decker observed that the administration has kept this type of information confidential. You can make input, but you do not get anything back. We will probably get a Deputy and Undersecretary in June. Wright asked if this is having any effect on the science. Decker noted that DOE is a complex place. The Secretary has to pay attention to the problem of the day. It is difficult to choose staff and do the business of the Department at the same time.

McRae asked if it makes any sense to look beyond SC for members of the Subcommittee. Decker replied, yes, for coordination of the Accelerated Strategic Computing Initiative (ASCI). The National Nuclear Security Administration (NNSA) would be receptive to participating. He suggested that he could talk to General Gordon about it. Kulsrud commented that that certainly is a significant part of the program, and Dahlburg noted that they have a university program. Connolly cautioned that you would want to look at the unclassified portion of that. Decker noted that SC should probably look at its in-house programs before it expands to these other agencies. Wright commented that not looking at activities in Defense Programs (DP) would be artificial, but it would be helpful to talk to them informally, at least. Andy White said that it would be helpful to talk to those in DP and other agencies. Connolly noted that the ASCI machines are the best around and they are everyone's future. Decker replied that that was true and that DOE has tried to maintain contact and coordination with the people there, especially at the workshop level.

Wright said that a prior charge to look at facilities had led to the establishment of a Subcommittee. **Jill Dahlburg**, the Subcommittee chair, spoke on this Subcommittee's outlook. She outlined the charge to the Subcommittee:

- to assess the overall quality of these facilities relative to the best-in-class in the United States and internationally,
- to determine how these facilities relate and contribute to DOE mission needs, and

- to determine how the roles of these facilities might change during the next 3 to 5 years to serve the missions of SC.

The members of the Subcommittee are Jim Corones (Krell Institute), Dahlburg, Paul Messina (California Institute of Technology), Washington, Wolff, and Kulsrud. The questions that they are putting to the facility representatives are:

- What does your facility encompass in terms of resources, such as hardware, software, and service personnel?
- How does your facility characterize its user base?
- What new user communities or types of users is your facility going to try to support in the next 5 years?
- How would you compare your facility to some of the other high-performance computer centers, such as the NSF Partnerships for Advanced Computational Infrastructure (PACI) leading-edge sites, the Ohio Supercomputer Center, and Goddard Space Flight Center?
- What is the most important trend in high-performance computing for the next decade?

Several responses to these questions are being made at this meeting. The Subcommittee will also look at the "frequency" (what support can be moved around) and will ask the users about the facilities. McRae suggested that another question the Subcommittee might address is how users actually get to access the facilities' resources.

Wright introduced **Walter Polansky**, the acting director of Mathematics, Information and Computational Sciences (MICS) to give an overview of MICS' computation and network resources. The mission of MICS is to discover, develop, and deploy the computational and networking tools that enable researchers in the scientific disciplines to analyze, model, simulate, and predict complex physical, chemical, and biological phenomena important to DOE. It has two strategies for moving forward: to perform research that fosters and supports fundamental research in advanced scientific computing (applied mathematics, computer science, and networking) and to develop facilities that operate supercomputers, a high-performance network, and related facilities. Users come from throughout SC.

Facilities funding has been flat or declining since 1994. On the research side, bumps are produced by initiatives starting up. MICS has three major facilities.

The first is the NERSC facility at Lawrence Berkeley National Laboratory (LBNL), which provides capability resources and professional user-friendly services to computational scientists on projects within the missions of DOE. It began in 1974 at Lawrence Livermore National Laboratory (LLNL) as a computing resource for magnetic fusion researchers and was transferred to LBNL in 1996. It moved to Oakland in 2000. It provides an open computing environment for nearly 2400 users on several computing platforms [parallel virtual processing (PVP) and massively parallel processing (MPP)] at the terascale (which technology is refreshed approximately every 3 years). It allocates computing resources competitively, with about 60% going to SC program managers and 40% allocated by a Committee appointed by the NERSC director.

Accomplishments enabled by NERSC include codes for calculating or simulating

- The reduction of turbulence in tokomak plasmas (which can reduce transport) by zonal flow zones by calculating the behavior of 400 million particles with 5000 time steps
- Computational accelerator design by calculating the behavior of 500 million particles and predicting the beam-pipe aperture for the SNS linac
- Kaon decay amplitudes, successfully computing and reproducing the observed effect for the first time

- A parallel climate model that combines state-of-art models for the atmosphere, land, ocean, and sea ice
- Mechanisms of enzyme catalysis that combine density functional theory and molecular mechanics to study catalysis pathways
- Models of molecular processes in the environment

Several NERSC users have been recognized by their peers for excellent science with such awards as the Presidential Early Career Award 2000.

The second set of facilities are the Advanced Computing Research Facilities (ACRFs), which are designed to provide pioneer capability computing for scientific efforts relevant to the Office of Science mission and to provide a testbed to examine critical computer science issues. Ongoing projects of the ACRFs are

- Chiba City, a 512-CPU (central processing unit) Linux cluster at Argonne National Laboratory (ANL),
- Falcon and Colt, Compaq AlphaServer supercomputers at Oak Ridge National Laboratory (ORNL),
- Nirvana, a 2048-processor SGI at LANL,
- the Probe high-performance storage system (HPSS) testbed at LBNL and ORNL, and
- the TERA evaluation at the University of California at San Diego.

The major accomplishments of ACRFs include (1) providing state-of-the-art computational resources for the Grand Challenge Program calculations and (2) evaluating the feasibility of innovative computer architectures (e.g., the IBM, SP, SGI Origin 2000, Paragon, CM-5, Kendall Square machines) to meet SC computational needs. In the future, the ACRFs will explore novel architectures and testbeds, evaluate topical applications, and nurture/expand technical and vendor bases for future hardware purchases

In the 1990s DOE started with the High Performance Computing Research Centers (HPCRCs). In FY92, HPCRCs were established at LANL and ORNL to investigate global climate and materials. In FY93, it initiated the Grand Challenge computational research program. In FY95, it established an HPCRC at ANL for applications testing and computer science. In the mid 1990s, it established ACRFs and upgraded hardware at LANL, focused ORNL and ANL efforts, coupled Grand Challenges to specific ACRFs, and allocated a portion of NERSC for Grand Challenges. The Grand Challenges were completed in FY 2000.

The ACRFs have made many accomplishments, including

- introducing novel approaches for parallel I/O in the portable system ROMIO, which is currently implemented on all ASCI machines and has migrated into the commercial sector and
- developing and enhancing the high-performance storage system that won an R&D 100 Award in 1997 and that is now operational throughout DOE sites.

Another facility is ESnet, a high-speed data-communications network that provides a highly capable and reliable communication infrastructure and leading-edge network services that support SC's and the Department's missions. Specifically, ESnet is a nationwide high-performance research network that supports science in DOE with advanced network services and that is operated in cooperation among DOE, ESnet management, and end sites, providing an extensive structure of domestic and international interconnects and conducting an Advanced Technology and Research Program.

It started in 1986 with a 9600-baud connection at NERSC. In 1991 (before the HPCRCs), ESnet became the official network for Energy Research (now SC). Its backbone was a combination of 9.6K and 56K satellite and terrestrial circuits, DECnet, and TCP/IP with 22 sites connected to a T1. In 2000, ESnet signed a new contract with Qwest Communications to connect 250 sites.

The mission requirements that drive MICS' computer-resource and network investments are those of (1) the base SC research programs that have to deal with distributed terabyte files that are rapidly increasing in size and have growing needs for high-performance computer and network resources; (2) SciDAC, an SC initiative that requires scientific-challenge codes, computing systems and mathematical software, collaboratory software infrastructure, and scientific computing hardware infrastructure ($37 million of its FY01 $60 million will go to MICS); (3) ACRF and NERSC resources; and (4) biotechnology, specifically three programs in BER: Microbial Cell, Computational Biology, and Genomes to Life (all considered to be "fields of opportunity" for MICS).

The strategy that underlines all this hardware infrastructure was SciDAC, which put out a call for the identification of needs. That call identified the need for (1) a flagship facility for general-purpose and production computing (NERSC), (2) topical center(s) with architectures tailored to specific types of simulations, (3) experimental computing facilities, and (4) a high-speed communications network (ESnet). What is emerging is a flexible and robust system that could be adapted for the computing needs of biotechnology.

Why would MICS want to get into biotechnology? Because of the exciting opportunities there. Biological systems directly convert energy into useable forms. Biological systems employ low-energy reactions that parallel metal fatigue and corrosion; insights into those processes might be gained. Biological systems adapt to their environment; they organize, replicate, and repair themselves. The behavior of biological systems has not been described from first principles (as is done for atmospheric diffusion by the Navier-Stokes Equation).

The computational needs in biology and bioinformatics can be summarized as follows:

Problem Component	Computing Speed	Storage
Genome assembly	>10 teraflops sustained to keep up with expected sequencing rates	300 TB of trace files per genome
Protein structure prediction	>100 teraflops per protein set in one microbial genome	Petabytes
Classical molecular dynamics	100 teraflops per DNA-protein interaction	10s of petabytes
First-principles molecular dynamics	1 petaflop per reaction in enzyme-active site	100s of petabytes
Simulations of biological networks	>1 teraflops for simple correlation analyses of small biological networks	1000s of petabytes

The bottom line is that biotechnology holds lots of opportunities in SC, and MICS is expected to be in a position to address those opportunities.

McRae asked Polansky how he saw the investment in hardware vs algorithms. Polansky responded that he thought that the algorithms have to be hit very hard.

Kulsrud asked if the Department was locked into a number of centers. Polansky said that he thought a case could be made for one flagship, two to three topical centers, and some research centers. The Department is not locked into any set number. An RFP should be put on the streets to see what possibilities emerge.

Connolly noted that Sandia and Celera plan to build a 100-Tflop machine and asked if this was part of the MICS effort. Polansky responded, no. IBM is working along these lines, too.

Wolff asked what the strategy was to model biological behavior. Polansky said the strategy is to identify a box in which stimuli and responses are modeled. Lester commented that electronic structure is another such area, and asked how he arrives at what fraction of support goes to different topics. Polansky said that the basic research programs are essentially flat. The base research program turns over projects regularly. The other work is determined by peer review of responses to RFPs.

Meza asked if ESnet is part of the Advanced Science research program. Polansky replied that they are separate but coordinated. Meza asked if anything is being done to increase the number of vendors. Polansky said, yes, the laboratories are working with IBM, Compaq, and other vendors to test their equipment.

Giles asked if there was a mechanism by which users could play a role in future facilities planning. Polansky said that the best example is NERSC upgrades, which are guided by user groups. Those user groups maintain a “green book” that is updated every 3 to 5 years and tells what the next set of requirements and upgrades should be.

Dahlburg asked how the rest of SC could be stimulated to get interested. Polansky responded that BER is heavily involved; OASCR and BER work as partners. Other offices will be invited to be similar partners. High-Energy Physics and Nuclear Physics are likely such partners.

Dahlburg asked what the most important trend in high-performance computing is. Polansky said that we need to get off this price-performance curve. Lester asked if he could please amplify what he was talking about to avoid any ambiguity. Polansky said that if we keep on the current price/performance curve, the hardware advances will consume all the budget for MICS. We cannot make the science case for $300 to 400 million of computer hardware.

Washington asked how they got feedback from users. Polansky said that they use feedback from the program managers at the laboratories that are operating the ACRFs; sometimes they get anecdotes from the users.

Kulsrud asked what portion of the budget is devoted to facilities. Polansky replied that the base program is funded at $166 million, of which $64 million or almost half goes to facilities.

Stechel asked if he was concerned what will happen to the physical sciences as attention shifts to the biological and environmental sciences. Polansky responded that the Department is not abandoning the physical sciences. Much of this investment will migrate back to the physical sciences.

Meza asked if security concerns influenced his operations. Polansky replied, yes. We do not want to be targets of opportunity for hackers. The need is formidable now. We need to have adequate cyber security *plus* an open system.

A break was declared at 10:38 a.m. The meeting was resumed at 11:10 a.m. with the introduction of **Horst Simon**, the Director of NERSC. NERSC’s vision is to be a world leader in accelerating scientific discovery through computation by providing high-performance computing tools and expertise to tackle science’s biggest and most challenging problems and by playing a major role in advancing large-scale computational science and computer science.

NERSC is a supercomputer facility within SC. It is an unclassified, open facility, serving more than 2000 users in all DOE mission-relevant basic science disciplines. It celebrated its 25th anniversary in 1999.

In 1995-96, DOE and NSF competitively reexamined the role of the supercomputing centers because of the rapidly changing technology, better local facilities everywhere, and growth of computational approaches in all disciplines. The new model that emerged was one that encompassed intellectual services plus a major facility. Under this new model, new algorithms and strategies would be developed in medium- and long-term collaborations with the scientific user community, and the center is the working interface between computer science and physical science. Such a model is necessary but not sufficient.

The most important thing is to have a system that balances the newest technology and production quality, balancing high-end parallel processors for capability; large storage systems; and a vector system, personal computer clusters, and networks. The NERSC system architecture typifies the hardware resources of such an integrated system and reflects a long history of new-technology introduction:

9/1996	NERSC-2 first deployment
10/1997	NERSC-2 full production
1998	NERSC-3 procurement
4/2000	NERSC-3 first deployment
2001	NERSC-3 full production

A plot of NERSC computational power vs Moore's Law showed NERSC outperforming the projections of Moore's Law. This result was achieved by getting more cycles out of the machine, which makes the system more productive.

NERSC 3, Phase 2, was delivered in January of 2001. The total system has 158 nodes of 16 CPUs each; 134 are dedicated to parallel computation; 140 nodes have 12 GB of memory; and 18 nodes have 8 GB of memory. With a total of 2528 CPUs and a 1.5 Gflops per second peak, the system operates at 3.792 Tflops. The total memory is 1.824 TB, the total user-accessible shared parallel disk size is 20 TB, and the total local disk size is 11.4 TB (used mostly for system purposes).

What computer scientists are interested in is sustained system performance, which will, of course, be less than peak performance but which will estimate the amount of scientific computation that can really be delivered. System performance is measured with the Nash Parallel Benchmark, which is run every month. As of November 2000, the benchmark put NERSC 3 in second place behind the ASCI White machine at LLNL on the list of the top 500 computing machines. NERSC has a lot of interaction with the ASCI people at LLNL, which is quite advantageous.

Storage capability has continued to be improved, moving from 70 TB to 1.3 PB. At the same time, the usage has changed, with the experimental data being placed on the machine increasing tremendously. This situation has forced an allocation of storage.

All of these capabilities have to be tied with other systems; NERSC must support a large variety of systems and platforms. Over the years, staff skills and backgrounds have changed. In 1996, NERSC moved from intellectual to hardware resources. The number of technical staff was reduced from 79 in 1994 to 59 in 1998. Currently, NERSC employs 64 FTEs (full-time-equivalent employees). The number of degrees held be staff members increased from 42 in 1994 to 53 in 1998; the number of PhDs increased from 11 to 24 during that same period. An organization chart of LBNL's National Energy Research Scientific Computing Division showed how NERSC fitted into that division.

The intellectual services NERSC provides include the development of new algorithms and strategies developed in medium- and long-term collaborations with the scientific user community. These new services require a change in the staff skills and background, provide innovative assistance, make NERSC a working interface between science and computer science, require a new model of scientific computing support, and develop new user communities.

In 1996, NERSC reinvented itself to meet the challenge of transitioning its user base from single-vector-process computing to highly parallel computing. This transition was highly successful. However, mapping offices within SC and their scientific disciplines 1:1 on specific computational technologies (e.g., partial differential equations and Monte Carlo techniques) was not successful. So, NERSC went to the Red Carpet Plan to make Grand Challenge projects successful. That plan achieved success by adjusting system limits as needed (e.g., long 512-way jobs), raising priorities for jobs at critical times, developing software modifications to allow for large calculations, presenting specialized talks at conferences for specific user areas, porting software that otherwise would not have been available (e.g., CERNLIB), and parallelizing public-domain software and optimizing it for the center's platforms (e.g., NetCDF).

The Red Carpet Plan produced the first breakthrough at the 1-Tflop level of computing. That accomplishment won the 1998 Gordon Bell Prize for best performance of a parallel supercomputer application for a team of collaborators from DOE's Grand Challenge on Materials, Methods, Microstructure, and Magnetism.

One way NERSC developed new user communities was the porting of CERNLIB to the T3E. It was the first porting of this library to a highly parallel platform.

A lot of NERSC's material has transitioned to the Web. It has also installed classroom training and set up a group to provide innovative assistance. From LBNL, it got the Oakland Scientific Facility, which is a 20,000-sq-ft computer room with 7000 sq ft of office space. They have a 10-year lease with three 5-year options. The construction costs for the computer room were $10.5 million.

For the past 5 years, the budget at NERSC has been flat or declining. The FY01 budget was allocated 47% to computational systems, 11% to storage systems, 31% to staff in high-performance-computing, and 11% to staff in advanced development. These funds have been leveraged by maintaining a close interaction with all other OASCR-funded projects at LBNL and by attracting $3 million in computational laboratory-directed research and development (LDRD) per year from LBNL. NERSC is well-integrated into the Bay area intellectual community through the University of California (UC) at Berkeley, UC Davis, the Mathematical Science Research Institute (MSRI), and International Computer Science Institute (ICSI). It is also involved in strategic collaborations on a national scale with three SC laboratories: ANL (PC clusters, computational grids, and visualization), ORNL (PROBE and distributed storage), and Pacific Northwest National Laboratory (PNNL) and with ASCI at LLNL (another IBM platform).

The usage of NERSC can be looked at several ways. By number of *users*, the FY00 usage breaks down to the following percentages:

computational science and mathematics	6
environmental sciences	6
life sciences	5
earth and engineering sciences	2
chemistry	9

materials sciences	13
fusion energy	13
high energy physics	4
accelerator physics	6
nuclear physics	11
other	25

By actual *usage* (time), it breaks down to the following percentages:

computational science and mathematics	4
environmental sciences	10
life sciences	7
earth and engineering sciences	1
chemistry	9
materials sciences	15
fusion energy	19
high energy physics	13
accelerator physics	8
nuclear physics	9
other	5

By institution, these two sets of figures break down into

Users:	DOE labs 59%	other labs 2%	industry 3%	universities 36%
Usage:	DOE labs 57%	other labs 4%	industry 4%	universities 35%

NERSC has an allocation process, which was established in 1999. Two new boards help guide NERSC: the NERSC Policy Board (NPB) and the NERSC Program Advisory Committee (PAC). The most important impacts of the new allocation and guidance rule are the recognition of NERSC as a DOE facility and the peer review of facility use, which assures the highest quality of science and balanced use of this unique facility and which counters criticism of NERSC as a “closed shop.”

NERSC’s technical accomplishments include:

- cluster computing and its technology evaluation (for example, Lenny Oliker won the best-paper award at SC99 for his comparison of an MPI, OpenMP, and multi-threaded implementation on the T3E, Origin 2000, and Tera MTA);
- the development of a vision for a DOE Science Grid that will provide a shared and, when possible, uniform view of resources in order to facilitate access to and construction of distributed systems;
- Visapult, which tied together resources in DOE with grid technology (it won the bandwidth challenge with a 140-GB/sec bandwidth); and
- increasing the utilization on the Cray T3E from 50% to >95% by direct collaboration with SCI Cray.

The impact of NERSC on the DOE mission is exemplified by:

- the computational accelerator physics Grand Challenge Project, in which simulations of the Next Linear Collider (NLC) resulted in an improved linac design with a higher acceleration gradient, saving $100 million over the original design;
- the Numerical Tokamak Turbulence Project (NTTP) Tokamak Grand Challenge in which all but a very small percentage of the results were obtained on the T3E at NERSC; and
- bringing the high energy nuclear physics (HNEP) community back into the supercomputing community.

NERSC's strategy for the future is to continue providing high-end systems and comprehensive scientific support. But we see a fundamental change in the way DOE SC is approaching computational science: the formation of science-challenge teams (e.g., SciDAC). We will make all efforts to support these teams. Also, because bandwidth is increasing so quickly (much more quickly than computing power is increasing), building a DOE science Grid and integrating it into NERSC are very important. NERSC will have to become the biggest supercomputing node in this science Grid. We see this, the so-called unified-science environment, as an excellent opportunity to provide mechanisms for integrating all of the DOE SC resources, such as the ACRFs and other facilities, and for building an environment that connects all of this.

In summary, NERSC has established an excellent track record in acquiring, installing, and maturing high-performance-computing technology, has had a major positive impact on computational science during the past several years, and has taken maximum advantage of its intellectual resources to advance the state of the facility and to increase its value to DOE.

Connolly asked how soon, if ever, Linux clusters will be production ready. Simon answered that he would assign a 75% probability to NERSC-4 (FY03) having a cluster of symmetric multiprocessors (SMP) and a 20% probability to a PC cluster. LBNL is looking at this possibility, and NERSC is monitoring that effort. It might be better known in 6 months.

Meza asked what the difference was between Moore's law and the NERSC performance. Simon cited the move to parallelism and NERSC's interaction and collaboration with vendors. Dahlburg asked what that was in real flops. Simon responded that it was almost a factor of 10, going from 1 million resource units to 10 million.

Norman Kreisman asked if he was assuming that Cray will not be able to market the NEC SX series of machines. Simon replied that Cray said they were not focusing on marketing NEC technology to government clients.

McRae asked how NERSC found out who else is doing cluster experiments and what are they doing. Simon answered that this is done through contacts with other laboratories and by participating in workshops. The question NERSC is interested in is whether these clusters can handle 200 users.

Washington asked how he would use an additional $5 million. Simon responded that NERSC has a flat budget and increasing staff costs. People get more expensive, so technology gets cut into. To put NERSC on the Grid and to support the challenge teams takes a lot of human resources. In terms of hardware, when SciDAC is online, there will be a demand for more hardware.

Stechel asked how the staff works with the scientists. Simon responded that NERSC selects the right people who were interested in computational science and the collaborations they can get involved in. Also, the proximity of the area's universities and laboratories helps. Right now he is comfortable with the balance between computational science and, say, physics.

Kulsrud asked if the knowledge gained at NERSC can be carried over to the commercial sector. Simon replied, yes, because of our collaboration with the vendors.

Wolff asked if an increase in bandwidth *requires* NERSC to get on the Grid. Simon said that more bandwidth does not mean that everyone will link their PCs together. It will lead to more centralized computing. The bandwidth allows us to integrate simulation, observation, storage, and theory. If DOE builds a Grid, it would lead to a productive environment that NERSC would need to participate in. A 5-year plan is being developed that incorporates the implementation of Grid technology at NERSC.

Wright said that every time computing power goes up by a factor of 10, you have to rethink the structure of computing. How do you do this? Simon responded that NERSC expects that SciDAC will lead to large community codes, which is a new method or way of doing computational science. Examples are NWChem (Northwest Computational Chemistry Suite) and the community of high-energy physicists.

McRae asked how technology is transferred into industry. Polansky replied that DOE is still working on those issues and needs to know the administration's policy toward technology transfer. Simon responded that NERSC has a CRADA (cooperative research and development agreement) with Intel.

Meza asked how resources are allocated outside DOE. Simon said that applicants for resources have to list a DOE program manager. If they do not have a DOE contact, NERSC staff ask the appropriate program manager if this project should be supported.

Wright asked how the number of users will change over time. Simon noted that the number has been stable for several years. The number of principal investigators (PIs) has decreased, but the number of users has stayed the same. This is a manageable number. The strategic focus will be on the high end, but the legacy users will not be abandoned. A team is looking at how to do this.

Dahlburg asked what the ideal SciDAC team would look like. Simon replied it would have one great physicist, two great mathematicians, and a bunch of people who manage data storage and transfer. About 20 people at PNNL run NWChem.

Dahlburg asked where the validation of this new piece of technology is. Simon responded that NERSC may have to put more into software engineering than it has done in the past. NERSC can give the users tools, but cannot tell them how to use them. It *can* tell them how to get the best out of the tool.

McRae asked Simon to compare the cost-effectiveness of NERSC vs the NSF centers. Simon said that, compared to NSF's leading-edge facility, NERSC is comparable. It has a different philosophy, relying on entirely professional staff. He did not believe that the user gets the same level of service at NSF centers as they do at NERSC. The Department of Defense (DOD) centers have a much larger budget than NERSC does, but not that many more machines. Connolly asked if the $30 million/year budget includes hardware, and Simon replied that it includes everything.

Stechel noted that NERSC has 3% of its users in industry and asked if they were collaborators. Simon replied, yes, through DOE collaboration grants to industry. You have to be a user before you become a collaborator, though.

Meza asked if the California power crisis will have an effect on operations. Simon responded that NERSC has a contract that guarantees its power price to 2004.

A break for lunch was declared at 12:33 p.m. The meeting was called back into session at 2:00 p.m., and Wright turned the floor over to **Juan Meza**, the Chair of the Subcommittee on Biotechnology. He listed the charges to the Subcommittee, which asked advice on the following:

1. Areas in which the Advanced Scientific Computing Research (ASCR) program should target investments to have maximum impact on the underlying science. Possible examples of areas include specialized facilities for biological computation, basic research in underlying mathematical algorithms, or advanced computer science related to data management.
2. How to most effectively couple research supported by the ASCR program with discipline-specific research carried out by biologists.

He then introduced **Daniel Drell** to speak (instead of Ari Patrinos) on the computational needs and challenges in the age of genome-scale biology.

The Office of Biological and Environmental Research (BER) Program has a diverse research portfolio in which the genome interacts with structural biology, global environmental change, molecular nuclear medicine, and bioremediation. For biology to advance, an ongoing partnership between BER and OASCR that addresses scientific challenges at the interface of computing and global change research is vital. This new partnership must meet the needs of tomorrow's biology.

One task requiring such collaboration is the DOE Climate Change Prediction Program. It links climate research to terascale computing.

In the genome game, DOE is not the only player. The private sector is aggressively entering the field, as testified to by the headlines that announced that Compaq, Celera Genomics, and Sandia National Laboratory have announced an R&D alliance to design a 100-Tflops supercomputer specifically for life sciences use.

Tomorrow's biology will have as inputs high-throughput DNA sequencing, structural-biology facilities, and computation. Science has to explore biology in many ways that are different than in the past.

To make the same point again, computing will be a key component of BER's programs in structural biology, genomics, and Genomes to Life (a program for modeling the internal processes of a living organism, the next step in biology). Two solicitations went out last week (for modeling the microbial cell and modeling biological systems); these projects will start before the end of this fiscal year.

Genomes to Life is subtitled "accelerated biological discovery." The recently completed genome sequence is, essentially, a parts list. Now we need to know the functions and interactions of those parts. A flowchart illustrated the pathway from understanding the genome to an understanding of the microbial community.

Genomes to Life will use informatics and computation tools and the acquisition of experimental data to (1) map the machines of life, proteins and protein complexes; (2) map the regulatory networks to find out how cells control proteins and protein complexes; (3) integrate molecular machines into functional pathways and entire cells through the Microbial Cell Project; (4) characterize natural microbial genetic and metabolic diversity; and (5) use microbes to help solve energy/cleanup problems by producing tools to protect people from energy by-products.

It is obvious that Genomes to Life requires many areas of computational biology from assembly and analysis of genome sequences (a lot is presently available) to prediction of protein structure, to enzymatic activation of environmental mutagen, and finally to elucidating the amino acid synthetic pathways of a living cell. Each of these types of modeling requires advances in simulation techniques and computer hardware.

What is needed is a new relationship between biology and mathematics/computer science. The traditional relationship was one of service: What tools can the mathematicians and computer scientists give to biologists to solve their problems? The new paradigm is a partnership: What are the fundamental underlying mathematical and computational principles of living systems? With the genome sequence in front of us, we need to start building things with these parts.

High-performance computers will be needed for each component of cell-level simulations. Genome assembly will require more than 10 teraflops sustained speed to keep up with the expected sequencing rates. Computational biology is characterized by its need for continuous high-performance computing, rather than periodic large-scale simulations. Otherwise, scientists will never understand the data they have already accumulated.

Drell then turned the floor over to **Michael Knotek**. He introduced a document (*Genomes to Life*) as a roadmap for the next generation of biology. There were many boundary conditions imposed on the effort to produce this roadmap:

- how to serve as a core for all DOE biology,
- maintaining complementarity with NIH, and
- getting our arms around what computing environments will look like 10 years from now.

The biotechnology industry is already putting up large computing facilities (100-Tflop machines scalable to a petaflop). To perform the new biology, you must have a computing capability that does not look like anything available today. OASCR and BER can determine what biology will look *like* and *be* in the future. Genomes can now be determined quickly (at the rate of a microbe a day). The idea of this program is that proteins (almost) never work by themselves. They work with other molecular machines, thousands of which work together in a cell. The question is, how does a genome function in a cell? The goal is to explore the functions that microbial communities fulfill, looking at microbes that are used to sequester excess carbon, clean up the environment, produce and use energy, etc. Genomes to Life aims to achieve a fundamental comprehension and systematic understanding of life, employing a comprehensive strategy, high throughput, high data intensity, and genomic expression.

The only way to do this is with massive computer systems. These systems are made up of physical and chemical machines that move around, pull things apart, put them together, and recycle their components. Components are passed mechanically through the system.

Much of the genome's makeup is unspecified in function, as yet. These parts are the "junk" portion; they probably control what happens. Inside a cell are thousands of on-off switches; modeling these will be a huge job in itself. This cis-regulatory function is very important to understand and will be very difficult.

The first goal is to identify and characterize the molecular machines of life, seeking to find out what those machines are, what they do, and how they do it. Experimental data from Genomes to Life and other programs will be used to characterize proteins and their interactions and to measure their activities and functions. Informatics, computation, and theory will be used to analyze, model, and integrate all the data and information produced. These activities should lead to an understanding of protein complexes and machines, proteome and protein complex dynamics, pathways and cell processes, and proteome composition and structure as a function of cellular conditions.

The second goal is to characterize gene regulatory networks. Here, experimental data from Genomes to Life and other programs will be used to probe DNA sequences, transcriptome, and protein DNA and to conduct functional analyses, testing, and validation of regulatory networks. Informatics, computation, and theory will be used to identify regulatory elements, interpret data, model the networks, and design and simulate regulatory networks. These activities should lead to an understanding of comparative DNA sequences, regulatory-network components, regulatory-network architecture, and functionality.

The third goal is to characterize the functional repertoire of complex microbial communities in their natural environments at the molecular level. If you look at a large community, you can treat it as one genome that allows it to function, and each little niche in the world has its own community. Here, experimental data from Genomes to Life and other programs will be used to develop microbial-community sequence data; examine gene activities, protein structures and complexes; and probe community responses, functions, and pathways. Informatics, computation, and theory will be used to assemble, annotate, mine, and analyze the data produced and to model metabolic pathways and community responses to environmental changes. These activities should lead to an understanding of the whole-genome sequencing and genome-wide diversity of uncultured microorganisms; the cellular, biochemical, and ecological functions of organisms; and the metabolic capacities, regulatory networks, stability, and adaptation of communities.

What will it take to accomplish these three goals? Structural studies, mass spectrometry, and next-generation DNA sequencing (like the lab on a chip).

Michael Colvin stepped up to identify the advanced computational and modeling needs in the biological sciences. Biological science does not have just a single thing it wants to model or simulate. Computational analysis and simulation have important roles in the study of each step in the hierarchy of biological function. Starting with the DNA sequence, science needs to make sense of the protein sequence and regulation. Sequence annotation is a distinct computational challenge. It then needs to come up with a structure for that sequence; computers will be used to predict 3-D protein structure. All the analytical data will then need to be integrated in molecular modeling. Experimental data integration will lead to multiprotein machines. And network elucidation, pathway simulations, and organism simulations will lead to an understanding of bacterial communities and multicellular organisms.

The overall vision for biology is to provide a framework for integrating new biological data to create new understanding. Take a hypothetical example: If new data show that a gene forms a complex with a DNA-repair enzyme, the function is added to the genome database. Then, one could identify the regulatory sequence associated with DNA repair. Or predict the fold and assign the function, allowing one to fill in the component in the DNA repair pathway and update the model of microbe behavior and/or predict chemical inhibitors.

How? Predictive molecular simulations of most biochemical processes require new algorithms and computers (e.g., first-principles dynamical simulations of enzyme activity). The state-of-the-art method runs on an ASCI computer today, simulating about 600 atoms for 10^{-12} seconds, which requires 3840 processors for 12 days. The experimental data are available to simulate about 100,000 atoms for 10^{-3} seconds (a millisecond), but improvements of 11 orders of magnitude are needed in the computers and algorithms.

Another way biology could use computers now is in using homology-based protein "threading" to determine the structural characterization of many newly sequenced genes. The best method today matches patterns thread by thread onto templates, looking for the best match. The state of the art today would allow threading 5000 genes per day on a teraflop computer, allowing the structural assignments possible for about 30% of the new genes. What is needed is to develop new algorithms for finding more-distant homologies, high-throughput structural refinement methods, and automated management of the predicted-structure databases.

A third and final way would be the problem of coming up with phylogenetic trees to a common ancestor to find promoter regions etc. But this problem would require new computational algorithms to assemble rigorous phylogenetic trees from full genome data. The state-of-the-art method would allow using NP-hard (nondeterministic polynomial time) methods that are limited to about ten bacteria with about 3000 genes. What is needed are hundreds of organisms, which would need improved heuristic algorithms.

Thus, new developments in the computational sciences are central to the Genomes to Life initiative. Genomes to Life describes a new form of biological science that is heavily dependent on computations, is inherently multidisciplinary, and uses new types of computations. Much of the real science will occur in the computational sector, handling huge amounts of data. What mathematical tools will be needed to handle these time and size scales? Computational linkages include genome assembly, gene finding, comparative sequence analysis, kinetic analysis of regulatory pathways, molecular simulations of protein-DNA interaction, sequence-based homology identification, analysis of expression data, molecular simulations of protein-protein interaction, kinetic models of regulation and metabolism, metabolic-network analysis, comparative analysis of metabolic capability, microecological simulations, and analysis of expression and mass-spectroscopy data.

In Genomes to Life, the first goal is to identify the molecular machines of life. Computational needs will include improving the bioinformatics methods needed to analyze and integrate experimental protein-expression data, adapting and developing databases and analysis tools for integrating experimental data on protein complexes, developing algorithms for the integration of diverse biological databases, providing functional and structural annotations of protein-sequence data, and developing modeling capabilities for simulating the function of multiprotein machines in cell networks and pathways.

All models of microbial behavior will depend on information about the macromolecular machines that mediate function. Computation has roles at each step in deriving these data: raw DNA-sequencing reads, the assembled genome, sequence homologues, phylogenetic relationships, protein structure (this task is currently done on an ad hoc level), and protein function. It would be desirable to model all of these machines and their components down to a quantum-mechanical level.

The second goal is to characterize gene regulatory networks, most of which are extremely complex. The computational needs here are to take a DNA sequence, extract the regulatory elements (including operon and regulon sequences) with sequence-level comparative genomics,. simulate these regulatory networks with both nondynamical models of regulatory capabilities and dynamical models of regulatory kinetics, and predict the behavior of modified or redesigned gene regulatory networks. One example from the recent literature is the model of the *E. coli* phage-λ lysis vs lysogeny decision circuit. This type of task should not be a multiyear effort entailing many months of supercomputer runs.

The third Genomes to Life goal is to characterize the functional repertoire of microbial communities because most microbes exist in complex environments with many other microbes. The computational needs here are to facilitate multiple-organism shotgun-sequence assembly; to improve comparative approaches to microbial sequence annotation and gene finding; to use those approaches to assign functions to genes; to reconstruct pathways from sequenced or partially sequenced genomes; to evaluate the combined metabolic capabilities of heterogeneous microbial populations; and to integrate the regulatory network, pathway, and expression data into integrated models of microbial-community function.

The fourth goal is to develop computational methods and capabilities to develop a predictive understanding of biological systems through/from/with sequencing informatics, sequence annotation, structural annotation, functional annotation, new databases, data integration, microbial ecology, modeling and simulations, and visualization. The development of the computational infrastructure for such computational biology presents many challenges:

- Only primitive visualization methods are currently available.
- Many different types of biological data need to be integrated.
- Data are derived from many different methods and laboratories.
- The data-collection methods are evolving very rapidly.
- The size of biological datasets is exhibiting explosive growth.
- Few standards and many incompatible databases exist.

These are many of the same challenges that DOE experiences in its other programs. Strategies need to be worked out to produce near-term basic informatics and annotation; these strategies will require ~10 teraflop computers.

In addition, major computational challenges still remain in organizing, integrating, and visualizing biological data. It is desirable to integrate these capabilities into atomic and functional models. Such integrated multiscale molecular models can provide information on macromolecular interactions and function. Protein structure prediction resolves the problem down to one of molecular mechanics, which will allow the calculation of classical molecular dynamics, which will allow the development of quantum-mechanical simulations.

The necessary elements required for the advance of computational biology are already core strengths of the DOE: bioinformatics, advanced computing and algorithms, protein-structure prediction and macromolecular simulation, and molecular modeling and computational chemistry.

In conclusion, biology has many needs for advanced computer science and large-scale computing. Genomes to Life provides a framework for creating a new kind of biological research that is integrated with the computational sciences. A key goal at this time is to begin to build an effective partnership between computational and biological scientists. A computational-biology workshop is planned for this summer to develop a detailed roadmap for Genomes to Life computational components. In closing, Colvin noted that, about 50 years ago, J. Von Neumann was asked by the Navy to justify further work for computer research.

Wright inquired about the status of Genomes to Life. Knotek replied that it is in the 2002 budget. McRae noted that the budget is missing four orders of magnitude of the costs for what is being proposed and asked how this effort fits in with what is going on in NSF or NIH. Colvin said that the plan is to use the limited funds available to show what can be done, demonstrating that high-end computing can extend biological research significantly and using that demonstration to leverage additional funding.

McRae asked why they select microbial communities to model, given their complexity. Colvin replied that they are homeostatic; in modeling microbial communities, you average out a lot of functions. Given the level of fidelity or resolution at which the science would be operating, they are not that complex. Knotek commented that a recent workshop addressed the definition of a “minimal cell.” That task was scrapped because a cell is important only in the context of the community in which it lives. Drell pointed out that the vast majority of microbes do not cause disease. We can bring mission interests into play in terms of microbial communities. Knotek noted that this strategy also brings an institutional framework to this effort. The cost will be several times what it is today. The applications in health will be tremendous and will bring in private-sector interest. The effort will probably grow to $1/4 billion per year in 10 years. NIH will have to foot part of this bill, and they will want to.

Kulsrud noted that there is a place here for special-purpose computers to simulate these organisms and asked if commissioning such machines is part of the plan. Knotek replied that the way living systems function is not pretty. Putting together a computer that can model this complexity will require specialized machines. In addition, the methods required to take the high-throughput output and turn it into computerized data will be highly specialized. Colvin pointed out that the algorithms are rapidly evolving. Then the gene changes, and you have to revise the algorithm.

Giles commented that there should be a focus on the computer-science problems that need to be addressed and asked if there are any ethical concerns about the DOE effort. Drell replied, yes, this program will raise some ethical challenges. One that has come up is intellectual property protection for structural information. Program managers can do little about ethical concerns. A broader community must consider and resolve this topic.

Stechel asked if they had looked at partnerships with materials science and molecular simulation. Colvin responded that they had. The RFPs will appeal to materials scientists. They will be the ones who have the expertise. Stechel said that she was thinking of catalysis and coupling data from informatics, not borrowing from the other pieces of science. Colvin commented that, on a laboratory by laboratory basis, the people are already partnering with materials science people. DOE will look at involving the materials science people in the workshop.

Wright noted that a lot of the mathematics that is being talked about is not part of the common portfolio of mathematics.

McRae stated that the computing capability that was being talked about could play a large role in communicating the importance of biology to the general public.

Meza commented that the algorithms do not scale to go from 10^{-13} to 10^{3} seconds as was suggested in the presentation. Colvin responded that he had picked a very challenging example, but it is the type of contribution that will need to be made. That is where a lot of the fundamental chemistry and physics need to be done. Meza asked about the uncertainty that would be involved, and how would it be dealt with. Drell admitted that there is a lot that we do not know about yet, and it is information that is hard to get. But it is vital that it be obtained if we are going to model these systems at any resolution. A lot of work needs to be done here. There have been some sensitivity analyses done. Colvin observed one bit of good news: cells are remarkably stable; they reequilibrate after changes in pH or glucose concentration.

Wright said that statistics must be part of all this. Colvin agreed. Generally, in biology, we are underestimating the statistical requirements. Here, it has not even been prioritized, yet. Wright asked if the Committee should discuss where investments should be made for the greatest return; certainly, a large amount of algorithmic development will be needed here. Knotek responded that (1) no other agency besides DOE can bring together what biology will need during the next 20 years; (2) policymakers should not focus too soon, thus limiting options; and (3) DOE should get the other agencies involved, along with private industry. Colvin said that the mathematics has to be done, then the algorithms, then the data handling. The normal paradigm of putting postdocs on the problem after the data have been gathered will not work. Connolly commented that biocomputing specialists normally say, “We don’t even know how to do the experiments yet. We’ll do the computing later.”

Giles stated that this program is talking about transforming a field and who is a biologist. Colvin responded that it is widely recognized that computational biology is the field of the future.

A break was declared at 3:55 p.m. The meeting was called back into session at 4:25 p.m. Jim Corones, a facilities Subcommittee member, introduced **Rick Stevens** to speak about the computing facilities at ANL. That institution has been supporting computer science, mathematics research, and DOE Advanced Applications since the 1980s, accumulating 20 years of computing-facilities innovation. This history falls into two distinct periods:

1. During the ACRF period (1983-1992), the focus was on exploring parallel architectures, developing programming models and software tools and training a new generation of computer science (CS) researchers. During this period, they experimented with every major form of parallel computer architecture except one (the no-dataflow machine), and ANL’s ACRF served as an international center for parallel computing.
2. During the HPCRC period (1992-1999), ANL focused on production-oriented parallel

computing for Grand Challenges in addition to computer science.

Today it has an IBM SP with 144 nodes (which is ready for retirement); an SGI Origin 2000 with 128 CPUs and 12 infinite reality pipes; a 512-CPU IA-32 Linux cluster; and some special-purpose equipment:

- a 120-TB IBM 3495-based tape system with a 10-TB data front end;
- special-purpose clusters (IA-32 Linux plus NT and Alpha Linux) for visualization, data caches, software development, and CompBio;
- hundreds of PC/workstations and servers;
- high-performance networking and a QoS testbed;
- five Grid nodes and a development environment; and
- a Virtual Reality and Visualization Laboratory.

The available software supports CS primarily and application development secondarily. It includes a comprehensive suite of software development tools (e.g., compilers, libraries, debuggers, program viz, and performance tools), systems software for parallel systems (e.g., schedulers, systems management, file systems, merge/passing libraries, accounting, visualization, and data management), and Grid-development tools.

The staffing is lean: the Mathematics and Computer Science (MCS) Division has a systems-development and support staff of about 13 FTEs. Five are focused on large-scale systems, four on workstation computing environment, two on advanced network engineering and support, and two on advanced visualization and support. With these personnel, MCS facilities support three communities: computer science and math researchers (internal and external), computational-science developers, and semiproduction computational-science users (mostly internal). The active accounts include about 1000 users, of which about 300 are large-scale system users, about 500 are general MCS Division computer users, and about 200 are general collaborators outside these two categories.

The IBM SP usage is divided among the Physics Division (36%), biology (24%), climate (14%), chemistry at ANL (7%), engineering (5%), chemistry (4%), nanotech (2%), computer science (4%), energy research (1%), material science (1%), and other (2%). The SGI Origin usage is divided among ASCI/FLASH (35%), general circulation model (1%), nuclear physics 8%, superconductivity 7%, genomics 2%, actinide chemistry 1%, visualization 29%, and computational fluid dynamics (CFD) 17%.

Computer science in the OASCR Program needs interactivity to edit, compile, run, debug/run, and repeat (in many cases those runs are very short and very wide). It also needs flexible systems software: a specific operating system (OS), kernel, or a specific set of libraries and compilers. Because these requirements frequently conflict with some other user's needs, the ability to reconfigure the hardware is needed along with permission to crash the system and, in some cases, root access. The systems need the ability to test at scale. A nonrequirement is exclusivity; one just wants to know if something will work, so the machines can be shared.

Scalability is an unrecognized crisis. It is hard to achieve. The code must be augmented, and often there is a complexity and lack of scalability testbeds. ANL's solution to the problem of scalability is Chiba City, the Argonne Scalable Cluster with 256 computing nodes, 32 visualization nodes, 8 storage nodes, and a Myrinet interconnect. Its mission is to act as a scalability and open-source-software testbed. Planning of Chiba City started in November 1998, installation occurred in October 1999, development and debugging were carried out from November 1999 to February 2000, early users logged on in March 2000, and in August 2000 it went into production support mode and became available to research partners in computer science and computational science. In June 2001, the system will be opened up to scalable-system software developers. Availability is important. At some point, the testbed aspect will get oversaturated. Currently, it is running at 61.7 % utilization.

Grids will have a critical capability for DOE personnel to get work done. Networking and grids are driven by the dispersion of researchers and the need to connect disparate computational resources. *But* grids are not useful without *bandwidth*. The state of Illinois is funding a project to provide a dark-fiber network, called I-Wire, that will run many fibers to each participating institution and provide the ability to use any color. It will be highly integrated with the CS capabilities at ANL. DOE could link these testbeds together with dedicated land-line dark fibers.

The new user communities at ANL include biology and nanoscience (which constitute the strategic focus for growth) along with climate, high-energy physics/nuclear physics (HEP/NP), nuclear engineering, and energy systems. “Grids” will emerge as the mechanism for application delivery, so it is important that DOE computing facilities become Grid-enabled.

Biological computer-aided design (CAD) offers three challenges: (1) understanding biological systems from an information-systems standpoint; (2) modeling biological systems (genes, molecules, pathways, organelles, cells, tissues, organs and organisms, and communities); and (3) designing new biological structures and systems.

Comparing the ANL facility with other leading high-performance computing (HPC) centers, ANL’s focus has been primarily on enabling advanced CS research and at the same time supporting a limited set of applications (a dual-use center). It has less flops and bytes than NERSC, National Center for Supercomputing Applications (NCSA), or the San Diego Supercomputing Center (SDSC) but does more software development than most HPC centers. It has substantially fewer support staff than most HPC centers. It does not see itself competing with most HPC centers but rather complementing them and collaborating on enabling technologies. ANL believes it can take on more risky technologies than HPC centers that are primarily focused on production computing for applications. ANL believes that its facility’s quality is world-class and that it enables world-class CS.

The important trends that are perceived include the complete dominance of commodity technologies and architectures; how this trend is leveraged is the key HPC architecture issue. Other trends include the increasing availability of bandwidth and storage capacities, the continuing struggle with scalability in software, and the incredible importance of Grid-oriented infrastructure.

Corones then introduced **Thomas Zacharia** to speak about advanced scientific computing at ORNL. The ORNL Center for Computational Sciences (CCS) is a modest center with a 1.5-Tflop IBM and a 0.5-Tflop Compaq system. It encourages and supports long-running jobs. Such a lean center allows flexibility. It has never had more than 100 active users, and never has more than one-third of those using the machines at once. It works with other laboratories and is a production archive for the Atmospheric Radiation Measurement (ARM) Program. It constantly deals with increasingly larger applications and larger amounts of data, and its performance ratings are always high.

When the center was established 10 years ago, the staff found that developing applications made it imperative to use students. The center now has 20 postdocs who will be mentored jointly by professors and Laboratory researchers. The State of Tennessee has provided $12 million for the Joint Institute for Computational Science, which is being used to build a 40,000-sq-ft computer room.

The CCS was established in 1992 with an Intel Paragon XPS150, fastest computer in the world at that time, hosting application pilots in materials, HENP, groundwater, climate, chemistry, and the genome. It operated with multidisciplinary teams of scientific, CS, data, networking, and visualization experts. A strong component of its activities is the evaluation of new architectures; another is new programming paradigms for using MPP.

The CCS developed the first application to sustain 1 Tflop, winning the 1998 Gordon Bell Award. A longstanding climate-simulation milestone was first met on the CCS Compaq system in October 2000. It created GIST (the Genomic Integrated Supercomputing Toolkit), which aided in the assembly of human genome chromosomes 5, 16, and 19 and in their initial annotation. Today, CCS is one of the nation's leading unclassified computer facilities; it is unique in its partnership with OASCR, BER, and the Office of Energy Efficiency and Renewable Energy (EERE) and in the significant investments from IBM and Compaq.

The CCS has the first Tflops-peaking system in an SC laboratory and the largest unclassified computer in SC. It has been in production since April 2000 and complements NERSC. It is currently supported by SC-BER for biology, climate, and other projects. It is critical to the Accelerated Climate Prediction Initiative (ACPI) pilot project, Computational Biology, and Genomes to Life. It supports the astrophysics pilot project and offers key capability for SciDAC and other SC applications

The Compaq Alphaserver (a 0.5-Tflop machine) gives better performance for applications because of its memory, architecture, and bandwidth. An upgrade to a 1-Tflop system is planned. This machine came in for an early evaluation of its application performance and scalability, file-system technology, and dual-rail Quadrics switch.

The CCS has integrated teams in climate, materials, biology, and CSET (Computer Science and Enabling Technology) and a decade of leadership in early systems evaluation. More than 30 systems evaluations have been performed in the past 10 years; most recently, for the Compaq, IBM, Intel, and SRC machines. Evaluation results have affected the center's procurement decisions and coding strategies. CSS has done a performance comparison of the Compaq and IBM machines, showing a significant performance edge for Compaq on specific applications.

The center has a high-performance storage system, moving up to 1 TB of data stored per day, and it is funding a joint project with BER's ARM Program.

Historically, the Paragon has largely been used for materials science, high-energy physics, chemistry. The current usage of the IBM and Compaq machines is for climate biology, materials science, and fusion. Because it pushes large uses, the center is hitting 85% utilization. The SciDAC projects will stretch its capabilities.

The CCS systems are managed to meet user needs. The entire center is often dedicated to one application to solve large, time-critical problems. The user services group is integrated into the applications group. Some examples of what the Center has accomplished include the development of the parallel KKR-CPA alloy electronic structure code (originally an LDRD project), which won the 1990 Gordon Bell Award. It participates in a partnership in Computational Materials Science with Ames and BNL. And its Materials, Methods, Microstructure, and Magnetism (M^4) program also won the Gordon Bell Award in 1998.

Climate science that is enabled by the CCS includes ensemble simulations of the DOE Parallel Climate Model. Computational biology enabled by the CCS includes the assembling of chromosomes 5, 16, and 19 of the human genome, the annotation of all human genes and those for more than 50 microbials, and the recognition with PROSPECT of the most protein folds in the CASP4 competition.

In the future, the top 500 machines will continue to increase performance past the 1 million

Gflops level. For the time-critical problems, the demand for computational capacity is increasing, and the cost for computers is flattening out. There is a need to look at architectures that will lower the costs and increase capabilities. What likely will be looked at is terascale production systems; early evaluation of next-generation systems; and a focus on applications, specifically for climate, materials, and biology; and tuning machines and applications to work together. CCS is already working with IBM, Compaq, the national laboratories, and universities on next-generation architectures, collaborating in application areas like the life sciences, materials, and climate. A CRADA is in the final stages of development; it will provide for five joint postdocs, staff exchanges, and funding from BER and IBM.

In the future, high-performance knowledge extraction needs to be developed. The CCS will work in concert with NERSC and ANL with overlap and interaction in common software libraries and submission methods, easy file transfers, and shared training evaluation results.

The next jump in large-scale systems will be to 10,000 to 80,000 processors, but these machines will be affordable. This advancement is something that a DOE center and OASCR should do to meet the needs of SciDAC and other similar groups. Community codes need to be encouraged; CCS will host and maintain those codes.

In June, construction will start on two new buildings that will be completed next year. This project will have a 40,000-sq-ft computer room with 4 MW of power and piping for 1100 tons of cooling. It will provide office space for 250 staff members and classroom and training areas for center users and visitors.

Wright called for any public comment. There being none, she asked if there was any discussion of the ANL and ORNL centers. Washington asked how much interaction occurred among NERSC, ANL, and ORNL. Stevens replied that each laboratory seeks to find niches that it can excel at. Zacharia said that the most effective collaborations are those that occur when groups can work together on projects toward common goals. McRae asked what the common goal is. Zacharia said that it was to do science effectively, so one person can concentrate on one aspect, and another can concentrate on another aspect. Stevens said that the goal is to accelerate the advance of computational science and that there is a lot of crosstalk among the laboratories. McCurdy noted that this collaboration goes back more than a decade.

Wright observed that machines seem to achieve greater peaks, but the amount of the peak that can be used seems to be monotonically dropping, and asked what can be done about that. Stevens said that the reasons we have the problem are that the benchmarks are driving the marketplace. But those benchmarks do not reflect the needs of scientific computation. Science is not a big enough client to change that. We can correct some of it by better managing cache management, but the fraction of peak is not the metric you want to use.

Wright asked what SC should be doing. Stevens suggested that it (1) get together with vendors to have joint development efforts for better sustained performance and (2) get better benchmarks so architects can make better design decisions. Meza commented that it is not an unrecognized crisis. Stevens responded that the community has not been working on it long. There is no scalability testbed. To solve these real problems, you need to make investments. Other agencies are not doing it, either. You have to solve the scalability problem.

Wright observed that, when scalability was discussed in ASCI, the hard problems were seen to be nonlinear systems that do not scale. Stevens stated that that problem is addressed in SciDAC, but there is a scalability problem in operating systems, debuggers, etc. which will turn into a crisis if it is not addressed. DOE is the right place to get leadership.

Wolff asked about the ESnet and I-wire model? Stevens responded that DOE should investigate nontraditional options for gaining access to high-bandwidth networking, such as owning its own fiber, or owning its own wavelengths on the fiber, and then should solve some of the network-interface problems for the highly parallel machines.

McRae asked how many computer scientists care about scientific computing. Stevens replied that his center had no problem finding people interested in working on Chiba City. Young people can easily be recruited. It is not a barrier. Zacharia concurred. ORNL hired 30 people in his division last year. There should not be any difficulty in getting people, but the field is limited by budget.

Kulsrud asked how much of the $150 million budget for computer science at ORNL is DOE funds. Zacharia replied, 50 to 60%; a lot of work comes in from DOD and the Defense Advanced Research Projects Agency (DARPA). Wright asked how much they got from OASCR. Zacharia replied that this year they got $5.5 million from OASCR, a significant increase from last year for ORNL. Stevens replied that his center has gotten about $4 million a year for the past 5 years; this year it is lower.

Jim Corones asked if there was any understanding of the horizontal distribution of independent users across all the computing-center sites. Oliver said that there is a major intersection of users among all the computing centers. Zacharia said that, in the materials science community, that is not unusual to occur, that each user group has about 70 members, and that the individuals use each different facility for a different purpose.

The meeting was adjourned at 6:00 p.m.

Thursday, May 3

The meeting was called to order by Margaret Wright a 9:00 a.m., and she introduced **James Leighton** to give an overview of the ESnet Project. ESnet's mission is to provide an interoperable, highly capable, and reliable communications infrastructure and leading-edge network services that support DOE's missions. Its vision is to provide seamless and ubiquitous access via shared collaborative information and computational environments to the facilities, data, and colleagues needed to accomplish their goals. Its role is to be a component of the SC infrastructure, critical to the success of SC's research programs

ESnet is a nationwide high-performance "agency mission" network chartered to provide advanced network services to support scientific research in DOE. It is centrally funded by SC. It was rated both "outstanding" and "extremely cost-effective" at the last formal Program Review. It has an extensive structure of domestic (commercial and R&D) and international interconnects supporting DOE collaborative efforts along with a growing Advanced Technology and Research program. ESnet is really a network of networks; thus there has to be a lot of cooperation among network operators. It is regularly reviewed by DOE and the laboratory directors.

ESnet does not serve individual users; rather it serves institutions. About 10,000 to 100,000 researchers in the United States use ESnet, mostly for SC programs, such as HEP, NP, Fusion Energy Sciences (FES), Basic Energy Sciences (BES), BER, and MICS. Traffic is also carried for DP and others. It connects to essentially all the national laboratories, to hundreds of universities (through the Internet 2), and to hundreds of foreign institutions in cooperation with large and small collaborations. Many of the users are casual users unaware that they are actually using ESnet, performing many data-intensive and computationally intensive tasks.

ESnet is an enabler for DOE science, dramatically changing the way science is done. There has been an increase in scientific productivity with much shorter turnaround times for disseminating, assimilating, and testing new ideas. Innumerable meetings are held over the network via remote conference technologies, and the DOE community has become a leader in the use of the network for science. Network planning and deployment based on program requirements and technological opportunities.

The network is used for basic services (e-mail, file transfer, remote login, and distributed file systems); for teleconferencing [an integral part of the work flow (this is growing rapidly, particularly video conferencing)], planning, and coordination; for remote access to unique facilities; for collaboratories; and for distributed computing and particularly Grid computing.

One line of ESnet's ancestry can be traced back to the 1974 dial-up access to the Controlled Thermonuclear reactor Computer Center (CTRCC). Then MFEnet, HEPnet (i.e., DECnet), and ARPAnet were developed. A memo was signed in October 1986 to create the multiprogram ESnet. The February 1990 log shows a security problem with a hacker. In March 1990, all T1 trunks and routers for ESnet (1) were brought online. In February 1992, the ESnet (2) RFP was released (producing 2.5 years of protest). In June 1994, three T3s were turned on. In August 1994, a master contract was signed with Sprint for a fast-packet service. Between 1995 and 1998, the T3, OC3, and OC12 ATM connections brought online. In November 1995, the decision to move ESnet and NERSC to LBNL was announced. In June 1999, the ESnet (3) RFP was released, and a contract with Qwest was signed in December 1999. A major transition is currently under way.

ESnet has a somewhat complicated committee structure, but it makes everyone feel that they are part of the program. The FY01 budget is $16.77 million, and additional funds (roughly $2.5 million per year, mostly pass-through funds) are collected for special project or program support.

The staffing is broken down into:

	Career	**Contract**
Engineering Services Group	6.0	
Information and Services Group	4.0	
Technical Services Group	10.0	1.0
Operations		2.0
Administrative Support	1.0	
Management	2.0	
Total	23.0	3.0

Traffic on the network is increasing by a factor of 2 each year. Communications capabilities are doubling (and costs are halving) each 12 months. The umbrella services contract includes three major components: advanced services and technology for production network, high-performance facilities, and research collaboration with Qwest. It is a multiyear contract (3+2+2 years) for more than $50 million. It will overlap with the existing Sprint contract, with a nearly 2-year overlap possible but not expected. The contract was signed in December 1999, the first sites went online in August 2000, and approximately 90% of the initial transition is now complete.

ESnet's collaboration services (i.e., video conferencing) began in 1990-1991 with the HEP use of "excess" bandwidth on T1 trunks. It has now grown to the point that a typical month includes more than 1200 conferences lasting an average 2 hours apiece and roughly 100 conference-days of collaboration per month. A web-based reservation system (DCS) has been developed.

Although the system is unclassified, security has to be considered. That security responsibility falls under the auspices of the LBNL Computer Protection Program Manager (CPPM), and ESnet has its own Project Security Officer and a defined AUP (author user policy). ESnet's security responsibilities cover ESnet resources only and end at the site demarcation.

In the early days, ESnet developed its own research and advanced technology, such as multiprotocol routers (supporting IPv4, DECnet, OSI, and X.25 all at once at one point). It was an industry catalyst for ATM (asynchronous transfer mode) carrier deployment (a technology used for high-speed communication). It also worked on and deployed several advanced protocols, such as CIDR, BIG-4, MBGP, and Ipv6, and it is currently proposing to do some testbed work and to do work on the PKI/Directory.

Future initiatives that will act as drivers (pushing us with some very demanding requirements) include the Large Hadron Collider (LHC), SciDAC, and the SNS. Critical issues for the future include keeping ahead of the curve on domestic connectivity, monitoring the quality of connections to universities, continuing to improve international links, and deploying and supporting advanced services.

In summary, the ESnet project thrives in an environment that is experiencing phenomenal growth in usage, has an extremely rapid pace and broad spectrum of technology, has simultaneous user demand for performance and reliability, and is extraordinarily dependent upon wide-scale collaboration and interaction with peers. The program is highly service-oriented and enjoys the trust, confidence, and support of both its technical and end-user communities. It is a highly successful, effective, and cooperative effort in meeting the networking and data communications requirements of the agency's science research community for 15 years now.

Wright asked how the growth in ESnet's traffic compares with nationwide growth. Leighton said that commercial networks have experienced a doubling every 2 to 3 months, which, when adjusted for how you measure it, is about the same as what ESnet sees. It is not growing as quickly as the other networks. All of the major vendors have more work than they can handle, leading to delays in connection and deployment.

Kulsrud asked if there were any measures of user satisfaction. Leighton said that they have several levels of users. If you go to the sites (the local-area-network folks), ESnet's working relationships with them are very good, and their feedback is very positive. The major PIs are generally pleased with ESnet's services. Very little criticism is received, which is amazing when you consider that something is always breaking in the massive, complicated network.

Meza asked if security is a critical problem. Leighton replied that it was not a problem in terms of ESnet's customer set, but it is a critical issue internally. It is under control, but is not a done deal.

Connolly asked if ESnet pays the total cost of the international connections. Leighton responded that, actually, ESnet pays a very small portion of the cost for international traffic. Connolly asked what the OC/2 connection to CERN costs. Leighton said that he did not know; it is paid for by the HEP community. Because of deregulation in Europe, communication costs have plummeted to 20 or even 10% of what they were 10 years ago. Some links to Japan are paid for by others and managed by ESnet.

Wolff asked him to discuss the alternative to buying ATM service from Quest. Leighton said that ESnet looked at that; the budget would be $40 to 50 million per year.ESnet's costs are significantly less than that. In addition, because ESnet ran the system for everyone, the sites (e.g., national laboratories) have been able to reduce their wide-area-network staffs significantly. Also, the ability to connect with just one service avoids the redundant costs of having several

Internet service providers (ISPs).

Wright asked what would ESnet do if its budget was not flat. Leighton said that they would prepare to deal with the new requirements that users are going to place on us in the future; they would put some investment in our testbed capabilities. Washington observed that the centralization of data storage will actually reduce traffic on the network. Leighton agreed; there are different ways that people are doing science, and these different paradigms have different impacts on the network. The paradigm that Washington described still requires people to go to the data.

Wolff asked what their relationship with I-Wire was. Leighton said that ESnet has a committee that is emphasizing research activities within the community. That committee has suggested a joint effort between Esnet and I-Wire, and that possibility is being studied. Keep in mind that ESnet is a production, not research, network; it needs to operate reliably. It might make sense to bring I-Wire up on one of our testbed networks.

Wright called for an open discussion of facilities. Connolly started by saying that facilities facilitate. People outside use them. To see if a facility is working, you ask the users. He asked if these facilities poll their users and report the results, and, if so, would they share those reports. Simon responded that NERSC sends out questionnaires from which they get a 10% response. On a scale of one to seven, NERSC gets almost all sixes and sevens. Everyone says they need more access to resources. The center also has internal goals and metrics (which are in the report that was distributed at the meeting). The third level is the user groups. They often address special questions. Then there is always informal feedback. Zacharia said that ORNL's feedback occurs day-to-day from the research teams using the facility. The center then responds to the needs expressed. The ORNL CCS does not survey its users because that set of users is so small; instead, the center's staff directly interacts with the users. Stevens said that his facility is much the same. He gets monthly reports and tries to deal immediately with issues that are raised in those reports. It is a technical, interactive dialogue rather than a survey of users.

Washington asked if there was a clear articulation of the balance between the topical and the general centers. Oliver responded that the expenditures have been about 50% of the MICS budget for the past 5 years. The grand challenges plus other influences cause that ratio to vary with time.

McRae asked what the overall quality of these facilities is, and how one would justify a claim of "best in class." Stevens said that each facility has an external review annually. Each year, there is a comparison with other similar facilities, and the results are documented in the report of the review. The second way to justify such a claim is by the quality of the peer collaborations that are formed to use the facilities. The third way is to assess the accomplishments that come out of the laboratories. Zacharia said that the DOE outside review asked the same question. The centers exist to push scientific advance and accomplishments. In each of the areas in which they operate, the clients have gotten major recognition for the advances they were able to make in collaboration with the center. Simon said that NERSC is doing many of the same things. In addition, in its user survey, users have an opportunity to comment on the services received, and 80 to 90% of the comments state that NERSC's service is better than that of other centers. The center also pushes the science output of our users, as measured by the awards they have won.

Wright stated that users will use any machine on which they can get cycles, so the argument that users do good science seems hollow. Stevens asked, whom do they talk to about their code problems, etc.? When you determine that answer, you can probe who has the best relationships with their users. There are subtle differences in how the supercomputing centers are operated. Who has the lion's share of the leaders (as users) might be a better measurement. Simon said that the premise is wrong because people actually rely on software that was developed at and is maintained by a specific center. They cannot go elsewhere and get the same service. Zacharia noted that an ACRS is not just a machine but an agglomeration of personnel, skills, and expertise that push key strategic decisions related to the DOE mission.

McRae asked who the competition is. Simon said that DP is a major competitor as well as NSF. Zacharia was of the opinion that the centers are not competing with anyone because they are focusing on DOE mission-critical needs (biology, climate, and materials).

Connolly asked if the center leaders saw a change in direction or vision during the past 10 years. Stevens suggested a reduction in vision. The ASCI strategy is to pump down dollars to vendors who are not going to take any risk. The program could take the risk of architecture development. ASCI is not set up to do that, nor is DARPA. McCurdy said that this program should have a hardware development program to take those risks.

Stechel asked what people would lose if these facilities went away. Simon noted that three-quarters of NERSC users have no other supercomputer center to turn to if NERSC was not available. Stevens called attention to the fact that many software packages developed at ANL are used by users and incorporated into vendors' packages. Computer science software development advances would be lost if the centers went away, and vendors would have less software. The OASCR program is the crown jewel in this nation's software development. If it went away, the whole nation would suffer because a key source of software development would be gone. Zacharia noted that the science work that has been done at the centers has translated into tremendous advantages in technology that everyone takes advantage of. McCurdy said that, if you took away half of the MICS budget, you take away the scientists' MICS supports. The quality of the facilities makes a tremendous difference in the science produced.

Lester offered the opinion that, when users go to a variety of centers for services, it dilutes the services that DOE provides. He said that he gets one-quarter of the resources that he requests from the centers and goes to a range of other sources to get the rest of what he needs.

McRae said that all he hears is gloom and doom and pleaded for the center directors to make a case that could be takes to Congress to double the centers' budgets. Stevens objected that the presenters had been asked to review the past, not outline the opportunities. McCurdy stated that one could make a compelling case going through discipline after discipline. Stevens suggested that a positive discussion could be held in terms of bandwidth, cycles, storage, etc. Wright offered that, at the ASCAC meeting in October, presentations could be made on the facilities' opportunities, needs, and strategies. Stevens said that it would be good to describe the vision for the next 5 years.

A break was declared at 10:50 a.m. The meeting was called back into session at 11:15 a.m. to hear **Stephen Eckstrand** speak on the SciDAC Program. SciDAC is an integrated program to create (1) a new generation of scientific simulation codes that take full advantage of the extraordinary computing capabilities of terascale computers, (2) the mathematical and computing systems software to enable the scientific simulation codes to effectively and efficiently use terascale computers (developed by research team that are large and dispersed), and (3) a collaboratory software environment to enable geographically separated scientists to effectively work together as a team and to facilitate remote access to both facilities and data.

The infrastructure entails hardware infrastructure and software infrastructure, calling for operating system enhancements and other resources to provide data analysis and visualization, programming environments, scientific data management, and problem-solving environments.

The FY01 budget for SciDAC is $57,304,000. The SciDAC schedule has been accelerated to get funding into the budget in July and to pay for it out of the FY01 buget. Award recommendations were made May 15, 2001. More than 150 proposals were received, and the overall quality of the proposals was very good to excellent, but only about one-fourth to one-third of the proposals can be funded.

ASCR's plans for FY01 include funding the Enabling Technology Centers ($8 million each for Applied Mathematics and Computer Science), the Middleware and Network Research and Applications ($10 million, which includes collaboratories and Grid pilot projects), the Scientific Application Partnerships ($3 million), the Computational Science Graduate Fellowships ($1 million), the Advanced Computing Research Facilities ($2 million), and an upgrade of ESnet services ($1 million).

In the National Collaboratories and High-Performance Networks Program, 82 preproposals were received, and 48 proposals were received: 12 in high-performance networks, 11 in networking/middleware, 12 in middleware, and 13 for collaboratory pilots. For the Integrated Software Infrastructure Centers, 36 preproposals were received, and 16 proposals were received: 9 in computer science and 7 in mathematics.

The program for the development of scientific simulation methods and codes for terascale computers was funded at $1,931,000 to understand and predict the energetics and dynamics of chemical reactions and the interaction between chemistry and fluid dynamics. (This work is being carried out in BES.) The program in computational chemistry (also being carried out in BES) received 40 preproposals and 30 proposals.

BER climate-modeling plans for FY01 include $8 million for the Development of Scientific Simulation Methods and Codes for Terascale Computers Program. This effort will develop state-of-the-science coupled GCM-based climate models to simulate and predict the Earth's climate at both regional and global scales for periods covering decades to centuries, including the calculation of levels of certainty and uncertainty. It will also develop flexible, efficient, and extensible software frameworks to keep climate models at the cutting edge of scientific understanding and computational technology. BER's Climate Change Prediction Program received ~50 preproposals and 41 proposals.

In FY2001, Fusion Energy Science allocated $3 million to the Scientific Simulation Methods and Codes for Terascale Computers Program, funding work in three areas: (1) predicting microscopic turbulence and macroscopic stability in magnetically confined plasmas, including their effect on core and edge confinement; (2) predicting the electromagnetic fields, beam dynamics, and other physical processes in heavy-ion accelerators for inertial fusion; and (3) understanding basic plasma science processes, such as electromagnetic wave-particle interactions and magnetic reconnection. Fusion Energy Science received 13 proposals.

In FY01, High Energy and Nuclear Physics allocated $7 million to the Development of Scientific Simulation Methods and Codes for Terascale Computers Program, funding work in four areas: (1) simulating beam dynamics and electromagnetic fields in particle accelerators to predict and optimize the behavior of accelerator components; (2) developing hardware and software infrastructure for large-scale simulations of quantum chromodynamics (QCD), the fundamental theory governing strong interactions; (3) developing comprehensive models of supernovae explosions; and (4) implementing (with ASCR) collaboratory pilot projects for large HENP experiments. The HENP Program received 6 proposals; two of those proposals would implement parts of the vision for Grid computing.

Giles asked if the overarching theme of SciDAC survived. Eckstrand replied that there are lots of good pieces. The projects are large projects, averaging four institutions. There will be some smaller projects.

Washington asked how they arrived at the budget numbers. Eckstrand responded that the program was started to support climate change and combustion research. When that concept failed, the program offices got together and contributed funds for high-end computing. Between offices, it depended on individual circumstances how budget numbers were agreed upon. The bulk of the money had to go into the computing infrastructure.

Meza asked how the SciDAC funds from Genomes to Life and Climate Change are related. Oliver answered that they are both from BER but are separate monies.

Connolly asked how much money will go back to national laboratories and how much will go to universities. Eckstrand answered 55/45, probably; but we will have to see how the reviews turn out. Getting mixed proposals makes it easier to review; we got very few laboratory-only proposals. Connolly asked what the budget looks like for next year. Eckstrand said it looks flat. The projects funded this year are for three years. We should see a recompetition then.

Wright asked what the new money went toward. Eckstrand replied that the most common thing was hiring postdocs and then hiring grad students. A couple of projects added a new faculty member. Washington stated that this program will form a stronger bridge between the universities and the laboratories.

Giles asked if there is an impact on the facilities under discussion. Eckstrand said that all the applicants had to state the CPU and storage usage they expected, but other than that, he did not see much else. Only QCD researchers stated that they needed specialized computers with little storage. Polansky commented that the SciDAC proposals called for heavy collaboration and the introduction of new Pis. Fred Johnson observed that the mathematics proposals were strong on meshing and libraries. There were no major surprises, but there were some very strong proposals.

Wright asked what need for new mathematics might result from Genomes to Life. Johnson said that it would need new contributions from mathematics and computer science. In mathematics, DOE would want to have a separate announcement and do something different.

Wright noted that everyone agreed by e-mail that the week of Oct. 22-26 would be good scheduling for the next meeting. She suggested meeting on the Thursday and Friday of that week (Oct. 25 and 26). She noted that the Facilities Subcommittee will meet some time and do some site visits. In addition, the Committee will be forming another Subcommittee to respond to Jim Decker’s new charge letter, so that group will be convened and will be producing a report.

She called for public comment. There being none, she asked the Committee if there were any other questions or comments about facilities. Washington asked if DOE’s other computational facilities outside OASCR should be included in the committee’s considerations. Oliver said that the current charge does not include them, but the new charge the Committee is about to receive certainly will include them.

Wright thanked everyone for their participation and adjourned the meeting at 11:58 a.m.

Respectfully submitted by Frederick M. O’Hara, Jr., Recording Secretary, 5/23/01
Revised by Margaret H. Wright, ASCAC Chair, 6/1/01