Minutes for the Advanced Scientific Computing Advisory Committee Meeting
October 17-18, 2002, Hilton Washington Embassy Row Hotel, Washington, D.C.

ASCAC members present:

John W. D. Connolly, Vice Chair	Gregory J. McRae
Jill P. Dahlburg	Juan C. Meza
Roscoe C. Giles	Ellen B. Stechel
Helene E. Kulsrud	Stephen Wolff
William A. Lester, Jr.	Margaret H. Wright, Chair

ASCAC members absent:

Karen R. Sollins	Warren Washington

Also participating:
Steven Ashby, Director, Center for Applied Scientific Computing, Lawrence Livermore National Laboratory
Melea Baker, Office of Advanced Scientific Computing Research, USDOE
Melvyn Ciment, independent consultant
James Corones, President, Krell Institute
Dona Crawford, Associate Director for Computation, Lawrence Livermore National Laboratory
Jack Dongarra, Director, Innovative Computing Laboratory, The University of Tennessee
Edward Dunbar, Oak Ridge Institute for Science and Education
James Hack, Head, Climate Modeling Section, National Center for Atmospheric Research
Daniel Hitchcock, Senior Technical Advisor, Office of Advanced Scientific Computing Research, USDOE
John Houghton, Program Manager, Office of Biological and Environmental Research, USDOE
Fred Johnson, Program Manager, Computer Science, MICS, ASCR, USDOE
Gary Johnson, Program Manager, ACRT, MICS, ASCR, USDOE
Kathleen Kingscott, Director of Public Affairs, IBM, Washington, D.C.
Alan Laub, Director, SciDAC, ASCR, USDOE
C. William McCurdy, Associate Laboratory Director, Lawrence Berkeley National Laboratory
Paul Messina, Director, Center for Advanced Computing Research, California Institute of Technology (Thursday only)
William Miner, Program Manager, NERSC, MICS, ASCR, USDOE
Frederick O’Hara, ASCAC Recording Secretary
Raymond Orbach, Director, Office of Science, USDOE
C. Edward Oliver, Associate Director, Office of Advanced Scientific Computing Research, USDOE
Walter Polansky, Acting Director, Mathematical, Information, and Computational Sciences Division, Office of Science, USDOE
Jennifer Gregory, Oak Ridge Institute for Science and Education
Charles Romine, Program Manager, Applied Mathematics, MICS, ASCR, USDOE
Rick Stevens, Director, Mathematics and Computer Science Division, Argonne National Laboratory
Linda Twenty, Program Analyst, Office of Advanced Scientific Computing Research, USDOE
Mary Ann Scott, Program Manager, Collaboratory Research, MICS, ASCR, USDOE

John van Rosendale, Program Manager, Computer Science, MICS, ASCR, USDOE
Andrew White, Director, Delphi Project, Los Alamos National Laboratory
Thomas Zacharia, Director, Center for Computational Sciences, Oak Ridge National Laboratory

About 55 others were in attendance in the course of the two-day meeting.

Thursday, October 17, 2002

Chairwoman **Margaret Wright** called the meeting to order at 8:30 a.m. and reviewed what had happened since the Committee had last met. The Earth Simulator had been announced just prior to the most recent meeting, in May 2002. Two significant results have taken place in the intervening months, both in response to requests from Raymond Orbach, Director of the Office of Science (SC). First, the Facilities Subcommittee organized a workshop to discuss the Earth Simulator; James Corones will report on that workshop later in this meeting. Second, the committee was asked to prepare a short (two-page) statement about its views on a possible initiative in advanced scientific computing. This statement, sent to Orbach on May 29, 2002, stated that the Advisory Committee strongly supports a science-based initiative with a sustained commitment from the Department of Energy. Further, the Advisory Committee believes that any such initiative should be grounded in strong collaborations among scientists, computer scientists, computational scientists, etc. The Advisory Committee noted that it would be a mistake to focus solely on hardware in any initiative. No single system architecture will be optimal for every problem; the effort will need sufficient architectural diversity. Every element of the program should be subject to peer review, and cooperation will be an important characteristic.

Wright noted that this meeting will feature presentations by two thoughtful, leading experts in the field: Jack Dongarra and James Hack. It will also offer reports from the various subcommittees and from the Office of Advanced Scientific Computing (ASCR).

She noted that ASCAC is not a watchdog or a review committee, but at the same time, it is definitely not a rubber stamp or doormat. She asked the committee members to introduce themselves and then introduced **Raymond Orbach**, SC Director, to give an update on the Office.

He said that ASCAC had had a profound impact on the direction and agenda for ASCR, SC, and DOE. SC's thinking has moved away from a response to the Earth Simulator to what can be done in ultrascale computing. High-end computation, as embodied in modeling and simulation, is the third pillar supporting scientific advance. (Theory and experiment are the other two.) We need the ability to go backwards with simulations to understand the dynamics of a system. Ultrascale computing can play a critical role. Orbach reported that he had met with vice presidents at Boeing, General Motors, and General Electric, who said that in many instances industry can no longer afford the development of prototypes and therefore needs advanced simulation capabilities.

SC has moved from an approach of competitiveness with Japan to one of scientific accomplishment. Congressional staffers are understanding that one size does not fit all and that science must be the driver for any program in high-end computing. Within DOE, substantial support has been received from leadership for an ultrascale computing program in advanced scientific computing. Eleven workshops have been put together on what high-end computing can contribute to the advance of science; it is important to know what can be accomplished and what

Deleted: Dr.
Deleted: im
Deleted: Dr.
Deleted: s
Deleted: the
Deleted: of the Office of Science (SC)
Deleted: t
Deleted: .
Deleted: Dr.

level of sustained computing speed would be needed to make the next step in the broad range of science.

The commodity chip was sold as a scientific tool and so the industry could stay afloat. The competitive environment does not lend itself to the profit margin that these companies have enjoyed in the past. Virtual prototyping is needed by industry, but the computing capabilities are too expensive. Industry is very interested in being able to do virtual prototypes, and industry leaders have been asked to estimate the savings expected from such capabilities.

Deleted: that

The auto industry can invent a new gadget and put it in production in 18 months. It first appears in the Cadillac, and buyers pay dearly for it. When the R&D costs have been paid off, the gadget appears in the Chevrolet at a much lower cost. The computing industry needs to take note of this strategy.

The ultrascale machines will be expensive, and it must be determined what architectures will make sense. Initially, DOE will be working with different configurations and interconnects to create a user system like that in high-energy physics. That is the paradigm for the future in ultrascale computing. DOE will make large-scale computing facilities available to scientists on a peer-reviewed basis. There is an enormous interest from the National Nuclear Security Administration (NNSA) and from other government agencies, and they will be partners in this effort. The computing community needs to determine what the problem at hand demands and to design a machine to address that problem. It needs to build on the Accelerated Strategic Computing Initiative (ASCI) process, bringing together computational interests and leading the country forward.

Deleted: side of DOE

McRae noted that two other key drivers are people and algorithms and asked where all the people to do this will come from. Further, he observed, the level of investment in algorithms is small in comparison to that in hardware. Orbach responded that it *can* be hard to find computer scientists who are interested in computational science. There must be large investments in universities, more than just fellowship programs. The lack of support at the graduate level has discouraged U.S. students from going into computational science. University leaders will need to start up new programs in high-end computing and other aspects of computational science. Also, a market must be produced for those graduates by being more competitive in the scientific marketplace. People do not want to go into a field where they will be only number two or three in the world, where they are two orders of magnitude behind the leaders. Our prosperous nation must make an investment in R&D.

Dahlburg called attention to the fact that DOE is a mission-driven agency and that a new activity in ultrascale computing would have to be tied to the mission of DOE and the work DOE's scientists are doing. Orbach answered that DOE will have to be able to show, for example, that designs for the ITER [International Thermonuclear Experimental Reactor] are optimized. Scientists and politicians must have confidence that it will work before committing to it. The United States is behind in the computational power needed; it cannot dawdle.

Kulsrud commented that these 11 workshops should also identify what the bottlenecks are in the computational process. Orbach agreed and said that the question is not what the peak teraflop is but what you can do with it. The "time to solution" must be focused on.

Connolly said that most of the increase in speed is accomplished with the number of processors, but few problems scale up to thousands of central processing units (CPUs). Scientists need to work with computer scientists to learn how to get the most out of parallel architectures. Wright noted that everyone can come up with examples in which a careful choice of algorithm produces a greater speedup than an inappropriate algorithm run on a machine with faster raw speed. She asked if Orbach was finding sympathy for his views on computation; whether people understand ultrascale computing, the need for teamwork, etc.; and whether there was anything ASCAC could do to promote these concepts. Orbach said that some aspects of ultrascale computing are hardware-oriented, but success will depend on the right blend of hardware, mathematics, computer science, and applications. SC and ASCR were a lonely voice back in June; ASCAC's help made it possible to broaden the support for this position. SC needs a sustained effort from ASCAC; it cannot allow the message to dissipate. ASCAC members should explain to everyone they meet what needs to be done. Wright asked if ASCAC should be developing more position papers or if anything else would be useful. Orbach responded that he would be delighted to receive a roadmap if Ed Oliver felt that that would be helpful.

Deleted: that

Connolly noted that most of the workforce comes from universities but that many PhDs have found few job opportunities and have reconfigured themselves for careers in other fields. He suggested the creation of computational professorships at universities. Orbach responded that a broader approach might be to train people who can do the algorithms and understand how to create solutions for specific problems. They will be in both the universities and industry. Industry is doing beautiful things. The jobs are there; we do not have the machines to work with now.

Giles said that, to sustain discovery, both people and machines are needed. Measurement is needed to see that an environment that promotes discovery is, indeed, being produced. Orbach said that that is what the Office of Management and Budget (OMB) requires, but we do not know how to make such measurements. McRae suggested looking at how SEMATEC does such measurement.

Wright declared a break at 9:45 a.m. She suggested shifting the dates of future ASCAC meetings to July and February to have greater input and reaction to the budgetary process. She called the meeting back into session at 10:26 a.m. and asked **Jack Dongarra** to review the current status of the Earth Simulator.

The Earth Simulator is a tour de force for a supercomputer system; it is not a commercial product. With high-bandwidth connections, it has 640 nodes with 5120 CPUs producing 40 Tflop/s (theoretical peak) and consuming 7 MW of power (ASCI White uses 1.2 MW; Q uses 6 MW), which would lead to a power cost of $6 million per year. It has a footprint of four tennis courts and will be on top of the Top 500 list of the world's fastest computers for at least 5 years. It is centralized, proprietary, expensive, and not on the network.

Deleted: (not an enormous number for a high-performance computer)

The SX-4 (1995) was constructed on a number of boards with 2 Gflop/s, eight vector pipes, and a 125-MHz clock. The SX-5 (1998) was a one-board computer with 8 Gflop/s, 16 vector pipes, and a 250-MHz clock. The Earth Simulator (SX-6, 2002) has 8 Gflop/s, 8 vector pipes, and a 500-MHz clock; its LSI is a chip, not a board. A few days ago, NEC announced the specifications of the SX-7. It will have more processors per node (32) and 10 times fewer nodes than the Earth Simulator. It will have a peak performance of 18 Tflop/s and a high-speed crossbar. It is expected to ship in December of 2003. On the horizon is the NEC SX-8, a factor of 2 better than the SX-7.

The Earth Simulator (SX-6) has a classical vector architecture with the computations being done in the vector registers. The scalar unit of the arithmetic processor has superscalar-architecture registers but only one path to the main memory (a weakness). The Earth Simulator does have large scalability in moving data within the machine. That is the impressive part of the machine; it results from a fully connected crossbar switch that has a theoretical bandwidth between nodes of 16 GB (12.3 GB realizable). The machine's architecture is hierarchical. The parallel programming language portion employs message passing automatic parallelization (microtasking), OpenMP, the message-passing interface (MPI), and high-performance Fortran (HPF).

The Earth Simulator has 16 GB of memory per node or 10 TB of memory in the system. The memory bandwidth is 32 GB/s from memory to register per processor, 256 GB/s from memory for a node, or 163 TB/s from memory for the system. Data movement through the switch is 12.3 GB/s on the crossbar (bidirectional between nodes) or a cross-section bandwidth of 8 TB/s. With MPI_PUT, users realize 11.63 GB/s.

The LINPACK benchmark (achieved performance) for the Earth Simulator is 35.86 Tflop/s at 87.5% efficiency, solving a problem of size n = 1,041,216 with 5120 processors; this benchmark took 5.8 hours to run. This means that, for the June 2002 Top 500 (which is published twice a year), the performance of the Earth Simulator is about one-sixth of the performance of all the other Top 500 computers combined; it is greater than that of the sum of the next 12 computers on the list; it is greater than that of the sum of the top 15 computers in the United States; it is greater than that of all the DOE and Department of Defense (DOD) machines (37.2 Tflop/s; the sum of all the DOE Defense Programs (DP) and SC computers is 27.5 Tflop/s); and it is much greater than that of the three National Science Foundation (NSF) centers' computers (7.5 Tflop/s). Machine performance is increasing quickly because of parallel processing. The Earth Simulator has the largest factor of improvement from the previous year (5.0) that has ever been seen, despite having fewer processors (5120) than ASCI White-Pacific (7424), the previous year's leader. The Earth Simulator achieved this performance by operating at an efficiency (percent of theoretical peak) of 87%, compared to ASCI White's efficiency of 65%.

It is notable that IBM and Hewlett-Packard (HP) have a lot of machines in the Top 500 list. HP has 168 (12 in the top 100), and IBM has 164 (47 in the top 100).

The success of the Earth Simulator comes from its ability to move data. The NEC SX-6 can move from memory one word per cycle per vector pipe. With eight vector pipes per processor, there are 64 B/cycle or 32 GB/s or 4 Gword/s of memory bandwidth. This measure scales with the number of processors. With a peak of 8 Gflop/s, this scaling gives 32 Gword/s or the potential for 64 Gflop/s. The IBM Power-4-P690 high-performance computer has 16 processors sharing 12 GB/s (1.5 Gword/s) of memory bandwidth, and each processor has a peak of 5.2 Gflop/s to give a total of 83.2 Gflop/s.

The scalar side of this machine is a four-way super scalar; but on the scalar side, it moves around only 4 GB/s (a Pentium 4 moves 4.264 GB/s). When a lot of scalar instructions are being executed, the scalar performance of the Earth Simulator is relatively poor, 1 Gflop/s peak. With little vector processing, a one-processor Intel Pentium 4 produces 1.19 Gflop/s vs the Earth Simulator's 1.16 Gflop/s; however, with a lot of vector processing, the Earth Simulator gets 7.58 Gflop/s of a theoretical peak of 8.0 Gflop/s.

Looking at the future of Japanese computing, one sees that Fujitsu announced on August 22, 2002, that the PRIMEPOWER HPC2500 Supercomputer with a peak performance of about 85 Tflop/s would be available in January of 2003. At the Tokyo Institute of Technology, the Agency of Industrial Science and Technology (AIST) is putting together a very large commodity cluster with a peak performance of about 40 Tflop/s. The Japanese are also looking for funding for a National Campus Grid Infrastructure that would produce 300 Tflop/s.

Twenty years ago, the Lax Report (*Report of the Panel on Large-Scale Computing in Science and Engineering*, December 26, 1982) recommended

- increasing access for the scientific and engineering research community through high-bandwidth networks;
- increasing research in the computational mathematics, software, and algorithms necessary for the effective and efficient use of supercomputing systems;
- training personnel in scientific and engineering computing; and
- performing R&D basic to the design and implementation of new supercomputing systems of substantially increased capability and capacity (beyond that likely to arise from commercial requirements alone).

In high-bandwidth networks, we have seen significant improvements.

In research in computational mathematics, software, and algorithms, the Scientific Discovery Through Advanced Computing (SciDAC) program is an excellent start of cooperation among laboratories at about $60 million per year, but the original program called for an investment of $250 million per year. In 1999, the President's Information Technology Advisory Committee (PITAC) noted that "Over the past decade the U.S. has underinvested in research to create fundamentally new software technologies."

In the training of personnel in computational science and engineering, about 50 universities now offer a graduate computational science and engineering program. Undergraduate education is not at the right level today.

In terms of R&D basic to the design and implementation of new supercomputing systems, no well-established U.S. company remains with its primary focus on technical, high-performance computing. Today, the market is focused on a small and decreasing number of vendors. Research needs to come from the academic community in large amounts. The nation needs to invest in developing a cadre of researchers in this area. Vendors are not producing systems that are specifically tuned to the needs of the scientific market. They are building systems that optimize graphics for games, where the market is.

The country is not faced with a race, but it needs a program, and long-term planning matters. Our country needs a 10-year plan with strategic, rather than tactical, objectives. It needs to look forward to multiple generations of implementation and to realize that it is all right to make mistakes because the lessons learned will influence the following generations. The program should start with a 5-year project that has a new focus every year. It should not relabel SciDAC as an Earth Simulator response.

The issue is not just about hardware. Funding for research in architecture, algorithms, software, systems, *and* hardware is critical. The Earth Simulator path is based on tying together vector processors with a fast crossbar switch, which produces problems with applications that are scalar orientated. The U.S. path is based on tying together scalar commodity processors, which creates performance problems when applications could benefit from vector operations. For some problems, vectors are not the solution. The U.S. scientific community needs a combination of different architectures to handle a wider range of applications. The United States needs to direct the computer industry to develop equipment that can effectively solve our most challenging problems. There is a big difference between a Web server and a scientific computer. Our most challenging problems cannot be solved on a large bunch of Web servers. At the same time, there is not a market at the high end; incentives are needed. The research community needs to partner with the vendors and to engage with them at the beginning of the design stage. The nation must also attract and sustain a cadre of talented scientists and engineers in the development of these systems. All of this will require cooperation among the SC laboratories.

Twenty years ago, the Lax Report made some recommendations that have since been reinforced by the Branscomb Report (1993), the Hayes Report (1995), and PITAC (1999). These reports and developments in the field have made it clear that the current scalable, parallel, high-end computing systems are not well suited to many nationally important, strategic applications. Additional funding should be focused on innovative architectures, hardware technologies, software strategies, and algorithm development that overcome the limitations of today's systems. DOE has to take on the stewardship role of research in computers designed for computational science in addition to algorithms, software, and systems. Also, the relationship of hardware vendors and computational scientists must be changed.

Connolly said that vendors always point to balance sheets and the need for profits and asked Dongarra what types of incentives should be made. Dongarra suggested using the same method used to build aircraft carriers: put up money and seek proposals for prototypes; another strategy would be to appeal to the prestige and patriotism of the vendors' leaders.

McRae asked what the root problem was. Dongarra responded that vendors were sucked into profit gains in making large-market machines and then passed them off to scientific users. Also, we need to develop workers to design the large, specialized architectures needed for high-performance machines. It is a struggle to get money for such educational programs. McRae asked how one sells the program. Dongarra answered that the hardware is the sexy part; then one has to create the story that pulls it all together.

Dahlburg asked if people are trying the arguments suggested earlier by Orbach for substituting virtual prototypes for actual experiments. Dongarra replied that he did not see it being said often or loudly enough.

Connolly said that Congress wants to know what this is good for, not what problems it addresses. Dongarra replied that there are a lot of ideas out there, but not all of them are good. They need to be vetted against computational scientists; those scientists need to be involved before things progress too far and all they can do is make cosmetic changes in the machines.

Meza noted that none of the roadmaps that had been proffered had been successful and asked what they should contain. Dongarra agreed that the roadmaps are there, but policymakers often focus on short-term issues; what is needed is a long-term commitment, education, and funding. Meza asked what ASCAC and DOE can do about this. Dongarra suggested providing incentives and showing examples of how productivity can be improved. Wright said that the scientific community needs to jump in and do something; it should try harder and speak with one voice. She asked Dongarra if he saw prospects for this. Dongarra replied that he saw some promising signs (e.g., SciDAC) and noted that the physics community gets together, figures out how to cut up the pie, and then goes out to get the needed funds. Laub commented that it is hard to pigeonhole SciDAC; it got funded accidently. Its success is from dangling money in front of people.

Deleted: things

Wright then introduced **James Hack** to describe the GS40 Earth Simulator Initiative. He started with some caveats: he is not a spokesperson for the Earth Simulator R&D Center, the National Center for Atmospheric Research (NCAR) has no direct involvement with the Earth Simulator R&D Center, and NCAR is not presently partnered with the Earth Simulator R&D Center, although a partnership has been proposed.

Deleted: T

The Center's machines are aimed at specific, strategic targets with applications to atmospheric and oceanic science and to solid-earth science. The project team set out a schedule in 1996, and they hit the operating date on the mark. At the current time, the U.S. computational-science community does not fully understand the organization of the project and how it works. The Center's budget is split among several agencies that have different priorities.

The Earth Simulator has a MIMD-type [multiple-instruction/multiple-data] distributed memory parallel architecture. The goal was to run NCAR-type programs 1000 times faster. To do that, they would need an efficiency of 12.5% for a machine with a peak performance of 40 Tflop/s. A major factor in gaining that efficiency is the single-stage crossbar network which is connected to all of the 640 processor nodes. It uses 29,000 km of cable to make those connections. The Center is collocated with the Frontier Research System for Global Change in Yokohama, where its infrastructure support is reported to have cost roughly $50 million. The machines are located in a 50- by 65-m building that is electromagnetically shielded to prevent interference; light is brought in by mirrors. It is a remarkable engineering feat.

The system performance is what finally generated interest in the United States. The machine attained a LINPACK benchmark of 35.6 Tflop/s at 87.2% efficiency.

The AFES (Atmospheric General-Circulation Model for the Earth Simulator), developed at the Center for Climate System Research at the University of Tokyo and at the National Institute for Environmental Studies, was an important target application for the Earth Simulator. It is a global spectral model running with full physics with an implementation that allows for coupling with oceanic and other component models. They claim to have sustained an execution rate of 14.5 Tflop/s using 320 nodes (half of the machine) and 7.6 Tflop/s using 160 nodes.

A back-of-the-envelope analysis of the reported AFES benchmark performance indicates that a conventional high-resolution configuration of this kind would contain about 708 × 10^6 spatial degrees of freedom, about 3000 times greater than typical climate configurations, which translates to approximately 100,000 times the typical operation counts from the way these models are normally run. The turnaround rate for a typical reference model is approximately 5200 times real time at a sustained execution rate of 10 Gflop/s. A comparably configured high-resolution AFES configuration is estimated to execute about 70 times real time at 14.5 Tflop/s. To achieve the reported execution rates they are likely to be subcycling the dynamics, exploiting a high-performance Legendre transform library, and ignoring operation-count savings from hemispheric symmetry. If that is true, their actual turnaround rate may be even lower than 70 times real time. This analysis does not indicate that this is a viable production configuration (i.e., it would not enable routine century-long simulations) without significant algorithm changes. However, it *does* provide unmatched high-resolution modeling capability, which will allow researchers to explore and identify scientific problems pacing high-resolution global modeling.

The Central Research Institute of the Electric Power Industry (CRIEPI) ran the NCAR CCM3, the last U.S. fully vector code (1996-1997), on the Earth Simulator at the resolution of a 40-km transform grid. They were able to simulate a single day in about 72 seconds (1200 times real time) using 512 processing elements (PEs); which translates to a sustained execution rate of about 500 Gflop/s. For reference, typical climate configurations of the NCAR model (using a 300 km resolution) require 72 × 10^9 floating-point operations to simulate a single day. Using 64 IBM Power-3 processing elements, which sustain an execution rate of about 4.2 Gflop/s, can complete a one-day simulation in about 17 seconds (about 5000 times real time). With the Earth Simulator capability, a very-high-resolution model configuration would be a routine way of conducting scientific work. Machines available in the United States cannot match these capabilities. The Los Alamos National Laboratory (LANL)–NCAR parallel ocean program (POP) model is currently getting a turnaround of about 0.12 simulated years per day using 128 IBM SP nodes (512 PEs), or running about 45 times real time. CRIEPI estimates a turnaround of

Deleted: The Center falls under the auspices of the AIST, operates in parallel with the Frontier Research System for Global Change, and with that organization provides services to the Research Organization for Informa ... [1]
Deleted: -type
Deleted: and a
Deleted: boost
Deleted: and a
Deleted: that has 128 switches, each of
Deleted: and
Deleted: , although it has not be ... [2]
Deleted: a
Deleted: with
Deleted: with
Deleted: An
Deleted: system
Deleted: the
Deleted: produces
Deleted: the
Deleted: is
Deleted: about
Deleted: gives
Deleted: , and a
Deleted: e
Deleted: gives
Deleted: They
Deleted: shorter
Deleted: machine
Deleted: I
Deleted: allows exploration and ... [3]
Deleted: got a simulated
Deleted: with
Deleted: that would give
Deleted: With 64 Pes, it got
Deleted: for a sustained
Deleted: such as this becomes ro ... [4]
Deleted: touch it
Deleted: For t
Deleted: l
Deleted: , we are
Deleted: on
Deleted: projected

<10 simulated years per day on about 128 nodes (1000 PEs), or about 3600 times real time, on the Earth Simulator. This would enable routine multi-decadal global eddy-resolving ocean simulations.

The current assessment is that the Earth Simulator will enable fundamentally new simulation capabilities, such as efficient exploration of high-resolution parameter space, extensions to the physical climate system, exploration and use of ensemble techniques, and sophisticated treatments of processes with large uncertainties. Also, it is a dedicated facility, a focused computational resource that dwarfs anything available in the United States. However, access to the facility is limited to Japanese collaborative partnerships.

Deleted: is only going to occur through

Science opportunities fall into two categories: capability (high-resolution modeling, extensions to the physical climate system, and sophisticated treatments of processes with large uncertainties) and capacity (the exploration and use of ensemble techniques). The Earth Simulator provides plenty of both!

The scientific opportunities and questions that can be addressed in high-resolution global modeling include the physical and numerical sensitivity to resolution, internal variability as a function of resolution, feedbacks between the fine and large scales, and regional representation of climate properties. From CRIEPI, it is estimated that a global T170 (70-km) configuration of the NCAR Community Atmosphere Model (CAM) would run about 2100 times real time using 256 PEs on the Earth Simulator, or be capable of simulating about 6 simulated years per day. What is expected from the IBM Power-4 with 256 PEs is about 220 times real time, or a simulation rate of approximately 0.6 simulated years per day. Hack then went on illustrate how many important features in the climate system would be better simulated using much higher resolutions than are currently used in global climate models.

Deleted: running

Deleted: on the Earth Simulator

Deleted: give

Deleted: with

Deleted: e

Deleted: out of

Deleted: about

Deleted: Currently, high-resolution global modeling misses data because of truncation, and it misses subgrid-scale processes. As a result, parameterization is relied upon for a lot of the data produced.

Deleted: ¶

Ensemble techniques allow one to explore the dominant modes of variability in the (chaotic) climate system, which has a lot of noise. They also provide the opportunity to evaluate climate sensitivity, can be used to attribute the response of the climate system to a specific forcing factor, and can be used to evaluate modeled response on a hierarchy of time scales. Although ensemble techniques can consume large amounts of computing time, these techniques *do* allow the refinement of model results in the face of projection uncertainties, low signal-to-noise ratios, and large amounts of natural variability.

Deleted: consumptive of running time

Deleted: us to

Partnering to use the Earth Simulator involves certain costs and risks. An important cost is related to the maintenance of a large evolving software infrastructure on multiple architectures, where users would not have ready access to the vector architecture. Other costs arise from the coordination of experimental design, setup, and execution, and the logistics associated with migrating the simulation data back to the United States for analysis. Terabyte data sets are routine, and the costs of a dedicated data-analysis platform are not trivial. Access to the Earth Simulator is a particularly difficult problem; it is a closed facility, and an onsite presence is required. Moreover, the basic governance and management process for the Earth Simulator is poorly communicated and understood, even in Japan. This involves important issues such as allocation, scheduling and access policies related to the use of machine resources.

Deleted: . A lot of costs are involved because of

Deleted: and

Deleted: More costs would be involved in

Deleted: getting

Deleted: is

Orbach had discussed the ramifications of the Earth Simulator in terms of the United States' losing the lead in climate science research and the closing of this gap requiring an aggressive and integrated 5-year investment program. Hack said that he does not believe that this is an accurate assessment. The United States has *not* lost the lead in climate-science research; it continues to provide scientific leadership in climate research, but it is handicapped with respect to implementing new knowledge and ideas. The main "gap" is a lack of access to the best available

Deleted: s

tools because the architectural choices have been artificially constrained during the past 5 years and because dedicated resources have been relatively meager.

The Earth Simulator is an embarrassment for U.S. high-performance-computing (HPC) efforts. It is not a revolutionary machine. It *is* an extraordinary, evolutionary engineering accomplishment based on approaches that were pioneered and developed in the United States. It appears to be a well-balanced high-performance system architecture, and has exposed the flaws in the commodity approach to HPC. The machine presents an immediate scientific exposure in that it provides a significant scientific advantage to computationally-starved applications. Factors that limit this exposure are that these applications need to be immediately ready to fully exploit the opportunity (i.e., they can fully exploit a distributed-memory vector architecture), and that Japanese partnerships are required for access to the facility. The Earth Simulator's computational advantage, and the exposure it presents, is likely to exist for at least 3 years, meaning that there is a significant period of time during which opportunities for pioneering discoveries will be missed. Over the longer term, no scientific experience will be gained from operating in the new parameter space enabled by the Earth Simulator capabilities.

A number of questions are raised by the Earth Simulator. Should it be exploited through international partnerships? The answer is maybe, if the required investment is clearly worth the effort. There are many risks, but such partnering could have an enormous scientific payoff. Should infrastructure investment be sought in the United States? The only immediate response that could be made is to try to match the capability of the Earth Simulator. One way to do that would be to invest in large SX-6 or SX-7 systems. The path forward for such a strategy is not obvious which begs the question does the computational-science community really want to go back to the future? Another way would be to invest in large, dedicated, "scalable" commodity systems; but that may be an even worse idea because it could be perceived as rewarding mediocrity and dishonesty. Equally important, despite claims to the contrary, the United States cannot match the capability of the Earth Simulator in the short term using commodity high-performance systems. A longer-term investment strategy may be the best solution. That strategy needs to include a concerted effort to focus the U.S. HPC program in a comprehensive way that includes hardware, software, and people. The computational-science community needs to stop applying irrelevant economic models to HPC (one does not go to K-Mart to buy an aircraft carrier). And it needs to separate tinkering from investments that facilitate science. The computational-science community must stop pandering to the U.S. computer-hardware manufacturers. Instead of asking what the vendors are able to deliver, we need to be clearly defining what we need. This must include a willingness to pay for comprehensive, well-balanced computational environments, not distributed bits and pieces leveraged off of commercial products. The applications community needs to insist on responsiveness and to demand accountability from the vendors. The computational-science community needs to get off the peak-performance bandwagon, which continues to be an intellectually dishonest characterization of high-performance computing capabilities. Finally, there needs to be more investment in a comprehensive software infrastructure. Having a stable, functional, and robust scientific-application development environment is important to the scientific community. Another important component of this investment includes fundamental algorithm development.

In summary, the Earth Simulator is the real thing. It should not be a surprise; it went almost exactly according to plan. It is clearly the most powerful general-purpose HPC system in the world. It may not strictly be a one-of-a-kind system in the longer term. It is a real HPC effort, not a repackaging or demonstration project. It is a dedicated high-performance simulation capability

Deleted: has

Deleted: it

Deleted: Drawbacks that the United States must face include the need for

Deleted: the need for

Deleted: and

Deleted: a

Deleted: ,

Deleted: but

Deleted: s

Deleted: D

Deleted: It

Deleted: and it needs to pay for the whole package, not distributed bits and pieces.

Deleted: and to

Deleted: and

that can serve as a paradigm for other areas of socioeconomic or scientific importance. It highlights several problems with the U.S. HPC effort:

- You can't build a Lamborghini out of Volkswagen parts.
- An HPC overhaul is in order unless the United States is willing to relinquish its leadership role.

Deleted: no

Meza asked if his statement about U.S. leadership in climate science will still be true 3 years from now. Hack replied that he hoped so; under some scenarios, it might not be.

McRae asked about using adaptive meshes instead of finer meshes. Hack responded that how and where you do the refinements is the problem. The weather community has invested some effort in developing these methods. But given that you are looking for relatively small changes in the climate system, you need to be concerned about introducing systematic biases by how you discretize the system. McRae noted that global climate is a global problem and asked why the U.S. researchers should not just collaborate. Hack replied that there are a lot of collaborations. The way this center is set up does not allow outside access, hampering collaboration.

Deleted:

Deleted: money

Deleted: it is

Deleted: and

Deleted: can be induced

Connolly asked how much climate modeling uncertainty would be reduced with these more powerful machines. Hack answered that there are some climate processes that cannot currently be parameterized. One has to go to a higher resolution, and that requires bigger machines. Connolly asked if his calculations were limited to one type of machine. Hack replied, no; 10 years ago, they worked on building distributed-memory implementations of climate models through the CHAMMP [Computer Hardware, Advanced Mathematics, and Model Physics] Program, where one goal was to explore higher resolution configurations. The algorithmic and implementation effort was successful, but the machines did not provide adequate scalability. The climate scientists then opted for a decomposition strategy that exploited vector constructs in one dimension and some type of message-passing paradigm in the other dimension. Therefore, the models being run now have a one-dimensional decomposition that currently limits the scalability of the model. A more general two-dimensional implementation is in the works under a collaborative SciDAC activity.

Deleted: and modeled

Deleted: increasing the resolution

Deleted: at

Deleted: were not successful in terms of

Deleted: e

Deleted: .

Kulsrud asked how the Japanese expect people to use this machine if it's not accessable via the network. Dongarra conjectured that they may want to control how people use the machine.

Giles observed that the Committee and the scientific and policy communities have discussed a lot of ways to bring vendors, laboratories, users, etc. together and asked Hack for his reading on this. Hack replied that he hoped the comprehensive strategy would include such collaborations. Wright asked if he had any comments on what that collaboration should be. Hack said that he thought that a lot of sensible things were said (this morning). Leveraging fruitful technologies is not the way to go. The Cadillac approach is the way to go. Pay for the R&D up front and let the benefits trickle down. Dongarra responded that the consensus is that an effort will be put in place to build a better piece of hardware, but won't include software, human resources, and other factors. Mission-oriented efforts do not employ a long-term strategy. Wright commented that budget deliberations *are* a short-term operation.

Deleted:

Deleted: and not

Meza observed that computing is an integral part of how we do science, suggesting that scientific leadership could be compromised by not having access to something like the Earth Simulator. Hack said that he would feel a lot better about the Earth Simulator and its use if he saw a science plan in place to govern the use of that facility. A primary consideration should be making a facility run well for the applications that need to be pursued.

Deleted: ; we need it because we want to do science

Wright noted that these collaborations are hard, and they do not happen without money.

Andrew White commented that the response has to be a sustainable long-term program. We have failed at both the government and the hardware vendor ends during the past 10 years. What lessons have we learned? [No answer.]

Stevens asked how much the United States spends on climate related activities. Hack guessed that it might be as high as $10 billion[†], a substantial portion of which is on observation. Stevens asked if there could be a voice within the atmospheric science community to say that $1 billion was needed for a computational instrument to analyze all these data to make progress in the science. Hack replied that, years ago at NCAR, computing funding was taken out of the atmospheric science program for that purpose because the need for computational resources was so important. Such decisions could be made again in the future. Stevens noted that the decision to fly a satellite is made on the basis of the data payback. A similar consensus is needed about the value (payback) from a computing facility with a willingness to devoting, say, 10% of research funds on computing resources. Hack said that he thought that that is possible.

[†] This estimate is actually much closer to $1.8 billion, with approximately half of the resources associated with observational investments, such as satellite platforms.

Wright noted that where it gets hard is when someone has to make a decision about the allocation of budgets. She declared a break for lunch at 12:54 p.m. The meeting was called back to order at 2:15 p.m., and Wright called upon **Alan Laub** to present an update on the SciDAC Program.

SciDAC is the first federal program to support and enable real computational science and engineering (CSE). It is supposed to produce a new way of doing science. Its funding is small; it is a proof of concept for the first simulator machines, although the hardware component has not been funded, yet.

CSE is a widely accepted label for an evolving field concerned with the science and engineering of systems and methods to solve computational problems arising in science and engineering. It is not just programming. One starts with a physical system that needs to be understood better, develops a mathematical model, works out solutions, and identifies and exploits feedbacks. Simulation and modeling focus on the end result; CSE focuses more on how to get to the end result. One constantly tries to improve the analytical processes through the important feedback loops among modeling, CSE, and simulation. The science that is discovered is the ultimate measure of success. CSE is founded on mathematics, computer science, and applications. It is *not* a subdiscipline of any of these, but uses all of these disciplines heavily. Graduate education in CSE is just starting.

By its nature, CSE is team oriented, making it intrinsically hard to do. CSE usually requires teams with members and/or expertise from at least mathematics, computer science, and application areas. One person cannot know it all. This collaborativity brings about language and cultural differences. The usual reward structures, especially in academia, focus on the individual and are incompatible with the goals and methods of CSE. SciDAC helps break down barriers and can lead by example; the national laboratories of DOE are a critical asset. There are many benefits for DOE and science, both technical and sociological.

SciDAC has been under way for a little more than a year. The first principal investigator (PI) meeting was in Washington in January 2002. Its theme was an introduction to the integrated SciDAC Program and the initiation of team building. The second PI meeting will be held in

Deleted: experimental data collection

Deleted: replied, about

Deleted: this

Deleted: Wright then introduced **James Hack** to describe the GS40 Earth Simulator Initiative. He started with some caveats: he is not a spokesperson for the Earth Simulator R&D Center, t

Deleted: T

Deleted: he National Center for Atmospheric Research (NCAR) has no direct involvement with the Earth Simulator R&D Center, and NCAR is not presently partnered with Earth Simulator R&D Center, although a partnership has been proposed.¶
The Center's machines are aimed at specific, strategic targets with applications to atmospheric and oceanic science and to solid-earth science. The project team set out a schedule in 1996, and they hit the operating date on the mark. At the current time, the U.S. computational-science community does not understand the organization of the project and how it works. The Center falls under the auspices of the AIST, operates in parallel with the Frontier Research System for Global Change, and with that organization provides services to the Research Organization for Information Science and Technology. The Center's budget is split among several agencies that have different priorities.¶
The Earth Simulator has a MIMD-type [multiple-instruction/multiple-data-type] distributed memory and a parallel architecture. The goal was to run NCAR-type programs 1000 times faster. To do that, they would need an efficiency boost of 12.5% and a peak performance of 40 Tflop/s. A major factor in gaining that efficiency is the crossbar network that has 128 switches, each of which is connected to all of the 640 processor nodes. It uses 29,000 km of cable to make those connections. The Center is collocated with the Frontier Research System for Global Change in Yokohama, and its infrastructure support cost $50 million. The machines are located in a 50- by 65-m building that is electromagnetically shielded to prevent interference; li ... [5]

Deleted: 6

Deleted: Tflop/s at 87.5

Deleted: 2

Deleted: % efficiency, although it has not been as reliable as one might want.¶
The AFES (Atmospheric General-Circulation Model for the Earth Simulator), developed at the Center for Climate System Research at the ... [6]

March of 2003 in Napa, Calif., on the theme of assessing the progress of SciDAC (especially in making the team approach work).

As part of making a science case for SciDAC, one can point to examples of early success:

- Steve Jardin of Pacific Northwest National Laboratory (PNNL) said that SciDAC "is a significant factor in our productivity, comparable to that obtained by going to the next-generation computer." Codes are running 2 to 4 times faster because of new algorithms.
- Tony Mezzacappa of Oak Ridge National Laboratory (ORNL) said "The SciDAC Program is making possible a whole new class of supernova simulations."
- Rob Ryne of Lawrence Berkeley National Laboratory (LBNL) said that SciDAC algorithmic advancements and visualization in accelerator design enable us to "optimize designs to reduce costs and help ensure project success."
- Donald Batchelor of ORNL said that his fusion-energy codes run 10 times faster because of suggestions from the mathematician on the team.

Dahlburg asked where SciDAC should go next. Laub said that it should get greater funding for application development. Program management would like to broaden the number of applications funded and enlarge the teams on each project. In materials science, biology, and fusion, there are prime candidates for expansion. Dahlburg noted that the SciDAC people in fusion are having great success and that those not in SciDAC are feeling that they are left out.

Lester asked if there would be a new call for participants. Oliver responded that the office will need the money first.

Lester asked where the computer resources come from. Laub replied, the National Energy Research Scientific Computing Center (NERSC) and Oak Ridge. Lester asked how much computing time they got. Laub replied that they make requests for CPU time on, say, NERSC. Then the available time is allocated. More time will be available in the future on the Oak Ridge Cheetah machine. The program could be a victim of its own success and run up against capacity limitations. Also, it needs to take a multiarchitectural path.

Giles asked if SciDAC will spin off projects, absorb all science, or die out. Laub responded that it will spin off daughter programs, become a way to conduct science, and die. Giles asked if it will stimulate more R&D than it pays for. Laub said that he expected so, but that he sees it evolving slowly.

Meza suggested that, in addition to the flops, SciDAC should add networking, storage, and other resources. Laub admitted that some of the users are pushing and planning for that.

McRae asked what the future networking requirements for SciDAC would be. Laub responded that the High-Energy Nuclear Physics (HENP) people have a great need for networking and storage, as they detailed in the Transatlantic Networking (TAN) report [www.uscms.org/s&c/reviews/doe-nsf/2001-11/docs/TAN-report-final.doc.pdf].

Wright introduced **James Corones** to report on the activities of the subcommittee formed to consider a response to the Earth Simulator. This activity started at the May 2, 2002, ASCAC meeting in light of the Earth Simulator success. The Facilities Subcommittee report, which was presented at that meeting, was proposed to be considered as an interim report rather than as a final report; the effect of the Earth Simulator was to be added. Orbach specifically asked for a quick response to the Earth Simulator issue, and the Facilities Subcommittee was asked to organize a meeting on the Earth Simulator.

Deleted: to be

The Subcommittee held a workshop on May 15-16, 2002, to learn more about the Earth Simulator. Several speakers and panels made presentations on such subjects as the DOE science-initiative needs, technology assessment, and response options. Despite the short notice, a robust contingent of people attended the workshop. More information about the meeting, including talks and the report, can be found at www.ultrasim.info/esrr_meeting/index.html.

Some conclusions reached are:

- The Earth Simulator is a real general-purpose machine.
- It is focused on a specific class of scientific problems.
- It provides a significant scientific advantage to its users.

The core issue is leadership in computational science. Numerous areas of science and engineering are positioned to take advantage of the Earth Simulator class of machines. Workforce issues are a critical part of the response. Another set of conclusions within the nexus of the DOE environment is that ASCR is well, perhaps uniquely, positioned to lead a national response because of

- its previous and ongoing planning;
- DOE’s SciDAC experience, an important element;
- DOE’s experience working with vendors;
- serious underfunding of computational science in ASCR (in both absolute and relative metrics); and
- DOE’s good positioning in the interagency community.

About 10% of the NSF’s R&D funds are invested in computing, and 10% of NNSA’s. Only 5% of SC’s R&D funds are invested in computing.

In terms of technology readiness, there are no technology barriers to forming a national response, and close collaboration between government and vendors is necessary.

In terms of the shape of the response, the Earth Simulator is a challenge to the United States’ leadership in computational science, requiring a science-driven response. Short-, intermediate-, and long-term components are needed, including advanced architecture development, computational science and enabling technology research, and focused technology deployment in support of the DOE mission applications.

The report said that DOE laboratories and domestic computer vendors have a long history of successful collaborations in high-end computing. Those parties are ready and willing to respond to the Earth Simulator challenge. The mandate of ASCR is computational science in support of DOE missions. With adequate resources, DOE can compete in, win, and continue to dominate this space.

Connolly asked if the report will have an addendum on what the national laboratories are doing to respond to the Earth Simulator. Corones said that the 11 workshops are identifying specific responses, and their results are posted on the website. The broad response will be covered in the later presentation by Oliver and Polansky.

Wright declared a break at 3:05 p.m. She called the meeting back into session at 3:29 p.m. and asked **Walter Polansky** and **Edward Oliver** to review the response of ASCR to the Earth Simulator and the resulting ultrascale computing planning effort. Polansky reported that, in FY02, the Mathematical, Information, and Computational Sciences (MICS) Division of ASCR

- conducted a workshop and eight town meetings to evaluate the impact of the Earth Simulator;
- launched an early-career PI planning activity to strengthen the core research program;
- cosponsored with the Basic Energy Sciences Advisory Committee (BESAC) a workshop on theory and modeling in nanoscience;

- conducted in cooperation with HENP a workshop on networking requirements for the future of science;
- conducted Genomes to Life workshops on applied mathematics and computer science, continuing an 18-month collaboration with the Office of Biological and Environmental Research (OBER); and
- initiated an ESnet backbone upgrade from 622 Mbs to 10 Gbs to service increased networking requirements for science.

In addition, it is working in cooperation with Fusion Energy Science (FES) on a series of workshops on computation in fusion energy. In FY03, it plans to initiate reviews of applied mathematics and collaboratory pilot research activities, to initiate a review of the SciDAC portfolio, and to continue workshops and town meetings to assess ultrascale-simulation needs. The eight town meetings held so far have included only the core DOE research community. These meetings must, in the future, include vendors, also, as ultrascale simulation is discussed.

Increases in the MICS budget derive largely from the Genomes to Life Program and from nanoscience.

The Earth Simulator has revolutionized the field of scientific simulation. It was a wake-up call. The resulting workshop found that, without a robust response to the Earth Simulator, the United States is open to losing its leadership in defining and advancing frontiers of computational science as a new approach to science. This area is critical to both our national security and economic vitality.

Polansky reviewed the events related to the Earth Simulator as they unfolded, starting with the announcement of the Earth Simulator in the *New York Times* on April 20, 2002. The Earth Simulator rapid-response meeting was held May 15-16. A meeting was held with IBM, ORNL, and NCAR June 12. A meeting was held with the Office of Science and Technology Policy (OSTP) on June 14. A meeting was held with Cray, ORNL, and NCAR on June 20. The subcommittee visited NASA Ames and Silicon Graphics on June 19-21. The Society for Industrial and Applied Mathematics (SIAM) minisymposium was held July 8. The SAC meeting was held July 17. DOE visited the Earth Simulator July 2. Discussions were held with the fusion, chemistry, astrophysics, accelerator-design, network, and nanomaterials communities from August 5 to 30. The office also talked with NERSC users and the biological community September 5-17.

The town meetings typically ran 1.5 hours and involved a small slice of the DOE SC applications community. Typical questions posed were

- How can this science be advanced through simulations?
- Why are these advances important to the field and to DOE? What breakthrough simulations need to be performed in terms of the science?
- What computational and networking resources would be needed to perform these breakthrough simulations?
- When would the community of researchers be ready to use those resources?
- What challenge does the Earth Simulator pose to each field of science?

The town meetings were designed to stimulate discussions among peers about opportunities presented by ultrascale computing. It was hoped that the attendees would do self-assessments of the influence the Earth Simulator may have on simulations of physical, chemical, and biological systems. The attendees were asked to write white papers on building the science case for ultrascale simulation. Primarily, the desire was for them to hold further dialogue among themselves. An air-tight science case is needed, and the applications people are being counted on to supply that. Based on the white papers, a team in ASCR started fleshing out a proposed activity.

The team started rolling up the science case from the white papers on climate science, magnetic fusion energy, combustion science, environmental molecular science, and astrophysics. A performance-improvement factor of more than 50 or 100 was postulated by each of the white papers. They also identified the needed simulations and the significance to the average person of that simulation in their respective fields. The application-performance matrix will list the best performance on an application and what was learned from that application.

Some of the issues raised in the town meetings are the needs to deliver leadership-class computers for science; to couple application scientists with computer architects, engineers, and semiconductor researchers; to partner with industry on applications; and to partner with domestic vendors.

The scientific problems of strategic importance that were cited typically involved physical scales that ranged from 5 to 50 orders of magnitude, involved a number of scientific disciplines (e.g., chemistry and fluid dynamics to understand combustion); had to be addressed by teams of mathematicians, computer scientists, and application scientists; and used facilities that produce gigabytes of data shared among scientists throughout the world.

Deleted: generate

Within the MICS base program is an early-career PI activity. Its purpose is to identify exceptionally talented researchers early in their careers and to interest them in research programs relevant to DOE missions. It is designed for tenure-track regular faculty at U.S. academic institutions 5 years or less after receiving the doctorate or after completing a postdoctoral position. In FY02, ASCR received 132 applications and made 17 awards for a total of $1.6 million per year for three years. Of these, 7 were in applied mathematics, 8 in computer science, and 2 in high-performance networks. Other agencies cannot provide this type of training because the typical award includes a 1-month visit of the PI and a graduate student to a national laboratory. ASCR is waiting for the budget to be passed to issue the request for applications for next year.

Deleted: the

McRae asked how far down the list of 132 applications to the early investigator program one could go and still keep the same quality. Polansky replied, about 30.

Wright asked about the award to ORNL as an Earth Simulator response with no peer review and how such awards will be made in the future. Polansky responded that a detailed account of ASCR's plans will likely be made at the next ASCAC meeting. Those plans need help in the science case, in the planning process, and in selecting what is important. The Office would welcome input on formulating those plans. Oliver commented that they needed something like the Lax Report. Messina noted that, by next March, ASCR will be locked in; any advice must be offered before that. Wright said that she would like ASCAC to come up with a method to help in developing the overall plan for this activity.

Deleted: our

Deleted: r

Polansky returned to the topic of the award to ORNL and said that there was a press release (origin not revealed) in late July about an evaluation of the Cray X-1 for suitability for simulation. The press release said that ORNL was to be the recipient of a $100-million grant for that purpose. That report was false. The Cray X-1 evaluation is being done as part of another program at another advanced computing testbed [at Argonne National Laboratory (ANL)]. Peer review of the ORNL center as a testbed was conducted about a year ago. That review gave ORNL and its Advanced Computing Research Center's testbed high marks. As part of that program, ORNL and ANL are pre-approved to conduct evaluations of promising computer

Deleted: It

Deleted:

systems. A proposal to evaluate the Cray X-1 was received and reviewed and is funded to the amount of $3 million. It could be the first evaluation of an ultrascale-simulation application.

McRae noted that the time-zero of the Earth Simulator was actually 5 years before 2002 and asked how much is known about the Japanese planning process. Polansky responded that, early in the process, the Japanese selected a suite of applications that they wanted to pursue and then determined what breakthrough simulation they wanted to perform. The United States can learn from that. McRae asked if the United States should cut down the number of applications it considers. Polansky estimated that the number will be cut down to one or two applications that have the greatest promise for the country.

Dahlburg noted that the paradigm is shifting from hardware to team collaboration and asked what the Office is doing to respond to that shift. Polansky answered that the Office has started some important activities in a number of areas. They have to be knitted together into an integral unit.

Meza asked if any future meetings were planned with NNSA. Fred Johnson replied that, at a program-manager level, ASCR has a healthy, cooperative relationship with NNSA. Staff members had a tour of all SC and NNSA computer facilities this year. There was also a multi-agency workshop on intelligence communication [on computer facility needs] that SC was invited to participate in because of its relationship with NNSA.

Meza asked if the Office had thought about a midcareer PI activity. Polansky replied that it had not, but would be interested in hearing any ideas about it.

Wright noted that there was no mention of mathematics or computer science in the presentation. DOE is uniquely qualified to make sure that mathematics and computer science are brought into the planning and execution phases of R&D. Polansky replied that the point was well taken.

Messina commented that unsolicited proposals have sometimes instigated DOE action and asked if DOE is getting such suggestions in response to the Earth Simulator. Oliver replied, yes and no; the Office welcomes good ideas from any source.

Kulsrud asked where that $3 million for the evaluation of the Cray machine came from. Polansky answered that there is money in the FY02 budget for testbed activities. Dahlburg asked what had been tested in the past. Polansky mentioned the IBM Power-4 and Compaq systems. Hitchcock added the Paradigm, Kendall Square, Chiba City, full Linux, Alpha, and many more technologies.

Deleted:

Connolly asked what the rest of the $15 million in the FY02 budget was used for. Polansky answered, lease payments, machines that were bought, resources for SciDAC applications, and some small-scale evaluations.

Wright noted that competition among the national laboratories is a perennial problem and asked how DOE promotes cooperation. Polansky said that there were several ways to promote cooperation, including having as open a process as possible. Now, DOE is trying to get the lay of the land through these town meetings. All of the important players have to be involved.

Stechel said that industry can get excited about the use of computers to boost productivity; it requires simulations at all levels. She asked where the best point would be for DOE to couple with industry. She noted that they need theorists and simulators; they never think about those personnel needs. By just pushing the high end in machines, coupling with industry will never happen. Polansky noted that developing partnerships with industry is useful on two fronts: DOE has high-end problems that require high-end approaches, and those approaches may be useful to some industries. The second is that industry is a consumer of high-end computers. This activity

Deleted: it

will push the high-end computers to new advances that industry can benefit from. Stechel stated that the excitement of simulations must be pushed at all levels, not just at the high end. Wright called attention to a recent talk by a researcher from Procter & Gamble on the simulation of fluid flow in Pampers and air flow over a Pringle, showing that simulation is becoming important in consumer industries.

Meza asked what effects Homeland Security will have on ASCR. Polansky said that that is being debated in Congress.

Lester asked how the ASCR staff made judgments about the base program and SciDAC allocations. Polansky said that the available massively parallel processing hours available on NERSC were partitioned to include some (about 20%) for SciDAC. That portion then went to the SciDAC program manager for allocation. The base-program portion was divided among the program offices in accordance with their budgets, with priorities within each office being determined by the program managers. Lester asked if that was the best way. Polansky said that it offered the best opportunities for the SC portfolio. An unsolicited-proposal model should also be looked at. Dahlburg called attention to the fact that the ASCAC Facilities Subcommittee has recommended changes that would make it a national program. Polansky noted that this is the first year that there have been constraints in the cycles available. Miner explained that the program office said the science did not need to be reviewed again, so a core review was done, including a computer scientist, a NERSC person, a computer scientist, and an application person. They determined what the best use of the resources would be, and the program manager determined the time to be allocated.

Wright opened the floor to public comment. Ciment asked about the performance matrix and whether there has been an analysis of software performance from software updates. Polansky replied that ASCR does not do that, yet.

Ashby commented that the capability needs should go out 5 to 10 years. Dahlburg replied that that is being done by a series of roadmap subcommittees.

Crawford stated that the (false) perception in the community is that this is a race for having the biggest computer; there is a perception that this activity is not planned, balanced, sustained, and cooperative (among laboratories, program offices, and agencies). The word needs to be gotten out about what the plan is. Wright asked Oliver when the FY04 plan would be known. Oliver responded that they could not talk about it until the FY04 budget is officially announced by the White House.

Kingscott said that, as a vendor, IBM realizes that there is a need for various architectures to be pursued. It is important to lay out the long history and need for discussions between the vendor and user communities. One cannot lose sight of the marketplace, however. Failure in the marketplace means the revenues to fund these activities dry up. Human resources are important. We must grow the number of computational scientists, chemists, etc. The United States can do this, and she urged that it not be looked upon as just a DOE enterprise.

Meza asked Crawford if there was a particular committee that DOE and/or ASCAC should go to in NNSA to discuss these options. Crawford suggested getting the NNSA and DOE laboratory leaders to work on this question. Messina noted that NNSA has an advisory committee that could be a conduit for such discussion.

Wolff commented that it is easy to call for government-industry cooperation, but such cooperation is difficult to effect. He asked if there were any new ideas on how to do it. Polansky said that the office does not have any hard and fast plans, but it has had discussions with each of the vendors interested in ultrascale simulations and asked them to submit white papers. It expects them and others to participate in future town meetings. Giles asked if anyone had considered working with the National Institute of Standards and Technology (NIST), which has a mandate to work on such development. Wolff retorted that NIST is a whipping boy of Congress.

There being no further public comment, Wright adjourned the meeting for the day at 5:22 p.m.

Friday, October 18, 2002

Wright called the meeting to order at 8:27 a.m. She noted that SC has six advisory committees. Their chairs have met and visited with OMB, OSTP, and others to tell them why they should support science. They do not advocate for any particular program. There was tremendous interest after the announcement of the Earth Simulator. She urged other committee members to meet with these agencies, if possible.

She asked **Charles Romine** to speak about the activities in Applied Mathematical Sciences (AMS). In years past, research proposals from the national laboratories were sent out by mail to independent reviewers, but that did not give a good control over the disciplines. Romine divided the AMS portfolio into three groups of topics, and organized a review by six experts on October 1-2 of the first group. This process will be repeated during the next two years for the remaining two groups. The three groups of topics were

- linear algebra, optimization, predictability, and uncertainty quantification;
- differential equations and high-performance computing; and
- computational fluid dynamics, advanced meshing technologies, and others.

The PIs' presentations were made publicly; that is, representatives from all the laboratories were present at all the presentations. The written proposals as well as the oral presentations were reviewed; generally, the panel gave high marks to the laboratory research activities.

Romine is considering an emphasis on multiscale mathematics because many SC applications involve multiscale physics, biology, and chemistry. It is not that the computing power is too small but that the understanding of mathematics is incomplete. This assessment is true in a variety of disciplines: climate, biology, high-energy physics (all of which involves many length and/or time scales), combustion, and fusion. Current models assume separation of length and time scales, but this simplifying assumption fails in practice as you introduce greater amounts of physics.

Each case involves transfer of information across scales. Coupled equations are not known to be well-posed. The data they are based on are frequently unreliable. Current solution techniques are not adequate to handle multiple scales; techniques are brittle and do not scale up or down well. Computational power will be taxed; a brute-force approach will not work. However, the effectiveness of new multilevel methods may benefit from improved information.

Multiscale mathematics may need input from research in the following areas: variational principles, separation of scales, analogues of filtering techniques, the interplay between discrete and continuous mathematics (ASCR has investments in some discrete mathematics; techniques useful in some areas are sometimes helpful in others), and advanced statistical methods.

A roadmap is needed because of uncertainty about what areas of mathematics are needed to make progress. Programmatic incentives have to be put in place to ensure that the mathematics research is strongly coupled to applications. The critical mathematical and application areas have to be identified. The areas of research must be prioritized so that the critical areas can be tackled first. The roadmapping is just beginning now; no draft is on the table, yet. A series of general

Deleted: s

Deleted: We must prioritize

Deleted: we can tackle

and targeted workshops are planned to identify two or three application drivers. Comments and input on the plan from the Committee are welcome.

McRae observed that a revolution is taking place in sensor technology, producing a massive amount of data. Mapping data is a topic that should be included in the roadmap. Design optimization and uncertain propagation are other topics that should be pursued.

Dahlburg said that each of the two or three drivers should include experiments that produce the mathematical problem. In the ICF [Inertial Confinement Fusion] Program, the same program managers were in charge of both the theoretical and experimental aspects, making cooperation between theorists and experimentalists very effective. Giles commented that the researchers would then need access to an experimental computer facility. Also, he had never heard of a mathematics program in such a setting and asked if mathematicians supported such an idea. Wright suggested that it might be given a different name. As long as scales are separated, the problem is limited. But when scalars interact, the mathematics are often application-specific. Romine said that the hope is that some commonalities will emerge, at least commonality between or among a few applications. It should be recognized that DOE does support both algorithmic development and simulations.

Wright called upon **Ellen Stechel** to review the joint ASCAC-BESAC workshop on theory and modeling in nanoscience, which took place on May 10-11, 2002 and was cochaired by Stechel and William McCurdy of Lawrence Berkeley Lab. Stechel reviewed the makeup of the workshop steering committee and presented the operating definition of nanoscience used at the workshop: the study of structures, dynamics, and properties of systems in which one or more of the spatial dimensions is nanoscopic (1 to 100 nm), resulting in dynamics and properties that are distinctly different (often in extraordinary and unexpected ways that can be favorably exploited) from both small molecule systems and systems macroscopic in all dimensions. Interesting properties and behaviors can be exploited at this scale. The workshop attempted to identify challenges and opportunities for theory, modeling, and simulation in nanoscience and nanotechnology; investigate the growing and promising role for applied mathematics and computer science in meeting those challenges; and make the case for new investment in theory, modeling, and simulation in nanoscience and nanotechnology by DOE.

Deleted: Bill

The central challenge of the workshop emerged quickly: Within 5 to 10 years, there must be robust tools for the quantitative understanding of structure and dynamics at the nanoscale, without which the scientific community will have missed many scientific opportunities as well as a broad range of nanotechnology applications. Without investment, this promise will not be realized.

Theory and modeling have an impressive history in driving the development of technology. Giant magnetic resistivity, which was predicted by theory before it was experimentally revealed, is now a $6 billion per year business for hard-drive manufacturers.

The 55 attendees at the workshop were split between nanoscience and math/computer science. Written contributions were solicited from the national laboratories. The science was laid out, and the mathematics community responded about what it could contribute. The outputs from the workshop were the laboratories' white papers and a report, which did not make a case for nanoscience but for *theory and modeling* in nanoscience. The report is not a research agenda. It is science-driven and makes a case for investment in these techniques and areas. It makes the following points:

- Nanoscience emerged from the appearance of extraordinary experimental advances during roughly the past 15 years.

- Parallel revolutionary advances were made in theory, modeling, and algorithms in the same period:
  - Density functional theory for electronic structure,
  - Ab initio molecular dynamics (Car-Parrinello),
  - New methods for classical Monte Carlo simulation,
  - New quantum Monte Carlo methods for electronic structure,
  - New mesoscale methods including dissipative particle dynamics and field-theoretic polymer simulation,
  - Fast-multipole approaches,
  - Multigrid algorithms

The report identified a number of fundamental theoretical challenges:

- To bridge electronic through macroscopic length and time scales;
- To determine the essential science of transport mechanisms at the nanoscale;
- To devise theoretical and simulation approaches for nano-interfaces;
- To simulate with reasonable accuracy the optical properties of nanoscale structures and to model nanoscale opto-electronic devices;
- To simulate complex nanostructures involving soft biologically or organically based structures and hard inorganic ones as well as nano-interfaces between hard and soft matter;
- To simulate self-assembly and directed self-assembly;
- To devise theoretical and simulation approaches to quantum coherence, decoherence, and spintronics; and
- To develop self-validating and benchmarking methods.

Simulation is needed to understand these phenomena because sometimes there are no equations to lead the way; rather, equations are developed from simulations. There is a growing and promising role for applied mathematics: Applied mathematics has a strong, recent history of affecting the theory and modeling of molecules and materials through such techniques as mathematical homogenization, fast multipole methods, fast Fourier transforms, sparse linear algebra, multigrid methods, adaptive mesh refinement, and optimization methods. There are some clear and directly relevant opportunities, but some of the mathematics of likely interest (perhaps the most important mathematics of interest) is not fully knowable at the present.

Deleted: , so that

The uncontested consensus of the workshop was that collaborative efforts between applied mathematicians and scientists in nanoscience will yield significant advances central to success. The real role of applied mathematics is to make tractable the problems that are currently impossible: bridging length and time scales, fast algorithms, and optimization and predictability. These are mathematical and programming problems. The potential barriers are:

- Opportunities will be missed if new funding programs in theory, modeling, and simulation in nanoscience do not aggressively encourage highly speculative and risky research.
- Theoretical advances in separate disciplines are converging on this intrinsically multidisciplinary field.
- The traditional separation of experiment, theory, applied mathematics, and computer science and separation by subdiscipline will impede progress if the separation persists.
- A new investment in theory, modeling, and simulation in nanoscience should facilitate the formation of (multidisciplinary) alliances and teams of theorists, computational scientists, applied mathematicians, and computer scientists.

The consensus observations of the workshop were:

- The role of theory, modeling, and simulation in nanoscience is central to the success of the National Nanotechnology Initiative.
- The time is right to increase federal investment in theory, modeling, and simulation in nanoscience.
- Fundamental intellectual and computational challenges remain that must be addressed to achieve the full potential of theory, modeling, and simulation in nanoscience.
- New efforts in applied mathematics and computer science, particularly in collaboration with theorists in nanoscience, will almost certainly play a critical role in meeting those challenges.
- Many opportunities for discovery will be missed if the new tools of theory, modeling, and simulation are not fully exploited to confront the challenges of nanoscience.

SC is in a unique position to build a successful new program in theory and modeling in nanoscience because much of the nation's experimental work in nanoscience is supported by DOE, new nanoscience facilities are being built at DOE national laboratories, DOE supports the core portfolio of applied and numerical mathematics for the nation, and DOE has unique resources and experience in high-performance computing and algorithms.

Wright thanked Stechel and McCurdy for their great efforts.

Dahlburg asked what the experimental advances of nanoscience were so far that justified calling it a new discipline. Stechel said that there is a lot of investment in nanoscale sensor technology. McCurdy pointed to drug-delivery systems, autoassembly of stunning nanostructures, and heterogeneous nanomachines. He noted that the ASCR FY03 budget holds $3 million for theory and modeling in nanoscience, but no matching investment has, as yet, been requested in the Basic Energy Sciences (BES) FY03 budget request. Romine noted that he had met with his BES counterparts and they have discussed the need for this investment. Dahlburg asked whether any SciDAC projects are already doing this theory and modeling. Stechel replied, no; it is too new and complex.

Meza asked what "some of the math is unknowable" means. Stechel answered that it means that it is not known where in mathematics the advances will come from.

McRae asked how one would use nanoscience to design practical catalysts. Stechel said that mixed phases is one way that has a lot of variables, and optimization can be done with a theoretical approach. McRae asked if the workshop participants looked at manufacturing. Stechel said, no, except for recognizing self-assembly as critical.

Kulsrud asked where these mathematicians are going to come from. Stechel replied that, if the money and interesting challenges are there, it is expected that they will come. Wright noted that there is an adequate number of mathematics PhDs in this country, but most of them do not realize that DOE has an applied-mathematics program and interesting problems to work on. Connolly agreed that there are a lot of interesting mathematics problems here. A major one is scalability. Algorithms need to be made scalable. McCurdy noted that a whole section of the workshop was devoted to the scalability of algorithms.

Then van Rosendale asked whether Ford Motors has other nanoscience applications besides catalysis. Stechel said that Ford is mostly watching, but two promising areas are materials response and ubiquitous sensing. McCurdy stated that the MICS office shepherds work that will be essential for work in nanoscience. The ground is very fertile. It will bring great rewards and is ripe for investment.

Wright introduced **Mary Ann Scott** to report on the workshop on high-performance-network planning that was held in August. It is important to consider a strategy for SC network infrastructure now because the number of applications that need network infrastructure is

Deleted: workshop

growing and resources are limited. End-to-end performance will not be solved with added bandwidth; monitoring and other services are needed. With the deployment of grids, advanced services are needed. Our computational leadership is being challenged. Huge amounts of data have to be turned into a community resource. At the same time that all of this is going on, advances are being made in optical networking, and rapid, sweeping changes are occurring in the telecommunication industry in the form of alternative business models and price-performance opportunities.

Researchers and other users are scattered across the country. The method for serving their needs has to start by considering applications; one must know what science will look like in the future. To engage SC program offices in network planning based on a vision for future science, a workshop was convened to bring together visionaries from the multidisciplinary teams spread out across the country in the fields of high-energy nuclear physics, chemical science, fusion, etc., along with network providers and network and middleware researchers. The workshop was to consider

Deleted: entire

- what would be possible in the realm of science if it was unfettered by communication;
- how high-impact application requirements affect network provisioning, network research, and middleware research; and
- what alternative business models make sense in the context of the science scenarios.

The workshop report talks about the role of advanced infrastructure in realizing DOE's science visions, enabling middleware research, enabling network research, a network provisioning model, a governance model for bringing all this together, and a path forward. It can be found at doecollaboratory.pnl.gov/meetings/hpnpw/.

High-performance network infrastructure is critical because much of science is a distributed endeavor. Because of the considerable commonality in the services needed by this distributed endeavor, an infrastructure for distributed science can be defined. Services must allow science to scale in many ways: the number of participants; amount of data; diversity in data use; and ability to combine simulations, measurements, and analyses. As the science paradigm shifts toward large, distributed collaborations, an integrated, advanced infrastructure will be needed that is well beyond today's capabilities. Revolutionary shifts in how science is done can only arise from a well-integrated, widely deployed, and highly capable distributed computing and data infrastructure.

The high-priority middleware research areas are secure control over who does what, information integration and access, coscheduling and quality of service, and effective network caching and computing. The high-priority network research areas are ubiquitous monitoring and measurement infrastructure, high-performance transport protocols, Internet Protocol multicast (which is a fragile technology today), guaranteed performance and delivery, intrusion detection, and distributed systems vs firewalls.

SC needs an integrated network-provisioning model, and resources for high-utilization science and for network research are important. The major challenge is to create an integrated governance model to get this effort up to a production level. DOE needs to continue to analyze requirements to address the services the community needs for each element (production networking, resources for high-utilization science, and resources for network research), to evaluate multiple opportunities, and to develop an integrated roadmap across programs and among infrastructure components.

Wolff commented that all the middleware and network research is interesting and valid and asked where the money is coming from. Scott admitted that resources are limited. Wolff asked whether DOE is looking at using external, uncontrolled resources for DOE program needs. Scott said that some people in the field are looking at this. Wright asked what type of advice ASCAC could offer. Scott replied that the problem is that resources are limited, but the opportunities are out there. At some point in time, things are going to break. Currently, the community needs to recognize what the strategy should be and what resources will be needed. Wolff asked her if the group would like help in prioritizing research opportunities. Scott replied that this is not about the research but about the overall network strategy. Dahlburg requested that Scott present the Subcommittee's 5-year vision at the next ASCAC meeting. Scott agreed to do that.

Deleted: we

Wright observed that Karen Sollins would say that there is a large opportunity for network research and that resources should be made available. Wolff commented that there is much flux in the field. Who knows how long cheap fiber is going to be available? He asked if there was anything that could be done now. Scott replied that networking infrastructure research needs to be done. The current network does not do what some application users need it to do. Wolff asked if there was anything that was so time-critical that DOE needs something from ASCAC before the next meeting. Scott replied that there are opportunities that will not be available indefinitely. The user community needs to work across agencies, pool resources, and look at these opportunities jointly. Wolff asked about the National Terascale Facility. Scott acknowledged that there are some that are talking about such a facility, DOE is open to learning people's views about such a facility, but some things are just not doable.

Wright declared a break at 10:25 a.m. and called the meeting back into session at 10:43 a.m. to hear a report from **Jill Dahlburg** on the Integrated Simulation and Optimization of Fusion Systems (ISOFS).

Dahlburg stated that the world needs an alternative energy source in the next 45 years. The hydrogen economy will need cheap energy. The world is at a crossroads; it will not be using fossil fuels. A demonstration fusion reactor is about 50 years off. Research is aimed at containing plasma at a high temperature and density. The most advanced device is the DIII-D tokamak at Princeton University, which has 100 cosponsors. It is time to make the next step in fusion. It is not *if* but *when*. The basic physics are understood now. More personnel are needed as are "experiments" with low-cost simulators. The ultimate power plant may not be a tokamak, but the technology is ready to make a burning plasma in a tokamak. The pieces of the fusion program need to be brought together. A charge letter was sent to the Fusion Energy Sciences Advisory Committee (FESAC) to complete a roadmap on how to carry out this integration for a demonstration power plant; that roadmap was to be done by the ISOFS Subcommittee.

A workshop was held on May 23, 2002, and a report was written. It proposes a fusion simulation project to advance significantly (1) the ability to predict accurately the behavior of plasma discharges in a toroidal magnetic fusion device at all relevant time and space scales and (2) the capability to carry out virtual experiments of a burning, magnetically confined plasma. The project's goal would be to create a comprehensive set of theoretical fusion models, an architecture for bringing together the disparate physics models, and the requisite algorithms and computational infrastructure.

The next step in the fusion program *is* a burning plasma, probably ITER. A strategy was drawn up at a subsequent FESAC meeting in Austin. It calls for a suite of coupled, self-consistent models at a cost of $400 million over 15 years. This would allow going to a demonstration fusion power plant in 35 years. FESAC recommended that the United States join the ITER effort or go off on its own.

Collaboration is needed among fusion, computer, and computational scientists. Hardware is also needed: a big machine and a network to support collaboration at a distance.

A second workshop, held in San Diego in September, will produce a second report. A major problem is weaving together the differently scaled variables: sources, turbulence, X-MHD, 1.5-D transport, and materials by breaking them across four topics,

- the plasma edge,
- turbulence on a transport timescale,
- global stability, and
- whole-device modeling,

and by developing selected algorithms and software architecture that reflect all of the scaled variables for each of these topics. The deadline for the report is Dec. 1, 2002. The fusion community is seeking funding for this project. The joint subcommittee believes that, given this investment, the resulting knowledge will allow a demonstration plant to be built.

Wright asked if the Subcommittee would like anything from ASCAC. Dahlburg replied, comments from this Committee on the upcoming report before submission for FESAC approval.

Wright asked **Juan Meza** to present the final report from the Biotechnology Subcommittee. The charter of the Subcommittee was to identify the areas on which ASCR's program should target investments to have the maximum impact on the underlying science. Examples of possible areas included specialized facilities for biological computation, basic research in underlying mathematical algorithms, and advanced computer science related to data management. The Subcommittee was also to determine how to couple ASCR-supported research with discipline-specific research carried out by biologists most effectively.

Early on, the Subcommittee found that

- The use of computational science and mathematics is relatively new in most areas of biology.
- Where mathematical and computer-science techniques are used today, the algorithms and tools appear to be adequate, but their use is still rudimentary.
- A growing body of problems will require new mathematical and computational techniques that are not yet available.
- Biological systems are fundamentally different from the types of systems currently studied and modeled; they are inherently noisy, complex, and self-regulating systems.
- Experimental data to validate models is difficult if not impossible to obtain.
- In many cases, it is not understood whether all of the underlying components are being modeled.

The Subcommittee's first recommendation is based on the recognition that computational biology will drive new fields of mathematics and computer science. Therefore, the ASCR program should address these new areas through investments in fundamental mathematical and computer-science algorithms, including some for dealing with stochastic dynamical systems, parameter estimation, multiple time scales, model reduction, computational topology, hierarchy, and data interpretation. Some areas have a clear, quantifiable need for computational power, such as biomolecular simulations. The best argument for new computational capability occurs with molecular dynamics and quantum chemistry. New methods must be developed to address larger systems.

The second recommendation is that the ASCR program should continue to invest in biophysics and biomolecular simulations, which are already having an impact in the underlying science. This area in particular has the greatest (quantifiable) need for more computational capability. The development of new database infrastructures that could be used for computational biology is needed. Many of the databases used in the biological community are heterogeneous, distributed, and error containing. The scientists really need to be able to access, combine, and ask new questions of these biological data sets easily and efficiently.

The third recommendation is that the ASCR program should develop new database and data-management infrastructures that can be used for computational biology. Some possible areas include common data-interchange methods, common ontologies to describe these structures, and automated query access (to ask such questions as, "Is this similar to anything already known?").

The fourth recommendation is that ASCR should help to develop training programs for the next generation of computational biologists. For example, it could develop short courses aimed at mathematicians and computer scientists who want to start research in computational biology, or it could establish Computational-Sciences Graduate Fellowships targeted at computational biology to attract new students into this area. Few people bridge biology and mathematics. Training the next generation of computational biologists who can serve to bridge the gap between biology and computer sciences and allow a common dialogue between the two communities is an urgent need. Without any individuals with expertise crossing the discipline boundaries, little prospect exists that the necessary collaborations can be fostered.

The fifth recommendation is that it is premature to develop any specialized computational facilities in computational biology because the underlying research problems are still too poorly understood. At a previous meeting, ASCAC has expressed serious concerns about the establishment and funding of specialized computational facilities in other areas of science. The case against such facilities for computational biology is equally valid. The Subcommittee believes that the money would be better spent on fundamental mathematical and computer science.

Giles asked if the goal in data management was for DOE to get ahead of the curve. Meza said that some people believed that it was too big a task for ASCR; others said there was a lot of low-hanging fruit that could reward an investment.

Wright called attention to the fact that the committee had previously noted the need for discrete mathematics and serious computer-science algorithms. Meza acknowledged that need and said that that was in the first recommendation of the report but that it was buried under "data interpretation." Some Subcommittee members had pointed out that a lot of discrete mathematics goes into trying to understand the data coming out of experimental work. Wright pointed out that there is more to it than that, mentioning combinatorial techniques in protein folding, where there are exponentially large numbers of ways that things could fit together and where some clever algorithms are used to figure out which ways make more sense. Meza said that the Subcommittee did not mean to imply that data interpretation was *the* only thing that should be pursued.

Kulsrud asked whether biological databases are fundamentally different from other databases. Meza replied that they have a lot of noise and errors in them and that it is problematic how to make them consistent. Houghton said that the amount of data coming out of sequencing is increasing in a unique way. Also, different types of data are coming in from different sources, which lends to the noise and uncertainty.

Wright commented that many people call computational biology bioinformatics, which some people call databases. Biologists have not followed developments in computer science too carefully. She asked what the status was of the Genomes to Life Program. Gary Johnson said that a solicitation had been put out. A call was made for a large, integrated effort; that call got a large response. A call for FY03 is now being crafted. The program managers are in agreement with the recommendations of the subcommittee and are in the process of doing most of them. To bring these communities together, at least one “home” needs to be established for computational biology. The possibility of a facility for computational biology should not be taken off the table. Houghton said that an update of the Genomes to Life report is in the works and calls for taking proteomics to the next step, which will require a lot of computational skill and effort.

Meza said that the Subcommittee believed that it was still too soon to construct specialized computational facilities. Gary Johnson said that there are already specialized computational facilities and specialized machines for computational biology. Some equipment in use now could support computational biology. The “home” probably would contain computers but not *just* computers. Johnson stated that, in his opinion, the last recommendation of the Subcommittee is perhaps poorly cast. It should not be used to take facilities in general off the table. Dahlburg noted that, in other discussions, this Committee has said that specific machines for problems that are not yet understood are a bad idea. Stechel said that the intellectual environment of a “home” is important to foster discussion and understanding across disciplinary boundaries, but care has to be taken when suggesting specialized hardware. Corones asked if the word “specialized” could be changed to “dedicated” to avoid ambiguity. Meza said that he thought that the Subcommittee meant “specialized.” McRae said that the message from computational science in the past has been that building general facilities is much more effective than trying to build special machines with add-ons. There are two types of problems, some of which are well defined (e.g., quantum chemistry and electrodynamics) and suited for facilities, and others are not well defined (e.g., biological processes and metabolic rates). Dahlburg said that she thought that this statement could easily be misinterpreted.

Connolly said that there are two distinct types of problems. Some are well-defined (e.g., quantum chemistry, molecular dynamics, and genomics), and they need large-scale computing. Then there are all the others, where the parameters and processes are unknown. A specialized facility may make sense for the former but not for the latter. Stechel said that the precise meaning may be accurate but the interpretation may be erroneous. That is a legitimate concern. Van Rosendale seconded her statement.

Wright said that the worry is about designing machines for specific applications for problems that are, as yet, undefined. She asked if the statement could alternatively say “hardware” rather than “facilities.” Meza stressed that there is a priority that the Subcommittee is trying to get across here. The fundamental algorithms are what should get the funding.

Wright suggested wordsmithing this recommendation offline and having an e-mail approval in about two weeks. A consensus was reached on that point. She opened the floor to general discussion from the committee.

Connolly asked if a white paper from the committee on the Earth Simulator was still on the table. Wright said that she had a list of assignments to the committee members for white papers. She opened the floor to public comment. There being none, Wright adjourned the meeting at 12:00 p.m.

Prepared by
Frederick M. O’Hara, Jr.
Recording Secretary
October 29, 2002

Revised and submitted by
Margaret Wright
Chair
December 9, 2002

Page 8: [1] Deleted James J. Hack 12/5/2002 10:52:00 AM

The Center falls under the auspices of the AIST, operates in parallel with the Frontier Research System for Global Change, and with that organization provides services to the Research Organization for Information Science and Technology.

Page 8: [2] Deleted James J. Hack 12/5/2002 10:57:00 AM

, although it has not been as reliable as one might want.

Page 8: [3] Deleted James J. Hack 12/5/2002 11:07:00 AM

allows exploration and identification of

Page 8: [4] Deleted James J. Hack 12/5/2002 11:18:00 AM

such as this becomes routine, and the m

Page 12: [5] Deleted mhw 12/9/2002 11:41:00 AM

he National Center for Atmospheric Research (NCAR) has no direct involvement with the Earth Simulator R&D Center, and NCAR is not presently partnered with Earth Simulator R&D Center, although a partnership has been proposed.

The Center's machines are aimed at specific, strategic targets with applications to atmospheric and oceanic science and to solid-earth science. The project team set out a schedule in 1996, and they hit the operating date on the mark. At the current time, the U.S. computational-science community does not understand the organization of the project and how it works. The Center falls under the auspices of the AIST, operates in parallel with the Frontier Research System for Global Change, and with that organization provides services to the Research Organization for Information Science and Technology. The Center's budget is split among several agencies that have different priorities.

The Earth Simulator has a MIMD-type [multiple-instruction/multiple-data-type] distributed memory and a parallel architecture. The goal was to run NCAR-type programs 1000 times faster. To do that, they would need an efficiency boost of 12.5% and a peak performance of 40 Tflop/s. A major factor in gaining that efficiency is the crossbar network that has 128 switches, each of which is connected to all of the 640 processor nodes. It uses 29,000 km of cable to make those connections. The Center is collocated with the Frontier Research System for Global Change in Yokohama, and its infrastructure support cost $50 million. The machines are located in a 50- by 65-m building that is electromagnetically shielded to prevent interference; light is brought in by mirrors. It is a remarkable engineering feat.

The system performance is what finally generated interest in the United States. The machine attained a LINPACK benchmark of 35.86

Page 12: [6] Deleted mhw 12/9/2002 11:41:00 AM

% efficiency, although it has not been as reliable as one might want.

The AFES (Atmospheric General-Circulation Model for the Earth Simulator), developed at the Center for Climate System Research at the University of Tokyo and at the National Institute for Environmental Studies, was an important target application for the Earth Simulator. It is a global spectral model running with full physics with an implementation that allows for coupling with oceanic and other component models. They claim to have a sustained execution rate of 14.5 Tflop/s with 320 nodes and 7.6 Tflop/s with 160 nodes.

An analysis of the reported AFES system performance indicates that the conventional high-resolution configuration produces about 708×10^6 spatial degrees of freedom, about

3000 times greater than the typical climate configurations, which is about 100,000 times the typical operation counts from the way these models are normally run. The turnaround rate for a reference model gives 5200 times real time at 10 Gflop/s, and a comparable AFES configuration gives about 70 times real time at 14.5 Tflop/s. They are likely subcycling the dynamics, exploiting a high-performance Legendre transform library, and ignoring operation-count savings from hemispheric symmetry. If that is true, their actual turnaround may be even shorter. This analysis does not indicate that this machine is a viable production configuration without significant algorithm changes. It *does* provide unmatched high-resolution modeling capability, which allows exploration and identification of scientific problems.