Final Minutes
Advanced Scientific Computing Advisory Committee Meeting
Mar. 3–4, 2009, American Geophysical Union, Washington, D.C.

ASCAC members present:

F. Ronald Bailey	Anthony J. G. Hey
Jacqueline H. Chen	John W. Negele
David J. Galas	Rick L. Stevens
Roscoe C. Giles, Chair	Larry L. Smarr
James J. Hack	Robert G. Voigt

ASCAC members absent:

Marsha J. Berger	Horst D. Simon
Thomas A. Manteuffel	

ASCAC members who joined portions of the discussion by telephone:

Victoria A. White	Thomas Zacharia

Also participating:

Melea F. Baker, Office of Advanced Scientific Computing Research, Office of Science, USDOE
Jon Bashor, Computing Sciences Division, Lawrence Berkeley National Laboratory
Peter Beckman, Vice President of Engineering, Turbolinux
Arthur Bland, Project Director, Leadership Computing Facility, Oak Ridge National Laboratory
Charlie Catlett, Chief Information Officer, Argonne National Laboratory
Christine A. Chalk, ASCAC Designated Federal Officer
Paul Steve Cotter, Department Head, ESnet, Lawrence Berkeley National Laboratory
Brent Draney, National Energy Research Scientific Computing Center, Lawrence Berkeley National Laboratory
Steven Hammond, Director, Materials Sciences Center, National Renewable Energy Laboratory
Barbara Helland, Office of Advanced Scientific Computing Research, Office of Science, USDOE
Daniel Hitchcock, Office of Advanced Scientific Computing Research, Office of Science, USDOE
Patricia Hoffman, Principal Deputy Assistant Secretary, Office of Electricity Delivery and Energy Reliability, USDOE
Douglas Kothe, Director of Science, National Center for Computational Sciences, Oak Ridge National Laboratory
Paul Messina, Retired Director, Center for Advanced Computing Research, California Institute of Technology
Juan Meza, Department Head, High Performance Computing Research, Lawrence Berkeley National Laboratory
Frederick M. O'Hara, Jr., ASCAC Recording Secretary
Philip Overholt, Office of Electricity Delivery and Energy Reliability, USDOE

Dan Reed, Director, Scalable/Multicore Systems, Microsoft
Thomas Schulthess, Director, Swiss National Supercomputing Centre
Michael R. Strayer, Associate Director, Office of Advanced Scientific Computing Research, Office of Science, USDOE
Richard Strelitz, Computer, Computational, and Statistical Sciences Division, Los Alamos National Laboratory
Katherine Yellick, Director, National Energy Research Scientific Computing Center, Lawrence Berkeley National Laboratory

About 45 others were in attendance in the course of the two-day meeting.

Tuesday, March 3, 2009
Morning Session

Chairman **Rosco Giles** called the meeting to order at 9:05 a.m., welcomed the new members to the Committee, and asked each member introduce himself or herself. The next meeting will be held at the American Geophysical Union on August 11-12, 2009; and the meeting after that will be on November 3-4 in Oak Ridge, Tennessee. Public comments were to be allowed after each presentation at this meeting.

Michael Strayer was asked to give an update on the status of the Office of Advanced Scientific Computing Research (ASCR).

The vision for the Office is to be first in computational science and best in class in advancing science and technology through modeling and simulation with this vision being reflected in facilities, enabling technologies, and computational partnerships.

ASCR has had great success during the past 12 months in delivering science:

- Of the top ten computational science accomplishments for the year, six were Scientific Discovery through Advanced Computing (SciDAC) projects.
- The DCA++ code won the 2008 Gordon Bell Prize.
- An algorithm for unlocking new frontiers in density functional theory (DFT) won a 2008 Gordon Bell Special Prize. The two leadership-class facilities (LCFs) swept the high-performance computing (HPC) benchmark.
- Bert Debusschere won the Presidential Early Career Award for Scientists and Engineers (PECASE) for his Stochastic Dynamical Systems that introduced rigorous, mathematical methods capturing stochastic uncertainties in computational biology and providing a framework for simulation-based discovery.
- James Sethian was elected to the National Academy of Engineering for the development of efficient methods of tracking moving interfaces.

One of the outstanding accomplishments during the past decade is the compendium of software developed, including programming models, development/performance tools, libraries, system software, and visualization and data analytics.

The ASCR Base Program issued three solicitations on (1) Multiscale Mathematics and Optimization for Complex Systems, which garnered 426 letters of intent and expects 114 full proposals; (2) Petascale Tools, which got 97 proposals representing 34 projects; and (3) Next-Generation Networking for Science, which received 40 proposals.

SciDAC had a committee of visitors (COV) that recommended ongoing reviews. This year's reviews include those for (1) science applications (SAs) and science application partnerships (SAPs), (2) centers for enabling technology (CETs) and institutes, and (3) distributed systems.

The next allocation has begun for the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) Program. Climate research, materials science, astrophysics, and biological sciences have the largest percentage of the available hours.

The Cray XT5 at Oak Ridge National Laboratory (ORNL) is running jobs at a sustained petaflop (Pf) or more with a peak performance of 1,382 Pf, on 150,176 cores, with 300 TB of system memory, and with a disk bandwidth of 240 GB/s. The other LCF is Argonne's IBM Blue Gene/P, which runs at 556 TF. The other facility ASCR operates is the National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory (LBNL). It has several machines on which the allocation of time is programmatic. It currently has a Cray XT4 upgraded to 350 teraflops (TF). Proposals are currently being reviewed for the next upgrade, which will include a significant increase in power.

The Energy Sciences Network (ESnet) has recently undergone a transformation to a 40 Gb/s core with multiple redundancies and metropolitan area networks (MANs) to facilitate the transfer of data from CERN [Conseil Européen pour la Recherche Nucléaire, now the Organisation Européenne pour la Recherche Nucléaire]. It will be upgraded to 100 Gb/λ, a substantial increase in capacity.

ASCR's FY08 appropriation was $351 million; its FY09 presidential request was $369 million; the House markup provides $379 million; and its Senate markup provides $369 million.

ASCR hired an applied math program manager in the Research Division; an offer has been made for a data and visualization program manager in Computer Science; and positions will be posted soon for a computer scientist program manager and a collaboratories program manager. These actions represent a vigorous build out of staff.

What computing will be needed to enable the grand challenges in science in the next few years? ASCR is delivering Petascale Science Today, and it has been a real struggle. 80% of LCF time goes to INCITE. ASCR also works with Pioneer Applications to deliver scientific results from day one; the software needed for these applications needs to be determined. ASCR continues to nurture world-class mathematics and computer science research efforts and applications critical to DOE missions through SciDAC. It also is looking at how to realize the promise of the extreme scale, which entails visualizing the big facility, justifying it with science, socializing it, and building it if that is the consensus of the community.

ASCR held a number of workshops to identify the barriers to extending computation to the extreme scale. One was on scientific challenges for understanding the quantum universe, and another was held to identify the needs for ESnet.

Another barrier is posed by the multicore processors, where there are multiple parallel general-purpose processors (GPPs) and multiple application-specific processors (ASPs).

There have been seminal workshops on cybersecurity. And ASCR has been reaching out to the DOE community. Simulation and modeling at the highest scale can help the Office of Science (SC) and other DOE offices reach their goals and objectives. It will also engage other agencies in multicore computing [e.g., the National Nuclear Security Administration (NNSA), Europeans, and Japanese].

ASCR's research strategy is to broaden its research beyond SciDAC. SciDAC delivers computational tools and techniques to advance DOE science through modeling and simulation by collaborative teams to produce great science achievements. The community needs to discover methods and algorithms to fully describe understanding of nature over vast scales of time and space. That goal will require algorithms for solving complex, mission-relevant science problems. The mathematics that underlies prediction will need to be understood. Computer science will need to facilitate the use of emerging Leadership-scale computing resources. The implications of new computing architectures will need to be understood. Tools to extract meaningful information from petabyte data sets, including analytics, will need to be developed. An understanding will be needed of the management and performance of federated 100 Gb/s networks, enabling geographically distributed research teams to collaborate.

ASCR's facilities strategy is to provide the tools for high-end computing through high-performance production computing and leadership-class computing. An upgrade path for these facilities will be needed. Getting there will require investing in the future and linking it all together with ESnet.

Bailey asked about the follow-through with the recommendations made by past subcommittees. Strayer answered that one recommendation was to engage the Office of Fusion Energy Sciences (FES) in a follow-up to the Tokamak's simulation. Currently, there is no director of that office, and it has no budget to carry out that project. The effort needs to be managed as a capital project, and FES needs to take the lead. It is hoped that there will be activity in a year or so. The network report's recommendations produced networking research at the national laboratories. An early buildout of ESnet to 100 GB is being undertaken. Several more MANs have also been finished. A COV looked at SciDAC and vigorously followed through on its recommendations.

Smarr observed that, as the components of ESnet are put together, the end-to-end capabilities of the users are hindered by the bandwidth gap between the network and the campus. Middleware may produce a strongly coupled system. He asked how much DOE was doing on such connectivity for the end users. Strayer replied that computational endstations are being developed to allow such conductivity. Once the continuing resolution is overcome, DOE will begin that work. He introduced **Daniel Hitchcock** to present more information. Hitchcock said that DOE does not tell campuses what to do; it leaves that to Internet-2. Sending a petabyte of data to a campus research location is essentially a denial-of-service attack on the researcher's desktop. DOE is a looking at ways to get around that by putting the data where the campus researcher can get to it. It is also looking at hybrid architectures to put the analytic tools in the hands of the researchers. Smarr noted that collaboration is also needed. Without that high-speed capability, opportunities for collaborations will be ruled out unnecessarily. Hitchcock replied that such connectivity has already been installed for the Large Hadron Collider (LHC) campuses, and doing that for others is being looked at. The program scientists determine what is needed, but there are people who are potential users who are not included in these workshops. Smarr said that best practices are shared among users.

Giles pointed out that the leadership capabilities of ASCR effect the scientific output across DOE, and asked what the Committee can do to foster that leadership. Strayer answered that the government is constrained about how it can get advice. It gets advice through charges. Giles mused that perhaps the government could overhear what is said and formulate a charge. Stevens pointed out that when one goes outside SC, there are offices that may not have their own advisory committees for ASCAC to work with. Strayer and said that one could put together a charge for this Committee to work with such an office. It is clear that the other advisory committees in SC assess and weigh the activities of their respective offices.

Negele observed that, if one looks at computational science, there is an alignment of the interests of the Administration and of the Secretary of Energy. Computational science could have a leadership role in the solution to their problems.

Bailey asked if there were an anticipated change in SciDAC. Strayer answered that something might be started that is SciDAC-like to boost performance and efficiency.

Bland asked what the process was for filling the positions of Director of SC and Under Secretary for Science. Strayer said that the filling of those positions is the prerogative of the White House, and that President Obama has teams to fill those and thousands of other such positions.

A break was declared at 10:21 a.m. The meeting was called back into session at 10:45 a.m.

Giles noted that the "balance" report is still progressing under the revised charge. That report will give the Office some good advice. The computing COV will be chaired by Dona Crawford, and a progress report will be given in August.

Peter Beckman was asked to report on improving high-performance computing software.

The open-source community provides almost all of the HPC software on the world's big computers today. The types of software found on the extreme-scale computers are Linux Operating System, libc, Perl, PAPI, UPC, OpenMPI, ScaLAPACK, VisIt, and many more. This is quite a diversity. "The stack" is what will be remembered about this era of computing.

Many companies are hosting open-source HPC software stacks for small Linux clusters. This practice has not been successful at the large scale. For some markets, a closed-source business model continues to work well [e.g., for single-node optimized math libraries and compilers; debuggers for small clusters; and some queuing systems, parallel file systems, and hierarchical storage management (HSM) systems], although IBM is open sourcing essentially all of its software on Blue Gene/P.

However, the largest-scale systems are becoming more complex, with designs supported by large consortia. The software community has responded slowly to this model change. Significant architectural changes are occurring, and the software must dramatically change. Finally, the ad hoc community coordinates poorly, both with other software components and with the vendors; it needs improved development and coordination.

In times past, one person could develop the whole thing, but to build more capable, efficient machines, one must go to globally distributed design/build teams. An example of such a team is the Japanese "> 10 Pf" supercomputer to be built in 2012. It will be more open than the Earth Simulator. Another example is the Partnership for Advanced Computing in Europe, which is part of the European roadmap for computational science infrastructure.

Power and cost have taken off, but software has not responded. All components need to double their concurrencies every six months.

Multicore comes in a wide variety, and these types of architecture for parallelism are highly divergent. Accelerators are being put out to boost compute speed. The software stack will need to change to accommodate these accelerators.

Multicoring and 3D packaging (putting memory on the CPU) are changing the paradigms, and one must ask how system software will change. Operating systems and software libraries have to reflect the storage mechanism. Today, there are hardware features that are uncoordinated with software development. Today's software does not have power management, multicore tools, math libraries, advanced memory models, etc. Only basic acceptance-test software is delivered with the platform; all of the other software has to be ported to the machine later. Vendors often "snapshot" key open-source components and then deliver a stale code branch. Community codes are unprepared for this sea change in architectures. There is no global evaluation of key missing components.

The International Exascale Software Project (IESP) workshops are designed to improve the world's simulation and modeling capability by improving the coordination and development of the HPC software environment. They will build a plan for how the international community can join together to improve software available for high-end systems over the next 2 to 10 years. DOE, NSF, and the European Union have committed their support for the workshops. The first workshop will be in Santa Fe, April 7–8; white papers are encouraged.

This effort needs to be an international collaboration because of the scale of investment; because of the need for international input on requirements; and because Europeans, Asians, and others are working on their own software that should be part of a larger vision for HPC.

Apache creates a foundation for open, collaborative software development projects by supplying hardware, communication, and business infrastructure. There is a way for incubator projects to become Apache projects. There are 800 "committers"; the effort is very cohesive.

A plan could include working with vendors to create the HPC equivalent to the International Technology Roadmap for Semiconductors; getting the community working on software before machine becomes available; developing a community-proposed unified roadmap for exascale software; identifying the missing components for future architectures and working up a plan to address them; developing models for working more closely with vendors; identifying key application areas to drive development; producing community software development models; and establishing the funding and organizational models (e.g., Apache).

The hope is to improve the capability of computational science; build and strengthen international collaborations; build and improve R&D program developing new programming models and tools addressing extreme scale; develop the open source guided by a roadmap; conduct joint programs in education and training for the next generation of computational scientists; and encourage vendor engagement and coordination for more capable software supporting exascale science.

There will be three workshops during the next year. A broad engagement by the community is being worked on. There will probably be some initial reports in summer

2009. An immediate payoff is also being planned for. Activities are being described on www.exascale.org.

Smarr asked if they had looked at dividing up the writing of code by country. Beckman replied that issues will be surveyed at the first workshop; certain countries have developed different expertises. Smarr asked what the Field Programmable Gate Array High Performance Computing Alliance (FHPCA) was doing. Beckman answered that interest in it had waned.

Voigt asked what was being done in training. Beckman said that finding great graduate students is very hard. The open-source community needs to be organized better. Bailey pointed out that one concern is getting the application people involved. Beckman said that SciDAC users had been surveyed, and most are taking a wait-and-see attitude about open source. They need to be assured that all of the stack will be made available under the model.

Giles asked how international open source can be. Beckman replied that the community is already international. Many vendors are U.S. vendors. The community needs to be pulled together. The world community generally does not see the Japanese additions to the stack, and that needs to be addressed.

Negele suggested that the quantum chromodynamics (QCD) group would be a good one to nurture. Beckman agreed that the QCD groups has developed a lot of the needed components, not just the science.

Strelitz asked rhetorically, why has VisIt triumphed? Because it works with the users. Beckman agreed. A new graphics supercomputer is being installed, and the code team of VisIt has worked closely with the users.

A break for lunch was declared 11:45 a.m.

Tuesday, March 3, 2009
Afternoon Session

The meeting was called back into session at 1:10 p.m. **Arthur Bland** was asked to describe the petascale computing being done at ORNL.

Only 41 days after the assembly of a totally new 150,000-core system, the Jaguar had two real applications running at more than 1 Pf and had won two of the four HPC Challenge awards

Thomas Schulthess and his team received the 2008 Gordon Bell Prize by running an application at 1.352 Pf on the Jaguar. Other finalists were invited to run their codes on Jaguar. Those who did ran at the highest performances they had ever achieved.

Science Applications are Scaling on Jaguar, such as the DCA++ in materials, LSMS in materials, SPECFEM3D in seismology, WRF in weather, and POP in climate.

Jaguar had a four-phase acceptance test in hardware checkout, functionality, performance, and stability. The acceptance test was reviewed by a panel of peers. Since completing acceptance, it has been doing early-science projects on climate change, bioenergy, solar, energy storage, energy transmission, combustion, fusion, and nuclear energy, astrophysics, chemistry, nuclear physics, and turbulence, all getting early access to get their programs up and running. More than 20 projects are running today, each needing a petascale resource to achieve its goals. They are run by more than 100 of the best computational scientists from universities, national laboratories, government agencies, and private industry around the world. More than 500 million hours have been allocated for this early-science period, twice the entire INCITE 2008 allocation. Jaguar can deliver 3.6 million hours daily.

Some of the projects investigate chemical nanoscience at the petascale; direct numerical simulation of diesel jet flame stabilization at high pressure; the Hubbard model with disorder; high-resolution climate explorations with a new scalable dynamical core; charge patching of electronic structures and charge transports of organic and organic/inorganic mixed nanostructures; quantum Monte Carlo calculation of the energetics, thermodynamics, and structure of water and ice; decadal predictive skill with the community climate system model; lignocellulosic biomass deconstruction. A lot of the results coming out of the machine are enabling breakthrough science.

DOE's broad range of science challenges demand a balanced, scalable, general-purpose system. Jaguar is the world's most powerful computer designed for science from the ground up. Its system memory is 3 times larger than that of the next biggest in the world. A lot of studies were done on the science to be done on the machine and the needs of that science. This system's nodes have a 16-GB DDR2-800 memory, 25.6 GB/s direct-connect memory bandwidth, 6.4 GB/s direct-connect hypertransport, 8 cores, and 1 SeaStar2+. A node is 16 GB of memory operating at 76.3 Gf/s; a blade is 64 GB operating at 294 Gf/s; a rack is 1.54 TB operating at 7.06 TF/s; and the system is 300 TB operating at 1382 TF/s. 192 I/O nodes are spread throughout the system to make sure there are no hot spots. Each is connected to the disk farm.

A Center-wide file system ("Spider") is used to provide a shared, parallel file system for all systems at a demonstrated bandwidth of more than 200 GB/s. It has 192 storage servers and more than 3,000 InfiniBand ports. This is a scalable architecture. It is undergoing system checkout with deployment expected in summer 2009.

The combination of the XT5, XT4, and Spider with a login cluster completes Jaguar. Completing the simulation environment to meet the science requirements involves a data archive, an application development cluster, an end-to-end cluster, a remote visualization cluster, and the Everest Powerwall, all tied into the Scalable Input-Output Network (SION).

In 2003, a 40,000 ft^2 LEED [Leadership in Energy and Environmental Design] certified computer center was built. Today, ORNL's facility is among the most efficient data centers in the country. Its power utilization efficiency is 1.3; the national average is 1.83 (lower is better). The 13,800-V power in the building saves on transmission losses. The 480-V power run directly to the cabinets saves $1 million in installation costs. High-efficiency power supplies were used in the cabinets. A flywheel-based uninterruptable power system (UPS) is used for highest efficiency (99%). On the cooling side, a liquid-cooled design removes heat to the liquid before it leaves the cabinet, saving about 900 KW of power just in air movement and 2,500 ft^2 of floor space.

Lessons learned include:

- The design must be driven by science requirements.
- One must have a detailed plan. These are big projects, and project management is critical.
- When the inevitable problems occur, go back to the science requirements.
- Preparation of the applications on the XT4 started in the spring and summer before acceptance.

- The liaison model of supporting the INCITE projects is the *right* model.
- Applications are the long-leadtime item.

To summarize, DOE-SC and ASCR are leading in modeling and simulation for critical international problems. Important, time-critical problems in energy assurance, climate change, materials science, and basic science are running today, solving problems that were unapproachable just a few years ago. Early engagement with the vendors is critical.

Where do we go from here? A Mission Need Statement (CD-0) was signed by Raymond Orbach, calling for two systems to provide capability and provide architectural diversity.

Bailey asked how general-purpose Bland saw future machines being. Bland replied that he saw a variety of applications and needs. Peak is not the right measure; bandwidth and memory are the correct measures. Machines will have to go to some sort of acceleration and memory-bandwidth expansion.

Smarr asked what the source of the electricity was. Bland replied that the Tennessee Valley Authority has coal, nuclear, and hydro power. Smarr asked what portion of the electricity does not produce CO_2; that will be a metric of the future. Computers are becoming the largest producers of greenhouse gases. ORNL is in a great position to be a leader in using green energy.

Hey asked what real highlights could be put forward to indicate that the science being enabled is transformative, not incremental. Kothe replied that there will be huge steps in materials, catalysts by design, superconductivity, and nuclear energy. Strayer added that there are teams out now looking for examples of transformative research.

Thomas Schulthess was asked to describe the DCA++ project.

In superconductivity, the metals experience a change of state as they go from normal to super conductors. It took about 50 years to develop a theory to describe this phenomenon. However, 20 years later, there is no predictive power for the transition temperature (T_c) in known materials, no predictive power for the design of new superconducting materials, no explanation for other unusual properties of cuprates, only partial consensus on which material aspects are essential for high-T_c superconductivity, and no controlled solution for any proposed models.

Today, inhomogeneities in the material (Cooper pairs) are being looked at with the 2D Hubbard Model and the dynamical cluster approximation (DCA)–quantum Monte Carlo (QMC) method.

With DCA, one can integrate the remaining degrees of freedom out of the bulk lattice, leaving the inhomogeneic cluster in reciprocal space. One can then transfer the clusters back to real space with the DCA and solve the many-body problem on the cluster with QMC. The Hirsch-Fye QMC (HF-QMC) was used for the quantum-cluster solver.

Because the numbers are very large, the computation is very inefficient and slow. Delayed updates produce huge improvements in the time to solution.

If some of the code (the HF-QMC) is run in single precision while the rest of the code (specifically the cluster mapping) is run in double precision, the same results are obtained as when all the code is run in double precision. Again, time to solution is improved by the use of this mixed precision.

As the QMC computations are scaled up to a couple of thousand processors, the speedup is almost ideal. The computations scale well up to 50,000 cores, but only weakly beyond that.

Sustainable performance increases almost linearly to 409 TF/s on almost 50,000 cores on the Cray XT4 and to 1.35 Pf/s on 150,000 cores of the Cray XT5, where the introduction of mixed precision produced a speedup factor of 1.9 (almost twice as fast). This performance represents a capability enhancement of more than a factor of 1000 since 2004, a factor of 300 being produced by peak flops/s and a factor of 20 by algorithms. In the future, a new class of QMC algorithms, continuous-time QMC, is expected.

These methods and computational capabilities allow one to take a deep look into the mechanisms of high-Tc superconductivity, producing simulations of the superconducting transition in a model without phonons and showing that the "glue" of Cooper pairs is caused by spin fluctuations. DCA++ optimally mapped the DCA/QMC method onto today's hardware architectures and allowed the use of the HF-QMC with delayed updates and a mixed-precision acceleration to study disorder and nanoscale inhomogeneities in a highly scalable manner, based on a C++/STL generic programming model. These new algorithms, new computers, and innovative software design allow the solution of real simulation problems of high-temperature superconductivity.

The team, made up of physicists, software scientists, computational mathematicians, computer scientists, the center, and the hardware vendor, was not tightly managed; it could go after the solution of the problem.

Smarr noted that the pairing was attributed to the spin fluctuation, and asked what experimental tests verify that assertion. Schulthess said one can develop a form for the pairing strength. The spin susceptibility is being measured at the Spallation Neutron Source (SNS) and compared to magnetic properties. Maier at al. began this work and published it in 2007. One can also compare transition temperature with theoretical predictions.

Stevens asked if they had software support in the mixed precision. Bland replied that he did not know if they are using the libraries today. The work moves from application to application.

Juan Meza was asked to describe the linearly scaling 3-D fragment method for large-scale electronic-structure calculations. There is a diverse team of material scientists, physicists, mathematicians, and computer scientists.

About 6 years ago, nanoscience projects started being funded by SciDAC. Reformulating the question with better algorithms was found to often improve solver performance significantly. More than 100 papers came out of this project, including a new linear scaling 3D fragment method that took a divide-and-conquer approach to solving large systems, was able to model systems with more than 36,000 atoms with excellent scaling up to 160,000 processors, and received the Gordon Bell Special Award in 2008.

Nanostructures have different electronic structures than bulk materials do. To model them, one needs 1,000- to 100,000-atom systems and $O(N)$ computational methods. Also, algorithms were needed that ran well on parallel supercomputers.

Quantum mechanical calculations require N coefficients to describe one wavefunction, with N being proportional to the number of wavefunctions, M. All of the wavefunctions must be orthogonalized against each other, which requires, M^2 work, thereby producing an algorithm that scales as the third power of N. The repeated calculation of these orthogonal wavefunctions makes the computation expensive.

Previous work on linear-scaling DFT methods followed three main approaches: the localized-orbital method, truncated-density-matrix method, and divide-and-conquer method, all with advantages and disadvantages. Most approaches use the localized-orbital or truncated-density matrix method with lots of local minima. None of them scales to tens of thousands of processors.

The Linearly Scaling 3 Dimensional Fragment (LS3DF) method is a novel divide-and-conquer scheme with a new approach for patching the fragments together with no spatial partition functions. Rather, it uses overlapping positive and negative fragments. This new approach minimizes artificial boundary effects and is massively parallelizable.

In a 1D LS3DF example, a (1D) wavefunction is divided into fragments, which are then broken into smaller fragments that subsequently need much less computation. One can do the same thing to a 3D quantum dot. The fragments are broken into sets of groups, which are then all added up in the end.

The major computational components of the LS3DF method are the generation of fragment potentials, VF; the solution for fragment wavefunctions; the computation of total charge density; and the solution of the global Poisson equation. The most time-consuming part of an LS3DF calculation is for the fragment wavefunctions. Therefore, the standalone PEtot code was modified with the planewave pseudopotential, and BLAS3 was taken advantage of with an all-band algorithm. If the PEtot efficiency is greater than 50% for large systems, then it will be 30 to 40% for the fragments.

This method is a variational formalism, so the mathematics are very sound. The division into fragments is done automatically and is based on an atom's spatial locations. The typical large fragments have about 100 atoms, and the small fragments have about 20 atoms. The processors are divided into *M* groups, with *Np* set to set to 16 to 128 cores. *M* is between 100 and 10,000. Each processor group is assigned the same number of fragments, typically between 8 and 100.

The wavefunction calculation scales very well with the number of processors; scaling is not a big problem in data movement. The LS3DF speeds up almost ideally. For a typical system (3456 atoms), 1 hour is needed for a converged result. This is very reasonable. LS3DF is 400 times faster than PEtot. Near-perfect speedup is observed across a wide variety of systems (weak scaling) for Jaguar (Cray XT4), Franklin (dual core), and Intrepid. For a particular system, a ZnTeO alloy, one sees weak scaling for the calculations.

When the code was first run on intragroup fast-Fourier-transform communication, it was very inefficient; tuning that performance got 50% inside-group FFT and 50% inside-group double-precision general matrix multiply (DGEMM).

The code ran at 135 TF/s on 36,864 processors of the quad-core Cray XT4 Franklin at NERSC, a 40% efficiency. It ran at 224 TF/s on 163,840 processors of the BlueGene/P Intrepid at the Argonne LCF, a 40% efficiency. And it ran at 442 TF/s on 147,456 processors of the Cray XT5 Jaguar at ORNL, a 33% efficiency (the file system was not working). For the largest physical system (36,000 atoms), LS3DF is 1000 times faster than direct DFT codes.

In terms of the mathematics, the same convergent result of LS3DF is similar to that of the direct local density approximation (LDA) method. LS3DF accuracy is known to be determined by fragment size. More fragments lead to greater speed and to larger errors. A comparison to a direct LDA calculation with an 8-atom 1×1×1 fragment size division showed that the total energy error is 3 meV/atom or about 0.1 kcal/mol. These are very small errors that will not contribute significantly to the overall error.

In terms of the science, doping produces a stepping stone for the electron to get up to the next energy level. The questions were: Is there really a gap? And is it optically forbidden? The answers were: Yes, there is a gap, and O-induced states are very localized.

A problem was encountered: Certain nanostructures have dipole moments (experimentally observed) where theory says that no dipole should exist. The inequality comes from shape-dependent self-screening.

In summary, LS3DF scales linearly to more than 160,000 processors. It reached 440 TF/s. It runs on different platforms with little retuning. The numerical results are the same as those of a direct DFT based on an $O(N^3)$ algorithm, but at only $O(N)$ computational cost. LS3DF can be used to compute electronic structures for systems with more than 10,000 atoms with total energy and forces in 1 to 2 hours. It can be 1000 times faster than $O(N^3)$ direct DFT calculations. It enables one to produce new scientific results in predicting the efficiency of proposed new solar-cell materials

Negele asked how they dealt with orthogonality in the original problem. Meza replied that they used explicit orthogonalization of the wave functions.

Charlie Catlett was asked to speak on a scientific R&D approach to cybersecurity.

A series of summits was held in late 2007, where DOE laboratory directors and chief information officers discussed improvements and strategies in cybersecurity. Working groups were put together, one of which was to look at opportunities for SC to focus research efforts on improving cybersecurity. Three open workshops and one closed workshop were held in 2008. Authors from multiple DOE laboratories, other agencies, industry, and universities produced a report, *A Scientific Research and Development Approach to Cyber Security*, vetting it with industry leaders and with more than 20 agencies. The consensus of agency representatives was that the research plan outlined in the report would be extremely useful for their missions and that the work was both valuable and well-suited to DOE and SC needs and strengths. Industry feedback was consistent with multiple federal advisory reports that federal investment in cybersecurity research must be increased as a high priority. The National Science Foundation (NSF) spends about $300 million a year on cybersecurity R&D.

DOE has unique requirements: a national-scale civilian and classified infrastructure, assets, and programs and international science communities. It also has unique strengths: national laboratories with strong multidisciplinary programs and rich academic and industry collaborations and mathematics and computational-science programs coupled with leadership-class facilities. These strengths give DOE a good set of resources and experiences to deal with these problems.

Today, mathematics and computational science are generally untapped by cybersecurity efforts with only mathematics-based intrusion detection and limited use of modeling and simulation being exploited. Computer architecture, with its inherent trust among components, and treatment of data as passive assets, is anachronistic. Federal cybersecurity policy is also reactive and tactical; providing defense only against specific, previous tactics and using an awkward underlying model of "layered defense."

There is some urgency to the cybersecurity problem. Vulnerabilities are discovered almost simultaneously with the first attack exploiting the vulnerability. The working group focused on the need for fundamental, architectural rather than incremental improvements in cybersecurity. It looked at three focus areas:

- Mathematics for predictive awareness,
- Self-protective data and software, and
- Platforms that are trustworthy systems constructed from untrusted components.

These are three high-priority, high-return areas of focus for cybersecurity R&D that could be addressed by a DOE program funded at $30 to 50 million per year.

DOE, with its national-laboratory assets and their accumulated expertise is uniquely positioned to advance the field of cybersecurity.

It is recommended that DOE start thinking about the mathematics and use cybersecurity as a computational science and engineering challenge, leveraging INCITE. DOE's computer-science and computer-architecture programs should explore novel approaches to active data. DOE's system-software and architecture programs should pursue new operating-system, distributed-application, and platform architectures harnessing state-of-the-art capabilities, such as multicore.

SciDAC-scale multidisciplinary teams should be formed to address the problems of cybersecurity. An "X-Prize" style of approach might set clear targets and inspire broad competition to engage industry and to facilitate many "failures" to find the diamonds in the rough. ASCR should conduct proactive research collaborations with industry, other agencies (e.g., NSF), and other DOE programs.

Leadership computing, data analysis, and related infrastructure should be harnessed to support computational science as well as nearer-term needs, such as sensor data analysis and intensive software-vulnerability testing.

Negele asked if it would be unacceptable to have the same security protections at the different DOE facilities. Catlett replied that this is a sensible goal for similar facilities and risk levels and that this would be best done by simplifying DOE guidelines to instruct the national laboratories to implement National Institute of Standards and Technology (NIST) guidelines. That is where we want to be. A three-level form of protection has been pushed.

Hey said that he did not see any software engineering in the list presented. Catlett replied that that came up in several areas and is in the full report.

Smarr asked whether, if one sets the goals out in the open, one is giving the hackers lead time to overcome the countermeasures. Catlett responded that, if one wanted to do this at a classified level, one would eliminate a significant fraction of the expertise and creativity needed to adequately address the challenges. It is also generally the case that the application of the technology is what is sensitive rather than the technology itself.

Yellick underscored the need to recognize the different levels of risk and types of activities at the laboratories in any discussions regarding normalization of security protections.

Stevens asked what the strategy should be to have the ASCR program to participate. Strayer replied that the Office has startup money for FY09 for cybersecurity. The out-years cannot be discussed because of the embargo on future-year budgets. The Office has shared outstanding ideas with the community.

Barbara Helland was asked to discuss the ASCR facility-allocation policy.

Under the previous policy (until January 2009), up to 10% of allocable hours at all facilities were distributed by the Director of SC (SC-1); up to 10% went to the facility director; 70% of NERSC time went for SC mission computing; and 5% at Pacific Northwest National Laboratory (PNNL), 10% at NERSC, and 80% at the Oak Ridge LCF and Argonne LCF were used for the INCITE program.

The INCITE COV recommended that: The selection processes for leadership-class and DOE capability-class computing should be separated. A significant portion, but less than half, of INCITE computational resources should be allocated to high-end DOE capability-class computing using a similar INCITE-type process. Consideration should be given to increasing the frequency of INCITE calls for proposals or at least staggering the annual call with other relevant calls, such as the Energy Research Computing Allocations Process (ERCAP). An appeals process for INCITE allocation decisions should be implemented. And, in approximately 5 years, a formal review panel should be convened to assess the impact of the INCITE program.

Under the new ASCR allocation policy: Facility directors will retain up to 10% of the allocatable hours to support pilot or startup projects, to support code scaling, and to support petascale computer science and performance-metrics research. The majority (60-85%) of available processor hours on the leadership computing resources will be allocated through the INCITE program, but the LCF directors will be responsible for conducting joint reviews and selecting projects for their facility. There will still be a joint call. In addition, as part of the triennial on-site operational assessment of the respective LCF, a more in-depth review of the scientific accomplishments will be conducted by a subcommittee of ASCAC. The majority (60-85%) of available processor hours at NERSC will be for researchers working on SC-funded or SC-relevant projects. And the ASCR Leadership Computing Challenge (ALCC) will replace the SC-1 Director's Reserve. NERSC and PNNL will no longer participate in the INCITE program.

The ALCC will provide one-year allocations for

- Scaling the applications selected by the ASCR advisory committee to meet ASCR's computational science annual Program Assessment Rating Tool (PART) metric
- Out-of-cycle, proof-of-concept projects for other SC offices and federal agencies
- Broadening the community of researchers from other SC offices or Federal agencies able to take advantage of DOE's leadership-computing resources.
- Meeting the national needs/priorities. For example, ALCC resources have been previously been provided to The U. S. Army Corps of Engineers for storm-surge calculations after Hurricane Katrina.

There will be an ALCC reserve of between 5 and 30% of the resources at NERSC and the LCFs. A percentage will be established by the ASCR Associate Director with input from facility directors. Proposals requesting an allocation may be submitted to the ALCC program any time during the year. They will be reviewed for computational readiness and science. This approach amounts to 100 million hours (175 after the ORNL LCF comes online) to allocate.

Stevens asked if the number of projects awarded under INCITE would remain the same. Helland replied, yes. INCITE has few projects with large allocations

Chen asked if projects will have to be ready for one or all platforms. Helland responded that researchers can select the resources. She did not know how the allocators were going to deal with this.

Bailey noted that SciDAC had been cited as imposing double jeopardy and asked how that would be handled. Helland answered that that would be handled through the ALCC.

Giles asked how the problem of uneven reviews across disciplines would be handled. Helland said that these will probably be joint reviews.

Voigt asked if the LCFs would get any guidance between internal and external reviews. Helland responded that that has not been considered yet. Hack added that discussions are being held about how to set up this process, and the COV report will be used for guidance. The process will be uniform across facilities.

Negele noted that 60 to 85% is a large spread and asked how it would be determined what percentage was to be used. Helland replied that demand for a given facility will play a large role, and that demand will have to be estimated ahead of time.

The floor was opened to public comment. Bashor announced that the SciDAC 2009 meeting will be held in San Diego on June 14–18, 2009. Details can be seen at SciDAC.gov.

The meeting was adjourned for the day at 4:36 p.m.

Wednesday, March 4, 2009
Morning Session

The meeting was called to order at 8:30 a.m.

Paul Messina was asked to report on the scientific-challenge workshop series.

The purpose of these workshops is to identify grand challenge scientific problems in specific research areas that can exploit computing at extreme scales to bring about dramatic progress toward their resolution. The goals of the workshops are to

- identify grand challenge scientific problems;
- identify associated specifics of how and why new high-performance computing capability will address issues at the frontiers; and
- provide a forum for exchange of ideas among application scientists, computer scientists, and applied mathematicians.

The workshop chairs work with relevant DOE program offices and colleagues to identify key areas to cover. Panels are defined, and panel chairs are recruited. White papers for each panel are drafted and posted in advance of workshop. Priority research directions (PRDs) are identified by each panel. The panels are populated by domain science experts as well as by mathematicians and computer scientists. Observers from agencies and from the math and computer science communities are invited to each workshop. The panels report in plenary sessions in the middle and at the end of a workshop.

The first workshop was Challenges in Climate Change Science and the Role of Computing at the Extreme Scale, which was held November 6-7, 2008, in Washington, D.C. It was chaired by Warren Washington. Its goals were to review and identify the critical scientific challenges; prioritize the challenges in terms of decadal or annual timelines; identify the challenges where computing at the extreme scales is critical for climate change science success within the next two decades; engage international scientific leaders in discussing opportunities to shape the nature of extreme-scale scientific computing; provide the high-performance-computing community with an opportunity to understand the potential future needs of the climate change research community; and look for breakthroughs. It developed PRDs for

- Model development and integrated assessment
- Algorithms and the computational environment
- Decadal predictability and prediction
- Visualization and computing productivity

The second workshop was on Scientific Challenges for Understanding the Quantum Universe and the Role of Computing at the Extreme Scale, which was held on December 9-11, 2008, at the Stanford Linear Accelerator Center (SLAC). It was chaired by Roger Blandford, Norman Christ, and Young-Kee Kim. It identified PRDs in cosmology and astrophysics simulation and in accelerator simulation. It also addressed software to some extent, identifying low-level modules and libraries, community application codes, and interoperable data analysis as priority needs.

The third workshop was on Nuclear Science and was held on January 26-28, 2009, in Washington, D.C. It was chaired by Glenn Young, David Dean, and Martin Savage. It identified PRDs in the physics of extremely neutron-rich nuclei and matter, which involves computing the properties of nuclei that determine the r-process nucleosynthesis path in stars and the nucleonic matter in neutron star cores and crusts; the microscopic description of nuclear fission, which involves computing the paths to fission and its products; nuclei as neutrino physics laboratories, which involves computing properties of nuclei in double-beta-decay experiments and neutrino-nucleus cross-sections for supernova calculations and neutrino decays; “reactions that made us,” which involves computing (1) the triple-alpha process that produces ^{12}C and (2) $^{12}C(\alpha,\gamma)^{16}O$. Martin Savage observed that exascale computing resources are required to achieve the central mission of the Office of Nuclear Physics (NP). NP will be considerably unified by exascale computational resources. An organizational infrastructure is required to enable full utilization of exascale resources. Collaborations within and external to NP are vital. Prompt access to cutting-edge machines is essential.

The next workshop will be on Fusion Energy Sciences and will be held on March 18-20, 2009, in Washington, D.C. It will be chaired by Bill Tand and David Keyes. Panel topics will include plasma and fusion energy sciences and the cross-cutting areas of scalable algorithms; data analysis, management, and visualization; mathematical formulations; and programming models, frameworks, and tools.

The following workshop will be on Opportunities in Biology, which will be held on April 30-May 2, 2009, in Chicago. It will be chaired by Mark Ellisman and Rick Stevens and will include keynote speakers to provide a vision.

The next workshop will be on Nuclear Energy Systems, which will be held May 11-12, 2009, in Washington, D.C. It will be chaired by Ernie Moniz and Bob Rosner.

A workshop on Basic Energy Sciences will be held on August 12-14, 2009, in Washington, D.C. It will be chaired by Giulia Galli and Thom Dunning and will cover combustion; many body effects, correlation, and excited states; catalysis and electrochemistry; time-dependent processes; and nanoscale materials and technology.

The last workshop in the series will be on NNSA and Office of Science Missions, which will be held in September or October 2009 in Washington, D.C. It will be chaired by Alan Bishop and Paul Messina and will cover hydrodynamics; materials science; chemistry/biology; nuclear and particle physics; infrastructure, systems, and network simulations; and verification, validations, and uncertainty quantification and error analysis.

Some related workshops will be held defining the properties of the computing environments that will enable researchers to tackle the science challenges, guided by the findings of the eight "Science Challenge" workshops. In addition, two or three workshops will identify the key software environment and tools that are necessary for productive use of present and future leadership computer systems and viable approaches to developing them.

Michael Strayer had the vision to conceive these workshops and the other DOE-SC Associate Directors and their program managers have given their support. Walter Polansky and Lali Chatterjee have worked with the other DOE program offices to create a vision for the workshops and provided guidance on their scope and organization. Moe Khaleel and his PNNL team provide invaluable support and contributions at all levels. The community members who have organized and participated in the workshops are identifying exciting science challenges that could be tackled with future computer environments of much greater power than today's.

Smarr asked if this roadmap had been shared with the rest of the field of nuclear physics. Messina answered that none of the workshop reports have been finished, but they must be presented to the broad, applicable communities.

Negele noted that computational science played a central role in the most recent Nuclear Science Long-Range Plan. The major issue is building the full community to support the transition to exascale computing. The central role of computation is well understood in the field of nuclear science. That understanding should be inculcated in all of the communities whose topics are being explored by these workshops.

Giles asked whether, in the other workshops after fusion energy, the theme on how to work on algorithms and mathematics was emerging. Messina replied that more mathematicians and computer scientists are being included in the panels of the workshops. The follow-up workshops will have even greater computational science involvement.

Smarr noted that, with the new administration, a lot more money is coming into energy. The importance of computing should be stressed in areas that had not previously used modeling and simulation. Messina agreed.

Hey asked if one can offer incentives to computational scientists to develop codes for the exascale machines. Messina responded that that is done through SciDAC. Another approach is the University Alliance Program of NNSA. Each could produce community codes.

Yellick was surprised to see such an emphasis on DOE rather than on the full communities. Messina responded that, if he were Michael Strayer, he would want to build a case for DOE funding of computational science. There could be opportunities for other agencies or industry to see the benefits of supporting exascale computing. Strayer added that the climate workshop was international in scope and that the Office works closely with the Europeans and the Japanese and with industry representatives to advance progress toward the exascale and beyond.

Smarr asked if NSF had been invited to these workshops. Messina said that they had been invited and that Tony Chan is going to participate in the high-energy physics workshop.

Stevens noted that another team is organizing on banking and finance to address the need for algorithms and security felt by the Federal Reserve and the U.S. Treasury.

Steven Hammond (NREL) observed that there had been a workshop by ASCR and the Office of Energy Efficiency and Renewable Energy (EERE) on computational research needs for biofuels and other technologies.

Dan Reed was asked to speak on cloud computing.

A major transition is occurring in computing, and infrastructure is being built to enable discovery. Truisms for this transitional period include:

- Bulk computing is almost free, but software and power are not.
- Inexpensive sensors are ubiquitous, but scientific data fusion remains difficult.
- Moving lots of data is still hard because we are missing trans-Tb/s networks.
- People are really expensive, and robust software remains extremely labor intensive.
- Scientific challenges are complex, and social engineering is not our forte.
- Our political/technical approaches must change, or we risk solving irrelevant problems.

Commercial economics point to (1) many-core processors and accelerators and (2) software as a service and cloud computing. These practices will drive change in technical computing, just as did “killer micros” and inexpensive clusters. This is a chance to reinvent computational science.

Next-generation applications will occupy a mid-range position in the concurrency spectrum between local software and global services. Economics drives change. Moore’s “law” favored consumer commodities. Economics drove enormous improvements. Specialized processors and mainframes faltered. The commodity software industry was born. Custom HPC hardware largely disappeared because it is hard to compete against a 50%/year improvement. The implications here are that the consumer product space will define outcomes, and research environments will track commercial trends.

Doubling the number of cores cannot supplant increasing clock rates (at least without major innovation) because of sequential code, the lack of parallel algorithms, the difficulty of programming, and the availability of few abstractions. Parallelism is changing the computing landscape. If existing applications cannot use large parallelism, new applications and systems will arise, such as software plus services and mobile computing. Clouds are an obvious outcome, producing a new software architecture. When applications are hosted, even sequential ones are embarrassingly parallel with few dependencies among users. Moore’s benefits accrue to the platform owner; technical tradeoffs are not entirely one-sided because of latency, bandwidth, privacy, and off-line considerations on the one hand and capital investment, security, and programming problems on the other.

If one looks at the service continuum, one sees a fungibility of time and space, producing a rich user experience. Some cloud opportunities are service continuity, scale on demand, support for web applications, and reduced infrastructure and costs.

Cloud application frameworks have three components:

- Operating-system virtualization,
- Parallel frameworks, and
- Software as a service.

All cloud services have a virtualized compute environment based on Windows Server; durable, scalable, and available storage with abstractions; and automated management of the service lifecycle. There is a basic Azure Service Platform, which then makes available other high-level services (e.g., Live, .NET, SQL, SharePoint, and CRM). One can build a system that looks like it is running on one's desktop but with higher-level, interoperable services and a resilient manager.

Massive amounts of multidisciplinary data are rising rapidly. This data deluge has a number of social implications. For example, one can now pose the question, "What correlations can I glean from everyone's data?"

In Azure, data storage occurs at two levels: basic Azure storage and SQL database services. The data center is huge. Its site selection was driven by such considerations as the proximity of the user population, power pricing, environmental concerns, construction costs, tax climate, IT labor availability, heat dissipation, and corporate citizenship. A data center's costs are incurred for land, core and shell, architectural design, and mechanical/electrical construction (82%). The building blocks need to be right-sized.

There are more similarities than differences between exascale computing and cloud computing in terms of energy efficiency and carbon footprint, reliability and resilience, interconnect and photonics, memory stacking and access, exascale data management, many-core and processor functionality, system software and management, and programmability and models.

Clouds offer an opportunity to address a number of cyberinfrastructure components. The research infrastructure poses several challenges: insatiable demand, distributed acquisition/deployment, distributed cost structures, long-term sustainability, the user mix (mainstream users vs high sophisticates).

The major DOE science resources are smaller than one cloud data farm.

The computing continuum (parameter studies, hosted applications, data analysis and fusion, and desktop acceleration) will benefit from cloud computing. The way to think about clouds is that they provide high-level services at the desktop level.

There is an opportunity to partner via the Microsoft Engagement Proposal, which offers Azure cloud services, Internet2/ESnet connections, Tier-one support, major technical engagement team, hosted data sets, agency/university allocation, and fixed-price service. This arrangement will be available by the end of the year.

Azure cloud service is a Windows-based service with community tools. An engagement team could bring about community-code porting and partnership, including Linux-compatibility options.

Today is an inflection point because of the economic challenges and technology transition. We can change the game, just as we did before. Microsoft wants to engage the community.

Bailey asked what had happened to the hot topics of the past 10 years. Reed answered that one has to separate hype from reality. The broadband opportunities are better now. Price-performance has decreased. One has to look at explicit versus implicit costs. One does not see a lot of those costs. Applications have squeezed out a lot of costs. There are classes of applications for which cloud computing makes sense and some classes for which it does not. One needs to look at the cost per unit of computation. Bailey asked if Reed could cite some examples of scientific usage of cloud computing. Reed noted that Berkeley has been standing up a lot in parallelized applications.

Hey noted that scientific applications in the clouds are in the early stages but hold a lot of promise.

Stevens asked, given the differences between the models of cloud computing, how long it would take to transition applications to the clouds. Reed answered that some could be done in a matter of weeks. ISE applications already exist in the cloud that are instantly available.

Smarr asked what the status was of engagement with academics on prototyping experiments at universities. Reed said that these are the early days. Experiments are not being done right now. This is the time to influence how the system is set up. A year from now, it will be commercially fixed. The quality of change will be seen in large, complex applications.

A break was declared at 10:08 a.m., and the meeting was called back into session at 10:18 a.m.

Steve Cotter was asked to give an update on ESnet.

The ESnet 4 system is being built out on top of the Internet2 network; 25 routers, 9 new hubs, and 40 new waves have been installed. A lot of additional new connections have been made in the past 6 months. The MANs and their associated sites are being upgraded. These upgrades replace 6509s with MXs. ESnet4 is now a 10-Gb network with MANs, laboratory connections, international peerings, backup waves, the Science Data Network, and some 20-Gb links. The metro-area rings are in Long Island, Newport News, Atlanta, West Chicago, and San Francisco.

The 12-month circuit availability is very high. Site availability is even better than circuit availability.

ESnet traffic continues to grow about 70% per year. The traffic is being dominated by large flows. It is expected that the traffic in July 2010 will be about 10,000 Tb/month. Workshops are being held with the various communities to understand their data-transfer needs. Today, climate-model data are growing faster than high-energy physics’ data, the traditional leader in data transfer.

In 2010, the network will have its 10-Gb network and 20 to 50 Gb on the Science Data Network. Beyond 2010, 100 Gb will be needed, largely because of LHC Tier-1 and Tier-2 usage.

ESnet4 planning assumes technology advances will provide 100 Gb/s optical waves (they are 10 Gb/s now). The ESnet4 Science Data Network (SDN) switching/routing platform (Juniper MX960) is designed to support new 100-Gb/s network interfaces. ESnet is involved in a collaboration with Internet2, Juniper Networks (core routers), Infinera [dense wavelength division multiplexing (DWDM)], and Level3 (network support) to accelerate the deployment and help drive down the cost of 100-Gb components.

Network services offered include: (1) The Internet protocol (IP) Network, whose best-effort routing is simplistic, opportunistic, and resilient; however, it provides no assurances, consistency, or predictability. (2) The Science Data Network, where the On-

Demand Secure Circuits and Advance Reservation System (OSCARS) provides predictable and specific network-service performance that applications can demand and will be reliably provided through automated agents. The network is no longer just a cloud over which the user has little or no control, but a cyber-resource that can and should be directed by the application, just as other resources (such as compute/server cycles, storage resources, or workflow scheduling). (3) perfSONAR end-to-end monitoring service, which provides useful, comprehensive, and meaningful information on the state of end-to-end paths; it supports regularly scheduled tests and archiving of results, acting as an intermediate layer between the performance measurement tools and the diagnostic or visualization applications.

The OSCARS service requirements include guaranteed bandwidth with resiliency, traffic isolation, traffic engineering (for ESnet operations), secure connections, and end-to-end, cross-domain connections between Labs and collaborating institutions.

OSCARS started with Phase 1 [proof of concept and intradomain virtual circuit (VC) services] and evolved through Phase 2 (interdomain interoperability and pre-production ESnet VC services), Phase 3 (productionalizing OSCARS), and Phase 4 (extending service offerings).

The OSCARS takes a community approach to supporting end-to-end virtual circuits in the R&E environment is coordinated by the DICE (Dante, Internet2, Caltech, ESnet) working group. Each organization potentially has its own InterDomain Controller approach (though the ESnet/Internet2 OSCARS code base is used by several organizations (flagged OSCARS/DCN [Dynamic Circuit Network]). The DICE group has developed a standardized InterDomain Control Protocol (IDCP) for specifying the set up of segments of end-to-end VCs. Many organizations have implemented and deployed systems that are compatible with the DICE IDCP. Several "higher level service applications" have adapted their existing systems to communicate via the user request side of the IDCP.

Daisy chaining of change-path computation engines is expected. Fermi National Accelerator Laboratory (FNAL) and Brookhaven National Laboratory (BNL) changed the reservation workflow, added a notification callback system, and added some parameters to the OSCARS application program interface (API) to improve interoperability with automated provisioning agents, such as LambdaStation, Terapaths, and Phoebus. As of Dec. 2, 2008, there were 16 long-term production virtual circuits instantiated, all of which support the Office of High Energy Physics (HEP). Between Jan. 1, 2008 and Dec. 2, 2008, there were roughly 2650 successful HEP-centric virtual circuit reservations. OSCARS generated and managed virtual circuits at FNAL, one of the U.S. Teir-1 data centers. OSCARS is also used on the ESnet itself.

perfSONAR services include tools (SNMP Measurement Archive, Lookup Service, Topology Service, Circuit Status Measurement Archive, Status Measurement Archive, perfSONAR-BUOY, and PingER Services); visualization; and alarming. The hardware is currently being deployed across the network to support ad hoc measurements for debugging. perfSONAR BUOY deployment hardens the infrastructure. perfSONAR R&D activities involve scaling and robustness enhancements, such as visualization tools, integrating OSCARS circuits, and alarming.

For federated trust services, the DOEGrids Certification Authority is used. Vista – IE browser support, support for the ESnet 2-factor authentication token project, and cloning and geographical dispersion are being provided. The activity with active certificates continues to increase. DOEGrids Certification Authority and its key management hardware will be cloned and dispersed around the United States to improve continuity of operations and disaster recovery issues, improve availability to customers, and provide for future robust services.

During the past 6 months, ESnet has increased security by (1) implementing two-factor authentication for ESnet network engineers requesting privileged access to the network management plane, reviewing and redefining access to the network management plane, and (2) upgrading the Bro Intrusion Detection System. The ESnet Security Peer Review on Feb. 11-12, 2009, went very well; federal, research and engineering, and commercial experts reviewed ESnet security practices and procedures. In addition, disaster-recovery processes are being improved.

The goals of the website redesign are to produce better organization of the information, easier navigation, searchability (not everything in pdfs), a collaborative tool, and the integration of business processes into the site (i.e., a “My ESnet” portal for site coordinators and users and Exploring Google Earth or a similar network visualization).

Stevens asked how much tracking was being done of the LCFs’ ability to transfer data to or from users’ institutions. Cotter responded that such tracking was not being done in a formal way. Draney added that weekly meetings are being conducted in the LCFs to do just that.

Smarr asked for some examples of science being enabled by these supernetworks. Cotter replied that they had been looking into that. The climate community has several success stories. Videoconferencing is another example.

Hey asked if all of the input/output switches will be optical. Cotter responded, no.

Patricia Hoffman was asked to describe the use of high-performance computing in managing the electric grid.

The electric grid is a complex system. Data are collected in real time and used to model the system. The system has 30,000 transmission paths; over 180,000 miles of transmission lines, and 14,000 transmission substations, and 3,170 traditional electric utilities (239 investor-owned, 2,009 publicly owned, 912 consumer-owned rural cooperatives, and 10 federal electric utilities) that are governed under a mix of federal and state regulatory structures.

Much of the grid works well. A significant data-collection infrastructure exists. However, the existing systems must be fully understood before they can be improved. There are three interconnects: the Eastern Interconnect, the Western Interconnect, and the Texas Interconnect. What is wanted is to analyze the data at the interconnect level.

The electric power infrastructure consists of competitive power generation; the Federal Energy Regulatory Commission FERC-regulated (long-distance) transmission; state-regulated (local) distribution; and retail (i.e., residential, industrial, and commercial) customers, who may have their own microgrids. An analysis of this system looks at management, power quality, and distribution-system loads and electricity generation.

More generation capacity is going to come online, and some of it (e.g., wind power) will introduce complex patterns of power timing, level, and quality. In addition, a carbon price tax may complicate management of the portfolio of resources. At the same time, consumers are getting more energy intensive and volatile. Reserve margins will be tight. The issues that will need to be dealt with will include:

- Adapting the electric grid to a low-carbon future
- Consumer energy management (enabling energy efficiency, demand response, energy storage, distributed energy, and plug-in hybrid vehicles)
- The Smart Grid (deploying sensing-data collection, monitoring, and automation)
- Adding transportation (hybrids) to the load demand
- Security (ensuring availability and integrity through control system security, diagnostics, and recovery capabilities)

Computational capabilities and analytics need to be in place to manage these changes.

The driving forces behind grid advancement are (1) the historic events of the blackout of 2003 and the hurricanes of 2005 and (2) the concern about cyber-vulnerabilities.

A model for getting it right is Hawaii, which wants to introduce 70% renewables and use distributed resources to reduce peak power. The system has been modeled, but the model results were not realistic. An analysis has been performed of the information needed to model the system correctly.

The first step in developing a Smart Grid network is communication integration. Wide-area networks (WANs), substation controls, and energy-management systems for homes are needed.

Phasors are a key technology for the future. This is mature hardware that is supporting emerging networks and applications. It produces GPS-time-synchronized high-resolution data to characterize power quality (in terms of phase, stability, angle separation, synchronization, etc.) across the grid. Visualization of these data and controls are then used to enhance transmission reliability. What is envisioned is taking these data from throughout an interconnect. Therefore, new algorithms, new mathematics for characterizing uncertainty, new methods to enable efficient use of high-bandwidth networks, and new software architectures and new rapid development tools for merging legacy and new code without disrupting operation are needed.

What is needed is to conduct data analysis and modeling and simulation for decision support.

Stevens said that Hoffman seemed to be asking for faster-than-real-time analysis to analyze transients as well as transitions between steady states. Overholt replied that the FERC had always collected performance data on utilities.

Voigt asked if they had work ongoing. Hoffman replied that they are trying to install the phasors and are building the visualization capability. There are 200 phasors across the United States now; that number will increase to 1000. Models and analytical tools need to be built. A partnership is needed with the scientific community.

Negele asked who got and received the data if one successfully models electrical production, use, and transmission. Hoffman replied that the interconnect and then the system operators would receive those data. Overholt added that the interconnects already have models of the networks, but those models are very slow. A Reliability Center will monitor the entire web and try to provide the information needed to head off problems. The goal is to let operators deal with problems in real time.

Smarr asked what the state of play in modeling was. Overholt replied that they were working with NSF research centers that involve 13 universities. The work looks at energy, regulatory, environmental, and other inputs to produce a power-flow analysis. With such a tool, one could integrate wind with the grid. Hoffman said that her office has

$4.5 billion, part of which will go to colleges and universities for workforce development.

Bailey asked what their engagement was with the SC national laboratories. Overholt replied that they were working with ORNL, LBNL, Sandia, and PNNL. The level of engagement has been growing in the past few years.

Hoffman noted that the Office's annual appropriation is $400 million. It is getting $4.5 billion from the stimulus package.

Strayer said that ASCR would be delighted to work with [the office] on workshops to develop planning for modeling the grid.

Hey asked about E-ARPA (the energy version of the Advanced Research Projects Agency). Hoffman replied that DOE is still trying to figure out how E-ARPA is to work.

Giles asked if continuous sensor data were to be transferred. Hoffman answered that transparency and security need to be balanced in the data-transfer architecture.

The floor was opened to public comment. There being none, the meeting was adjourned at 11:35 a.m.

Respectfully submitted,
Frederick M. O'Hara, Jr.
Recording Secretary
March 25, 2009

Corrected:
Barbara Helland, March 26, 2009
Charlie Catlett, March 26, 2009
Arthur Bland, March 26, 2009
Juan Meza, Apr. 2, 2009