Draft Minutes
Advanced Scientific Computing Advisory Committee
March 31, 2014
American Geophysical Union, Washington, D.C.

ASCAC members present:

Marsha Berger
Vinton Cerf
Barbara Chapman
Jacqueline Chen (via telephone)
Jack Dongarra (via telephone)
Roscoe Giles (Chair)
Sharon Glotzer (via telephone)
Anthony Hey
Gwendolyn Huntoon (via telephone)
Juan Meza
John Negele (via telephone)
Vivek Sarkar (via telephone)
Victoria White
Dean Williams

ASCAC members absent:

Vincent Chan
Linda Petzold

Also participating:

Melea Baker, Administrative Assistant, Office of Advanced Scientific Computing Research, Office of Science, USDOE
Steven Binkley, Director, Office of Advanced Scientific Computing Research, Office of Science, USDOE
Laura Biven, Senior Science and Technology Advisor, Office of the Deputy Director for Science Programs, Office of Science, USDOE
Arthur Bland, Project Director. Oak Ridge Leadership Computing Facility, Oak Ridge National Laboratory
Keri Cagle, Oak Ridge Institute of Science and Education
Christine Chalk, ASCAC Designated Federal Officer, Office of Advanced Scientific Computing Research, Office of Science, USDOE
Susan Coghlan, Deputy Division Director, Argonne Leadership Computing Facility, Argonne National Laboratory
Vincent Dattoria, Computer Scientist, Office of Advanced Scientific Computing Research, Office of Science, USDOE
Patricia Dehmer, Acting Director, Office of Science, USDOE
Barbara Helland, Acting Associate Director, Office of Advanced Scientific Computing Research, Office of Science, USDOE
Alexandra Landsberg, Applied Mathematics Program, Office of Advanced Scientific Computing Research, Office of Science, USDOE
Carolyn Lauzon, Program Manager, Office of Advanced Scientific Computing Research, Office of Science, USDOE
Paul Messina, Director of Science, Argonne Leadership Computing Facility, Argonne National Laboratory
Frederick O'Hara, ASCAC Recording Secretary
Brian Plesser, Attorney-Adviser, Office of the General Counsel, USDOE
Andrew Siegel, Director, Center for Exascale Simulation of Advanced Reactors, Argonne National Laboratory

About 45 others were in attendance.

Before the meeting, Brian Plesser conducted the annual ethics briefing for the Committee members.

Morning Session

The meeting was called to order at 8:34 a.m. by the chair, **Roscoe Giles**.

Patricia Dehmer was asked to review the activities of the Office of Science (SC). She started by summarizing her testimony given a week prior before the Energy and Water Development Subcommittee of the House of Representatives, where she stressed that SC is pursuing areas critical to the United States, such as high-performance computing and the achievement of the exascale. Other countries are also pursuing this capability. The Department currently uses high-performance computers in the conduct of research on biology, medicine, nuclear fuels, and combustion/fuel efficiency. The next generation of computers will be more productive, especially in understanding complexity. The United States needs to be the first to benefit from the next generation of computers. SC's other priorities are selected increases for research and for instrument and facility construction; the optimal operation of its scientific user facilities (e.g., light sources and computational resources); and maintaining a balance between research and facilities. She had also pointed out that, in developing the proposed budget, choices were made on the basis of advice from advisory committees and user groups. There have been many major high-performance computing workshops in the past 5 years; the reports of these workshops were important because choices are made on the basis of what one knows.

Questions asked by congressmen on the House Subcommittee included:

- What does exascale mean?
- What are fast computers good for?
- Why do we need to be number one in the world?
- Does the collaboration with the National Nuclear Security Administration (NNSA) really exist?
- Has the Office of Advanced Scientific Computing Research (ASCR) reached out to the other offices of SC?

When one is dealing with appropriators, 95% do not have science and math in their backgrounds. The workshop reports are very technical. Decisions are being made by people who do not have the backgrounds of the workshop participants.

The Office of Biological and Environmental Research (BER) and the Office of Basic Energy Sciences (BES) have large appropriations for high-performance computing. The Office of Nuclear Physics (NP) is doing fine constructing a new facility. The Office of High Energy Physics (HEP) is thin on research and facilities. The Office of Management and Budget (OMB), the Office of Science and Technology Policy (OSTP), and Congress will not support anything that does not have community support. The two programs still awaiting budget advice, HEP and Fusion Energy Sciences (FES), have struggled.

The Secretary of Energy Advisory Board (SEAB) has made some decisions and studies on the exascale and on research facilities, the bioenergy research centers and Advanced Research Projects Agency-Energy (ARPA-e).

SC has two nominees awaiting confirmation: Marc Kastner, dean of Massachusetts Institute of Technology's School of Science, has been nominated to be the next Director of the Office of Science (SC). Lynn Orr, professor in Stanford University's Department of Energy Resources Engineering, has been nominated to be the next Under Secretary for Science of DOE. The Senate is working through their list of nominees very slowly.

Williams asked if there were some way that the Committee's reports could be made more understandable to Congress. Dehmer replied that the Committee's reports are aimed at a particular audience that does not include Congress. It is the person making presentations to Congress who needs to be able to make them understandable and appreciated.

Meza asked how the Office strikes a balance between research and facilities. Dehmer replied that an effort is made to get a budget whose funding reflects 40% research, 40% facilities, and 20% for continuing operations.

White asked whether, in discussions of U.S. competitiveness, long-term R&D was ever considered or whether the focus was more on immediate returns. Dehmer said that long-term R&D is discussed in a different venue. China and Japan have been investing in long-term R&D in high-performance computing, have begun making and using indigenous components and parts, and have started to leap-frog the United States' competencies. That is worrisome.

Cerf noted that sometimes one cannot do research without facilities. Demers said that that is true in every discipline. Cerf pointed out that the national laboratories were developed 60 years ago for particular reasons. There is a tension between relevance for R&D and jobs in funding facilities. Dehmer replied that there are studies of the laboratories currently going on, as requested by Congress. The national laboratories saw a tremendous change in the 1980s when they went from in-house facilities to open user facilities. Today, SC directly funds 28,000 people and serves an equal number of people who come in from academia and industry to use the facilities. National laboratories now interact much more with industry both local and large, a trend pioneered by ASCR. The Nobel Prize in chemistry is awarded in biochemistry every 3 years; the past four of those medals have gone to users of SC facilities.

Giles noted that the natural timescale for high-performance-computing development is a long-term process. The Committee has responded to one letter charge and will take on another one during this meeting on workforce development. Dehmer said that one can see the results of the charges in the Department's testimony before Congress. The charge on workforce development is different. There has been a consolidation of fellowships into the National Science Foundation (NSF), the Smithsonian Institution, and the Department of Education. Discussions were held with the Office's OMB examiner during the previous week about the Computational Science Graduate Fellowship. The terms "fellowship" and "traineeship" will not be used by SC in budget requests; workforce-development activities will have to be related to a national laboratory and be mission-relevant. The workforce-development activities for teachers and faculty have been skewed toward high-performance computing. That is about the best that could be done. Hot topics need to be identified where personnel have been poached from the national laboratories. It is hoped that a list of workforce-development needs will be received from each associate director of SC. Giles pointed out that the Computational Science Graduate Fellowship has been extremely productive over the decades. To do away with that fellowship and the fellowship model in general does not seem rational. Dehmer agreed that the computational-science fellowships have been outstanding.

Hey asserted that going from the petascale to the exascale will be the foundation of a new industry because petascale computing cannot just be scaled up.

Cerf pointed out that computations can be done better (e.g., in terms of precision) if one has more computational power.

Giles said that high-performance computing and networking play an important role in the collaboration of personnel across the national laboratories.

Steven Binkley was asked to review the recent activities of ASCR.

ASCR is currently operating under a total FY13 appropriation of $417.8 million. The enacted FY14 appropriation is $478.1 million, and the President's request for FY15 is $541 million, a 13% increase over the FY14 figure at a time when the overall SC budget request reflects an increase of 1%.

The FY15 ASCR budget priorities are to conduct R&D and design efforts in hardware, software, and mathematical technologies that will produce exascale systems in 2022; to prepare today's scientific and data-intensive computing applications to migrate to and take full advantage of emerging technologies; and to acquire and operate more-capable computing systems. These priorities are reflected in increased requested funding in applied mathematics, computer science, next-generation networking for science, research and evaluation prototypes, and high-performance network facilities and testbeds, producing a $62.9 million increase in ASCR's total annual funding.

Several initiatives or themes cut across budget line items. The total amount of the exascale crosscut is $91 billion, up about 20% from FY14. Cross-cutting funding for data-intensive science increases almost $10 million. And funding for facilities increases $30.4 million, split between new platforms and exascale technology development.

Secretary Moniz has re-instituted the Secretary of Energy Advisory Board (SEAB). In December 2013 it established a task force to examine drivers for next-generation, high-performance computing. A week before this meeting, it reviewed and voted to approve final reports on hubs and tracking for oil and natural gas production and received updates on task-force activities, including the *Quadrennial Energy Review*, the nuclear nonproliferation study, and high-performance computing. It will receive and review the final report on high-performance computing at its June 2014 meeting. In a charge letter from the Secretary, SEAB was specifically asked to validate the need for next-generation computing. The task force established to respond to that request has had two meetings. The first focused on industry uses of exascale-class computing (in aviation, petroleum, search applications, and finance), and the second focused on the intelligence community and stockpile stewardship. SEAB's next meeting on April 1 is a working session on superconducting technologies and quantum information and computing. It is to deliver its final report to the Secretary by June 30.

A re-organization of the Department of Energy has brought over to SC some applied-technology offices (Fossil Energy, Energy Efficiency and Renewable Energy, Nuclear Energy, and Electricity Delivery and Energy Reliability). This realignment will produce a greater synergy among these offices and with the other offices of SC; they always had a lot of crosstalk. This co-organization also allows closer cooperation.

Six "teams" are being organized by the Office of the Undersecretary for Science and Energy to enhance cross-organizational participation and to inform the FY16 budget formulation. The six teams are advanced computing, the grid, manufacturing, subsurface science, supercritical carbon dioxide, and water–energy.

The advanced computing team has a strong focus on simulation. This team has been around since 2012; it has membership from across the Department, including NNSA and its Office of Defense Programs. It is converging on a set of goals, focusing on cataloging anticipated applied-program requirements (with a focus on simulation rather than demonstration), documenting available advanced-computing resources, and identifying gaps to inform investment recommendations. The team will be co-chaired by an applied program and SC. Its membership will be cross-cutting (including NNSA and Environmental Management), participatory, and self-nominating. Its next steps will be to finalize its charter and to conduct an initial workshop in April 2014.

Giles asked if one were talking about the petascale or exascale in the use of high-performance computing in other SC offices. Binkley replied that, originally, the petascale will be used; most of the other programs need to be bootstrapped up to the petascale. Cerf noted that, about 5 years ago, DOE and the Department of Commerce announced the development of the smart grid, with the conduct of R&D and the construction of a micro-grid. He asked if this activity were still going on. Binkley replied, yes, mostly in the Office of Electricity Delivery and Supply Reliability with assistance from ASCR. The two national laboratories most heavily involved in this effort are Pacific Northwest National Laboratory (PNNL) and the National Renewable Energy Laboratory (NREL). High-performance computing is being considered for real-time control systems.

Progress continues to be made in the extreme-scale effort by SC and ASCR, focusing on the software and algorithms underpinning exascale operations. Mathematical, computational, and computer-science research is being conducted on uncertainty quantification, extreme-scale advanced architectures, and codesign. High-performance computing and network facilities are doing R&D on platforms and critical technologies. FastForward investments in critical technologies and DesignForward investments in system-level engineering are being made with high-performance-computer vendors.

FastForward is promoting innovative R&D on the critical technologies of memory, processors, and storage. Participants are AMD, IBM, Intel Federal, NVIDIA, and Whamcloud. The FY15 increase takes FastForward research to the next level.

In domain-science programs, BES is receiving an additional $24.2 billion for computational materials sciences to develop research codes for the design of functional materials, and BER is receiving an additional $29.0 million for climate model development and validation.

The leadership and production computing facilities are made up of Titan at Oak Ridge National Laboratory (ORNL), Mira at Argonne National Laboratory (ANL), and Edison XC30 at Lawrence Berkeley National Laboratory (LBNL). These facilities make up about 25% of the SC user portfolio; this percentage was much smaller 15 years ago.

In networking, ASCR operates the ESnet, a 100-Gbs production network, which is seeing a significant growth in data requirements.

Large increases are being seen in the budgets of the ASCR computing facilities. The high-performance production computing of the National Energy Research Scientific Computing (NERSC) operates optimally with more than 90% scheduled availability. The LCFs are initiating a postdoctoral training program for high-end computational science and engineering and are investing in high-performance network facilities and testbeds. Cerf asked if this was a single- or multi-channel 100-Gbs network. Dattoria replied, multi; there is a single-channel, single-lambda backbone with multi-channel droplines to the national laboratories. Those droplines will be upgraded.

SC is often viewed by many as not being a player in “big data.” However, examples of big-data production and use abound within SC. New light-source experiments produce big data. Statistics on systems performance attest to the high data-transfer rates. A major trend is the capturing of data at the experiments’ end stations, the processing of that data in near real time, and the immediate guidance of future experimental activities. There is an extreme-scale science-data explosion in genomics, high-energy physics, light sources, and climate.

Projections have been made of the data drivers from DOE scientific use of facilities to guide budget requests. The next steps are (1) developing concepts for an SC-wide data architecture, originally including NERSC, ORNL, and ANL and then expanding to Brookhaven National Laboratory (BNL) and PNNL; (2) identifying a small number of large-data demonstrations to validate the approach to big data; (3) developing an initial baseline plan for system-wide architecture; (4) expanding two or three of the FY14 demonstration projects into pilot studies; and (5) initiating the development of system-wide middleware tools and new algorithms.

A new report has been issued on *Applied Mathematics Research for Exascale Computing*.

FY15 is the third and final round of Scientific Discovery through Advanced Computing (SciDAC). Future directions in this partnership area are being discussed with other SC offices.

Glotzer asked what the budget meant where it said that there was an additional $29 million funded in the domain-science program. Did this refer to principal-investigator grants or to centers? Binkley replied that domain grants are for collaborative work done with BES, BER, NP, and FES. The primary focus in BES will be on software that can be used on functional-material sciences. Program managers are formulating how that will be done, through centers or through funding opportunity announcements (FOAs). ASCR will be providing support to BES. The primary focus in BER will be to adapt the Community Climate Model (CCM) to the exascale.

Negele asked whether the current SciDAC might be continued. Binkley replied, no. There will likely be a new version adapted to the changing conditions of the world. The sister offices have to be willing to cofund the effort. Those offices seem eager to continue funding SciDAC.

Hey commented that there seems to be a lack of expertise in data analytics. Binkley agreed but expressed a belief that the situation is changing. There have always been big data sets produced by simulations and experiments. However, large amounts of less-structured data are now being produced. One has to chart one’s way very carefully and coordinate with other parts of the government on this topic and on how to ferret information out of data. This topic is being looked at intensively.

Dongarra asked what the output would be of the SEAB Task Force and how that output would affect the Office. Giles answered that the Task Force is looking at all architectures and data-science needs. Binkley added that the Secretary’s objective is to defend a major initiative, so he is vetting the program through SEAB. If the Task Force endorses the pursuit of the exascale, the Secretary will be able to defend the initiative before Congress more easily. Also, the core applications include not only the traditional applications but also a number of other high-technology areas that are increasingly important to the national economy. Congress generally wants to know what the one or two killer applications are. Giles said that one needs to be able to say:

1. Here is what can be done with exascale computers and
2. Without exascale computers, here is how the nation will suffer (by not understanding climate, not developing advanced materials, etc.).

A break was declared at 10:23 a.m. The meeting was called back into session at 10:44 a.m.

Arthur Bland and **Susan Coghlan** were asked to describe CORAL [the Collaboration of ORNL, ANL, and Lawrence Livermore National Laboratory (LLNL)]. Bland started the presentation.

DOE's LCFs provide the computational and data-science resources required to solve the most challenging scientific and engineering problems. There are two centers with two architectures to address the diverse and growing computational needs of the scientific community. Mira at the ANL LCF has a peak performance of 10 petaflops, 50,000 compute nodes, 768 TB of system memory, a 35-PB file system, and a bandwidth of 240 GB/sec. The Titan at the ORNL LCF has a peak performance of 27 petaflops, 19,000 hybrid compute nodes, 710 TB of system memory, a 32-PB file system, and a bandwidth of 1 TB/sec.

These machines have been accepted, so it is time to buy a new machine. DOE/SC and the NNSA have signed a memorandum of understanding, agreeing to collaborate on exascale research and pre-exascale and exascale acquisitions. CORAL is a collaboration of the Oak Ridge, Argonne, and Lawrence Livermore national laboratories to acquire three systems for delivery in 2017/2018. This collaboration grouping was done on the basis of common acquisition timings. It reduces the number of requests for proposals (RFPs) for vendors, allows pooling of R&D funds, supports the sharing of technical expertise among laboratories (all of which have world-leading experts to write the RFP), and strengthens the SC–NNSA alliance for exascale computing.

The approach seeks competitive systems for two architecturally diverse systems with a single RFP leading to two nonrecurring engineering contracts and three procurement contracts. Once the machines are selected, multi-year, laboratory–awardee relationships will ensure the delivery of the best performance. Both engineering contracts will be jointly managed by the three national laboratories. Each national laboratory will manage and negotiate its own computer procurement contract so they can exercise options to meet their specific needs. It is understood that the long procurement lead time may affect the architectural characteristics and designs of the procured computers. Technology will evolve, and the collaboration will need to work closely with the vendors.

With a combined RFP, the national laboratories get to see what is going on at more than one vendor, and the vendors will get to see the requirements of more than one computing facility. Working together on one RFP improves the quality of the technical requirements by having a larger group of experts provide input; facilitates responding to vendors' questions; and provides risk mitigation if something catastrophic happens to one vendor.

Having system diversity provides many advantages. It promotes price competition, which increases the value to DOE. It promotes a competition of ideas and technologies, which helps provide more-capable systems. It helps promote a rich and healthy high-performance computing ecosystem. It leads to a leadership position for DOE in high-performance computing. And it reduces risk that may be caused by delays or failure of a particular technology or shifts in vendor business focus, staff, or financial health. The CORAL team worked with the high-performance-computing vendor community to ensure that the responses had sufficient diversity to provide the above advantages.

The RFP has been issued and has produced two winners that will lead to three contracts. Cerf commented that getting all three parties to agree maybe a headache for the vendors. Hey asked where Las Alamos and Sandia fit in. Coghlan responded that they are covered in a separate procurement, Trinity, which is a little ahead of CORAL.

Coghlan continued the presentation: The statement of work listed technical details, such as power requirements and performance. Another document was issued to provide the proposal evaluation and proposal-preparation instructions. The statement of work had mandatory requirements, mandatory options (things that may or may not go into the contract), technical options (things that are important but are not required to be responsive to the RFP), and target requirements (features, components, performance characteristics, or other properties that are important but will not result in a nonresponsive bid if they are omitted). The target requirements have three levels: highly desirable, like-to-have, and nice-to-have features.

The highest-level system requirements are that the machine produce

- A speedup over current systems of 4 times on scalable benchmarks and 6 times on throughput benchmarks
- An aggregate memory of at least 4 PB and at least 1 GB per message-passing-interface (MPI) task
- A maximum power consumption of less than 20 MW
- A mean time between application failure that requires human intervention of more than 6 days
- Data-center capabilities
- A delivery date in 2017 with acceptance in 2018

Principles guiding the drawing up of the technical requirements included:

- Minimizing the number of mandatory requirements and allowing consideration of the widest range of architectural solutions
- Focusing on requiring science and throughput performance while avoiding overly prescriptive explicit speeds and feeds
- Agreeing on common technical requirements across all three national laboratories (except for power, cooling, and maintenance requirements)
- Requiring vendors to describe the available options to adjust system size and configuration in the future to meet individual site needs and/or budgetary constraints

Systems can vary from one another in many different dimensions: architecture, interconnect, input/output (I/O) subsystem, density, resilience, heterogeneity of nodes, memory and processor architectures, HPC stack, operating system, input/output, file system, programming environment, administrative tools, etc.

The CORAL benchmark categories represent DOE workloads and technical requirements. Those benchmark categories are scalable science benchmarks; throughput benchmarks; data-center benchmarks; skeleton benchmarks (platform characteristics); and micro-benchmarks, made up of small code fragments that are useful for testing programming methods and performance at the node level and for emulators and simulators. Three levels of benchmarks (the highly desirable, like-to-have, and nice-to-have features) are used to characterize the scalable science, throughput, data centricity, and skeleton, with the broad goal being an average performance improvement over today's systems of 4 to 8 times for scalable science applications and of 6 to 12 times for throughput applications.

The RFP was released in January with proposals due in mid-February. A two-step evaluation process is currently being carried out. The technical evaluation is being conducted by eight teams of experts with three people from each of the collaborating national laboratories plus one representative from either Los Alamos National Laboratory or Sandia National Laboratories to represent their users. These eight technical teams reviewed the proposals for three weeks ahead of a 2-day face-to-face meeting to assess each of the proposals against the draft statement of work and evaluation criteria. A separate evaluative step was conducted by the buying team, which consisted of the management, technical, and procurement leadership of the three collaborating national laboratories. They met to select the set of two proposals that provided the best value to the government. They looked at the best overall combination of solutions, technical proposal excellence, projected performance of the applications, feasibility of schedule and performance, diversity, overall price, supplier attributes, and risk evaluation.

The procurement approach is to develop a long-term contractual partnership with vendors. In addition, all three national laboratories are continuing to collaborate on both the engineering contracts and the build contracts. Targets are performance levels that both parties reasonably believe can be achieved, depending on the engineering. Those charges will be converted to traditional mandatory requirements after a go/no-go decision is made. The engineering contracts are tightly coupled to the build contracts. An announcement of evaluation results cannot be made until the contracts have been negotiated in late FY14.

Williams asked what the requirements and outcomes of the CORAL collaboration were in comparison to the NERSC–Trinity requirements. Coughlan replied that one cannot talk to outcomes. The requirements were for a not-as-large system with a different focus, and the acquisition was projected for a nearer term.

Cerf asked if R&D at the vendors was being paid for as a shared risk. Bland replied that the collaboration is paying for R&D that the vendors would not necessarily do. Cerf asked if there had been a request for information (RFI) prior to the RFP. Bland replied, yes. Cerf stated that he was much more satisfied with how this collaboration is being organized. He asked how many responses to the RFP had been received. Bland replied, more than two.

Chapman asked if the benchmarks were pure MPI [message-passing interface]. Coughlan responded, no; there were Open Agent Communication Language (Open ACL) and others. Chapman asked how the applications were specified. Coughlan said that the researchers were asked what they would need in a given timeframe.

White asked how risk played out if one vendor goes conservative and another goes out on a limb. Coughlan replied that both of these situations were desired; the vendors have to do sound science but also have to reach (e.g., in bandwidth). Bland added that an RFI was issued and a year of discussions was held before the RFP so there was a good idea of what the vendors were capable of. Microprocessor design cannot be changed, but stretch goals can be put in the RFP in some areas.

J. Chen asked if there were any consideration of using the work coming out of the codesign centers in terms of benchmarks etc. that are forward looking to the exascale. Bland replied that the consensus was that those benchmarks etc. were not far enough along to put into the RFP; however, they may be used in the acceptance criteria.

Roscoe Giles stood in for **Robert Lucas** in presenting the final version of the report on the Top 10 Exascale Research Challenges, which version reflected the changes discussed during the previous meeting when the report was accepted by the Committee.

The top-10 challenges are to achieve

1. The exascale with a machine that uses 20 MW or less of power
2. A high-performance, energy-efficient interconnect
3. New memory technologies, including stack memory, nonvolatile memory, and processor-in-memory
4. A software system to manage the power and resilience of the millions of nodes in an exascale system
5. Programming systems for exascale systems that will have billion-way parallelism and frequent faults
6. Improvements in data management
7. Changing algorithms to run with billion-way parallelism, which will require redesigning or even reinventing the algorithms used in them and potentially reformulating the science problems
8. Methods and software to efficiently carry out uncertainty quantification and optimization on complex multi-physics problems
9. Resilience and correctness in the form of programming tools, compilers, debuggers, and performance-enhancement tools while at the same time increasing programming productivity
10. Programming tools, compilers, debuggers, and performance that enhance how productive a scientist is when working with an exascale system

The findings were that

- Exascale computing is critical for executing the DOE mission
- U.S. national leadership is at risk
- The U.S. *does* have the technical foundation to create exascale systems
- An evolutionary approach will not be adequate to achieve the exascale
- The U.S. Government's continuous leadership and investment are required to create exascale systems insofar as the Government has a history of pushing the envelope and high-performance computing and insofar as the exascale is not a terminal goal but a qualitative leap

The recommendations were that

- DOE should invest in a program of continuous advancement in high-performance computing
- DOE should invest in the U.S. industrial base to catalyze the foundation for exascale systems

- DOE should invest in exascale mathematics and system software responsive to DOE missions and other U.S. Government requirements
- DOE should create an Open Exascale System Design Framework to enable cooperative hardware and software advancement, building on the idea of codesign

The report was accepted with prospective editorial changes at a prior ASCAC meeting. Those changes were made, and the report was delivered to ASCR and to SEAB.

Giles thanked Robert Lucas and the extensive list of task force members.

Steven Binkley was asked to present the charge letter on workforce development.

Binkley presented to ASCAC a charge letter requesting assistance in identifying disciplines in which significantly greater emphasis in workforce training at the graduate student or postdoc level is necessary to address gaps in current and future mission needs. The Committee was asked for its expert assessment of

- Disciplines not well represented in academic curricula;
- Disciplines in high demand, nationally, and/or internationally, resulting in difficulties in recruitment and retention at U.S. universities and at the DOE national laboratories;
- Disciplines identified in the previous two bullets for which the DOE national laboratories may play a role in providing needed workforce development; and
- Specific recommendations for programs at the graduate student or postdoc levels that can address discipline-specific workforce-development needs.

A written response back to SC is required by June 30, which is before the next scheduled ASCAC meeting in August 2014. A teleconference will be set up in mid-June to discuss and approve the response.

Cerf pointed out that this request has the implicit assumption that, if one had the right curriculum, one would have the right workforce. However, in the real world, there is now a longer-living workforce and an evolving technology. That situation is not amenable to a point-source education at the beginning of a career.

Giles said that learning is building through experience. Experience in ASCR shows interdisciplinary education is needed along with experience at the national laboratories. This charge asks about post-doctorates and the other normal educational modalities. Cerf wanted to be assured that the Subcommittee would not just focus on the graduate-postgraduate period alone.

Glotzer pointed out that retention is difficult in the computer sciences community and should be addressed.

Berger said that there should be continuing training, also.

Giles said that another topic that should be included is how DOE is uniquely positioned in this need for workforce development. A subcommittee will be established to respond to the charge. Cerf stated that the Subcommittee should include human-resource representatives from the national laboratories. J. Chen said that people who run applied-mathematics programs at the national laboratories should be involved. Giles stated that someone involved in high-performance-computing education at the NSF centers should be included. Sarkar offered that industry representatives should be involved. Hey said that there are university data centers that should be represented. Bland said that ORNL has hired many people through the Computational Science Graduate Fellowship Program and that the leadership computing facilities (LCFs) do a lot of training; they should be involved, as well. Glotzer suggested that former recipients of fellowships now working at the national laboratories should be invited to join the Subcommittee.

Giles said that the Subcommittee will be formed in the next two weeks and go from there.

A break for lunch was declared at 12:01 p.m.

Afternoon Session

The meeting was called back into session at 1:34 p.m.

Jack Dongarra was asked to discuss the development of a new metric for ranking high-performance computing systems.

The LINPACK benchmark was started 36 years ago. In the late 1970s, the fastest computer ran LINPACK at 14 Mflop/s; floating-point operations were expensive compared to other operations and data movement. The matrix size was n = 100 because that was what would fit in memory. The LINPACK code is based on a “right-looking” algorithm.

The benchmarks have evolved from LINPACK to High-Performance LINPACK (HPL) to Top500. Over time, the LINPACK benchmark went through a number of changes. It began with Fortran code, and the code was run as is with no changes. In 1989, the benchmark was realigned toward peak performance with an unrestrained matrix size. In 2000, an optimized implementation of the benchmark, called High Performance LINPACK or HPL, was introduced. The user just needed to supply basic linear algebra subprograms (BLAS) and a message-passing library. In 1986, Hans Meuer started a list of supercomputers around the world and ranked them by theoretical peak performance. The listings were merged into the first Top500 list in June 1993.

There were a number of rules for the HPL and Top500: the machine had to compute the solution to a prescribed accuracy. The benchmark excludes the use of a fast matrix multiply algorithm like “Strassen’s Method.” They had to be algorithms that computed a solution in a precision lower than full precision (64 bit floating point arithmetic) and refined the solution using an iterative approach. The authors of the Top500 reserved the right to independently verify submitted LINPACK results and exclude computers from the list that were not valid or not general-purpose in nature (i.e., the computer system had to be able to be used to solve a range of scientific problems). Any computer designed specifically to solve the LINPACK benchmark problem or have as its major purpose the goal of a high Top500 ranking was disqualified.

HPL is a widely recognized and discussed metric for ranking high-performance computing systems. When HPL gained prominence as a performance metric in the early 1990s there was a strong correlation between its predictions of system rankings and the ranking that full-scale applications would realize. Computer vendors pursued designs that would increase their HPL performance, which would, in turn, improve overall performance for many applications. Today HPL remains valuable as a measure of historical trends and as a stress test especially for leadership-class systems.

However, HPL performance of computer systems is no longer so strongly correlated to real application performance, especially for applications governed by partial differential equations. There is a gap between what the benchmark predicts and how applications perform. That gap will increase in the future. A computer system with the potential to run HPL at an exaflop may be very unattractive for real applications. Future architectures targeted toward good HPL performance will not be a good match for most applications. This leads one to a think about a new benchmark.

The good aspects of HPL are that it is easy to run, understand, and check results; it stresses certain parts of the system; there is a historical database of performance information; it is a good community outreach tool; and it is understandable to the outside world. However, it has bad aspects, as well: It is 36 years old; the Top500 (HPL) is 20.5 years old; it is floating-point intensive; it is no longer so strongly correlated to real apps; and it reports only one value, peak flops, which encourages poor choices in architectural features. The overall usability of a system is not measured. It is used inappropriately as a marketing tool, leading to decisions on acquisition being made on the basis of one number, a crazy way to buy a system. And benchmarking for days wastes a valuable resource.

The really ugly things about HPL are that it does not probe the architecture; it produces only one data point; it constrains the technology and architecture options for high-performance-computing system designers; and such a floating-point benchmark is not quite as valuable to some as our data-intensive system measurements. There are many other benchmarks.

The goals for this new benchmark are to augment the Top500 listing with a benchmark that correlates with important scientific and technical apps not well represented by HPL. It would be good to encourage vendors to focus on architecture features needed for high performance on those important scientific and technical apps; to strike a balance of floating-point and communication bandwidth and latency issues; and to reward collective operations, point-to-point messages, local memory system performance, and intra-node parallelism. It is also desirable to provide an outreach/communication tool that is easy to understand;

optimize; and implement, run, and check results. It would be good to preserve the historical database of performance information. Any new benchmark should also have longevity.

The proposed new benchmark is the High-Performance Conjugate Gradient (HPCG), which solves a large, sparse, computed matrix with a known solution with dense and sparse computations; a dense and sparse collective; and data-driven parallelism. The benchmark needs to have strong verification and validation properties.

The model problem is a synthetic discretized 3D partial differential equation. It is a single-degree-of-freedom heat diffusion model with zero Dirichlet boundary conditions, synthetic right-hand side, and a single solution. The benchmark should have relevance to a broad collection of important apps; be simple with a single-number solution, few user-tunable parameters and algorithms; an algorithm that is a vehicle for organizing; and easy to modify.

The execution of the proposed benchmark is carried out in seven phases: problem setup, validation testing, reference sparse matrix-vector and Gaus-Seidel kernel timing, reference CG timing and residual reduction, optimized run set up, optimized computer graphics timing and analysis, and report results. The preconditioner is a hybrid geometric/algebraic multigrid with grid operators generated synthetically, grid transfer operators, a smoother, and a bottom solve. There are 50 iterations per set. The total benchmark time for an official result is a 1-hr run; anything less is reported as a "tuning" result. The coarsening is 2x – 2x – 2x (8x total); the number of levels is four; and there are one pre-smoother sweep and one post-smoother sweep.

The merits of HPCG are that it includes major communication/computational patterns; rewards investment in high-performance collective operations, local-memory-system performance, and low-latency cooperative threading; detects/measures variances from bitwise reproducibility; and executes kernels at several (tunable) granularities.

Tuning runs on the K computer showed that some optimization produces a speedup of a factor of 10. Other tests indicated that multi-node scaling is very smooth.

There is no proposal to eliminate HPL as a historic metric. HPCG will serve as an alternate ranking of the Top500 to reflect more what these machines can do with real applications. Feedback has been solicited from various people at the national laboratories. Discussions have been held with and results received from every HPC vendor and with most LCFs. Discussions were also held at the Intel-sponsored SC14 Workshop on Optimizing HPCG. Several reports have been issued on the topic.

Giles asked how Dongarra expected this benchmark to affect the LINPACK relative order of systems. Dongarra guessed that it would reorder the top 25, rewarding computers that dealt with latency; the Japanese K and Blue Gene computers would likely rise in the rankings.

Berger pointed out that, years ago, Phil Colella had suggested seven types of applications to benchmark computers and asked why a suite of applications was not used or why different benchmarks were not used for each part of the machine. Dongarra replied that that had been considered and other benchmarks had followed that path. However, one ends up with a collection of numbers, and what one wants is just one number as the output from the benchmarking effort so the computers can be ranked.

Hey said that, with the success of LINPACK, he had learned that a one-number output is good. The system described is a good, more realistic expansion of LINPACK.

Cerf pointed out that no one measure is going to indicate what machine to buy. Any expansion of the concept of performance is welcome.

Giles noted that the fact that both versions are expressed in flops may lead to a false equivalency. Dongarra said that anything that is measured in flops is not good, either.

Paul Messina was asked to describe the Argonne Training Program on Extreme-Scale Computing.

The Argonne Training Program in Extreme-Scale Computing (ATPESC) provides training on (or at least exposure to) the key skills, approaches, and tools necessary to design, implement, and execute computational science and engineering (CS&E) applications on current high-end systems and the leadership-class systems expected to be available in in the next 10 years. The training features lectures by leading computer and computational scientists in the topics covered and extensive hands-on exercises on leadership-class computers.

With the challenges posed by the architecture and software environments of today's most powerful supercomputers (and even greater complexity on the horizon from next-generation and exascale systems) there is a critical need for specialized, in-depth training for the scientists poised to facilitate breakthrough science and engineering using these computational resources.

Computational research at the high end requires expertise in many areas. The architecture and software environments of today's most powerful systems will feature far greater parallelism and complexity of architectures than is extant today. Computational research projects tackle ever more complex projects.

The ATPESC covers many of the key skills that computational scientists and engineers need to carry out their research on leadership-class systems. In two weeks it is not possible to present detailed material on all topics; in some cases we aim to make participants aware of the existence of certain approaches, methods, and tools.

What does one need to teach people in high-and computing? The list includes computer architectures, mathematical models, numerical algorithms, programming methods that are effective across a variety of today's supercomputers, multiple approaches on unifying concepts and levels of abstraction that provide migration paths and performance portability, and approaches to building community codes.

Curriculum tracks include architectures, programming models and languages, mathematical algorithms and software, software engineering in scientific computing, performance and debugging tools, visualization and data analysis, data-intensive computing, community codes, and application case studies.

The training program is held at the Pheasant Run Resort in St. Charles, Illinois, which is fairly isolated, older, and inexpensive. The participants are kept hostage for two weeks. Participants pay no registration fee; and domestic airfare, meals, and lodging are provided for the participants.

The program begins mid-afternoon on a Sunday and runs from 8:30 in the morning to 9:30 at night every day except for free time on Saturday afternoon and Sunday. The program ends at noon on the second Friday.

The target participants are doctoral students, postdocs, and computational scientists. Developers of programming models, algorithms, and tools for leading-edge systems may also apply. Applicants must have substantial experience in MPI and/or OpenMP programming, have used at least one HPC system for a reasonably complex application, and plan to conduct CS&E research on large-scale computers. About 60 participants are selected.

Applicants submit through a website. In 2013, 160 applications were received. A committee of five selected 63 participants, including 1 undergraduate student, 32 graduate students, and 12 postdocs. There were 11 national-laboratory employees.

Slides for all lectures can be downloaded from the 2013 agenda page. Videos of the dinner talks will be online soon. Series of lectures were presented on supercomputing architecture trends, programming models for high-performance computing, accelerator and GPU programming, mathematical software and numerical algorithms, toolkits and frameworks, community codes an application case studies, visualization and data analysis, and data-intensive computing and I/O.

Extensive hands-on access was provided to leadership-class computers, including Vesta, Mira, Tukey, Titan, and Keeneland KIDS. In addition, during evening hands-on sessions, lecturers and their assistants provided individualized help and tutoring to the participants.

A dinner talk was presented each evening, touching on such topics as computing on demand; how I learned to stop worrying and blow up stars; large-scale visual analysis; software carpentry; teaming at the DOE national laboratories; dawn of the era of quantum computation; the Urban Center for Computation and Data; Argonne National Laboratory; and big data plus extreme-scale time to compute produces actionable insights.

The participants repeatedly said that they planned to use what they learned to inform their research and to transfer knowledge to their home teams.

Giles asked if there really was an exam on the last day. Messina replied, yes: write a few pages on the three most important things you got out of the program that you will apply to your work.

Cerf asked if this type of training session would be suitable for massively open online courses (MOOCs). Messina did not know. Interaction is a main component of the program.

Glotzer said that this program sounds fantastic. Every summer for the past 5 or 6 years, the University of Michigan has been running several 3- to 8-day summer schools on high-performance computing topics. A hands-on experience was found to be important. Several institutions were enlisted that could host six or seven students and provide a local teaching assistant. She asked if this strategy could be used to scale the Argonne program. Messina replied that it might, promising to look into it.

Meza asked if there had been any thought given to forming a community of alumni. Messina said that some alumni had set up a Facebook page, but the effort is very informal.

Andrew Siegel was asked to describe the Center for Exascale Simulation of Advanced Reactors (CESAR) program on codesign at ANL.

CESAR works with industry and DOE research partners to influence the design of future hardware architecture, system software, and applications based on the key algorithms underlying computational nuclear engineering. The physics are fairly general. CESAR also seeks to develop a new generation of underlying algorithms that enable the solution of significant outstanding nuclear engineering problems by leveraging exascale resources.

In nuclear engineering, the key physics are computational fluid dynamics (CFD) and neutron transport, optimized in parameter regimes relative to reactor simulations. CESAR's nuclear reactor coupled neutronics/hydraulics problem spans scales from 1 cm to 14 m. It is a big heat-transfer problem. The data structure needed to be stored is huge. The neutronics and CFD have to be meshed differently with vastly different properties.

The work was started with foundational codes that had been worked on for decades. There are two foundational codes for neutron transport: Monte Carlo and UNIC, a deterministic unstructured neutron-transport code. Nek5000, an incompressible spectral-element code, is used for CFD.

The codesign process consists of

- Working with the community and vendors to develop/leverage exascale straw-man models/hardware simulators
- Identifying application bottlenecks in depth as they exist on current platforms with basic extrapolations to exascale-type architectures
- Developing proxy applications to abstract key performance features, continuously evolve, continuously verify, identify/document verification bounds, be user friendly, be distributed robustly to the community
- Working with vendors to port to next-generation simulators and guide the design changes

The process is dynamic. One cannot simply throw something over the transom. One must participate actively in algorithm codesign. Three critical issues arise: as bottlenecks are encountered, one must identify where algorithmic compromises can be made; one must verify the relevance of results for the full application; and one must be open to new mathematical formulations.

Monte Carlo methods make one feel like one knows what one is doing very quickly. One can get incredible production quickly. But there are 50 billion particles being tracked that have 35 to 40 interactions that the user has to calculate for 50 GB of cross-section values over 9 orders of magnitude in energy.

In these programs, about 85% of the code is just looking things up. On existing systems, one cannot do a Monte Carlo analysis of a reactor.

Open Monte Carlo does not scale perfectly, even from 1 to 16 cores. One spends a good deal of time looking up macroscopic cross-sections. It is simple (only about 1000 lines of code) and easily run.

If one can get XSBench to run well, one will get Open Monte Carlo to work well, too. This process mimics real-world computation without the burden of large data files and obfuscating "last-mile" science. It provides a simplified basis for assessing/driving exascale architectures, refining and optimizing the existing cross-section lookup algorithm, and facilitating the development of heterogeneous cross-section algorithms. The faster the processor, the more CPU cycles will be wasted while waiting for outstanding reads. Energy savings can be achieved simply by reducing the CPU clock speed.

The flop/load ratio is altered by adding “dummy flops” to every microscopic cross-section lookup. NVIDIA, AMD, and IBM got very interested in this mini-app.

Anthony Scudiero of NVIDIA ported XSBench to CUDA to be run on their current generation graphics processing units for general-purpose computing (GPGPUs), carrying out “naïve,” “mainline,” and “export” ports. A speedup factor greater than 10 was achieved with a bandwidth of 128.5GB/s. He recently ran everything on the Gryphon simulator with very interesting results indicating that the platform may be easier to optimize than anticipated. Sergey Blagodurov of AMD showed that using TLMSim can more than double the bandwidth by identifying good data structures for caching.

In coding, for full-scale reactor simulation, 100 GB of cross-section data may be required. XSBench has XL and XXL modes that model cross-section grids of this size. AMD also has expressed interest in evaluating the feasibility of storing tally data (more than a terabyte in size) in NVRAM and has requested a mini-app as a basis for testing.

All of this work led to considering whether there are fundamental algorithmic changes that would be a better fit for exascale-type nodes. The question arose of using a large number of flops per second to replace random-memory lookups. It is a losing proposition on single-core traditional nodes, but what about candidate exascale-type nodes?

This led to a whole new class of ways of solving this problem. One does not need to store all of the data. One can rebuild the cross-section rather than storing and finding it. One can store only poles and residues for each resonance (about 100 Mb for 400 nuclides).

The ACE [A Compact ENDF] method was used to treat 21,200 neutrons per second. On the basis of that work, a new mini app, RS Bench, was developed to abstract key performance characteristics of the multipole-on-the-fly method. It is an open question whether this will become part of the codesign process.

Hey asked what was a candidate exascale-type node. Siegel replied that the amount of memory will be much less. There are several straw-man architectures. There is no new technology. The use of the parameter space will tell one where there are going to be problems.

Giles asked if resiliency were addressed. Siegel replied, yes. The Monte Carlo process allows one to throw out a particle if the neutron data are bad. It has to be taken care of at the middleware level.

A break was declared at 3:37 p.m., and the meeting was called back into session at 3:50 p.m.

Alexandra Landsberg was asked to present the program response to the Applied Mathematics Committee of Visitors (COV).

A committee of visitors (COV) surveyed (1) the efficacy and quality of the processes used by the Applied Mathematics Program to solicit, review, recommend, and document proposal actions and (2) the overall breadth and depth of the Applied Mathematics research portfolio. The COV reported a summary finding and a series of recommendations. The Program has now considered and responded to all of those recommendations.

The summary finding of the COV was that the Applied Mathematics Program is highly effective in its process to solicit, review, recommend, and document proposal actions. The Applied Mathematics program managers do an excellent job of monitoring all aspects of their portfolios. The overall breadth and depth of the Applied Mathematics portfolio is excellent. ASCR has a long tradition of supporting some of the best applied math research in the nation and also maintains an international leadership position in certain key mathematical areas.

The recommendations of the COV and the responses of the Program were as follows:

The COV recommended that the program managers continue to look for ways to enhance the program’s ability to attract new investigators, while seeking to maintain the overall excellence of the program. The Program responded that ASCR will continue to look for ways to attract new investigators. Currently, ASCR attracts new investigators through participation at DOE workshops and professional conferences such as Society for Applied and Industrial Mathematics (SIAM). Outreach to other professional societies such as American Statistical Association (ASA) has recently been initiated. Another mechanism to enhance awareness for researchers unfamiliar with ASCR is through participation on

ASCR review panels and electronic reviews. Enhancing ASCR webpages can also help better inform new investigators regarding the DOE Applied Mathematics program.

The COV recommended that program managers be allowed to travel as needed to scientific meetings. The Program responded that ASCR agrees that program managers should travel to scientific meetings. Given that travel funds are limited, program managers provide a prioritized list of the conferences and technical meetings they would like to attend. ASCR senior management then manages the travel funds.

The COV concurred with the 2011 report in recommending that the focus of the Computational Science Graduate Fellowship Program be expanded and funding doubled during the next 5 years. The COV also recommended that the program remain within ASCR and not be moved to the National Science Foundation. ASCR thanked the Committee for this comment.

The COV recommended that award rates for Applied Mathematics Program solicitations be made publicly available. ASCR will update the Applied Mathematics web pages and provide this information for solicitations in 2013 and thereafter.

The COV recommended adding an annual center-directors' meeting in order to enhance linkages among the three Mathematical Multifaceted Integrated Capability Centers (MMICCs). ASCR will add an annual center directors' meeting.

The COV recommended instituting the use of a standard reporting format for the annual progress reports, specifying the length and description. ASCR agreed that a standard reporting format is useful. A federal-wide standard format is being developed, and ASCR will implement that format as quickly as possible. For the DOE national laboratories, an 8-page standard format has been developed for the DOE national laboratories' progress reports.

The COV recommended that ASCR develop a short-term visitors program with the MMICCs, with concomitant funding, to enable promising researchers to develop collaborations with center members. ASCR will consider this recommendation, discuss it with the MMICC center directors, and examine other visitors' program to determine an appropriate implementation of such a program and the associated costs. Assuming strong interest from the MMICCs and availability of funds, ASCR may implement this recommendation in FY15.

The COV recommended investigating the addition of a new interdisciplinary program of applied mathematics−statistics–computer science−facilities that could drive the next generation of fundamental research broadly applicable to the analysis of experimental/observational facilities data. ASCR will consider this recommendation as part of ASCR strategic planning activities. Since the COV, a solicitation on mathematics and statistics for DOE data research projects was released, and six projects were funded. A workshop was held for these researchers with personnel from five SC user facilities.

The COV recommended that ASCR continue its outreach efforts to professional societies and research communities through sponsored workshops and conference attendance. ASCR will continue to pursue outreach to the applied mathematics community pending available funds for travel and for DOE-sponsored workshops.

The COV recommended that the review and annual reporting process for the MMICCs include a listing of awards and accolades received by the project participants to highlight the leadership role of the MMICCs. ASCR agrees with this recommendation, and an annual review of the MMICCs was implemented around the same time as the COV. These recommendations will be incorporated into the annual reporting requirements of the MMICCs.

The COV recommended that the Applied Mathematics Program develop a set of key mathematical areas that will have the greatest impact on the DOE mission and in which they can either currently claim or plan to develop international leadership. Since the COV, ASCR's Applied Mathematics has defined a set of key mathematical areas (i.e., core strengths), which are considered to have to the greatest impact on the DOE mission: numerical partial differential equations; linear algebra and scalable solvers; multiscale, multiphysics, and multicomponent methods and models; optimization; stochastic systems and uncertainty quantification; and mathematics and statistics for DOE data. The Office is currently looking for leadership metrics to numerically support the claims to leadership.

Meza stated that, on the last step, who one considers one's peers would indicate one's standing in the international community.

Hey asked what the term of the grants was. Landsberg answered, 3 years.

Berger asked how often the center leaders would convene. Landsberg replied, once a year in person; three times a year via telephone. Berger suggested that that might be burdensome. Meza replied that the Subcommittee was thinking of one or two face-to-face meetings.

Giles asked what the picture was of ASCR's ability to interact with the community. Binkley replied that there is a line item for travel in the budget. There has been downward pressure from the General Services Administration (GSA) and OMB for the past 5 to 7 years. He was of the opinion that the corner has been turned on that issue. The program managers need to get out and interact with the community. Barbara Helland pointed out that participation in conferences by federal employees (including those at national laboratories) all rolls up into one number. Binkley noted that the Office was seeing about 600 requests for travel a year. In the past 18 months, a lot of conference restrictions have been codified into law.

Stephen Binkley was asked to present the proposed ASCAC charge for a COV to SciDAC.

Version 3 of SciDAC is being wrapped up. It has been 7 years since the program has been reviewed.

To ensure the integrity of this research program and to ensure that it is meeting the challenges of the DOE mission, the lender charges the Committee to assemble a COV to review the management processes for the SciDAC portfolio. The COV should provide an assessment of the efficacy and quality of the processes used to solicit, review, recommend, and document proposal actions and those used to monitor active projects and programs during the fiscal years 2011, 2012, and 2013 for both the DOE laboratory projects and the university projects. The COV is also to comment on how the award process has affected the breadth and depth of portfolio elements, the degree to which the program is anticipating and addressing emerging challenges from high-performance computing and DOE missions, and the national and international standing of the program with regard to other computational-science programs. The COV is to report back at the August 2014 ASCAC meeting.

Giles announced that a subcommittee would be formed, and he asked for volunteers. There should be good representation from other SC offices. This group may be larger than the usual five-member COV.

Cerf commented that it might be good to have someone familiar with peer review from NSF on the Subcommittee.

Carolyn Lauzon (ASCR) was asked to review the ASCR Leadership Computing Challenge (ALCC).

ALCC originated in 2010. It allocates time at the ASCR computing facilities, the LCFs and NERSC. ALCC complements the other allocation mechanisms, the Innovative and Novel Computing Theory and Experiment Program (INCITE) and the Energy Research Computing Allocation Process (ERCAP). ALCC has a DOE focus. It allocates 30% of the time at the LCFs and 10% of the time at NERSC. It is a peer-reviewed process managed by ASCR. Awards are for 1 year. There were 39 awards in 2013 with a mean project size of 47 million CPU hours. The median request size has grown from 20 million CPU-hours in 2010 to 62,000,000 CPU hours in 2014. The largest request has grown from 223 to 800 million CPU-hours during that same time. Both the requests and the available time have increased about 60% during that time, also. In 2013, 58% of the requested hours were allocated.

A study was conducted of the science that drives the growth of the ALCC. The proposals submitted were binned by topic area and by year: five bins for the DOE applied offices, six bins for the basic science offices, and an "other" bin. Crosscutting proposals were assigned to the best match. If a proposal identified a DOE funding source, the funding was used to bin the proposal. In all other cases, PI affiliation, institution, and proposal topic area were used to bin the proposals. Fundamental science is the dominant science driver.

In the applied-science topics, the "other" category proposals were related to turbomachinery, aeronautics, and applied turbulence. ARPA-e proposals were related to heat capture for concentrated solar panels and to protein engineering. Nuclear Energy (NE) proposals were related to the Consortium for Advanced Simulation of Light Water Reactors (CASL), the Nuclear Energy Advanced Modeling and Simulation (NEAMS), and reactor safety and efficiency. NNSA proposals were related to laser plasma fusion and turbulence. Energy Efficiency and Renewable Energy (EERE) proposals were related to wind turbines, additive manufacturing, hydropower, and engine efficiency. Fossil Energy proposals were related to carbon capture and sequestration. The NNSA was the biggest user, NE was the second biggest user, EREE is growing, and ARPA-e is just showing up.

In the basic-science topics, ASCR proposals were largely related to SciDAC; HEP proposals were largely related to LHC modeling; FES proposals were largely related to ITER [International Thermonuclear Experimental Reactor] simulations; BES proposals were largely related to materials science and to chemistry; BER proposals were largely related to climate and to subsurface modeling; NP proposals were largely related to quantum chromodynamics; and "other" proposals related to basic science turbulence, earth science, weather, protein modeling, and blood flow. The biggest user was BES with FES and BER growing.

Since the ALCC's beginning, 77 PIs have been awarded time. The number of new users awarded time is decreasing. 13% of ALCC PIs later received INCITE awards. Historically, most of the users came from national laboratories, with university users increasing rapidly. In 2013, universities received 26 awards, national laboratories received 9, and industry received 4.

Giles noted that one objective of ALCC was to increase the opportunity to gain access to the LCFs by having more than one solicitation per year and another was to bring in more users. He asked if these objectives were being pursued. Lauzon replied affirmatively.

Meza asked what size jobs the users submit. Lauzon replied that those data are currently being gathered.

Cerf asked if the reports produced by this research were available to the public. Binkley replied that all ASCR publications are on the web. Chalk added that there is a software repository at the SciDAC centers. The Information Bridge of the Office of Scientific and Technical Information (OSTI) provides access to all DOE literature.

Giles asked if there were anything coming out about the availability of DOE data. Biven said that a memo had come out from OSTP that all proposals must have a data-management plan and that data have to be archived with public access assured. On the publication side, publishers have to make the manuscript of record available within 12 months of publication. OSTI has an agreement with the publishers. OSTI would use a proposed tool, PAGES [Public Access Gateway for Energy and Science], to obtain metadata from publishers about DOE-funded publications and to provide links to the online versions of those original publications. As designed, PAGES allows keyword (including author name) searches to retrieve abstracts, metadata, and links to full articles (the version of record) when available. A prototype of PAGES has been developed.

The floor was opened to public discussion.

Hey commented that some publishers are not making publications publicly available. Biven answered that OSTP is happy with the manuscript of record being made available by publishers, authors, authors' institutions, OSTI, or any other information-sharing mechanism or institution.

There being no additional discussion, the meeting was adjourned at 4:57 p.m.

Respectfully submitted,
Frederick M. O'Hara, Jr.
Recording Secretary
April 21, 2014