Draft Minutes
Advanced Scientific Computing Advisory Committee
September 20-21, 2016
Holiday Inn Capitol, Washington, DC 20024

ASCAC Members Present

Martin Berzins
Keren Bergman (September 20 only)
Barbara Chapman
Silvia Crivelli
John Dolbow
Jack Dongarra
Thom Dunning
Sharon Glotzer
Susan Gregurick
Anthony Hey
Gwendolyn Huntoon
David Levermore
Juan Meza
John Negele
Daniel Reed (Chairperson)
Vivek Sarkar
Dean Williams (via telephone)

ASCAC Members Absent

Vinton Cerf
Linda Petzold

Also Participating

Wes Bethel, Lawrence Berkeley National Laboratory
John Steven Binkley, Associate Director, Office of Advanced Scientific Computing Research, Office of Science, DOE
Linda Blevins, Office of Science, DOE
Christine Chalk, ASCAC Designated Federal Officer, Program Manager, Oak Ridge Leadership Computing, Office of Advanced Scientific Computing Research, Office of Science, DOE
Tiffani R. Conner, Oak Ridge Institute for Science and Energy
Ping Ge, Office of Science, DOE
James Glownia, Office of Science, DOE
Robert Harrison, Stony Brook University and Brookhaven National Laboratory
Barbara Helland, Director, Facilities Division, Office of Advanced Scientific Computing Research, Office of Science, DOE
Mark Johnson, Office of Energy Efficiency and Renewable Energy, DOE
John Labarge, Director, Office of Lab Policy, Office of Science
John L. Sarrao, Los Alamos National Laboratory
Deniese Terry, Oak Ridge Institute for Science and Energy

Tuesday, September 20, 2016
Morning Session

The U.S. Department of Energy (DOE) Advanced Scientific Computing Advisory Committee (ASCAC) meeting was convened at 8:30 a.m. EST on Tuesday, September 20, 2016, at the Holiday Inn Capitol by **Chair Daniel Reed**. Reed reviewed the agenda and asked ASCAC members on the telephone to introduce themselves.

Deneise Terry made safety and convenience announcements and announced that the meeting was being recorded.

Steve Binkley was asked to report on the activities of the Office of Advanced Scientific Computing Research (ASCR) and provided ASCAC with an overview of the agenda topics. Binkley then turned his presentation to the budget situation and ASCR updates. The Fiscal Year (FY) 2017 budget will operate on a continuing resolution for at least three months due to the impacts of House and Senate Energy and Water Development subcommittees' FY2017 Marks. The ASCR investment priorities continue to focus on exascale computing, facilities, big data, and research and development (R&D) for the Post-Moore era.

Support for basic and applied research for the Office of Science (SC), Exascale Computing Project (ECP) and Scientific Discovery through Advanced Computing (SciDAC) partnerships will be recompeted in FY2017. The Leadership Computing Facilities continue their preparations for the planned 75-200 petaflops (Pflops) upgrades at each site. The National Energy Research Scientific Computing Center (NERSC) will begin operation of the NERSC-8 supercomputer. There is a modest effort in R&D for post-Moore's Law computing, a modest effort in support of the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative included, in collaboration with Basic Energy Research (BER) and Biological and Environmental Sciences (BES) offices. Finally, the Computational Sciences Graduate Fellowship will be funded at $10 million.

The ASCR's FY2017 budget in the President's request to Congress is $663,080 million (M), a 6.8% increase over the FY2016 enacted budget. The FY2017 House Appropriations Mark is $621 billion (B), down approximately $40M, and the Senate Appropriations Mark is $656B, down about $7M. The FY2017 budget contained prescriptive language, from both the House and Senate, stipulating that certain conditions had to be fulfilled, thus focusing how the SC can balance cuts to areas of the overall program. ASCR is currently waiting on the House and Senate to resolve the differences. If the budget ends up in between or close to the House Mark then the research activities could be significantly hampered. The SC Associate Directors held meetings last week with the House and Senate Energy and Water Development subcommittees discussing the impacts across the entire SC portfolio with respect to the current budget Marks and made recommendations back to the subcommittees as to the appropriate courses of action.

Binkley paused for questions about the budget from ASCAC members. Levermore suggested that in tighter times activities that benefit the long-term health must be protected. Levermore asked if Binkley wanted advice from ASCAC on the priorities. Binkley indicated the allocations will depend on the restrictions placed on SC. For example, if SC has direction or guidance to fund certain activities departmental policy dictates SC must follow that. Reed suggested that there are other people who could lobby for an outcome and Reed encouraged folks to take that observation to heart.

Dr. Abani Patra is the new program manager for Applied Mathematics. Dr. Patra was most recently a professor at the University at Buffalo and has been a Program Director for the Office of Cyberinfrastructure at the National Science Foundation (NSF).

Cori (NERSC) and Theta at Argonne National Lab (ANL) are both in acceptance and deliveries have occurred. Phase I and Phase II of CORI are being merged and acceptance testing is occurring on the Theta system at ANL.

SciDAC-4 Institutes have been extended through March 2017. There are SC internal discussions concerning Funding Opportunity Announcements (FOAs) in FY2017. Active discussions have begun with domain programs (BES, BER, Fusion Energy Sciences (FES), High

Energy Physics (HEP), and Nuclear Physics (NP)) concerning SciDAC-4 relative to SciDAC-3 and are taking stock of Sci-DAC's history and the best formulation of a partnership.

Bill Harrod is leading the ASCR long-range planning (LRP) in conjunction with eight labs that have ASCR activities. The LRP is a precursor to developing an ASCR strategic plan which will look more holistically at the ASCR program. Binkley reiterated the moratorium on travel that is in effect during the first quarter of FY2017 and assured ASCAC that it does not affect travel by national lab personnel or grantees. Critical Decision 0 (CD-0), for the ECP, was approved by Deputy Secretary Sherwood-Randall on July 28, 2016. The Deputy Secretary gave guidance to 1) develop options for earlier delivery of exascale systems and 2) to complete CD-1/3A by November 2016. The CD-1 pre-review was conducted September 13-15, 2016 at ANL and indicated that the cycle cost analysis needed additional details. However, major project features including the project approach, project team, and quantitative cost estimates were endorsed.

A series of workshops have been held on requirements gathering beginning in June 2015. There was one workshop per SC Program Office concerning the implementation of Exascale Requirements Review; the last one is the ASCR review and will be held at the end of September 2016. The final review is similar to a capstone event and will be considering information from the previous reviews to pull everything together. Binkley mentioned ASCAC could look forward to an update at the December meeting.

There is a new position in SC listed at USAJOBS. This position is the Director, Office of Advanced Simulation and Computing (ASC) and R&D National Nuclear Safety Administration (NNSA). The posting will be open for 21 days and Binkley asked ASCAC to encourage qualified colleagues to apply.

Gregurick indicated concern that the moratorium for federal travel will impact strategic planning capabilities. Binkley said that there is some flexibility. SC has modified the approval authority up to Pat Dehmer, DOE, Deputy Director for Science Programs in the Office of Science, who issued guidelines on what she will approve as the situation is austere given its seriousness. The moratorium is not only impacting planning activities but also the ability for the staff to visit the National Laboratories (labs) and attend scientific meetings, both of which are critical to keeping abreast of developments.

Sarkar asked if other organizations within the U.S. government who are involved with the National Strategic Computing Initiative (NSCI) will face similar constraints and what impact that might have on the NSCI as a whole. Binkley was aware that the NSF did not experience similar issues in their budgets and the parts of NSCI that NSF supports should not be impacted. Requests in the President's FY2017 budget for NSCI participating agencies were up across the board. Sarkar said information on other organizational budgets would be helpful.

Reed asked what ASCAC needs to do to collectively advance the NSCI collaboration across agencies. Binkley said the interagency working group, who is overseeing NSCI, completed the implementation plan, which evolved into a strategic plan that was issued last summer (July 2015). Beginning October and running through FY2017, the focus is on developing a better methodology for public-private partnerships, considering innovative approaches to expand out into partnerships with industries and universities. Specific ideas will be shared at the December ASCAC meeting, at which time it would be useful for ASCAC to critique those and make recommendations.

Alexander Lazelere (audience) asked about the status of the potential collaboration between National Institutes of Health (NIH) and DOE. Binkley referred to Rick Stevens' presentation at the April 2016 ASCAC meeting concerning the cancer DOE-National Cancer Institute (NCI) pilot. Funding is in place for the pilot programs and funds are coming from both DOE sources and the NCI. The cancer pilots have been extended in the direction of involving industry in the sense that GlaxoSmithKline is now a partner and SC is working with the vendors.

Kathy Yelick was asked to provide information on the Superfacility for Data-Intensive Science. The old school scientific workflow progressed through individual that required a scientist to travel to the experimental site, conduct tests and gather data, then return home to analyze the data and write an article. With regard to high performance computing (HPC), the old school model focused on simulation which is still seen in the language in the NSCI. Old school scientific data search now requires search tools to find scientific data.

The new scientist is outside the normal system. New scientists have access to computing, figure out how to develop advanced algorithms, and have access to scientific data which can be downloaded from the internet. In new science the ability to remotely conduct experimental science is available and new models for HPC include data analysis. The new science has a vision of automated search and automated meta-data analysis, and HPC facilities need to focus on On-Demand Simulation.

By 2024 ESnet will be sending 100 exabytes (EB) of data per year. A Superfacility demonstration was conducted at the SLAC Linac Coherent Light Source (LCLS) where NERSC and ESnet tripled the network traffic, indicating that the network must be provisioned to handle the datastream.

Regarding the overall architecture for a Superfacility, the efficiency achieved from centralized computing centers only happens if the data can get in and out of the facilities for the experimental/observational data sets. The efficiency garnered from the centralized computing centers as well as configured systems for data-intensive science must be considered. NERSC in combination with Cray has developed the ability to containerize software.

In applied mathematics, scientists are considering designing mathematical algorithms to allow real-time analysis (a front end processing problem), new algorithms to automate analysis problems (such as machine learning algorithms), inventing new mathematical models to match new technologies (such as in the light sources and experimental measurement devices), and multi-modal analysis (combining data sets from different sources). There are similarities between data analysis and simulation and following a spectrum from memory intensive to computationally intensive problems.

Research problems exist in computer science to get the scale needed to send large amounts of data across the internet. In analytics, people want interactive access to data and to play with the data while running experiments. The Jupyter project allows people to build sophisticated interactive notebooks and embed code that can be executed.

Data analysis does not necessarily require new machines. Yelick discovered that the algorithms in genome assembly can run on an HPC system with a low latency low overhead network and the ability to do one-sided communication. It is possible to scale up the algorithms to tens of thousands of cores and the running time up by a factor of 100 or more.

Jack Dongarra and Daniel Reed wrote an article, “Exascale Computing and Big Data” (doi: 10.1145/2699414) on the differences in the software stack between simulation and data analysis. One popular software stack for data analysis is SPARK. To make the SPARK analytics stack run on the HPC system there must be a good network implementation and a fast Input/Output. A disadvantage of the HPC system given the SPARK model is there is no local disk, thus the experiment and check pointing occur on a remote filesystem. Despite the disadvantage, the failure rate is much lower.

The scientific process is poised to undergo a radical transformation based on the mathematics available in machine learning, other algorithms, and data rates coming from experimental devices.

Levermore asked if there has been consideration of how someone might take advantage of the network that was not anticipated. Yelick stated that while network security is a concern, the complication is that the network cannot be shut down or closed off because that defeats the purpose of a Superfacility. The strategy at Lawrence Berkeley National Laboratory (LBNL) is to monitor and prevent people from corrupting the data or taking over the system, but not to have hard firewalls that keep people from getting into the systems. In the execution plan for the Superfacility calls for picking a science problem or a facility at a time and build a computing, networking, and experimental facility for that specific problem.

Sarkar asked about the willingness to share data across different domains of science. Yelick indicated there are different cultures expressing different level of willingness to share data.

Chapman asked about organizational challenges and if any challenges have been foreseen. Yelick indicated organizational challenges are probably the biggest impediment.

Reed called for a break at 10:03 a.m.

The meeting was reconvened at 10:19 a.m.

Doug Kothe provided an update on activities in the ECP Application Development focus area.

DOE is a lead agency within NSCI and SC and the NNSA will execute a joint effort on advanced simulation through a capable exascale computing program emphasizing sustained performance on relevant applications and data analytic computing. The ECP is the vehicle for that effort. To achieve capable exascale requires a holistic approach that includes Application Development, Software Technology, Hardware Technology, and Exascale Systems. The ECP mission need includes to support DOE science and energy missions, to meet national security needs, and to address key science and technology challenges.

Exascale application driver(s) for Wind Energy is wide-scale deployment of unsubsidized wind plants hampered by large plant-level energy losses, currently at ~20%; for Nuclear Energy are understanding and predicting fuel failure and core damage in severe reactor accidents, near real-time load-following core simulator supporting onsite operating plant decisions, and engineering scale predictions of nuclear fuel performance and barriers to higher burnup; for Magnetic Fusion Energy is to prepare for and exploit ITER and other coming international major experiments such as Joint European Torus –Deuterium and Tritium (JET-DT), Japan Torus-60 Super Advanced (JT-60SA), and Wendelstein 7-X (W 7-X); for Advanced Manufacturing are advance quality, reliability, and application breadth of additive manufacturing (AM), accelerate innovation in clean energy manufacturing institutes, and capture emerging manufacturing markets; and for Combustion Science and Technology is optimizing combustion-based systems for energy efficiency and reduced emissions.

The scope of Application Development is to deliver science-based applications able to exploit exascale for high-confidence insights and answers to problems of National importance. The mission need is to create and enhance applications through: development of models, algorithms, and methods; integration of software and hardware using co-design methodologies; improvement of exascale system readiness and utilization; and demonstration and assessment of challenge problem capabilities. The objective is to deliver a broad array of comprehensive science-based computational applications that effectively exploit exascale HPC technology to provide breakthrough modeling and simulation solutions for National Challenges in scientific discovery, energy assurance, economic competitiveness, health enhancement, and national security.

The first request for information (RFI) for the Application Suite was released on May 31, 2015 and focused on all 17 DOE Labs yielding 135 responses received. The second RFI was issued September 15, 2015 to NSF and NIH to identify scientific research topics in need of 100-fold HPC performance on scientific applications. The remainder of the presentation focused on the 135 responses.

The process for selection of the initial ECP applications included gathering and analyzing concept white papers, reducing the candidate set of exascale application concepts to a candidate set of application projects, and selecting ECP applications from list of candidate application projects. The final candidates were asked to submit development plans which included the challenge problem, the exascale system driver, expected outcomes and figures of merit, project team, challenges and risks, resource allocation, data management, and integration. Applications were reviewed based on quality of the application's challenge problem(s), quality and makeup of project team, the technical plan, the need for capable exascale systems to address the challenge problem, and the technology Impact outside of its domain of focus.

Fifteen application projects and seven seed efforts were initially selected. The 22 projects crossed six strategic pillars including national security, energy security, economic security, scientific discovery, climate and environmental science, and health care.

There were 13 Application Motifs, or Algorithmic methods, mentioned in the applications. New data motifs included combinational logic, graph traversal, graphical models, finite state machines, dynamic programming, and backtrack and branch-and-bound. The applications focused mostly on classical motifs.

Eleven Application Co-Design (CD) Centers are currently being considered. CD Centers are expected to emit best practices, lessons learned, and methods for the applications. Next generation community libraries at a higher level than Linear algebra are desired from the committee. The review criteria for CD applications are similar to criteria for application development.

There are exascale application development projects for current DOE Science & Energy projects (10), current DOE Science & Energy and Other Agency projects (5), and the seven seed projects. An aggregate of technologies were cited in all the ECP applications were in five categories: programming models and runtimes tools (debuggers, profilers, compilers), mathematical and scientific libraries and frameworks, data management and workflows, and data analytics and visualization. Applicants were asked to prioritize the technologies listed in the application and to rank the technology in three categories: use, explore, or curious.

In projectization milestone-based management can impart a sense of urgency and enforce accountability, projectization can improve breadth and depth of science output, projectization forces communication when it is needed, projectization helps to mentor and train next generation leaders, and projectization requires active risk management when it is often an oversight.

Risks and challenges included programming models: what to use where and how, developing and integrating co-designed motif-based community components, infusing data science applications and components into current workflows, achieving portable performance, multi-physics coupling, and integrating sensitivity analysis, data assimilation, and uncertainty quantification technologies.

Next steps in the ECP Application Development activity are to solicit exascale application ideas and project plans from academia and industry by targeting specific application areas representing gaps in current project suite relative to ECP objectives, increase application focus on national security and data science problems, accelerate current application projects through startup and into a productive cadence, identify application team "weak spots" for targeted training and productivity improvement, publish initial version of exascale application requirements, select projects for CD centers of key application motifs, continue to engage community and stakeholder agencies/programs on priorities and impact, proactively engage data science and AI application community, and develop a sense of community.

Berzins asked about resilience. Kothe indicated the team is focused on a technology watch from the application perspective and application throughput and are more concerned about application integrity at this point.

Meza asked where verification and validation fit in to the development plans. Kothe stated verification aspects are the responsibility of the project and the ECP team will make sure it is being done adequately. Validation is not sufficiently recognized in the scope, but each project listed current or future data sets against which to validate the project. However, more cost sharing with the programmatic stakeholder agency will be needed.

Bergman asked if there was any need for more points of collaborations for CD centers between application groups and innovations on the architecture system side. Kothe said those interactions are crucial, but ECP will try to manage those interactions in a scalable way. The CD centers are a nexus for interactions with hardware. Network application teams will not be precluded from interacting with the hardware vendors. Once path forward selections are completed the path forward teams will work together with the applications teams.

Hey commented on his experience concerning projects and complimented Kothe's expression of the right milestones. Kothe said the ECP team wants to understand that things are progressing.

Levermore inquired if there were any DOE applications that leaped out as on the tipping point in the paradigm or incremental by comparison. Kothe described predictive simulation moving along multiple axes and some applications moved along one axis. The multi-physics codes concerned Kothe most because the robustness of the coupling algorithm is called into question in terms of robust, accurate answers. Levermore added that Kothe mentioned hypersonic reentry and one of the key things to model there is ionization processes, which requires the use of physics modeling.

Reed asked about the scale of threats given the human element. Kothe said the human element is one of the top three risks if not number one risk. Recruitment is doable but the retention will be hard. Kothe noted that retention will require the work to be exciting and fun, students and post-docs are learning, the projects can do agile, flexible R&D, and the team staff has career development opportunities.

Dolbow asked about the risk of simulations moving into regimes where the material data is not available. Kothe indicated that the NNSA applications are being very proactive; in fact the NNSA experimental programs are in large part guided by the application development activities. But from the material response side there are still a lot of open holes with regard to data.

Crivelli asked about computational steering. Kothe stated humans should be in the loop but that is something the team has just started to think about.

Jack Dongarra presented on the state of supercomputing today and international HPC activities. In the eyes of the Top500 (https://www.top500.org/) data there are 95 systems with greater than one Pflop) high-performance LINPACK (HPL). The interesting thing is approximately 50% of the Top500 computers are being used in industry. Exascale is in many countries and regions and of these Top500 computers Intel processors are used in 91% of the machines and Advanced Micro Devices (AMD) has another 3% thus x86 architecture is in 94% of the machines on the list.

The number one supercomputer on the Top500 list, the Sunway TaihuLight, is in China. TaihuLight is three times more efficient than the Top10 machines on the list. For the first time the U.S. does not have the majority of machines, China now has 167 supercomputers while the U.S. has 165. By comparison, China had zero machines in 2001 and now China has 1/3 of the systems, while the number of systems in the U.S. has fallen to the lowest point since the Top500 list was created.

The US and the European Union have relatively parallel performance over time. Japan's performance dropped (2006-2011) in performance deployment and then rebounded. China is the most striking because China had zero machines in 2001 and today the sum of the performance of those machines in China exceeds what has happened in the U.S. Additionally, the average age of systems in the U.S. was about 1.25 years and today it is about 2.25 years, in Europe it is about 2.75 years, in Japan it is about 3 years, but the age of machines in China is at 1.6 years, the youngest population by far.

The leading vendors of the Top500 machines are HP (25%), Lenovo (17%), and Cray (12%). Sugon, Inspur, and the National University for Defense Technology (NUDT) are Chinese vendors that collectively have 15% of the machines. In terms of performance share by company HP only has 10%, Lenovo has 8%, and Cray has 20%. The Chinese machine in Wuxi has 16% of performance share.

Although the U.S. has dropped from having a major presence in China in 2013 (~50%), to having almost no machines in 2015, the U.S. has 30 machines that are greater than one Pflop in performance, Japan has 11, and China has 10.

In February 2015 the U.S. Department of Commerce (DOC) cited concerns about nuclear research on the Intel systems and placed a blockade on Chinese groups receiving Intel technology. Since the DOC action there has been an expanded focus on Chinese-made hardware and software. In the latest "5 Year Plan" in China, the government is pushing to build out a domestic HPC ecosystem without using any U.S. microchips.

The TaihuLight machine http://bit.ly/sunway-2016 is built with 2 nodes on a card, 4 cards per water cooled board, 32 boards per supernode, 4 supernodes per cabinet, and 40 cabinets in the room. The TaihuLight system runs applications in earth system modeling/ weather forecasting, advanced manufacturing (CFD/CAE), life science, and big data analytics. Thirty-five applications are running on the machine, 6 applications running at full scale – 10M cores, 18 applications running on half the machine, and about 20 applications with a million cores or more. Twenty people help with optimizing applications to run on the system. The Chinese have three of the six finalist papers for the ACM Gordon Bell Prize, the other three are from Lawrence Livermore National Laboratory (LLNL), Imperial College, and Japan.

The Chinese are expanding the focus on Chinese-made hardware and software, using anything but from the U.S. There are three separate threads developing HPC, the Jiangnan (Shenwei), NUDT (Tianhe), and CAS-ICT (Godson/Sugon) machines with an exascale prototype due by the end of 2016. The Chinese government is pushing to build out a domestic HPC ecosystem coupled with indigenous semiconductor development and acquisition. Finally, there are plans for a smaller version of this machine.

Japan has a project called the Flagship 2020 Project, the post K computer built by Fujitsu with an ARM processor. The hardware and software systems are being designed by Fujitsu with RIKEN AICS is in charge of development. The machine is scheduled to be ready between 2021 and 2022. The Post-K machine will be 50x bigger than the K computer for capability and 100x bigger than the K computer for capacity. The Post-K system will be housed in the same location as the K computer. Japan has deployment plans for 9 other supercomputing centers in Hokkaido, Tohoku, Tsukuba, Tokyo, Tokyo Tech, Nagoya, Kyoto, Osaka, and Kyushu. By the end of 2016 the Oakforest-PACS at 25Pflops and will be the #1 machine in Japan when K goes away.

The European project Partnership for Advanced Computing in Europe (PRACE) is an international not-for-profit association under Belgian law with its seat in Brussels. PRACE Hosting Members are France, Germany, Italy and Spain who collectively have six world-class systems in place, three in Germany, one in Italy, one in France, and one in Spain. In PRACE there are 465 scientific projects, 12.2 thousand million core hours awarded since 2010, 21 Pflops of peak performance, and 530 M€ (euros) of funding for 2010-2015.

Moscow State University (MSU) has a machine based on Intel Xeon 5 processors together with a company called T–Platforms. The MSU supercomputer "Lomonosov-2" has 256 Intel processors plus Nvidia in a rack, they have several racks in place, and that machine comes in at about 3 Pflop/s in terms of its theoretical peak performance.

Levermore asked how many bits make up a flop? Dongarra said everything is 64 bit floating point operations. Levermore asked if the Chinese are thinking of increasing that. Dongarra said they are actually thinking of reducing it down to 16.

Glotzer asked if the embargo on Intel chips to China accelerated development of the Godson chips. Dongarra believes China was already on its way, but the embargo emphasized acceleration was necessary. It is clear that the U.S. will have no machines in China by the next round of the Top500 list.

Sarkar asked if the slogan, "anything but from the U.S." also referred to software, and requested information on the Chinese software strategies in general. Dongarra explained that the operating system has the look and feel of Linux but the Chinese did put it together themselves. The Chinese are redesigning packages to be used effectively on their machines and they are trying to port ANSYS onto the machines. Dongarra did not have a clear picture of the software plan.

Dunning asked if the Chinese were planning any effort to develop new applications. Dongarra said there are four software centers and believes those centers will be instrumental in developing and transitioning to their own set of applications for Chinese hardware.

Bergman asked about the use of DDR3 and the energy efficiency plans. Dongarra stated he was unsure of their design plans, but DDR3 memory is too limiting. Dongarra stated the Chinese are well aware of the energy efficiency issue. Dongarra noted that the TaihuLight is not a stunt machine. There are real applications running on the machine and there is considerable effort invested in optimization to effectively use the architecture.

Hey stated that in the European efforts there is another initiative led by Luxembourg, called an Important Project of Common European Interest (IPCEI), a partnership of Luxembourg, Italy, France, and Spain. IPCEI desires to have a European ecosystem comprising the whole value chain – hardware to architecture to software development. Hey noted there was supposed to be something done in September 2016, but was unaware of the status. Dongarra stated it was strategically put in Luxembourg. Hey elaborated that the chairman of the European Commission is from Luxembourg. The IPCEI is not funded by the European Commission because the UK and Germany would not support it. It is a separate way of funding an HPC initiative. Hey said it was created for the same reasons mentioned as the Chinese.

Berzins asked Dongarra to comment on the cultural changes at the government level and leadership level driving the Chinese. Dongarra admitted he unfamiliar with the top end but the five-year plan points out that HPC is a target.

Reed commented that one thing that struck him was the 28 nm feature size, there is a lot of headroom with just technology shrink and capability of the machine and suggested this as an interesting conversation for lunch.

Glotzer asked where the development talent in China was trained and educated. Dongarra stated the people had some training outside of China; from the UK and the U.S. Hey commented that his impression was that there are now lots of credible submissions from China.

Karen Remington (via Zoom) asked what is exciting or surprising about these developments in China? Dongarra indicated what is exciting and interesting is the Chinese having a machine with 10 million cores and they have applications running at scale on those machines getting good performance numbers. What's interesting accomplishment is moving from zero to being dominant in terms of numbers of machines in a brief period of time.

Buddy Bland, from Oak Ridge National Laboratory (audience), asked if there is some other explanation for the growth numbers in China and the U.S. Dongarra stated that Lenovo was an IBM company, but is now a Chinese company. But, at the top end, the machines that have a Pflop or greater of performance, there were 10 of them in China, 11 in Japan, and 30 in the U.S. which is a sizable number of systems.

Kothe commented on the large code bloat and questioned if it was a one-off tedious port. Dongarra stated he thought there was an army of people working on the codes. To get the application to run will require a team and people who understand the architecture and the application to make the match-up work correctly, and it needs a fair amount of effort to get the performance.

The meeting was adjourned for lunch at 12:16 p.m.

Tuesday, September 20, 2016
Afternoon Session

The meeting was called back into session at 1:41 p.m.

Daniel Reed led ASCAC in an open ended discussion. Reed posed possibilities for how an open discussion might be run. Suggestions included discussing technical issues, workforce development issues, ways ASCAC could be more effective helping DOE or a thematic question from DOE. Chalk instructed members on the phone to unmute the phone in order to participate.

Levermore began the discussion asking for areas in which SC sought advice. The SC has experienced the most difficulty in defining partnerships that are outside of the government space. The U.S. is competitively constrained for collaborations among entities in ways that some of its international competitors are not; there is a legal issue. Reed explained that there are anti-competitiveness laws which constrain what companies are allowed to talk about. Inferences can be made through public pronouncements, technical meetings, and so forth. Crivelli asked if all the partners, within exascale, have nondisclosure agreements with each other in order to share information. Reed said they have some.

Sarkar asked if the focus needed to be on partnerships with industry rather than collaborations across agencies. SC is currently struggling with partnerships with industry, but the need applies equally well to universities and other government agencies. Binkley indicated that working with other federal agencies is much easier for SC; the difficulty is finding match-ups with common interests. Sarkar commented that universities have played a major role in getting industry to work with the computer science community.

Reed explained his statement further. In some ways the constraints are related to what the U.S. response ought to be to China. In the U.S. semiconductor industry's transformation Semitech was formed, Micro Commercial Components (MCC) was formed, and there were different kinds of collaborative relationships created to mount a U.S. response. Binkley referred to the U.S. reduction from having three capacity fabrication facilities for microelectronics to only having two fabrication facilities. Hey asked if these are both owned by Intel? Binkley stated that one of them is owned by Intel while the other one is owned by Micron Technology, Inc. GlobalFoundries, owned by The Emirate of Abu Dhabi, holds the former IBM foundry. There has been discussion about the potential of new types of partnerships that meet legal boundaries and allow progress. Glotzer posited**,** if ASCAC knew what is and what is not allowed now then recommendations could be made. Binkley asked ASCAC to consider policy recommendations as well as other ideas that have been overlooked.

Hey stated that collaborating parties require a win-win situation. Hey referred to a statement by Bill Gates that all the advances in machine learning would be made by companies like Google, Facebook, Microsoft, and Amazon because they have huge proprietary data sets, large numbers of smart people in there R&D divisions, and essentially infinite computing. Hey suggested that since DOE has lots of complicated and multifaceted data sets these could be used as an experimental playground to enable university computer scientists and machine learning people to make serious progress that would be a win-win. Binkley said the idea is worth following up on and requested further discussion with Hey. Senior people in Microsoft have said that DOE is virtually a non-player in areas like machine learning and DOE is years behind everyone in the commercial sector. Hey suggested considering an ASCAC Laboratory Directed Research and Development (LDRD) on the topic.

Binkley offered a second request for ASCAC to provide ideas on how to influence the direction of university programs to encourage cross-disciplinarity in the emerging fields. Bergman commented on non-competitiveness, stating there are good examples of consortia with industry and academia and government doing something that is precompetitive, such as the manufacturing initiatives. She continued that she was stunned by Dongarra's presentation, how fast the U.S. fell from its number one position. Binkley added that a lot of the discussion is occurring in DOE and elsewhere in the government focused on what path forward the U.S. should begin to take with respect to China. Glotzer asked Binkley to describe areas for cross-disciplinarity. Binkley said biotechnology is a significant area and another is China's rapid development of its supply chain for making microelectronics.

Dunning, reflecting on SciDAC, advocated that SC could find enthusiastic support from other associate directors and develop a program in quantum computing. Several SC offices are working closely together on quantum computing. However, working with the Office of Management and Budget (OMB) is not as easy today as it was when SciDAC came about.

Sarkar sought clarification from the tone of the discussion asking if ASCR no longer wants to continue to lead in the Moore's Law area. Binkley assured Sarkar that was not what he was saying. The belief from SC and DOE is that through application appropriate technologies the reach down the roadmap can stretch farther than it is today. Binkley indicated SC could contribute to issues related to fabrication, which is not typically an ASCR purview. Moore's Law is not over and the question is can the U.S. actually move any further down that pathway over the next decade. Sarkar added that experience has shown that a combination of effort and innovation is needed across the system stack and at all levels of system hardware and software.

The numbers of the FY2017 budget numbers concerned Meza because the research side was significantly impacted. Meza wondered if an advisory committee could help Binkley make the case that it is important to maintain the basic research funding.

Levermore turned the conversation to the culture at universities. He suggested targeted opportunities for universities to create interdisciplinary centers. Training students in such an environment enables them to step into the new technologies and positions that are coming up in DOE.

Hey commented, concerning Moore's law DOE, NSF and Defense Advanced Research Projects Agency (DARPA) must all work together to compete. Hey inquired if any of the recommendations from the Committees of Visitors (COV) were considered by SC or ASCR. Chalk spoke and assured Hey that the COVs recommendations were not ignored; they were looked at within ASCR, SC, and OMB. Chalk pledged that while every effort was made to meet the COV recommendations there are several other factors at work with regard to the ASCR budget. She continued that as exascale was pulled into a separate project ASCR was struggling with the components that are left behind and considering sustainable investment.

Sarkar added that another area for discussion among ASCAC members was the workforce pipeline in terms of the budget realities. He mentioned that one of the implications is if that pipeline slows down it takes a long time to rebuild.

Berzins agreed with Sarkar and interjected that DOE is competing with companies who are prepared to put a lot of resources, in terms of money and people, into projects. He indicated that the rewards, in the labs and DOE, are not there compared to industry. Berzins stated that order to compete there must be equivalent rewards.

Reed closed the discussion stating that the purpose of the conversation was to talk about how ASCAC, in the most constructive way possible, can help make a case for the resources that DOE needs to remain competitive. He thanked ASCAC members for their perspective and continued with the agenda.

John Labarge explained the LDRD which is available at all DOE labs. LDRD is R&D work of a creative and innovative nature that is selected by the Director of a national laboratory for the purpose of building new capabilities, identifying and developing potential applications, and formulating new theories, hypotheses, and approaches that advance the DOE's missions.

LDRD is founded on three Statutes: The Atomic Energy Act of 1954, as amended; Title 42U.S.C.2103, 2051, and 2053, An Act for Authorizations and Appropriations for the Energy Research and Development Administration for FY 1977 (P.L. 95-39); and The Consolidated Appropriations Act, 2014 (P.L. 113-76).

LDRD has been studied and audited multiple times, including reports concerning the need for LDRD (Packard Panel 1983, The Energy Research Advisory Board (ERAB) December 1985 Guidelines, and The Commission to Review the Effectiveness of the National Energy Laboratories, 2016). Audits include 2001 Government Accountability Office (GAO) review (GAO-01-927) which yielded one recommendation, 2004 GAO review (GAO-04-489) which yielded no recommendations, and IG audits (2005 and 2009) which were terminated after the initial survey phase because there was nothing to find.

Performance metrics (FY2015) showed that LDRD funds supported 48 postdocs and 130 researchers in Energy and Environment, 458 postdocs and 987 researchers in NNSA, and 413 postdocs and 2158 researchers in SC.

A laboratory's maximum allowable LDRD funding level is 6% of its operating/capital equipment budget. DOE Order 413.2C spells out details about LDRD projects, that they must support the DOE/NNSA mission, be relatively small, and be limited to a maximum of 36 months. LDRD awards may not be used to increase funding on Congressional or DOE limited tasks, projects that require non-LDRD funds to accomplish, or general purpose capital expenditures. LDRD's are designed to enhance and develop the scientific and technical capabilities of the laboratories to support DOE mission needs, are chosen through competitive, peer-reviewed proposal processes that focus on their scientific merit, and are checked for consistency with the requirements of the LDRD Order.

DOE oversees LDRD and approves LDRD projects. Within DOE Headquarters, a Cognizant Secretarial Officer (CSO) approves each laboratory's annual LDRD program plan and maximum funding levels. SC has responsibility for the overall DOE policy (DOE Order 413.2C) and is the primary contact for policy clarification and issue resolution. DOE Field/Site Offices review and provide recommendations on the laboratories' annual LDRD program plans, certifies LDRD accounting methods, monitors compliance with LDRD policies/procedures, and assures the laboratories submit annual LDRD Program Reports to DOE. DOE Field/Site offices also review and concur on each proposed LDRD project prior to the obligation of funds and the start of work.

Improvements for LDRD and recurring concerns from Congress included develop better methods for measuring and communicating the value and impact of LDRD to key stakeholders, and in addition to the annual program reviews, conduct an independent review of LDRD using the federal advisory committee model. Hey asked if Congress was hostile to LDRD. Labarge indicated Congress is concerned about what LDRD is being used for. Congress is concerned LDRD takes away already limited resources from mission work, LDRD provides laboratories with a vehicle to expand their missions, and LDRD programs are not taxed evenly. DOE has and will address Congressional concerns by sharing stories where LDRD projects led to breakthroughs that support mission needs, ways LDRD investments strengthen core capabilities needed to achieve DOE missions, the fact that DOE reviews and approves annual program plan and all projects, and that DOE annually certifies that LDRD funds are accrued consistently and equitably.

Reed reminded ASCAC that they will be reviewing LDRD for all of DOE.

Dolbow commented that one of the important aspects of LDRD is retention, especially with young scientists. He added that mapping where LDRD dollars are going, in terms of partnerships outside the labs, might help allay some of Congress' concerns. Labarge supported Dolbow's observation about partnerships referring to the difficulty populating the review committee of people's relationships to the labs and LDRD work within the last 10 years.

Gregurick added that one benefit of LDRD is funding high risk projects and new innovations in science. She asked if new fields that have been explored or new ideas that came through LDRD funding have been captured. Labarge assured Gregurick that highlighting such projects is possible and the LDRD brochure is meant to be the venue.

Sarkar commented on the Congressional concern about using LDRD funding to expand the DOE mission. He stated that scientific advances by exploring new topics are within the core mission rather than an expansion of the mission. Labarge said that while scientists agree basic science is needed sometimes Congress does not see that.

Levermore added that LLNL was one of the first labs to have an LDRD program and that he was stunned that LLNL still has over half of the NNSA LDRD programs in terms of numbers.

Steve Binkley briefly covered the new charge for ASCAC to conduct an independent review of the LDRD program for DOE. Binkley stated that this charge came out of the Secretary of Energy Advisory Board (SEAB) Task Force action. The charge requests ASCAC review the LDRD program processes and the impact of LDRD at four of the DOE Labs, to include at least one SC Lab, one NNSA Lab, and one of the applied energy Labs that have had LDRD programs for at least ten years.

Martin Berzins provided an update on the LDRD subcommittee. There are approximately 1,700 LDRD projects per year with an average spend of $300,000 per project. LDRD funds support approximately 650 (2005) to 900 (2015) postdocs and on average about 30% of all laboratory post-docs were fully or partially supported. There were a higher percentage of postdocs supported at Los Alamos National Laboratory (LANL), LLNL and Sandia National Lab (SNL) with the majority of LDRD projects including early career researchers. Between FY2008 – FY2012 across the three laboratories 49% (LANL), 74% (LLNL), and 77% (SNL) of LDRD supported post-docs were converted to full-time staff members. Productivity of LDRD (2015) showed that at LLNL 50% of the patents, 50% of the records of invention, 30% of the copyrights, and 25% of the publications came out of LDRD.

Berzins pointed out that the interesting part of the directive was that the committee was asked to look at processes to determine funding levels, at processes to determine lab-specific goals and allocated resources amongst the goals, at processes to select specific projects, and at processes to evaluate the success and impact of LDRD.

The committee membership was created based on nominations from chairs of the SC program advisory committee chairs. The committee members all have the experience and the depth to do a proper analysis of this activity.

The subcommittee discussed how best to address the committee charge using available information and lab visits. The subcommittee formulated a detailed set of questions for the four labs based on the committee charge. Three lab visits are scheduled for National Renewable Energy Laboratory (NREL) Colorado, LLNL and LBNL California, and Oak Ridge National Laboratory (ORNL) Tennessee in late 2016 or early 2017. The draft report and comment period will be in spring 2017 with a planned final report in late spring 2017.

Lori Diachin (audience) of LLNL asked about the committee composition in terms of security clearances. Berzins assured Diachin there are members of the subcommittee with necessary security clearances.

Reed asked ASCAC members to forego the afternoon break and encouraged them to excuse themselves from the meeting as needed.

Robert Harrison briefed ASCAC on the report from the Committee on Future Directions for NSF Advanced Computing Infrastructure. The report was to set the stage and direction in advanced computing over the next 3 to 5 years and is available from https://goo.gl/skS6La.

The committee charge was to provide a framework for future decision making about NSF's advanced computing strategy and programs. The framework addresses such issues as how to prioritize needs and investments and how to balance competing demands for cyberinfrastructure investments. The report emphasizes identifying issues, explicating options, and articulating trade-offs and general recommendations. The study will not make recommendations concerning the level of federal funding for computing infrastructure. Advanced networking was not included as part of the charge because it was considered too large a scope. The committee conducted briefings at committee meetings in 2014 and 2015, SuperComputing 14 – Birds of a Feather (SC14 BOF), and December 2014 workshop (Mountain View, CA) which yielded 60 comments in response to questions posed in interim report.

There are four goals outlined in the report, each containing observations and recommendations. Goal one is focuses on positioning in science and engineering. Large-scale simulation and data analysis are revolutionary, systems that support a wide range of advanced computing capabilities are needed, university computing resources cannot support the need, and more specialized systems may be needed. Recommendations are that NSF should sustain and seek to grow its investments in advanced computing, and NSF should pay particular attention to providing support for the revolution in data-driven science along with simulation.

Goal two focuses on resources. Demand is growing and changing rapidly, there is no systematic or uniform planning process, ongoing and regular structured process is needed, and NSF must secure access to capabilities with respect to individual applications. NSF should construct and publish roadmaps, and NSF should adopt approaches that allow investments in advanced computing hardware acquisition, computing services, data services, expertise, algorithms, and software.

Goal three refers to aiding the scientific community to keep pace with changes. Better software tools, technical expertise, and more flexible service models can boost productivity. NSF should support the development and maintenance of expertise, scientific software, and software tools, and NSF should invest modestly to explore next-generation hardware and software technologies.

Finally, goal four discusses the infrastructure. The recent strategy of acquiring facilities and creating centers relies on irregularly scheduled competition among institutions, and the cost-sharing models presents challenges because of mounting costs and budget pressures, there is a risk that proposals will be designed to win a competition rather than maximize scientific returns, and there is no long-term support to develop and retain talent. The recommendation states that NSF should manage advanced computing investments in a more predictable and sustainable way.

Levermore asked if there was anything in the report about how cyber infrastructure should be overseen. Harrison noted that this was exclusively excluded from the committee charge and there is no comment on oversight in the report. He added that there has been an Advanced Cyber Infrastructure (ACI) initiated conversation on this topic in partnership with the Director of Mathematical and Physical Sciences (MPS) and there is a call for input. Dunning added that this is currently being discussed within NSF and Fleming Crim is in charge of that process.

Hey sought clarification, asking if NSF researchers are funded to use the large DOE facilities like synchrotrons and neutron sources. Harrison replied that not all of the large facilities are sponsored by DOE. Instead NSF and other agencies fund beam lines, the neutron photon sources, and 50% of NSF's supercomputer time is given to other agencies, like many of the DOE facilities the NSF computing facilities are open to others.

Crivelli asked if there is an estimated amount of computer time that goes to students. Harrison said he did not think they are consciously allocated, but students are eligible to apply for time. Berzins added that a good proportion of NSF computer time goes to students. Harrison indicated that students are the ones running on the machine, in terms of a formal allocation. Berzins stated the challenge for the allocations committee is that time constraints make it much harder to secure time on the supercomputer.

Linda Blevins shared information about the Early Career Research Program (ECRP), which began in 2009 with Recovery Act funds. The purpose of the program is to support the development of individual research programs of outstanding scientists early in their careers and to stimulate research careers in the disciplines supported by the SC. Proposals are invited in the ASCR, BER, BES, FES, HEP and NP program areas

Both Universities and national labs are eligible for the ECRP. University grants are at least $150,000 per year for 5 years for summer salary and expenses, and lab awards are at least $500,000 per year for 5 years for full annual salary and expenses. ECRP plans to have about 300 active awards in steady state (200 university awards and 100 lab awards). There is roughly $80M in funding for new and ongoing awards each year with ~ 60 new awards (40 university and 20 lab) per year.

To be eligible for the ECRP, a principal investigator (PI) must have received their PhD no more than 10 years prior and no less than one year of the deadline for the proposal. For DOE labs the PI must be full-time, permanent, non-postdoctoral employee. For U.S. academic institutions PIs must be an untenured Assistant Professor or Associate Professor on the tenure track. And an employee with a joint appointment between a university and a DOE lab must apply through the institution that pays his or her salary and provides his or her benefits.

The ECRP Merit Review includes six criteria, 1) scientific and/or technical merit of the project, 2) appropriateness of the proposed method or approach, 3) competency of applicant's personnel and adequacy of proposed resources, 4) reasonableness and appropriateness of the proposed budget, 5) relevance to the mission of the specific program to which the proposal is submitted, and 6) potential for leadership within the scientific community. ECRP discourages about 20% of the proposals.

General rules for proposals include a preproposal requirement, disqualification if the preproposal does not address research topics identified in the solicitation, no co-PIs, only one proposal per competition per PI, a limit of 3 submissions of the same proposal, and no letters of recommendation.

Rules for DOE labs state that the proposal must contain a focus on transition at the end of five years in order to integrate researchers into the labs, a letter from the lab director confirming that the proposed research idea fits within the scope of SC-funded programs, the lab scientist must charge at least 50% of their time to the award, execution of funding is at the PI’s discretion, and the employing lab addresses funding transition issues when the award ends.

Since 2010, ECRP programs have made 411 awards (270 university and 141 lab). The overall success rate is about six percent. There have been 42 states represented, 15 labs, and 109 universities. Approximately 25% of the proposals are from female PIs (105 women and 306 men). ASCR has made 37 awards (21 university and 16 lab), since 2010. Eleven percent were to support female PIs (4 women and 33 men).

The process of encouraging proposal submission based on preproposal fit began in FY2011. Proposal submission is encouraged for 85-90% of preproposals. Proposals are received from about 80% of those encouraged to submit. Blevins stated that the number of awards was low during FY2014 -FY2016 because of the transition to full funding, which is expected to take two more years (FY2017 and FY2018).

ECRP is the largest, single funder of the physical sciences in the country. The academic awardees come from a few department types including physics, engineering, biology, and mathematics. Most of the ECRP winners received their PhDs 4-6 years before getting an award, and about half of the winners are first-time applicants.

Glotzer inquired as to the number of applications by women in ASCR. Blevins indicated the demographic feature had just become available, and approximately 40% of the PIs answered the demographic questions, however this information is being collected in the profiles within

Portfolio Analysis and Management System (PAMS). Chalk assured Glotzer that she will provide this information to ASCAC the following day.

Meza asked if the ECRP kept the same numbers for underrepresented minorities. Blevins said this information is collected from individuals who create a new account in PAMS. Glotzer asked if this information could be compiled retroactively. Chalk interjected that she was confident the number for gender could be obtained, however, ASCR has not been reporting on underrepresented minorities.

Sarkar inquired about DOE's track record and the ASCR candidates being selected for Presidential Early Career Awards for Scientists and Engineers (PECASE) awards relative to other agencies. Blevins said that typically the White House dictates how many people SC can nominate. Hey asked how many awards the White House gives. Blevins stated they give about 100 awards; DOE gets 13, SC gets 6, the technology programs get 4, and NNSA gets 3.

Reed asked if Blevins had any sense of how the funding for young faculty members engages them with the laboratories. Blevins said lab management had provided feedback and indicated that ECRP has resulted in real changes in lab practices. Some of the labs, motivated by ECRP, have created LDRD programs for early career scientists. In terms of retention and recruitment, Blevins had heard that ECRP is an important pool for hiring for the early career PIs; ECRP awardees immediately hire one or two postdocs to work on the award with them at the lab. ECRP projects are exciting; they are on the cutting edge. The award is for five years of sustained funding at a substantial amount and PIs are able to recruit very good postdocs. In fact the labs' reports highlight hiring some of the postdocs that the ECRP recipients recruited.

Hey asked if the University awardees get recruited to the labs or just retain a close relationship when they finish. Blevins said that the university awardees do not have to have a formal connection to a lab to get this award. University awardees have reported the benefits of becoming involved with the PI meetings and workshops at DOE because they meet lab scientists.

James Glownia and **Ping Ge** spoke about the Office of Science Graduate Student Research (SCGSR). The Office of Workforce Development for Teachers and Scientists (WDTS) helps ensure that DOE has a sustained pipeline of science, technology, engineering and mathematics (STEM) workers; the WDTS has a workforce development mission, but not an education mission. There are four programs in the WDTS, Science Undergraduate Laboratory Internships (SULI), Community College Internships (CCI), SCGSR Program, and the Visiting Faculty Program (VFP). Over 1,100 students and faculty at DOE labs and facilities have participated in these programs. SCGSR operates the Albert Einstein Distinguished Educator Fellowship for K–12 STEM teachers, administered by WDTS for DOE, National Aeronautics and Space Administration (NASA), and NSF (P.L. 103-382), and the National Science Bowl®(NSB) finals competition in Washington D.C.

The goal of the SCGSR program is to prepare graduate students for STEM careers critically important to the SC mission by providing graduate thesis research opportunities at DOE laboratories. The research opportunity is expected to advance the graduate students' overall doctoral thesis while providing access to the expertise, resources, and capabilities available at DOE laboratories.

The SCGSR Program Priority Research Areas for the 2016 Solicitation 2 are based on SC's Assessment of Workforce Development Needs and evolving needs of SC programs. The SCGSR program is managed by the WDTS in collaboration with 6 SC research program offices and 17 participating DOE national laboratories (except National Energy Technology Laboratory (NETL) because it is federally funded). Oak Ridge Institute of Science and Education (ORISE) provides support for program administration. Since 2014, 207 awards have been given. Approximately 6% of the awards have been for ASCR research areas.

The SCGSR Program provides supplemental awards (no tuition or fees) to outstanding graduate students to spend 3 to 12 months conducting part of their doctoral thesis/dissertation research at a DOE laboratory in collaboration with a DOE laboratory scientist. U.S. Citizens and Permanent Residents can apply. There are two solicitations annually. Applicants must be enrolled in a qualified graduate program and be a PhD Candidate. The graduate research must be aligned with an SCGSR priority research area. And the applicant must have an established collaboration with a DOE laboratory scientist at the time of application, which SCGSR can facilitate.

Completed applications submitted to the SCGSR Program before the application deadline are evaluated using the DOE SC standard merit review processes including the scientific and/or technical merit of the proposed research, and the relevance of the proposed research to graduate thesis research and training.

Chapman asked Ge about developing relationships with universities and what ASCAC could do to help. Ge shared information about SCGSR outreach activities to university campuses, to faculty members, and to lab scientists. Glownia added information about the undergraduate programs. He stated WDTS will announce opportunities directly to the career placement offices and giving money to labs to develop outreach tools. Ge stated that the most effective communication and outreach is getting to the students multiple times, repetition is the key to remembering. Glownia added that SC drove development of a STEM portal that uses metadata and filter techniques to allow interested applicants find out what is available.

Dolbow remarked that the eligibility requirements indicate the SCGSR is limited to doctoral candidates and asked for demographic data in terms of the average applicant's year in program. Ge said that about 75% of applicants are in their 2nd to 4th year of study, the other 25% is in their 5th or 6th year, depending on the discipline. Dolbow asked how many applicants are in the fourth and fifth year versus their third. Ge said no more than ~25% are in the 4th and 5th year.

Sarkar asked about bridging from SULI to SCGSR. Sarkar suggested that exposure to something like SULI and receiving an NSF graduate fellowship in the first year of a computer science student's PhD could be helpful. Glownia agreed with Sarkar that the goal is to get students into programs and connect the dots. Sarkar asked if a senior who is applying to graduate school can apply for SCGSR. Ge said they cannot apply for SCGSR. Glownia pointed out the WDTS Graduate Student Fellowship program was sunset. Sarkar emphasized that when computer science undergraduates get a graduate school offer and a great job offer fellowships sometimes sway their decision. Glownia agreed that the NSF fellowships are open to anybody. Ge stated that for every cycle there are 3 to 4 NSF Fellows receiving an SCGSR award.

Crivelli revealed that she gave a presentation to a nearby community college and was surprised that students did not know about any of the WDTS programs. Crivelli encouraged scientists at the lab to visit area community colleges and universities and discuss what they do as well as these programs. Glownia added that WDTS is on the cusp of completing their online system and are starting to look at some of the geographical information. Pacific Northwest National Laboratory (PNNL) has the most participants from the congressional district that surrounds PNNL. Anecdotally SNL-Livermore's community college program tends to be very local. SULI is a national program, but CCI tends to be much more localized. There are over 1,200 community colleges in the U.S. and there is a lot of outreach opportunity.

Levermore indicated the actual advising of students is through Dean's Offices, departmental Dean's. If the outreach was aimed there, Levermore argued there might be an increased response from colleges and universities, especially small ones. Glownia added that at HBCUs, the Dean's Office is absolutely critical.

Labalette (phone) asked about the fraction of the awardees hired by the labs. Ge stated they do not have that data.

Reed offered the audience an opportunity to make comments or ask questions. The meeting was adjourned at 5:05 p.m.

WEDNESDAY, SEPTEMBER 21, 2016

The ASCAC meeting was convened by **Chair Daniel Reed** at 8:32 a.m. Chalk addressed Glotzer's request, from Tuesday; the average number of women who applied to the ECRP from 2009 to 2016 was 14% to 15%. She stated that Blevins' statistic of 11% is reasonable in this case.

Reed stated that Barbara Helland was sitting in for Steve Binkley and then introduced the agenda for the day.

Wes Bethel presented information on the Report of the DOE Workshop on Management, Analysis, and Visualization of Experimental and Observational Data: The Convergence of Data and Computing (Contract No. DE-AC02-05CH11231, Report LBNL-1005155) held September 29 – October 1, 2015. The purpose of the workshop was to get the science user facilities' perspective on problems they are facing and the coming multiple exabytes of data per year. The Experimental and Observational Data Workshop objectives were to better understand data-centric issues facing SC Science User Facilities, identify how to meet those needs and science objectives, and foster dialogue between Experimental and Observational Science (EOS) projects and ASCR. Workshop attendees represented 11 science user facilities as well as NSF and UK E-Science Grid.

Gaining scientific knowledge from experimental data is increasingly difficult; data-and computing-centric needs are increasingly intertwined and symbiotic; and there are acute and urgent data-centric needs in these programs, the ability to create data exceeds the ability to analyze and store it. The report issued eight findings and nine recommendations.

Finding one is all EOS projects struggle with a flood of data, multiple exabytes per year. Data are increasing in volume, complexity, variety. The opportunity cost in not dealing with this is a potential loss of science.

Finding two is all EOS projects have difficulty using large-scale, HPC facilities, as the needs of the EOS workloads are different than traditional HPC workloads. The challenge is that many HPC facilities' operational policies are optimized for high-concurrency, batch workloads. However, EOS projects consist of a complex, elaborate data life cycle and their use cases are divergent from how HPC facilities have been designed.

Finding three is that EOS projects have unmet time-critical data operations. The challenges is to be able to adjust on-the-fly (OTF) to accommodate data movement to gain access to a lot of cores and memory and data, for which HPC facilities are not designed.

Finding four is that there is a risk of unusable data and there are no adequate tools to create metadata. Bethel said that data usability is an issue but there is also the Executive Order for data dissemination (Executive Order -- Making Open and Machine Readable the New Default for Government Information, May 2013) to consider.

Finding five states that collaboration and sharing are central to EOS projects; sharing means giving data away and also being able to find it. Individual science user facilities are attempting to solve the problem alone. Participants noted it is often left up to the individual science user who travels to the facility to figure out how to deal with this problem.

Finding six states that the EOS data lifecycle needs are not being met and the only archival process is that provided by the publishing journal. The EOS projects have significant and complex data lifecycle needs that go well beyond what is provided by ASCR facilities and the ASCR research portfolio.

Finding seven is that software has a central role in EOS projects and the need for software is ubiquitous. The challenge is how to create needed software for improving experimental process and for implementing beam like data movement and data reduction workflows to perform preliminary quality assurance.

Finding eight focuses on workforce development and retention concerns. EOS programs require an interdisciplinary workforce. Specialized and multi-disciplinary knowledge is rare. Data scientists are in high demand in industry and there are insufficient or inadequate career paths within the labs.

Recommendation one is to address challenges posed by growing data size, rate, complexity, to modernize the data infrastructure, thinking about the EOS use cases, and thinking about the software infrastructure needed to actually make this happen.

Recommendation two is to evolve HPC facilities to include a focus on EOS needs by identifying and prioritizing EOS-centric operational and research needs for HPC computational facilities and networking infrastructure.

Recommendation three is to develop solutions for EOS data-centric workloads by considering the systematic intake and understanding time critical EOS needs.

Recommendation four is to improve EOS productivity with resilient, automated data pipelines by considering data movement, processing, and sharing and promoting reusability of workflows and methods across EOS projects and HPC facilities.

Recommendation five addresses metadata needs of the EOS community and focusing on developing a systematic understanding of how data are used.

Recommendation six is to expand capabilities for collaboration and sharing of EOS tools and data. Determining what is being shared.

Recommendation seven is to identify and fill gaps in data lifecycle, reproducibility and curation.

Recommendation eight is to expand efforts focusing on the EOS software ecosystem and to cultivate software R&D projects for EOS that follows best practices in software engineering.

Recommendation nine is to develop and nurture a data science workforce by prioritizing the role of data science activities in multidisciplinary teams, finding ways to recognize and reward the roles and skills of this group, and providing career paths.

All EOS projects face significant, complex data challenges, EOS projects are tackling these issues "on their own", there is no program-wide, coordinated effort to address the issues, and there are urgent and acute needs and EOS projects are not ready.

Levermore stated that very often the people collecting the data do not indicate how they filtered the real data to infer the data that they deposit. The question is not just storing the data, but also storing how it is processed. Dunning clarified that Levermore was meaning the metadata, adding often more metadata than is available is needed. Bethel added that the climate community has agreed on a community centric metadata model.

Berzins asked Bethel to comment on a workforce report that expressed there would always be people wanting to work in the labs despite the attraction of the commercial companies. Bethel said people who have a multidisciplinary background are a rare breed and they must be cultivated. Berzins suggested that labs need to rethink the culture that exists in terms of how people are used and in what conditions they work. Bethel stated that places like Google, where it is possible to do cool stuff and work on science projects, pays a lot more than the labs. Berzins added that Google is moving away from giving people freedom to do what they want, that the 20% they used to provide is vanishing very fast. Berzins wondered if there is a generational phenomenon and the shift will take much longer than supposed in terms of how people think about and deal with observations. Bethel referred to the participants comments that in the next few years they expect this much data coming out of the facilities.

Meza asked if workshop participants talked about developing a set of test problems that people from outside the labs could use. Bethel was unaware of any set of test problems available. He added that would be a great outcome for a science-facing data-centric project. One of the objectives is to produce data products along with the code and have them accessible, encapsulated, and presented in a way that facilitates and fosters research. Meza pointed out that the ASCI system did something very similar, developing computational kernels and attracted a lot of the good researchers from outside the labs to work on the problems.

Crivelli suggested that the EOS community and Office of Scientific and Technical Information (OSTI) continue the discussion together; to find the money to deal with software, to deal with data.

Gregurick asked if there is a way to exchange ideas and information on a larger scale than just DOE science user facilities, to include the biomedical and health science communities as well. Hey posed the question to Barbara Helland asking if OSTI is formally tasked with helping the labs with metadata. Helland said she believed that the labs have to start putting all of their information in OSTI. Steve Hammond (audience) interjected that there is a goal for OSTI to interface with the labs to make more transparent retrieval of the publications from all the labs. Hey reiterated his question if there is any mandate to assist the labs with data collection and metadata. Hammond said not the data, not yet. Helland added she thinks DOE is starting small.

Reed said the universities are starting to grapple with this as well because they fear an unfunded mandate from the federal government for data retention.

Bland, ORNL, added that for many years ORNL has been struggling with the data problem. ORNL has pilot projects working with OSTI to store data long-term, ORNL actually owns the data and issues DOI’s from OSTI. The labs need a mandate and a way of funding the activity that does not eat into the existing mission.

Reed read Williams’ (via Zoom) question asking for the top three things Bethel thought should be prioritized on the list for ASCR. Bethel said he would start by spending some time with this report because it contains a lot of information.

Reed asked about a rational process to decide what to throw away. Being thoughtful about what is discarded is perhaps as important as being thoughtful about what is kept. Bethel added in Alexander S. Szalay’s chapter 20 in the report he talks about creating reference data sets and over time the more important ones emerge while others will become less important.

Helland reminded ASCAC that SC just completed the Requirements Workshops with all of the other program offices, and they were asked to look at all of their data from the facilities, the labs, all the way through. SC, with the Facilities Division, plans to start looking at the data issue and develop a strategic plan, because SC thinks the facilities are going to have to really change their operations in the next 10 to 15 years.

Levermore commented that there was a National Research Council report for National Security Administration called Massive Data Sets which dealt with some of these issues.

John L. Sarrao discussed the role of HPC in stockpile stewardship. Sarrao explained ASC, and provided a brief history of LANL and computing before discussing stockpile stewardship.

ASC is a mission driven program, central to U.S. national security, as a provider of computational surrogates for nuclear testing. ASC supports the shift in emphasis from test-based confidence to simulation-based confidence.

LANL has been at the forefront of computing since 1943. Computing has always been a core component of the weapons program but the landscape changed in 1992. From 1943-1992 the U.S. conducted underground nuclear tests which provided validation. From 1992-present validation relies on simulation versus tests. The end of testing drove the adoption of the Accelerated Strategic Computing Initiative (ASCI) content in 1998.

Weapons science mission-directed goals require increases in both simulation scale and model complexity, a problem solved by computing, but the fundamental challenge is to understand the underlying physics. Code validation through experiments and test history remain essential, moving from nuclear tests to calibrated codes to predictive codes. Nuclear tests drive confidence in part by simulations and predictive codes increase confidence derived from simulations, but there is a fundamental challenge in the validation of code.

Historic trends indicate memory requirements increase dramatically with physics fidelity. Prediction fidelity over the last two decades is dominated by increased resolution, increased dimensionality, and sub-grid model complexity. Successful experimental validation of models is universal at higher resolution and full dimensionality and tuned by relying on locally valid sub-grid models. Increased resolution in simulation is changing the picture of how weapons work.

There are two tracks to consider regarding computing, advanced technology systems (the largest systems that do the largest most significant calculations) and commodity technology systems (day-to-day operations that are vitally important), getting that balance right is important.

Co-design of relevant (to ASCR) applications (e.g., mesoscale materials) remains an important frontier. The development of validated models will reduce uncertainty in integrated codes and provide predictive descriptions of newly manufactured materials and components. MaRIE will provide critical data to inform and validate advanced modeling and simulation to accelerate qualification of advanced manufacturing moving from “process-” to “product-based”.

LANLs stewardship strategy relies on HPC and co-design. The nuclear weapon mission moved from building the stockpile (cold war), to stewarding the stockpile (nonproliferation/ counter proliferation), and now to responsive stockpile (intelligence and data analysis). During the ASCI era, the nuclear global security mission focused on computation, modeling and experiments, and manufacturing was forgotten.

For nuclear performance exascale will allow 3D boost simulations with multiple coupled physical processes at unprecedented resolution, detailed highly resolved 3D nuclear safety simulations, and uncertainty quantification (UQ) performed in 3D at lower resolution with sub-grid models to capture unresolved physics. In weapons science exascale will allow resolution of important length scales with appropriate fidelity possibly *in situ* with performance simulation.

ECP is a key element of ensuring the future role of HPC in stewardship. From a national security perspective, exascale is important and urgent. ECP is the right modality for success because of lab lead projects spanning SC and Defense Programs. A holistic approach focused on capable exascale is the right approach; Labs are committed and putting the best talent forward.

Dunning remarked that one of the essential tensions in ECP is between resolution and fidelity. The problem of a trade-off between memory and the compute capability of a node is common. Dunning stated more memory than is currently feasible is required to use the current algorithms, and the only other approach is to develop new algorithms. Sarrao emphasized there is a need to break away from the pattern of using the same algorithm for "just one more paper". He encouraged projectization of R&D in the stewardship program. Sarrao said real milestones that dictate needing a result of X quality by X date call the question of resolution and fidelity more urgently. Levermore added that when the capability for more resolution is available it introduces physical regimes that are not modeled well. He encouraged a discussion that includes models as well as algorithms because stockpile stewardship is a complicated, multi-layer issue.

Berzins asked about the challenge of portability with known and unknown architectures and the general approach to address this challenge. Sarrao stated Advanced Technology Development and Mitigation (ATDM) is trying to balance an evolutionary approach and a revolutionary approach. Activities in next-generation code development are relatively revolutionary. The evolutionary approach is trying to update the legacy codes OTF.

Reed asked about new architectures from a workforce perspective and the sustainability issues. Sarrao stated that sustainability of the workforce is a key challenge. People doing this work are milestone-centric and focused on getting the next thing done; they struggle to separate themselves and play with ideas.

Vivek Sarkar provided an update on the X-Stack PI meeting. The central focus of the X-Stack program is to bring the best science minds from academia, laboratories, and industry together to address the scalable system software and programming system challenges to map exascale algorithms onto future hardware. The research has a significant focus on understanding how the software mapping needs to address challenges related to energy efficiency, interconnect technology, memory technology, resilience, and scientific productivity while keeping in mind the context of new algorithms. The X-Stack Program Manager is Sonia R. Sachs, and the DOE ASCR Research Director is William Harrod. There are currently 10 projects in the X-Stack Portfolio, including DEGAS, TraleikaGlacier, D-TEC, XPRESS, X-Tune, CORVETTE, SLEEC, PIPER, Vancouver 2, and ARES.

The X-Stack PI meeting, held April 6-7, 2016 at LBNL, had 109 participants equally representing labs, industry, and university. Over four years there have been approximately 400 X-Stack publications and roughly 200 software releases. A facilities panel was an open discussion to discuss what is needed to adopt new software. Observations included facilities staff install new software all the time but can only justify doing so when there is a critical mass of users, and simplified tool chain (e.g., libraries) and training resources are important enablers for transitioning research software to use in facilities.

Five breakout groups provided feedback on specific topics. Breakout 1 concerned "Future synergies in programming systems for data-intensive and compute-intensive science". Approximately 20 participants expressed that exploiting synergies between data-intensive and compute-intensive computations is critical for future HPC applications. Research opportunities included HPC data models and implementations, programming systems for integration of simulations and analytics, addressing new scalability and performance bottlenecks, and scalable integration of customized software stacks for science workflows.

Breakout 2 focused on "Programming systems for data centric computing". The 15 participants provided application examples including Smart Grid, Smart Cities, Smart Mobility, BNL nanotube optimization, magnetic fusion, experimental data management/analysis, Compact

Laser Accelerator such as Berkeley Lab Laser Accelerator (BELLA), NNSA High Energy Density (HED) physics facility such as National Ignition Facility, and astrophysics simulation. Research opportunities included programming systems for data-centric applications, schedulers/resource managers that address dynamic behavior of data/systems, application portability across HPC/Cloud platforms, constraint-based software stack: real-time, interactive, and computational.

Breakout 3 covered "Programming systems support for post Moore's Law computing" and asked how to exploit new technologies for science. Suggestions included QC/QA, field-programmable gate array (FPGA), Optical, Molecular computing, Neuromorphic and brain-inspired computing, Probabilistic and stochastic computing, and new memory technologies. The 18 participants also discussed how to incorporate new technologies into DOE's science processes.

Breakout 4 covered "Ensuring correctness for exascale and beyond". Correctness issues included inaccuracies and inefficiencies, detection approaches, and accuracy and reproducibility approaches.

Breakout 5 focused on "High performance languages". The group discussed risk factors in adopting a new language, which includes new syntax and semantics, requires new skills to be learned, an immature tool chain and developer ecosystem. New languages are adopted all the time in commercial/enterprise computing. The participants discussed how to lower the barrier to entry for adopting new languages in HPC. Suggestions included hackathons to try new languages on real-world problems, funding for 3-month mini-projects (micro-interactions), embedding computer science graduate students within application teams, and embodying programming model in the library instead of the language.

Computations and communications will be irregular due to heterogeneous processors characteristics, deep memory hierarchies, and much higher rate of faults than current platforms. Programmability and performance portability becomes increasingly important and challenging. With domain specific languages (DSLs), application code size can be reduced by 100x and the code generated and optimized has about 3x faster execution time. Unprecedented parallelism, asynchrony, and irregular computations and communications are well suited to task-oriented dataflow programming models. And continued advances in programming systems, runtime systems, and tools are critical for the Moore's Law "end game" (even beyond exascale).

Berzins remarked that the computer scientist will not understand the applications unless continuous mathematics as well as discrete mathematics is taught. There is a resistance to this in many mainstream computer science departments. Sarkar felt that is why the relationships are so important. Success has been achieved when a PhD student finishes the University requirements and spends extensive time in the labs. Berzins added the problem is it goes against the culture in many computer science departments, but multidisciplinary scientific computing makes it possible. Berzins made a second point that DSLs pose a barrier between the DSL people and the language people. Berzins stated the performance improvements were positive, but wondered if any of the technologies were applied at scale on real problems. Sarkar responded that they went up to tens of thousands of cores and even a thousand plus cores for an interactive demonstration at the meeting. Referring to Berzins' second point about DSLs, the researchers are exploring and trying to push for a new language. However, many of the researchers prefer the library approach because of advances in C++ with the lambda construct. Sarkar stated that the point is to be careful about exploratory research and determining the right path to transition.

Chapman asked which activities and efforts would be most important to carry forward.

Sarkar suggested the relationships where students and projects are well connected with the labs and universities. Relationships ensure the research is more relevant because the lab partners will provide real applications in which to demonstrate their research.

Gregurick asked Helland about DOE’s plans for continuing this research given the budget constraints. Helland indicated that much of the X-Stack funding moved into the ECP. She thought they were reviewing proposals now and hoping to find about 30.

Huntoon commented that ASCAC had heard more this meeting about the critical issue of workforce training in strategic areas, the funding that goes to it, and attracting people to take positions within the DOE lab.

Mark Johnson discussed the role of HPC in advanced manufacturing. Overall tying HPC with advanced manufacturing is responding to a trend in manufacturing in general. For example it is possible to purchase a hand-held teraflop computer for approximately $500. To a manufacturer, this means that computational power, on an affordable basis, can be placed on the factory floor and used proactively to control energy, to control productivity, to control workflow, to control quality; affordability changes everything in how computational power is used.

There is a need for much better modeling of high fidelity models. Johnson quoted a colleague from Alcoa; “as soon as you start transitioning any manufacturing problem from being a process problem to being an information technology problem you have the ability to innovate your product sector at the pace of Moore’s Law”.

Approximately 1/3 of overall energy goes into manufacturing sector, creating a mission to use that energy more effectively, but that energy is also money which affects the economy. An efficient manufacturing sector means a more productive manufacturing sector, which drives the base economic growth. As a Brookings Institute study, *America’s Advanced Industries: What they are, where they are, why they matter*, February 2015 indicated, for every 1 job in advanced manufacturing, 4 jobs are created outside of advanced manufacturing. On a productivity-adjusted basis the U.S. is one of the most cost-effective places to manufacture. Technologies from managing energy, which is about 25% of average manufacturing costs, drives economic growth and security, therefore having manufacturing in the U.S. is a basis of our economic security, and political stability and security.

Two broad areas the Office of Advanced Manufacturing (AMO) to work in are how to develop a robust U.S. clean energy economy where products are developed and manufactured in the U.S., and to make the entire U.S. manufacturing sector more competitive by making it more energy productive. This refers to the mission of the AMO and two issues from the AMO legislative authority, one is how to have energy efficiency more effectively in the manufacturing sector and the other is referred to have as “lifecycle energy impacts” as a result of manufacturing.

Energy intensive industries, Petroleum Refining, Chemicals, Wood Pulp & Paper, Primary Metals, Food Processing, and Glass & Cement, average 3/4 of the total energy use in manufacturing.

The AMO technology assistance programs are a key component of President's Better Buildings Initiative to improve energy efficiency of commercial and industrial buildings by 20% by 2020. Superior Energy Performance™ is a certification program helping facilities meet the ISO 50001 energy management standard and verify savings they achieve. Combined heat and power technical assistance partnerships is a country-wide program providing outreach to the community and industrial assessment centers that conduct energy assessments and provide training.

AMO has two ways of conducting R&D; through consortia and through individual R&D

projects. Consortia are needed because the manufacturing sector it is disaggregated. There is no longer a company like Ford who owns everything from the iron mines to the dealerships. Instead, manufacturing is a tiered system in the industrial sector; therefore people must work across and down tiers for everyone to receive benefit (e.g. public private partnerships).

AMO has been the leader for DOE in the National Network for Manufacturing Innovation (NNMI) with 11 Manufacturing Innovation Institutes launched to date. NNMI's are federally funded at $70M over five years matched by at least 50% from industry. Four NNMI regional centers of excellence were highlighted including IACMI in Knoxville, TN, Smart Manufacturing at UCLA/SMLC Lead, Critical Materials Institute, and Modular Chemical Process Intensification. AMO is developing manufacturing demonstration facilities such as the one in Knoxville, TN which has shared innovation space, supercomputing capabilities, and scientific infrastructure.

AMO is also involved with individual R&D projects, which include materials development, process development, dewatering, and scale up for electrolytes for batteries. AMO is working with SC on the acceleration of materials development and the Materials Genome challenge, moving from discovery to scale up. The project takes data at the research level and predicts what happens when it is scaled up to the kilograms scale and then the ton scale, thus reducing the scientific uncertainty of the scale up point allowing private capital to flow in and solve the problem. The High Performance Computing for Manufacturing (HPC4Mfg) program was established in 2015 with LLNL to apply HPC to critical manufacturing challenges. Each round offers $3M in funding and applications are due every 6 months. Next Generation Electric Machines (NGEM) which focus on developing energy efficient, high power density, integrated medium voltage drive systems which has the potential to save 1.6% of the total U.S. electricity consumption each year.

Dunning mentioned small businesses lack expertise using computers as well as software. He asked if the AMO has run into the same problem. Johnson indicated AMO has a range of companies with whom they are working and have noticed the lack of expertise to be true. AMO is working with intermediaries to train and assist small businesses.

Reed asked how AMO is addressing the paperwork challenge that occurs when small businesses work with government or universities. Johnson said that to the extent possible this is the good fight to fight. Johnson provided an example where the process was examined with people on both ends. In one instance, a full loop in the approval process was identified and AMO addressed this through memoranda signed by key decision makers.

Reed opened the floor to the audience requesting any public comment. Hearing no requests, Reed thanked ASCAC members and adjourned the meeting at 12:13 p.m. Chalk announced that the next ASCAC meeting will be December 20-21, 2016.

Respectfully submitted,
Tiffani R. Conner, PhD, PMP
Oak Ridge Associated Universities
October 20, 2016