Meeting Minutes
Advanced Scientific Computing Advisory Committee
December 20-21, 2016
American Geophysical Union, Washington, DC 20024

ASCAC Members Present

Martin Berzins
Keren Bergman
Barbara Chapman
Jacqueline Chen
Silvia Crivelli
John Dolbow (teleconference)
Jack Dongarra
Thom Dunning (teleconference)
Tim Germann
Susan Gregurick
Anthony Hey
Gwendolyn Huntoon (teleconference)
Richard Lethin
David Levermore
Satoshi Matsouka
John Negele (teleconference)
Linda Petzold (teleconference)
Daniel Reed (Chairperson) (teleconference)
Krysta Svore (teleconference)
Dean Williams

ASCAC Members Absent

Vinton Cerf
Vivek Sarkar
Juan Meza

Also Participating

Christine Chalk, ASCAC Designated Federal Officer, Program Manager, Oak Ridge Leadership Computing, Office of Advanced Scientific Computing Research (ASCR), Office of Science (SC), Department of Energy (DOE)
Jody Crisp, Oak Ridge Institute for Science and Energy (ORISE)
Tiffani R. Conner, ORISE
Hal Finkel, Argonne National Laboratory (ANL)
William Harrod, Director, Research Division, ASCR, SC, DOE
Barbara Helland, Acting Associate Director, ASCR, SC, DOE
Earl Joseph, Program Vice President for High-Performance Computing, IDC Research Inc.
Paul Messina, ANL
Cherry Murray, Director, SC, DOE
Ceren Susut-Bennett, ASCR, SC, DOE
Deborah Wince-Smith, President and CEO, Council on Competitiveness (USCOC)

Attending

Adolfi Hoisie (Pacific Northwest National Laboratory, PNNL)
John Sarrao (Los Alamos National Laboratory, LANL)
James Hack (Oak Ridge National Laboratory, ORNL)
Vients Klees (Broohhaven National Laboratory, BNL)
Louis Terminello (PNNL)
Randall Laviolette (DOE)
Altaf Carim (Office of Science and Technology Policy)
Claire Cramer (DOE)

Abani Patra (DOE, ASCR)
Alexander Lazalere (USCOC)
Candace Culhane (LANL, Contractor)
Beth Kaspar (LANL, Contractor)
Carolyn Lauzon (DOE, ASCR)
Sandra McLean (DOE, SC)
Ben Kallen (Lewis-Burke)
Laura Biven (DOE, SC)
Leland Caghan (Lewis-Burke)
Arthur Maccabe (ORNL)
Arthur Bland (ORNL)
Justin Whitt (ORNL)
Suzy Tichenor (ORNL)
Susan Coglan (ANL)
Robert Voigt (Krell Institute)
Steve Hammond (Lawrence Berkeley National Laboratory, LBNL)
James Ang (Sandia National Laboratory, SNL)
Ron Brightwell (SNL)
Paul Hovland (ANL)
Eliana Perlmutter (Lewis-Burke)
Jeff Nichols (ORNL)
David Brown (LBNL)
Jeff Hittinger (Lawrence Livermore National Laboratory, LLNL)
Douglas Kothe (ORNL)
Lori Deachen (LLNL)
Richard Carlson (DOE, ASCR)
Angela Thevenot (DOE, ASCR)
Lucy Nowell (DOE, ASCR)
Mike Ignatowski (LLNL)
Julie Stambaugh (DOE, ASCR)
Tameka Morgan (DOE, ASCR)
Rishi Khan (Extreme Scale Solutions, LLC)
Thuc Hoang (National Nuclear Security Administration, NNSA)
Vince Dattoria (DOE, ASCR)
Mike Schulte (AMD Research)
David Etim (NNSA)
Michael Papka (ANL)
Sonia McCarthy (DOE, ASCR)
Steve Raasch (AMD Research)
Richard Gerber (NERSC)
Punita Sinha (LLNL)
Jack Wells (ORNL)
Jonathan Carter (LBNL)
Karin Remington (Biological and Environmental Research Advisory Committee)

Tuesday, December 20, 2016
Morning Session

The U.S. Department of Energy (DOE) Advanced Scientific Computing Advisory Committee (ASCAC) meeting was convened at 8:39 a.m. EST on Tuesday, December 20, 2016, at the American Geophysical Union by **Anthony Hey**, standing in for Chair Dan Reed. ASCAC members introduced themselves.

Cherry Murray, DOE, Director of the Office of Science (SC) provided background on the SC. DOE has four mission areas: Energy, Science, Nuclear Security, and Environmental Cleanup. SC operates six programs: Advanced Scientific Computing Research, Basic Energy Sciences, Biological and Environmental Research, Fusion Energy Sciences, High Energy Physics, and Nuclear Physics. Approximately 60% of the SC budget supports facilities across the 17 DOE National Laboratories (Labs).

Typically the Federal Budget Cycle is operating on three budgets at one time. This year is complicated because there have been two continuing resolutions (CR), one until December 9, 2016 (H.R. 5325) and the second a 6 month CR into 2017 (H.R. 2028). The FY17 SC budget request was up 4% over FY16.

The ASCR budget is split between research and facilities. There is bipartisan support for increasing the Science budget by 3%. The Senate mark for ASCR FY17 budget is $656M, an increase of $35M, while the House mark is $621M, the same level as FY16.

SC created budget scenarios for the transition to the new government. Office of Management and Budget (OMB) requested a scenario called "current services", a 2% growth rate per year over 5 years. SC also prepared an internal "requirements scenario" with an increase of 7%. The requirements scenario would fund all requirements except exascale computing.

DOE's transition team is led by Tom Pyle, President of American Energy Alliance. The team consists of 11 members with briefings beginning on December 7, 2016. The Secretary of Energy, Dr. Ernest Moniz, met with nominated Secretary of Energy, Governor Rick Perry on December 15, 2016. Steve Binkley will become the Acting Director of SC-1 and SC-2 on January 17, 2017 until a new SC Director is nominated and confirmed.

Williams asked how DOE's visibility and standing in technology and science could be improved within the new government. **Murray** indicated the new government will appoint 4,000 people. Some of the new appointees are familiar with DOE activities, but many are not, therefore constant communications are required. *Lab Day on the Hill* for the Labs is extremely helpful. It is not well known that DOE funds all of the major research universities.

Berzins asked about the international arena with regard to equivalent funding of programs against which DOE competes. **Murray** stated that in Japan and Australia the budgets will be flat, China is showing exponential budget expansion, and Europe and the United Kingdom (U.K.) are unclear. British Prime Minister Theresa May announced there would be funding for Science and Technology in the U.K. but it is expected to be on the technology side. China is the major competitor to the United States (U.S.). **Hey** added that the U.K. is discussing ~£500M for infrastructure.

Barbara Helland, SC, Acting Associate Director for ASCR shared that a third division, Advanced Computing Technologies, is being added within ASCR and that division is seeking a Director.

ASCR FY17 budget has two proposed additions including research and development (R&D) for post-Moore's law computing, and support of the Brain Research through Advancing Innovative Neurotechnologies® (BRAIN) Initiative. However, CRs mean no new starts for projects. ASCR's budget request to Congress was 6.8% higher than the FY16 enacted appropriation. The two CRs decreased the ASCR budget; H.R. 5325 held budgets at the FY16 level and included a 0.5% cut, H.R. 2028, in effect through April 2017 lowered the budget cut to 0.2% instead of 0.5%.

Helland provided updates on ASCR's research. Research areas in ASCR are Applied Mathematics, Computer Science, Computational Partnerships, and Next Generation Networks for Science (NGNS). Scientific Discovery through Advanced Computing (SciDAC) is moving to SciDAC-4, and NGNS is focused on what Energy Sciences Network (ESnet) needs to evolve.

The Early Career program currently selects 2 new awardees from Labs and Universities each year. ASCR anticipates selecting 3 Lab and 4 University awardees for the 8th cohort.

ASCR is exploring Artificial Intelligence (AI) and sponsored a workshop in 2015 on machine learning which led to a Funding Opportunity Announcement (FOA) in 2016 titled "Machine Learning and Understanding for High Performance Computing Scientific Discovery" (DE-FOA-0001575). ASCR is currently in the awards process for this FOA.

Mathematical Multifaceted Integrated Capability Centers (MMICCs) were created to address long-term mathematics research challenges. A program review of MMICS was conducted

in October/November 2016 with positive outcomes that showcased the benefits of the MMICs program.ASCR plans to recompete the MMICs centers in FY 2017.

The Center for Advanced Mathematics for Energy Research Applications (CAMERA) at LBNL began as a Lab Directed Research and Development (LDRD) project. ASCR ran a pilot project which led to the current joint funding between Basic Energy Sciences (BES) and ASCR.

A workshop was held December 8-9, 2016 in Rockville, MD on Smart Networks. The purpose was to bring stakeholders together to discuss emerging opportunities and challenges for the network of the future.

ASCR provides high end computing facilities to users; these include NERSC, Leadership Computing Facilities at Argonne National Laboratory (ALCF), and Oak Ridge National Laboratory (OLCF). ASCR facilities and project updates included the Leadership Computing Facilities, National Energy Research Scientific Computing Center (NERSC), ESnet, Exascale, *PathForward*, and Cross-cuts.

Cori, at NERSC, started stability testing in November 2016 and will go into early science period in January/February 2017 if the testing is successful. The Cori upgrade was NERSC-8; the NERSC-9 project, an Alliance for Application Performance at Extreme Scale (APEX) partnership, is to acquire and deploy a pre-exascale high performance computing (HPC) system in 2020. A request for proposal (RFP) has been issued for NERSC-9.

ESnet6's equipment needs to be upgraded; the last upgrade was supported by American Recovery and Reinvestment Act (February 17, 2009) funds. The focus of ESnet6 will be on designing a network to support exascale computing with attention to exponential science data growth, reliability, and cyber resiliency.

ASCR exascale requirements reviews have been completed. The workshop reports for exascale requirements will be completed in Mid 2017. By March 2017 a first-draft response to findings will be organized and shared with ASCR Headquarters. A cross-cut workshop will be held in March 2017 with a cross-cut report completed by June 2017.

The Exascale Computing Project (ECP) mission need statement has been approved at the Deputy Secretary level. The Deputy Secretary asked if ECP could be sped up. There have been verbal approvals for CD1. ECP has used the Linac Coherent Light Source-II (LCLS-II) model of multi-lab involvement. There are 6 labs (DOE and NNSA) involved in ECP management and the ASCR/NNSA Memorandum of Understanding has been updated with additional management details.

There are 6 SciDAC FOAs available in FY 2017 and SciDAC-4 Institute FOAs are planned for release in late Spring 2017. SciDAC FOA/LAB 17-1674 topic is on modeling and simulation of nuclear fuels via HPC and is open to Labs, Universities, and Industry. Letters of Intent (LOI) are due February 2017 and proposals are due in April 2017.

There are several planned workshops and symposia in 2017 including the Quantum Testbed Stakeholder workshop in February 2017. A workshop on the Energy Consequences of Information Processing is being co-sponsored by the Air Force Research Laboratory, the Air Force Office of Scientific Research, and DOE ASCR in late February 2017.

Helland segued into the Committee of Visitors (COV) Charge for ASCAC to examine the whole ASCR research portfolio with a report expected at the ASCAC fall 2017 meeting.

Reed said ASCAC would assemble a committee and address the charge.

A morning break was called at 10:04 a.m.

The meeting was reconvened at 10:21 a.m.

Paul Messina, Argonne National Laboratory, provided an update on the activities of the ECP. The ECP aims to transform the HPC ecosystem and make major contributions to the nation. ECP has 4 project goals: foster application development, create usable software, enable rich exascale ecosystem, and ensure U.S. HPC leadership. ECP uses co-design and integration encompassing applications, system software, hardware technologies and architectures, and workforce development.

A capable exascale computing system delivers 50x more performance than today's 20 petaflop systems, operates at 20-30 megawatts, is sufficiently resilient, and includes a broadly supportive software stack.

ECP was originally designed as a 10-year project. However, the Deputy Secretary of Energy requested a shorter delivery time yielding a 7-year project schedule. The shorter timeframe will require advanced architectures to be developed by vendors.

The 4 key challenges of exascale are parallelism, memory and storage, reliability, and energy consumption.

There are 6 labs involved in ECP management, 3 from DOE and 3 from NNSA.

ECP applications will be chosen based on mission need. Fifteen application projects and 7 seed efforts were initially selected. ECP wants to increase university participation in the project and focus on big data and machine learning. ECP currently has 3 Industry partners, 16 Lab partners, and 12 University partners. ECP is funding 4 co-design centers: CODAR: A Co-Design Center for Online Data Analysis and Reduction at the Exascale, Block-Structured AMR Co-Design Center (Block-Structured AMR), Center for Efficient Exascale Discretizations (CEED), and Co-Design Center for Particle Applications (CoPA). CODAR is addressing online data analysis and reduction motifs. Block-Structured AMR is addressing structured mesh, block-structured AMR, and particle motifs. CEED motifs include unstructured mesh, spectral methods, and finite element methods. Finally, CoPA will address particles, involving particle-particle and particle-mesh interactions. ECP training is explicitly outlined in the work breakdown structure (WBS). ECP inherited the ANL training program on extreme scale computing. ECP training will also focus on software engineering in an agile fashion.

In terms of software technology, ECP will build a comprehensive and coherent software stack by extending current technologies to exascale, performing R&D, coordinating with vendor efforts, and developing and deploying high-quality and robust software products. Currently there are 8 DOE Labs, 15 Universities, and 2 Vendors working on software technology projects. A request for information (RFI) is expected in FY17-18 focused on universities and vendors. A gap analysis will identify missing aspects of the software stack. The results of the gap analysis will be targeted RFIs and RFPs aimed at closing the gaps.

The objective of the hardware technology focus area is to fund R&D to design hardware that meets ECP's targets for application performance, power efficiency, and resilience. Two programs support this effort, *PathForward* and *LeapForward*. LLNL is managing the procurement of *PathForward*. In August 2016, 14 *PathForward* proposals were received, 6 of these proposals were selected for Statement of Work negotiations, and contract awards are expected in January 2017. *LeapForward* is a new funding mechanism with the aim of improving the quality and number of responses to the Advanced Architecture Exascale Systems RFP. *PathForward* reduces the Technical Risk for non-recurring engineering (NRE) investments in the 2023 Exascale System. In contrast, *LeapForward* accepts Technical Risk in return for game-changing impact on the U.S. computing ecosystem and potential for early deployment in a 2021 exascale system.

SC and NNSA will procure and install the exascale systems. NRE contracts have been added to incentivize vendors to address gaps in their systems. ECP will operate testbeds for users each year throughout the project.

ECP has met milestones and has planned dates to achieve others. CD-0 was achieved in July 2016 and CD-1 and CD-3a were attained in November 2016. CD-2 and CD-3 are projected to be realized in September 2019 with CD-4 reached in September 2023.

Dongarra asked about additional costs with each year of schedule reduction. **Messina** said that the 7-year plan is between $3.9B and $5.9B more expensive than the 10-year plan.

Chapman was interested in the perceived risks and how those would be handled. **Messina** mentioned four high level risks that have been noted: 1) the availability of people with sufficient expertise and skills; 2) that *PathForward* investments will have an impact; 3) that the software stack will be integrated and meet the needs of applications; and 4) the fidelity, functionality, and capability of the applications. ECP has a mitigation strategy for each of these.

Bergman asked about the top 2-3 compromises expected between the 7-year and 10-year plans. **Messina** indicated there were no conscious changes to innovation or risk levels. The biggest compromise is that activities originally planned during the 10-year project will have to be performed outside the project in the 7-year plan. These activities include applications and software testing as well as final development activities.

Dunning asked about the cost increase due to software development. **Messina** said most of the increases will be for applications, software, hardware, and co-design. The intent is to provide more funds to the teams to offset the costs.

Matsouka compared the cost of ECP machines to the Post-K machine in Japan. He asked if there is an inherent mechanism to control costs of the ECP machines to make them affordable to the Labs. **Messina** indicated that *PathForward* and *LeapForward* are not control mechanisms but something that could pay off to develop systems that will be more affordable. The first 2021 system is estimated to cost substantially more than $200M. Investments with vendors are also a viable option for reducing costs.

Lethin asked how ECP planned to protect intellectual property and cybersecurity. **Messina** said they were putting controls in place for cybersecurity but ECP has a policy of open-source where possible. A lot of intellectual property will be available because it will be through software.

Levermore offered a reflection that a fundamental change in the kind of algorithms that will be used, data science, and machine learning. To get the full potential of ECP machines and the benefits of investment, the ECP community has to rethink the algorithms and the demands of the algorithms on the hardware. The bottleneck may be data pushing around and communication. **Messina** said in software technology there are elements aimed at these issues such as data storage, data transfer, analysis, and visualization. **Levermore** responded that he was illustrating an algorithm interplaying with the modeling problem, not a software problem. **Messina** added that the ECP approach is that as applications of these newer types are developed teams will identify what is missing and that includes algorithms. The mathematical libraries teams are being encouraged to have applied mathematicians as part of their team.

Hey asked if ECP had looked at different motifs. **Messina** confirmed they had.

Chen asked if there is a way to get more information, from a co-design perspective, regarding the abstract machine models out to the ECP community well ahead of the 3-year time window for the *PathForward* and *LeapForward* projects. **Messina** mentioned that the

PathForward quarterly reviews will be available to the ECP community and some of the necessary information is expected to be in those reports.

Crivelli asked if ECP has a plan to use associations with universities to contribute to workforce development and for a projection of how many people will be needed and are currently available. **Messina** said universities are heavily involved; 75% of the participants (advanced graduate students and post-docs) are getting training. Continued use of universities to do the training is anticipated. Messina speculated that more than 200 people will be needed over the next couple of years.

Dongarra asked if ECP planned to transition software to the vendors. **Messina** hopes to do that and said it is happening in some project funding because of long-term, pre-existing relationships. Where the transition does not happen, there will be a good repository of data.

Negele asked about the interplay between ECP and SciDAC in two areas: competition for the same human resources, and financial support of SciDAC. **Messina** agreed that workforce availability is an issue that would affect SciDAC, but had no solutions. However, a lack of funding and maintenance support for Aurora, Summit, Cori, etc. will not be the case because ECP software that is developed will be implemented on those machines. **Helland** added that the SciDAC partnership should be fairly robust. SciDAC-4 is focused on the Aurora's, Summit's, and Cori's but ECP is focused beyond that.

Deborah Wince-Smith, Council on Competitiveness (USCOC), discussed the need for new public-private partnerships in microelectronics. The USCOC was established by John Young, former CEO of Hewlett-Packard, who chaired President Ronald Reagan's Commission on Industrial Productivity. USCOC's vision has been to bring together CEOs from all sectors of the U.S. economy to join with university presidents, labor leaders, and national laboratory directors to understand next-generation drivers of productivity, maintaining a high standard of living for U.S. citizens, and to succeed globally. The goal of USCOC is to develop and implement policy recommendations.

There are currently four scientific revolutions occurring: digital, biotechnical, nanotechnical, and cognitive.

The digital revolution is converging with the physical world in all dimensions. USCOC believes the U.S. must maintain its leadership in semiconductor manufacturing, the fundamental building blocks of the new digital age. This leadership includes design, next generation R&D development, manufacturing, global supply chains, job creation, and new infrastructure in the U.S.

Semiconductors are a fundamental enabler and could add more than $15T to global gross domestic product (GDP) by 2030. Semiconductor products are the 3rd largest class of U.S. exports behind aircraft and automobiles. U.S. companies account for ~50% of the world's semiconductor market by revenue, and the industry employs about 180,000 workers with nearly 1M associated jobs. The jobs in this industry are high wage and highly skilled, but the industry is under constant pressure to innovate and invests heavily in R&D.

The semiconductor industry is at a critical juncture as the digital revolution unfolds and faces strong competitive, economic, and technological challenges. The three challenges presented were 1) commercial semiconductor fabrication migrating to Asia, 2) capital intensity of the semiconductor industry, and 3) profound technical barriers for manufacturability. Implications of these challenges are stalling of microelectronics innovation, constraints in the ability to improve the U.S. military, homeland security, intelligence gathering systems, and impacts on energy supply.

From the perspective of a public/private partnership, maintaining U.S. leadership in semiconductors will require bringing together all agencies in U.S. government, industry, academia, and national labs. The partnerships have to look at multi-scale problems starting with fundamental science moving up to devices and systems.

National security depends on fielding the most advanced weapons, communications, and cryptographic systems, all of which rely on microchips, semiconductors, and imagers that have not been compromised by tampering or counterfeits. Microchip manufacturing for all of the security sensitive applications such as smart grids, power plant controls, autonomous vehicles, and manufacturing plant controls, has migrated to Asia. The microchips are coming from the countries that lead the U.S. intellectual property theft and counterfeiting.

Four leading edge microchip technology companies in the world can provide the products needed: Taiwan Semiconductor, Global Foundries, Korea Samsung semiconductor, and Intel. New financial and public/private models that take the U.S. into this new world of next-generation microelectronics, trusted foundries, and ensures U.S. leadership in the development, design, manufacture, and global supply chain are needed.

Hey mentioned that the highest selling microchip in the world, the Acorn RISC Machine (ARM) microchip, has a strategy of outsourcing their microchip manufacturing. **Wince-Smith** said just focusing on the design and development, but not the manufacturing and production of these devices and systems, is not a long-term, competitive, national security strategy.

Lethin asked for concrete recommendations for ECP with regard to the manufacturing of components for the exascale machines. **Wince-Smith** said USCOC will be developing those sets of recommendations in 6-12 months in terms of needing the public/private partnership and new financing models. USCOC wants to involve the broadest sector of leaders in the country across Industry, Labs, and Universities.

Hey adjourned ASCAC for lunch at 12:23 p.m.

Tuesday, December 20, 2016
Afternoon Session

Hey reconvened the ASCAC meeting at 1:46 p.m.

Ceren Susut-Bennett, ASCR, provided a preliminary plan for computing “Beyond Moore’s Law” based on the National Strategic Computing Initiative (NSCI) strategic objective #3, “Establishing, over the next 15 years, a viable path forward for future HPC systems even after the limits of current semiconductor technology are reached (the "post-Moore's Law era").”

ASCR first discussed Quantum Computing at the DOE/ASCR workshop in February 2015. The consensus of the workshop was that quantum computing had reached a level of maturity warranting consideration of the impact to the DOE mission in the near and long term. Quantum Algorithms, Quantum Simulation, Models of Computation and Programming Environments, and Co-Design Approach were the four research opportunities listed in the report.

The Quantum Information Science (QIS) Interagency Working Group was formed in 2014. An RFI was released to solicit industry views on QIS in 2015, a public report was released in July 2016 and a White House Forum was held in October 2016. The public report summarized developments in and impacts of QIS and reviewed existing QIS programs. The report identified 5 impediments to progress and recommended a path forward with 3 components including core programs, strategic investment, and close monitoring of QIS.

The October 2015 Nanotechnology-Inspired Grand Challenge for Future Computing spearheaded the Grand Challenge Committee whose objectives were to highlight potential areas of Federal R&D focus and investment and to describe technical challenges, opportunities, and potential applications. Game-changing capabilities from the Grand Challenge were expected to yield emerging computing architectures, intelligent big data sensors, machine intelligence, and cybersecurity.

Both BES and ASCR have been involved in Neuromorphic Computing since the launch of the Grand Challenge. ASCR will explore "beyond Moore's Law" technologies with an approach to advance new hardware and software computing paradigms to extend Complementary Metal-Oxide-Semiconductor (CMOS) and Non-CMOS computing.

Quantum Simulation Teams will bring ASCR expertise to Quantum Computing. Quantum Testbed Facilities will evaluate the utility and facilitate technological development of quantum computing resources, and actively explore neuromorphic computing, and leverage HPC activities.

Several workshops for community engagement have occurred since 2014. A Quantum Testbed Stakeholder workshop will be held in mid-February 2017 in Washington, D.C.

Helland introduced a second Charge to ASCAC to form a subcommittee to review opportunities and challenges for future HPC capabilities.

Hey asked if companies that have performed quantum computing activities, such as Microsoft and IBM, were invited to the meetings. **Svore** confirmed that Microsoft, IBM, and leading experimentalists have been involved in the meetings.

Svore asked what was envisioned for the quantum simulation teams, if the simulation approach was in line with adiabatic quantum computing, quantum annealing, or the circuit model. **Susut-Bennett** indicated these teams have not yet been fully defined.

Bergman asked what other technologies under the umbrella of post-Moore's law are being considered outside of quantum computing. **Susut-Bennett** recognized the post-Moore's Law umbrella was large and said what makes sense and what the niche is for ASCR needs to be defined. **Helland** said ASCR is asking ASCAC's help on these definitions.

Lethin sought clarification on the budgeting for quantum computing in light of the CR restriction on new starts. **Susut-Bennett** noted that most of the requests for quantum computing are listed as new items in the FY17 request, however ASCR is in the planning stages now and there may be ways to move quantum computing forward.

Bergman mentioned that at Super Computing 17 (SC17) the first workshop on post-Moore's Law emerging technologies interested the community.

Matsouka asked about the prospects of other technologies that have a higher potential for solving mission-critical applications. **Susut-Bennett** indicated these prospects are part of the charge to ASCAC and that ASCR is open to ideas.

Hey asked if ASCR was working with NIST and others. **Susut-Bennett** confirmed ASCR was doing so.

An Online participant asked Susut-Bennett to describe the coordination with the National Science Foundation (NSF) HPC initiative. **Susut-Bennett** said ASCR was in contact with NSF colleagues on various aspects of quantum computing. Denise Caldwell, NSF Division Director, Division of Physics, is one of the co-leads in the QIS interagency working group.

Chalk closed the discussion indicating that ASCAC would form two subcommittees on the Charges and asked that ASCAC members interested in participating email Chalk and copy

Reed. Chalk clarified that a subcommittee needs to have 1-2 ASCAC members but the majority must be outside experts.

Hal Finkel, Argonne National Laboratory, discussed Field Programmable Gate Arrays (FPGAs) for supercomputing. The FPGA field has transitioned to a point of discussion about what is needed to move forward. FPGAs can serve as effective accelerators for workloads such as in bioinformatics, machine learning, and neural networks.

FPGAs are effective because of efficient data movement and parallelism. Compute energy decreases because data is moved less often. A central processing unit (CPU) moves data around the microchip while FPGA data flows through the microchip making better use of the transistors on the microchip.

Over 50% of the power on a CPU goes to fetch and decode functions and a small amount of energy goes to computation. On FPGAs the fetch and decode functions can be eliminated.

Modern FPGAs contain a large number of digital signaling processing (DSP) blocks. For example, the Intel Stratix 10 will have up to 5,760 DSP blocks and 11,721 blocks of random-access memory (RAM).

Industry is beginning to pair FPGAs and CPUs together. Intel/Altera is producing Xeon + FPGA system and Xilinx is producing ARM + FPGA systems.

FPGAs can easily support custom data formats, on-the-fly data compression, and encoding, for example.

Programmability is the sticking point for FPGAs, but there are tools which are usable now. Programming techniques range from low risk/ low user difficulty to high risk/high difficulty such as vendor-optimized libraries, overlay architectures, high-level synthesis, and Verilog/VHDL.

A key challenge in programmability is compile time, but overlay architectures can help. Overlay architectures take advantage of built-in blocks and can run at near peak clock rate.

Challenges remain for FPGAs but are relatively well-defined. Open Multi-Processing (OpenMP) is in its infancy, high-level synthesis is just now starting, CPU + FPGA systems are new, high-performance overlay architectures do not yet target HPC workloads, and no one has yet created an HPC-practical toolchain.

In summary, FPGAs offer promising direction towards higher floating-point operations per second (FLOPS) per watt. FPGAs will naturally fit into the accelerator-infused HPC ecosystem. FPGAs can compete with CPUs/ graphic processing unit (GPUs) and overlay architectures can address FPGA programming challenges.

Williams asked about debugging. **Finkel** said debugging is great in theory, but not in practice. Current debugging tools are far too low-level for effective use in an HPC environment. There's a lot of work to do in this area.

Karin Remington asked for an example where a bioinformatics team used FPGAs. **Finkel** said there are FPGAs in commercial genomic sequencing hardware. For example, Fernandez, Villarreal, Lonardi, and Najjar (2015) found that FPGA hardware accelerated sequence-mapping tool (FHAST) compared to Bowtie mapping tool generated an acceleration 70x faster than CPUs in DNA sequencing instructions.

Chapman asked about the correct programming high-level approach if OpenMP can work. **Finkel** said OpenMP can work. An OpenMP to Open Computing Language (OpenCL) translator has been built and can translate from OpenMP into OpenCL and then OpenCL can translate into FPGAs. The programming challenge has been to effectively take advantage of on-board resources, specifically the DSP blocks.

Lethin mentioned recent interest from Microsoft Catapult and Amazon FPGA, both available in the cloud, and asked why FPGA are now viable and more interesting now. **Finkel** mentioned 3 reasons: 1) the market is pushing towards having combined CPU/ FPGA packages, 2) FPGA companies, especially Intel, are now encoding Floating Point logic into the FPGA fabric itself which gives throughput advantage, and 3) programmability. The market drivers were not present 10 years ago. High-level synthesis technology has changed significantly and received a lot of academic attention, the FPGA companies have performed extensive R&D, and FPGAs have moved from an academic curiosity to a deliverable product in the last 3-4 years.

Levermore asked how FPGA performance changes at 64-bit. **Finkel** explained that FPGAs at 64-bits are limited by the amount of logic needed to shift things on the microchip. 64-bit multiplication needs 3-5 integer multipliers at various configurations, which cuts down the raw throughput by a factor of 3 to achieve the peaks at 64-bit rate. 64-bit is still much more power efficient than a CPU or GPU, but the raw throughput per device goes down.

Dongarra asked if this is an IEEE floating point operation FPGA and if it performs IEEE. **Finkel** confirmed the FPGA is IEEE and performs IEEE.

Dongarra asked about the cost of scaling down to 16-bit or lower. **Finkel** said 16-bit is reasonable. Applications such as machine learning can use much lower precision. Using iterative solvers also allows starting with lower precision.

Ang mentioned that 10 years ago resilience was FPGA's Achilles heel and asked if this was still an issue on a large scale. **Finkel** said market pressure has helped address resilience.

ASCAC was adjourned for the afternoon break at 3:04 p.m. and reconvened at 3:19 p.m.

William Harrod, ASCR, provided an update on the ASCR Long Range Planning (LRP) activity. The purpose of LRP is to create a necessary new vision, to showcase computing is fundamental, to establish ambitious goals, and to enhance U.S. leadership. The LRP team consists of 13 members from the Labs and ASCR.

Discussions on the LRP development process began in November 2015 and a preliminary LRP document was completed in December 2016. Two workshops were held in 2016 and one additional workshop will occur in 2017. The strategic plan for LRP will be based on the LRP document and developed by DOE Headquarters and ASCR.

The ASCR Vision has changed over time from simulation science to data-driven advanced computing.

There are six goals and four major research initiatives in the LRP. The goals focus on responsiveness, return on investment (ROI), productivity, computational technologies, HPC workforce, and U.S. competitiveness. The research initiatives focus on convergence, AI, new paradigms, and ubiquitous use.

Cross-cutting directives, which are the common thread across all the research thrusts, are complexity, productivity, and software sustainability.

Levermore asked Harrod to describe the HPC ecosystem mentioned along with U.S. competiveness (Goal 6). **Harrod** said the focus is on development that can be deployed in a facility and released as part of the HPC ecosystem. **Levermore** stated that in the HPC environment, DOE has a higher level of robustness than other government agencies or industry because of its use of hierarchies of models. **Hey** added that the differences between computer science programs are pronounced and DOE recognized that there was innovation in continuous funding for software sustainability.

Chen asked about programming models for data intensive workflows. **Harrod** wants a way to automate the workflow process to be easier thus enabling and encouraging scientists to come up with experiments.

Gregurick encouraged DOE to pay attention to AI and Large Scale Data Integration of heterogeneous data types and data security, privacy, and encryption. **Harrod** indicated the AI thrust includes multi-modal data analysis and that other agencies are heavily involved with cybersecurity, but ASCR will focus on data integrity in the 4th LRP goal.

Crivelli suggested pursuing the integration of citizen scientists.

Matsouka asked if there are opportunities for DOE investment in AI. **Harrod** indicated that since there is a lot of investment in AI now DOE may be able to take advantage of the research in AI and integrate it, which in turn might grow AI overall. DOE can do well in implementing some of the AI algorithms with demonstrations on the HPC system, which is very much part of the data convergence problem.

Berzins asked if LRP is thinking about a multi-disciplinary approach in computer science research. **Harrod** stated that the data convergence strategy involves using co-design-like centers and multi-disciplinary teams.

Levermore suggested that there are new algorithms where Density Functional Theory calculations are using AI. He added there is a European community considering these algorithms, that they are relatively DOE centric, and that they should be on the radar screen.

Martin Berzins provided ASCAC with an update on the progress of the LDRD subcommittee. The LDRD program yields approximately 2000 papers and 400 inventions per year and supports 30% of all Lab post-docs. The LDRD subcommittee was charged with evaluating the processes and impact of LDRD at four DOE Labs.

Memberships on LDRD subcommittee were from across all program areas in SC.

Through six committee teleconferences and ~700 emails, the LDRD subcommittee discussed how to address the Charge to perform an independent review of the LDRD program across DOE (May 19, 2016). Available lab reports were divided among the members and resulted in a guidance document with a detailed set of questions for the labs to respond to.

The LDRD subcommittee will visit the four Labs (Lawrence Berkeley National Laboratory, Lawrence Livermore National Laboratory, Oak Ridge National Laboratory, and the National Renewable Energy Laboratory) in January and February 2017. A draft report will be produced by February 2017 and there will be a comment period on the initial report in March 2017. The final report from the LDRD subcommittee will be delivered to ASCAC in April/May 2017.

Bergman asked if the subcommittee was seeking a uniform assessment of LDRD across Labs or if differences between the Labs will be taken into account. **Berzins** said the key was to show that money is well spent and that the Labs understand how the money is being spent. It is important to recognize the uniqueness of the Labs and their different missions. The LDRD subcommittee does not expect to see a uniform approach, but expects to see that the Labs measure some of the same things and that each process is managed with care.

Williams inquired about the metrics being used such as how many LDRDs were initiated, how many technology transfers occurred, and how these fit into ASCR's future. **Berzins** said the LDRD subcommittee is also interested in knowing who is paying attention to the process and what comes out of the LDRD program. The impact for research may not be seen for quite some time, therefore the metrics mentioned are more difficult to track and report.

Gregurick suggested looking at the trajectory of the people LDRD has supported since the LDRD is meant to fund innovative, high risk research as well as innovative, new people. **Berzins** agreed that this is extremely important. There is a process in place for understanding that LDRD works, even if confidentiality makes it difficult to obtain personnel information. **Hey** asked Gregurick if NIH had something similar. **Gregurick** shared that NIH has early career stage awards for external research communities and internal LDRD-like awards.

Remington mentioned it was worth pointing out that that this Charge is beyond ASCAC; it is cross-cutting across all of DOE and the subcommittee reflects membership from across all of the DOE FACAs.

Earl Joseph, IDC, Inc. discussed economic models for financial ROI and innovation from HPC investments. IDC is creating a new economic model based on new data sets. Traditional economic models use existing data. The project is key to DOE because scientific and innovation leadership are more dependent on HPC, economic leadership results from enterprise application of supercomputers, there is global movement to gain leadership in innovation and economic progress by the broad application of HPC, and it supports the NSCI initiative.

In 2013, IDC conducted a pilot study testing 3 approaches to set the economic models. 208 cases of scientific innovation and industrial ROI were used to populate the models, and the pilot project created a new innovation index. The current 3-year study, sponsored by DOE and NNSA, will refine the models, collect more data points, and publish the findings. The three financial models are: 1) ROI based on revenues/ gross domestic product (GDP) generated, 2) ROI based on cost-savings and/or profit, and 3) ROI based on jobs created.

Project personnel found that data collection is more difficult than expected in part because over 50% of the examples have been rejected and collection requires multiple face-to-face visits. IDC has addressed these challenges by hiring a full-time person in China, hiring a part-time person in Japan, and adding a program to motivate HPC center directors to increase participation. Because the term "innovation ROI" was misunderstood, IDC is now using Return on Research (ROR).

Three innovation indexes are used in the project; one based on the importance of innovation, one on the broad impact on organizations, and a combined score of the first two to create innovation "Class" levels. There are 8 innovation classes ranging from Class 1 – top 2-3 innovations PLUS useful to over 10 organizations, to Class 7 – top 50 innovations PLUS useful to at least 2 organization, and Class 8 encompasses the remaining innovations in the study.

There are currently 673 examples in the database, 148 financial ROIs, and 535 innovation RORs. The top countries by accomplishment type are U.S. (304), China (126), and the U.K. (120). Academic and Industry sectors represent the most accomplishments in the database.

The results thus far indicate substantial ROI in HPC. For every dollar invested in HPC $515 of revenue is generated and $52 of profit or cost-savings is realized. The average HPC investment cost per job created was $270K; 2,335 jobs were created across the financial ROI projects with the majority in manufacturing and financial industries.

The average cost of Innovation ROR projects was $12.7M. The database has 525 examples of basic and applied projects, the majority in academia, followed by industry and government, respectively. The three indexes of innovations yielded a rating of 3.43 (Importance index), 4.54 (Impact index), and 3.64 (Class index) out of 5.

IDC would like to expand the models to be used as a predictive tool for forecasting the value of a new supercomputer and to measure the value of an existing HPC system. These future research ideas have been tested with the Post-K machine in Japan and the Riken K-computer.

The next steps in the study are to 1) obtain more data points/ ROI cases especially from Germany, China, France and Japan, and 2) to distribute the results more broadly.

Levermore expressed concern about the early publication of the metrics used and those being applied incorrectly for unintended purposes. **Joseph** indicated some of the concern is controllable through the data structure which protects identifiers.

Berzins asked what the definitions mean since the terminology does not seem precise. **Joseph** explained the team's process which includes the researcher, a steering committee, and a technical committee where needed. First the researcher explains his or her research. Second the steering committee evaluates the written explanation and makes a decision. If the steering committee determines an expert is needed, a technical committee member reviews the application. If the steering or technical committee changes any of the ratings by more than 1, the technical committee's rating is accepted; if the rating changes by more than 2 points, the researcher is contacted again. Currently the steering committee rejection rate is around 40% and the technical committee rejection rate is 20%-25%.

On the ROI side, the process is similar. However, whenever there is a question about returns, the proposal is sent to an economic committee consisting of people in each domain who understand financial returns.

Berzins asked about future ROI estimates and developing a robust methodology to evaluate those returns on a sound financial level. **Joseph** stated that the majority of items that have a future value to them are rejected. In general, the ROI items included in the database are those where the company has already generated revenues or have very tight tracking on what they are going to sell. Joseph's team engaged the financial community on the front end because every dollar made and spent is tracked.

Hey worried about ROI in China, speculating that China must have invested billions in developing their microchips. **Joseph** clarified that if HPC research is conducted to make the HPC center more productive it is not included in the study. The team is looking for external outcomes.

Ang asked if the team discriminated against different types of investments, such as applications versus system software versus hardware, or open source software versus closed source software. **Joseph** explained that the investment is the individual researcher's cost for doing a project; what did the project cost that researcher, sitting in their country, in their industry, using that system.

Finkel asked how sampling bias towards more successful projects is avoided. **Joseph** acknowledged that the current sample is biased towards successful projects and emphasized the need for a higher quantity of data. The team has a proposal for the next stage of research to address the sampling bias. Joseph noted that the team wanted a more unbiased sample but found that nobody wants to disclose their failed research.

Hey adjourned ASCAC at 5:07 p.m.

Respectfully submitted,
Tiffani R. Conner, PhD, PMP
Oak Ridge Institute for Science and Education (ORISE)