ADVANCED SCIENTIFIC COMPUTING ADVISORY COMMITTEE
to the
U.S. DEPARTMENT OF ENERGY

MEETING MINUTES

January 13-14, 2020

Holiday Inn Capitol
550 C Street, Washington, DC

ADVANCED SCIENTIFIC COMPUTING ADVISORY COMMITTEE

The U.S. Department of Energy (DOE) Advanced Scientific Computing Advisory Committee (ASCAC) convened on Monday and Tuesday, January 13-14, 2020 at the Holiday Inn Capitol, 550 C Street, Washington, DC. The meeting was open to the public and conducted in accordance with the requirements of the Federal Advisory Committee Act. Information about ASCAC and this meeting can be found at http://science.osti.gov/ascr/ascac

ASCAC Members Present

Daniel Reed (Chairperson)
Keren Bergman
Martin Berzins
Barbara Chapman
Jacqueline Chen
John Dolbow
Jack Dongarra
Thom Dunning
Tim Germann
Susan Gregurick
Anthony Hey
Sandy Landsberg
Richard Lethin
David Levermore
Satoshi Matsouka
John Negele (online)
Vivek Sarkar (online)
Krysta Svore (online)

ASCAC Members Absent

Silvia Crivelli
Gwendolyn Huntoon

Also Participating

Steve Binkley, Deputy Director for Science Programs, Office of Science (SC), Department of Energy (DOE)
Christine Chalk, ASCAC Designated Federal Officer, Advanced Scientific Computing Research (ASCR), SC, DOE
Barbara Helland, Associate Director, ASCR, SC, DOE
Richard Carlson, DOE
Clair Cramer, DOE
Lois Curfman McInness, Argonne National Laboratory (ANL)
Lori Diachin, Deputy Director, Exascale Computing Project (ECP)
Ian Foster, ANL
Roscoe Giles, Boston University
Bruce Hendrickson, Lawrence Livermore National Laboratory (LLNL)
Travis Humble, Oak Ridge National Laboratory (ORNL)
Doug Kothe, Director, ECP, ORNL
Robinson Pino, DOE, ASCR
Fred Streitz, LLNL
Tom Uram, ANL

Attending

Frank Alexander, Brookhaven National Laboratory (BNL)
James Ang, Pacific Northwest National Laboratory (PNNL)
Louis Barbier, NASA
Jaydeep Bardhan, GlaxoSmithKline
Laura Biven, DOE
Arthur Bland, ORNL
Ben Brown, DOE
David Brown, LBNL
Suren Byna, Lawrence Berkeley National Laboratory (LBNL)
Jonathan Carter, LBNL
Leland Cogliani, Lewis-Burke
T. Reneau Conner, Oak Ridge Institute for Science and Energy (ORISE)
Tricia Crumley, DOE
Paul Doucette, Battelle

Robert Eades, IBM
Al Geist, ORNL
Raul Grout, National Renewable Energy Laboratory
Mark Guiton, Cray, Inc.
James Hack, ORNL
Aric Hagberg, Los Alamos National Laboratory (LANL)
Ronald Hawkins, EP Analytics
Carol Hawk, National Energy Research Scientific Computing Center (NERSC)
Jeff Hittinger, LLNL
Thuc Hoang, DOE/National Nuclear Science Administration (NNSA)
Paul Hovland, ANL
Michael Hundley, LANL
Fred Johnson, Retired, DOE
Ben Kallen, Lewis-Burke
Alex Larzelere, Retired
Steven Lee, ASCR
Arthur Maccabe, ORNL
Scott McKee, U.S. House Committee on Appropriations, Subcommittee on Energy & Water
Sandra McLean, DOE
Juan Mesa, National Science Foundation
Paul Messina, ANL, Retired
Kathryn Mohror, LLNL
Karl Mueller, PNNL
Esmond Ng, LBNL
Mike Parks, Sandia National Laboratory (SNL)
Grace Penguin, National Institute of Biomedical Imaging and Bioengineering
Mary Ann Picone, IBM
Timothy Proctor, SNL
Jini Ramprakash, ANL
James Ricci, DOE
Mark Rockwell, Federal Computer Week
Mike Rosenfield, IBM
Bettina Schuffert, German Research Foundation
Arjun Shankar, ORNL
Bill Spotz, ASCR
James Stewart, SNL
Ceren Suset, DOE
Valerie Taylor, ANL
Tamas Terlaky, Lehigh University
Hikmet Terzic, DOE
Angie Thevenot, DOE
Georgia Tourassi, ORNL
Louis Terminello, PNNL
Carolyn Vea Lauzon, DOE
Scott Weidman, DOE
Julia White, DOE
Justin Whitt, ORNL
David Womble, ORNL
Carol Woodward, LLNL
Xiu Yang, Lehigh University

Monday, January 13, 2020

OPENING REMARKS FROM THE COMMITTEE CHAIR, Dan Reed, ASCAC

Reed welcomed attendees to the meeting and introduced Dr. Binkley.

VIEW FROM WASHINGTON, Steve Binkley, Deputy Director of the Office of Science

Binkley discussed two topics, the FY20 budget and the Office of Science reorganization.

Larger budgets in the FY20 budget, approved in December 2019, have enabled the start of construction projects. SC has a good reputation on the Hill as evidenced by the larger budgets. In FY20, $71M was allocated across SC for Artificial Intelligence (AI) and Machine Learning (ML); $195M toward Quantum Information Sciences (QIS) activities ($75M for QIS research centers, $120M for basic research); and $242M for ITER. A QIS funding opportunity announcement (FOA) was released January 10, 2020. Two major construction projects are the Electron Ion Collider (EIC) and the Stable Isotope Production and Research Center (SIPRC).

Dr. Fall is making top-level changes to the Office of Science organization. The Principal Deputy Director is a newly created position, generated by the abolishment of the Deputy Director for Resource Management position. The new position is going through the DOE approval process. The new Deputy Director for Science Programs is Dr. Harriet Kung. The Deputy Director for Field Operations position is currently vacant.

Discussion

Reed asked for clarification on the effect of the reorganization on SC reporting lines. **Binkley** said Helland will report to Kung who reports to Binkley. Binkley plans to stay very engaged in the six programs.

VIEW FROM GERMANTOWN, Barbara Helland, Associate Director of the Office of Science for ASCR

Helland commended the ASCR History subcommittee, and acknowledged those who contributed content and those who "lived it."

In the FY20 appropriations DOE is encouraged to expand its collaborations with the National Institutes of Health (NIH). SC is instructed to work with the National Science Foundation (NSF) as much as possible on the QIS centers. ASCR will provide $10M for memory advancements regarding accelerated architectures in AI. Within the FY20 budget, the Leadership Computing Facilities (LCF) were allocated $150M for Argonne, $225 for Oak Ridge, and $110M for NERSC. Additionally, ASCR has $90M for ESnet, $39M for research and evaluation (R&E) prototypes, $10M for Computational Science Graduate Fellowship (CSGF), and $155M for Mathematical, Computational, and Computer Sciences research. The budget for Exascale has decreased in FY20 as expected.

In September 2019 the reorganization of ASCR began. As of November 2019 ASCR has three divisions, Facilities, Computational Science Research and Partnerships, and Advanced Computing Technology (ACT). The budget for ACT is primarily the R&E projects budget. Helland's visualizes ACT as an opportunity for co-design through partnerships (cross-agency, cross-program, and international). ASCR has been utilizing partnerships to work on core algorithm development for QIS. Applied Math (AM) and Computer Science (CS) are reaching out to other agencies and programs in SC on data management. ASCR is also asking what leadership computing will mean in the future. ACT will focus on basic research in novel architectures, microelectronics, cyber security, test beds and prototypes, ECP, and CSGF.

The QIS Centers FOA was released January 10, 2020. Awards will be $10M-$25M up to 5 years; full applications are due April 10, 2020. This FOA is unique in terms of SC breadth (all 6 program offices), scope (built on community input), integration, teaming, and leveraging other federal agencies' interests. Two projects for the AI/ML/data analytics co-design, lab-only call were awarded: Advanced AI Systems using multimodal data, and ARtificial Intelligence focused on Architectures and Algorithms (ARIAA). Three awards were given for the Scientific ML (SciML) to investigators at ORNL, LLNL, and University of Michigan.

Four AI Science Town Halls were held in 2019. There were over 1,000 registrants across all 17 DOE National Labs, representing 39 companies, 90+ universities, the 6 SC Offices, NNSA and EERE (Energy Efficiency and Renewable Energy). There is a new charge to ASCAC to examine program outputs to use as strategies, address AI challenges, and deliver on opportunities.

The ECP project review was held in December 2019; ECP has moved from R&D to delivery. Recommendations from the review included proceeding to CD2/3, customizing the Dashboard, and publishing a Community Outreach Document for each application. A meeting with Under Secretary for Science, Paul Dabbar, is scheduled for February 2020 to finalize the approval. The ESnet review yielded two recommendations: to examine the project documentation and eliminate inconsistencies, and proceed to CD2/3.

Advancing ASCR projects to CD2 required both Titan and Mira to be shut down. Only Summit, Cori, and Theta are available and this will affect the allocation programs.

The SC Distinguished Fellowship program began in FY19 and five awards were given in October 2019. This fellowship was authorized by the America COMPETES Act. ASCR's awardee in 2019 is Dr. Ian Foster.

Discussion

Dolbow inquired as to the vision of the ASCR reorganization and the decision to associate CSGF with ACT. **Helland** stated one strength of CSGF is its ability to anticipate the "next big thing". Locating CSGF in ACT will ensure that ASCR is aware of what is on the horizon and that a workforce is ready to support novel technologies and can easily move forward. The fellowship was recently changed slightly to add AM and CS.

Chapman expressed enthusiasm for the new effort to develop collaborations and partnerships across DOE and other agencies. She asked about plans to interact with other agencies, both national and international, in a similar way to the NIH Workshop. **Helland** explained that the current focus is on NIH, but ASCR has an existing partnership with the Department of Defense, and partnerships across agencies on the quantum initiative.

Gregurick asked if ASCR will be looking for science drivers through QIS, especially in quantum biology. **Helland** said all six SC programs are involved including biology for quantum sensors. The Request for Information (84 FR 22834) identified five technical scope areas in quantum (communications, sensors, computing, chemistry and materials science, and foundries).

Landsberg asked about workshops, solicitations, and collaborations with other agencies through the ACT Division. **Helland** said there was a Basic Research Needs (BRN) workshop for Microelectronics in 2018. BES (Basic Energy Sciences) has funding for microelectronics in FY20. The key point from the BRN for Microelectronics was a co-design effort. As new microelectronics are developed, the algorithms and applications that use them will be taken into account; ASCR also has to consider processes due to, and uses of, microelectronics. ASCR will continue to work with others DOE programs on cybersecurity; a workshop on cybersecurity was held in 2018; Carol Hawk, the new NERSC program manager, has experience in cybersecurity and Robinson Pino serves on the Cybersecurity Working Group.

Reed posed a philosophical question; how does ASCR balance a focus on state-of-art AI versus domain-driven (mission-driven) AI? **Helland** explained ASCR will focus on advancing the foundational role of AI; domain science should focus on the data and understanding data needs. BES does not own the data, the investigators do. The question becomes how ASCR tools can help BES collect data, make the data findable, and learn from HEP (High Energy Physics) – on their data solution. ASCR should do foundational research in AM and CS and build off of SciDAC (Scientific Discovery through Advanced Computing) work to assist domain scientists. **Reed** noted the BRN discussion brought out the interplay of new detector electronics and feeding sensor architecture, in materials science, into data analysis as well as AI issues. **Helland** said ASCR has to pay attention to the detectors, what is in the detector electronics, and what is available at the detectors. Ideally a subset of the ECP-developed software stack will work in the detectors, but ASCR must be able to scale. Not only do the facilities want a smarter detector, they want to ship data to an ASCR facility to ensure the detector is correctly set up.

Levermore added that the AI connection plays a growing role in developing equations-of-state and complex scientific and engineering computations. Progress has been made on those problems over the years, but it has been led by investigators outside of DOE. In the past, equations-of-state could only be calibrated with a certain number of experiments or theory. AI can be helpful developing equations-of-state when doing something exotic such as putting normally separate materials together. **Helland** noted two research thrusts are foundational research on AI and using AI in algorithms. **Levermore** expanded on his point stating AI in its native form is relatively amorphic. In the physical setting, structure of the physical symmetries exist; building those symmetries helps AI do its job. **Helland** said the SciML workshop identified that as a key priority research area.

Chen inquired about funding opportunities for researchers, across the SC programs, to interact on designs of new architectures, new systems, and new algorithms. **Helland** assumed the PathForward types of investments with vendors would focus on these areas. Like ECP, the operations of the whole organization will need to be outlined.

Matsouka asked about the mechanism(s) being instituted to assess future needs in AI. **Helland** explained the subcommittee on AI for Science is the first mechanism. It will be followed by the facilities' regular requirements reviews with the SC program offices. Other areas, identified in the FY21 budget, will be shared later.

REPORT FROM SUBCOMMITTEE ON EXASCALE TRANSITION, Roscoe Giles, Boston University

Giles shared the subcommittee's charge, activities, and findings and recommendations. Activities have included stakeholder interviews, community meetings, reports reviews, and subcommittee meetings. There are four categories of findings and recommendations.

The findings in Category A. Advancing and Building on ECP, focus on ECP's success, distributed teamwork, software ecosystem, and collaboration with facilities and industry. Recommendations are shared-software stewardship, engagement in future software needs, collaborative applications support, industry and academic engagement, and modern project management tools.

Category B. Advancing ASCR Research, findings include AM and CS necessity, and base research program constraints. The recommendations in this category focus on reinvestment in ASCR research, stable environment for basic research, and software distribution.

The findings for Workforce (Category C) focus on a skilled, diverse, and motivated workforce, career paths, and diversity, equity, and inclusion (DEI). Category C. recommendations include supporting "blue sky research", workforce retention, ties to universities, career paths for software professionals, and support for DEI.

The two findings in Category D. National and International Leadership, are US leadership of the exascale computing effort, and impact on DOE's mission. Recommendation topics are maintaining leadership, and engagement and collaboration with stakeholders across other agencies.

Discussion

Dongarra asked if the report will include any guidance on funding levels or allocations. **Giles** indicated the subcommittee is currently focusing on ascertaining the stakeholder-identified activities and products that should continue.

Dunning mentioned transition plans in other DOE projects, indicating transition to operations is standard operating procedure; DOE has a process and it should be no different for software than for hardware. **Giles** explained ECP has such a plan that was released in November 2019. The transition plan and this team's report are consistent. ASCAC can call for a new program within ASCR, which ECP cannot do. ECP's transition to operations plan is directed at DOE, while this report is for the Office of Science. The difference between the two plans is the idea that ASCR should take some responsibility for the software.

Berzins referred to recommendation A.4 about broadening academic engagement. He stated it fits well with the new ACT Division in ASCR. The ECP applications will be limited by data transmission speeds. The challenge is making the wide application space perform. There is huge potential for enabling faster learning and for co-design. The opportunity, when processes stall, to bring the broad applications forward and enable faster learning and faster science should be utilized. **Giles** said that part of Berzins comments are consistent with the research recommendations to reinvigorate the base, encourage partnerships in co-design, and interact with the broader computing community; it connects to applications and facilities – what is available in the application space and co-design. **Bergman** recommended building on ECP's current engagements with those vendors, who were part of the advanced R&D programs, to do more than simply deliver the machines. An advanced, technical community has been built with those who were involved. **Giles** was unsure the report specifically calls out the PathForward program participants, but the Workforce section includes comments about building on the expertise developed and the impact on staff.

Reed said resources will be challenging given the architectural diversity and uncertainty. Remaining relevant in the software base and conducting basic R&D is critical. **Giles** welcomed direction from ASCAC about how much to comment on the level of resources. Resources are a fast-moving target and outside the subcommittee's area(s) of expertise. Giles was reluctant to indicate the fraction of ASCR's budget that should be allocated to software stewardship versus research. **Reed** agreed that the subcommittee should be budget realistic within the range of what might happen. It is best to avoid suggesting specific directions that are impossible given the expected budget. **Giles** noted that the recommendations are targeted to ASCR, not to other agencies or even all of SC.

Hey asked about estimates on the resources required to maintain open source software. **Giles** explained there has been no direct estimate. However, supporting open source software will require coordination with entities beyond the facilities regarding porting ECP-developed applications.

Berzins explained elements of workforce competition and noted that universities experience the same challenges of lower salaries. Young people in the labs are passionate about what they do, but also recognize they are at a considerable financial disadvantage. The leading edge of ECP will only persist if the workforce stays in place or if more people are recruited who can do that work, there's a danger of losing those people. It would be foolish to underestimate international competition. The present leadership must be maintained, nurtured, and pushed forward aggressively. He championed a balance between uncovering and exploiting the next exciting areas, such as AI, and base leadership that presently exists, saying neither one of these can afford to be lost. **Giles** agreed and added that individual work choices are based on more than just monetary factors.

Larzelere asked about the future of high performance computing (HPC)-enabled modeling and simulation. **Giles** indicated the current changes will alter our sense of what it means to be a computer. The computer will be reimagined and the value of the computation will have to be considered differently. **Levermore** agreed and added what is important to Congress and the public is the scientific impact the new machines provide. The scientific mission of the agency must be addressed, but it may be tackled differently in the future. The community must consider how the impacts of future computers and scientific computing are presented. **Berzins** suggested the issue is time to solution. The faster a machine can compute the faster learning occurs – that has to be the metric.

Matsouka was concerned that as ECP transitions to operations there will be a loss of workforce. He asked about the next challenge that will engage and retain the ECP people. **Giles** said there are prototype applications created by ECP that need to have impact across the board. With the exascale technology, heterogeneous computing has to be applicable at many scales and for different problems. Certain challenges will still exist while others will be based on new initiatives such as incorporating AI into existing codes. **Matsouka** emphasized the investment in computing has had profound impacts in the broader IT (information technology) space. Without the early investments of *x*Forward to rejuvenate software initiatives ECP applications would not be where they are now. It is important that DOE has had these broader impacts in IT. He inquired why DOE investments are not being recognized. **Giles** explained the subcommittee is considering an additional observation on the impact of ECP on the broader society.

Bergman said the current workforce issue is more than reshuffling funding and creating exciting activities and topics; it requires DOE and ASCR to think more creatively about keeping and attracting talent. Such approaches could include the next great idea or challenge, the uniqueness of DOE and basic research, the long-term vision and mission, the participation in co-design, etc. These are things companies cannot offer. The importance of the workforce cannot be underestimated. **Giles** added additional considerations such as the ability to make significant contributions to leadership and their teams' efforts, ways investigators are treated in multi-disciplinary teams, and who gets to do interesting work. **Levermore** stated the subcommittee has met with a lot of young investigators and what came across was most are driven by their interest in a scientific mission; they want meaningful work that contributes to society as a whole. Exascale is the tool to achieve that goal. Workforce is a major bullet in the report because we have to refocus on the national mission. What makes people want to come to work is the impact they will have on a scientific goal.

Chen noted that ECP developed state-of-the-art applications and system software, many of which address science and engineering missions. AI/ML integration into computational and experimental workflows is becoming clearer. ECP has advantaged the community due to co-design efforts to make algorithms for solvers and AI/ML to work on complicated heterogeneous machines. This group of individuals can take the next leap and they are excited to move forward.

Negele asked about the number of personnel supported by ECP overall. **Suset** indicated the number varies depending on the partner or office involved.

QUANTUM USER PROGRAM AT THE OLCF, Travis Humble, Oak Ridge National Laboratory

OLCF established the Quantum Computing User Program (QCUP). Access to the quantum resources are based on a merit review of proposals and user agreements. The program is managed by OLCF and is supported by ASCR.

User access to the program begins with a project request by the principal investigator (PI), OLCF reviews the project and then notifies the PI of the decision. The PI is evaluated as a potential system user and they authorize additional user accounts, OLCF vets all users and notifies them of their account creation.

To date, the QCUP has a diverse user base (130+ users) and research portfolio. The current projects largely represent ASCR, BES, and HEP. Use case categories include physical sciences, data sciences, and applied sciences. The priorities of QCUP are to enable research, evaluate technology, and engage the community.

Discussion

Landsberg asked about the model to buy into the vendor resources and how the QCUP might support QIS in the future. **Humble** explained that users' queue time is becoming an issue. Vendors are responding by providing more systems. For example, IBM initially offered a single system, but due to the need, IBM now has 14 devices. There are term-limited contracts in place that are paid on a monthly schedule. Using vendor-owned systems has been the most difficult part of QCUP. Because OLCF is reselling access to the commercial system, the user agreements get more complicated; vendors are involved in the user agreement process.

Gregurick inquired about the largest molecule being simulated; is it possible to go larger than sodium hydride. **Humble** said the largest molecule was measured by the number of electrons (14 electrons requires a minimum of 28 qubits). Going larger pushes the performance limit of the devices and the noise floor takes over. There is a lot of theoretical analysis that goes much further to optimize programs for larger molecules. **Gregurick** asked if people are using dynamic field theory methods. **Humble** explained there is a lot of work in hybrid modelling. The Dynamic Mean-field Theory and other similar approaches are considering how to decompose a quantum model into two parts: classical and quantum.

Dunning asked if OLCF is developing quantum expertise at the same time as user access. **Humble** said QCUP has been stood up using the existing user assistance infrastructure. Vendors fill the gap of technical assistance; they also give seminars and workshops. Over the long run expert staff will be needed in-house. OLCF is presently mitigating the gap by pairing users with individual researchers. User assistance staff requires a certain mentality, a certain patience, and a willingness to help.

Berzins requested a high-level explanation for the jump shown between the raw data and filtered results in the presentation. **Humble** explained that the raw data is largely contaminated by noise in the devices, meaning when a quantum program is run filtering out the noise improves the signal, in this case the purified results. Quantum computing (QC) requires the use of algorithmic tricks to mitigate against the noise (error mitigation) and that is occurring. The differences between raw and purified means the purified results take advantage of those algorithmic tricks. **Berzins** exclaimed that the noise does not look like noise in the conventional sense. **Humble** added that the algorithms are finding the optimal energy value for the configuration of the particular problem, it is optimization. There is a trend it follows; it more or less looks like the real solution. However, the magnitude of that offset represents the noise. Our ability to control the device is why it cannot get to the actual solution.

Reed inquired if anyone is considering how to marry conventional and quantum in an outer-loop model to take advantage of different kinds of functionality for complex, multidisciplinary problems. **Humble** said the integration of quantum computers with traditional computers is exactly what is on their minds at OLCF. The idea that a quantum CPU (central processing unit) could be an accelerator for HPC is an important model, but understanding the algorithmic workflow precedes the technical design. Users are reporting the types of problems they can run and the types of algorithms they can use with the current devices. At the moment cloud seems sufficient, but with the next generation of fault-tolerant, error-corrected computers, a very tight infrastructure integration will be needed – in terms of proximity and memory systems for communication.

Reed dismissed ASCAC for lunch at 12:20 p.m., reconvening at 1:30 p.m.

40TH ANNIVERSARY ACCOMPLISHMENTS SUBCOMMITTEE REPORT, Bruce Hendrickson, Lawrence Livermore National Laboratory

Hendrickson reminded ASCAC of the charge and the decision, by the subcommittee, to create two documents, a detailed history and an accessible document. Activities since the last progress update have been to confirm von Neumann's role in setting up the math program and add John Pasta's call to begin the math program. The introduction and computational science sections have been reworked to emphasize software and the integrated impact. The architecture section was reorganized, adding a new story on Josh Fisher (representing ASCR's role in the invention of instruction level parallelism) and parallelism at the microelectronics level. Also included was a discussion of spectral deferred correction methods to the applied math section, and a sixth challenge. Finally the document was polished for consistency and clarity.

High-level lessons learned are ASCR's continued compelling and consistent vision, diverse funding models, workforce investments, partnerships, and testbeds and platform access. The challenges in the coming years will be technology distractions, funding balance, software support model, broader partnerships, workforce, and new roles, new opportunities, and new demands.

Over the next two months the subcommittee will gather and respond to additional input, locate and integrate additional imagery, and focus on design and layout.

Discussion

Levermore inquired if some citations will accompany the document. **Hendrickson** explained the editorial decision was to avoid citations and keep the document at a higher level. **Levermore** asked if there was a middle ground. **Hendrickson** articulated his concerns explaining a single paragraph in the document can include decades of work by multiple research teams. The breadth of coverage means that a comprehensive bibliography would be enormous. **Levermore** acknowledged this concern, and suggested having citations for material that might be difficult to find; adding citations would strengthen the activity's goal of laying out a history.

Lethin recollected that an archive was mentioned in the September ASCAC presentation. **Hendrickson** said while the subcommittee does not have the authority to demand an archive, a suggestion was made for ASCR to index and archive the materials. Paul Messina began gathering anecdotes; there is a vast amount of raw material that should be saved.

Dongarra asked if Levermore would like to see compendium of DOE document citations, observing that an annotated citation list could be as long as the report itself.

Levermore stated there should be some citations, especially those references critical to the decisions about ASCR. **Reed** noted it is a balancing act, but a number of artifacts are very hard to find; the report would benefit from citations.

Dunning offered a middle ground idea – an appendix with a list of some of the more important documents that tie into the text. **Hendrickson** said there are many scientific stories already covered in journals. He speculated that perhaps those more at risk of disappearing could be pulled out. He said he will let the team determine what can be done.

Landsberg said younger researchers may want to have the background material as an entry point to the primary literature. She also inquired about the comprehensiveness of the math section, stating it seems to end around 2000. **Hendrickson** pointed out the math section was written by three people. Phil Colella wrote an academic history of numerical partial differential equations; Jack Dongarra contributed content on numerical software; and Hendrickson added material on iterative methods, uncertainty quantification, and optimization. The latter two authors both included significant post-2000 material. **Dongarra** asked Landsberg to send additional material to the subcommittee to amend the section.

Dunning suggested adding a statement to the introduction expressing that the significance of advances do not become apparent until many years later – particularly in mathematics.

Lethin particularly liked the sections on commercialization, industry, and small business. However, he requested a highlight on the impact to climate science. **Hendrickson** explained the report mentioned BER (Biological and Environmental Research) and ASCR partnerships on climate modeling, numerical methods essential to model climate, and ASCR's support of the Earth Systems Grid Federation activity. ASCR has played an important role in climate science, however, this was not pulled together in a single story; a sidebar can be added to reference these contributions collectively. **Dunning** mentioned an additional figure will be added to the Computational Science section on BER's discussion of the impact ASCR has had on their climate studies.

Reed asked if ASCAC was comfortable voting on the report or if they wanted to see another version before voting. ASCAC members conducted a voice vote; all approved the report.

Berzins exclaimed there must be a way of adding the references that is workable; he referred the subcommittee to the book "History of Scientific Computing" as an example. **Chen** asked if ASCR can support an archive for the material. **Chalk** will contact OSTI (Office of Scientific and Technical Information) about the archive.

DISTINGUISHED SCIENTIST, Ian Foster, Argonne National Laboratory

Foster shared his research areas in terms of opportunities for scientific mechanisms. Since the 1980s Foster has conducted research in fast inference, parallel computers, networked systems, science services, and AI-first. Examples of the computing continuum include automated reasoning, wide area networks at large scales, computing on demand, automation PETREL (permissions management), and commercial cloud services.

There are large gaps to fill over the next decade, including revolutionary challenges to the computing platform. The question is what will people be computing on? The platform will not be a workstation, grid, cloud, or exascale/post-exascale computer, rather a combination of all of these. Devices will be able to connect in new ways; this should encourage thinking of computation in new ways that take device location and speed into account. Foster's concern is identifying a mechanism that will bridge the gap. Project Celerity, Foster's proposed research for the Distinguished Scientist Fellowship, will be performed at ANL to identify new mechanisms needed to bridge the gap between the emerging applications and the emerging data/computing continuum via a process of experimentation, discussion, and debate. New mechanisms may be fast inference, task parallelism, workflows, remote computing mechanisms, and cloud services. Mechanisms to break down barriers that prevent computing on multiple architectures and accelerators and platforms at the same time include a function fabric that will allow functions to flow to wherever computing is fastest, cheapest, etc.; a data fabric to allow access to data and compute on that data, manage it, and move it; a trust fabric that will allow investigators to determine who is allowed to compute, balancing certainty versus the cost of determining trust; and mechanisms for creating estimates on the state of the dynamic system. Elements of an open solution to coding the continuum include Parsl (writing programs), DLHub (model registry), funcX (function fabric), Globus data services and Globus Auth (data and trust fabrics), etc.

The mechanisms are designed to create a more agile infrastructure at the core of the computing platforms so that all elements will flow fluidly to the right place for the right activities. Activities will involve the use of AI methods for model training, and model and surrogate model creation. There is an intersection program to be faced by everyone – thinking about ways to re-architect and rethink efforts ASCR has been working hard on in a way that will enable the best use of a new, more fluid infrastructure and will support new applications.

New mechanisms will empower discovery. Over the past three decades progress has been made on identifying mechanisms and new ways of deploying them – starting with libraries and moving to cloud services. The interesting challenge is thinking what new mechanisms will be needed to enable effective use of new computing platforms for new applications. Mechanisms like Fortran, MPI, and Globus are effective when they are trivial to use and when they can be relied on to persist into the future. Looking forward the community needs to think about ensuring mechanisms are created, experiments are conducted, and work is performed to ensure these mechanisms are available.

Discussion

Reed asked for the single biggest technical challenge for the computing continuum. **Foster** said breaking down the barriers that prevent linking parts of the continuum. New mechanisms are needed to build applications that will automatically take data from one location, train a model in a separate location, and deploy it in yet a third location.

Hey stated that it seems Foster is defining the super facility of the future – one that is AI-enabled. **Foster** explained that at the moment only 1% of computing happens on supercomputers – the rest is on workstations. New methods and new data volumes means everyone will be doing HPC of a sort. A new computing infrastructure or computing continuum needs to be created to allow everyone to make use of the best computing capabilities to meet their needs. Foster speculated the new system will be a federated and heterogeneous model.

Lethin inquired how Foster envisions optimization to take place, over what representation, and to what objectives. **Foster** said he is using simple, active learning techniques to proactively gather information on the performance of applications on different platforms. He is refining those models over time using Bayesian inference methods to choose the next best experiment to perform. That works well with no analytical representation of the performance of the applications, but there are other methods to use in other contexts.

UPDATE FROM SUBCOMMITTEE ON ARTIFICIAL INTELLIGENCE, Tony Hey, ASCAC

Hey explained the new AI for Science charge to ASCAC. The charge directs the subcommittee to assess opportunities and challenges of AI/ML and identify strategies to address these challenges and opportunities. The proposed committee is made up of representative of: the other SC programs' Federal Advisory Committees; the new AITO Office; academics in HPC, AI, and Big Data; and a National Science Board delegate. The subcommittee will also gather inputs about AI applications and roadmaps from: academics; members of the Pharmaceutical, industrial, and IT industries and other interested stakeholders. A draft report will be delivered to ASCAC by May 2020 for preliminary comments with a final report in August 2020.

Discussion

Lethin expressed concern over the quantity of material from the Town Halls. **Hey** indicated he would ask the Town Hall conveners to sort the material initially.

DOE INTEREST FOR SUPPORTING 5G ENABLED ENERGY INNOVATION, Robinson Pino, ASCR

Moving into the future, data has incredible demands, networking capabilities will continue to increase, and 5G offers a lot of bandwidth with 1 millisecond of latency. 5G's impact on DOE's mission is being explored by different program offices. BES is leading the investigation of microelectronics; ASCR is seeing the rise of heterogeneous computers, AI, quantum computing, and new forms of computation. The question to be answered is, in combination with the existing and coming infrastructure, would 5G offer a tipping point for exploring advanced wireless networks?

A workshop on 5G Enabled Energy Innovation will be held in Chicago, March 10-12, 2020. The workshop will discuss R&D and innovation opportunities enabled by 5G and similar technologies. It will deliver a community-based report to help SC understand the challenges and opportunities in this arena. There are 10 technical focus areas that will form the basis for breakout group discussion. Whitepapers must be submitted by January 31 with registration due in early March 2020.

Discussion

Reed inquired about interagency collaboration in advanced wireless networks. **Pino** said there are potential interagency collaborations. To date there is no area in which even current wireless networks, like 4G, are being used within scientific experimentations. The goal of the workshop is to gauge the potential areas where advanced wireless networking could be used.

Gregurick asked if the workshop will explore risks in manufacturing, especially those countries who are ahead of the US in 5G? **Pino** said the main focus is on science drivers and determining if it makes sense for science to dig deeper into some of those questions. There is some overlap with microelectronics, some of the technologies are still in development and that could offer synergies. Having a better understanding, prior to exploring leveraging advanced wireless network capabilities, is necessary.

ARTIFICIAL INTELLIGENCE TECHNOLOGY OFFICE (AITO), Fred Streitz, detailed to AITO from Lawrence Livermore National Laboratory

Streitz described several drivers and developments in AI including defining how DOE thinks of AI. AI is disruptive and multi-disciplinary, AI trains computers to perform tasks that mimic human intelligence, AI continuously learns from non-ordered data, and AI drives economic growth and productivity. AI will add 14% to the US economy and 26% in China. The Chinese government is outspending the US in AI by ~10:1. AI is enabled by HPC. AI has a different technology path and purpose, and AI technologies must be built to combine AI and HPC in the future. DOE will focus on the application and advancement of AI broadly as AI that is strongly rooted in the private sector will not address DOE issues. The point of the AITO is to make coordination happen faster.

The plan of action is to create a DOE AI strategy, to institutionalize an AI Exchange, to develop and implement AI Leadership Training, to conduct workshops, and to prioritize and foster AI partnerships. There are ~600 projects across DOE doing AI; the future requires advances on hardware fronts and AITO is encouraging these explorations. DOE is in a unique position to make an impact on the AI landscape. Using HPC and AI together will lead to remarkable innovations. Deep partnerships are necessary and AITO is organizing across DOE to establish and maintain US leadership in AI.

Streitz shifted focus and provided an update on the Cancer Moonshot pilot projects, concentrating on Pilot 2 which he leads for DOE and is looking at cancer at the molecular level and connecting HPC to AI. Pilot 2 is exploring dimerization of the RAS protein – is dimerization a must or does is occur accidentally. If RAS has to dimerize to signal continuous cell growth there may be a way to disrupt the mechanism. Answering this question is a problem of length and timescale. To understand what is happening at a membrane level requires microns of material for microseconds of time. To understand what is happening at the protein interface requires microseconds of time across nm of information. Pilot 2 will build a macroscopic model to evolve the RAS protein on a membrane at the macroscopic scale (not invented yet), then build a protein model with molecular dynamics at near atomistic resolution to investigate the protein behavior (this is known). Pilot 2 will do this so all proteins talk to each other (unknown). Using the entire HPC simulation "important elements" will define what to study. To run the simulations Pilot 2 cut Sierra up, in ways IBM did not envision, and developed the Flux Scheduler (available open source) to handle macroscopic simulations. In total, Pilot 2 ran a 1 μm^2 simulation on the macro model for 152 µs with 300 KRAS, investigated 2.2 million patches, chose 116,000 simulations (each of 140,000 particles), 200 µs aggregated (which used all of Sierra, upwards of 90% efficiency on the machine), and ~ 10 trillion molecular dynamic steps were taken. Combining AI and HPC is how science is going to be done in the future, and teams like this, across agencies and domains, will become the norm. This work won "best paper" at SC19.

Discussion

Lethin inquired about the starting point for the US and China if China's investment is 10:1. **Streitz** explained 80% of all industries in China are actively deploying existing AI technologies, whereas the US is averaging 50%. China is not necessarily ahead of the US in terms of developing new technologies, but they are ahead of the US in terms of applying the technologies. **Lethin** asked for AITO's perspective on exports of AI technology. **Streitz** said much of the technology ends up being open source. The question is what is of value. Is it the construction of a deep neural net or the weights inside the deep neural net that is of value; one is easy to protect while the other is hard to protect. No one knows the answer as to what is being protected; people are just now understanding there is even a question. AI is an extraordinarily dual-use technology, it is extremely powerful in many different venues.

Berzins requested AITO's remit; is it to look at AI inside DOE, to look at what is going on in the whole of government, or to help the whole of government by looking inside DOE? **Streitz** explained AITO will look inside DOE while partnering with peer organizations across government. **Berzins** stated what is absolutely unique for DOE is the data, the HPC expertise, the workforce, and the ability to develop software but not necessarily to use it. He proposed thinking about what is unique as a way round the spending imbalances, and exploiting that uniqueness to quickly move ahead on DOE-related problems. The challenge is to figure out what DOE can do better than anyone else. **Streitz** remarked there are limited resources to apply to the maximum benefit; partnerships are important. The US has leverage, DOE has a unique spot, and industry is starting to realize that. **Landsberg** expressed excitement on the parallels with the Department of Defense; the Army AI Task Force is addressing the same problems and needs partnerships.

Chen asked if AI/ML was used for mapping and scheduling of tightly coupled codes, the molecular dynamics and macroscopic codes. **Streitz** explained the molecular dynamics had to run entirely on GPUs (graphics processing unit) making that decision easy. The rest was hand tuned.

QUANTUM COMPUTING RESEARCH TESTBEDS, Claire Cramer, ASCR

The ASCR Quantum Testbed Stakeholder Workshop was held in 2017. SC needs information to make informed QC facilities investments, including understanding the DOE-relevant application space, relationship between quantum processor characteristics and application performance, and characterizing protocols, metrics, and benchmarks.

Quantum Testbeds Pathfinder provides funding on projects that will provide decision support for future investments. In FY17, two awards were made to teams led by LBNL and ORNL. In FY18, funding was awarded to three teams led by Virginia Tech, University of Maryland, and SNL. Current QCs are too noisy for useful computations. Understanding how and why devices fail is necessary for progress to be made; this understanding will show a path forward. Quantum Testbeds for Science is a funding source to advance comprehension of how to use QC resources. In FY17 two Quantum Testbeds for Science awards, totaling $56.3M over five years, were given to teams led by LBNL and SNL.

LBNL's Advanced Quantum Testbeds' (AQT) mission is to integrate the current superconducting quantum processing units to help develop extensible quantum systems. A stakeholder meeting will be held April 29-May 1 at LBNL. The meeting will include pre-meeting workshops, speakers and panel discussions on QC and superconducting circuits, and in-depth discussions of the AQT program.

Discussion

Hey asked if the teams are building their own quantum simulator to run the programs before it is put on the hardware. **Cramer** explained the teams are focused on putting together a flexible, extensible, and capable hardware platform rather than simulation of the system. **Hey** clarified that he meant the programming environment. **Cramer** said it depends on the meaning of a programming environment. Software stacks that focus on the lower levels are being developed. For example, the SNL team is building an assembly language and will build interpreters for all of the publicly available programming languages (i.e., Qsharp). Given the resource limitations the teams decided it was more important to have high-quality hardware capabilities with staff that know how to run the hardware; the user experiments will be more like a scientific collaboration.

PUBLIC COMMENT AFTER WHICH ASCAC WILL ADJOURN FOR THE DAY

Robert Eads suggested the AI for Science subcommittee have an industry connection to energy. **Hey** said the list of names was an initial suggestion. He agreed to consider the proposal.

Tuesday, January 14, 2020

EXASCALE UPDATE, Doug Kothe, Oak Ridge National Laboratory; Lori Diachin, Lawrence Livermore National Laboratory; Lois Curfman McInnes, Argonne National Laboratory

Kothe covered the ECP activities since September 2019 in applications development (AD), software technology (ST), hardware and integration (HI), review for approval of CD2/3, and outreach events. Activities through FY23 were reviewed and included access to the exascale machines beginning in late FY20. ECP is proactively accounting for critical external dependencies in facilities, with HPC vendors, and open source software. Recommendations and key comments from the CD2/3 review included the Dashboard and Community Outreach Document (recommendations) and key comments on AD, ST, HI, Project Office, and Project Management.

Diachin discussed the AD-ST Dependency Database. ECP is managing dependencies within the project and with the facilities. To manage the complexities and critical dependencies, ECP created a database using Jira. Jira tracks dependencies and allows for real-time data analysis showing the level of dependency, consumer and producer information, and cross-products. The Jira dashboard allows high-level and detailed information to help manage critical dependencies. Example statistics from an analysis of the database indicated that MPI and OpenMP dominate the programming models, CUDA codes will need to migrate, there are large dependencies on C++, and ¼ of the application codes depend on Fortran.

Curfman McInnes shared information on the Math Libraries. The math portfolio goals are advanced algorithms, performance, and improving library sustainability and complementarity. There are six math projects – xSDK, PETSc/TAO, STRUMPACK/SuperLU/ FFTX, SUNDIALS/hyper, CLOVER, and ALExa/ For Trillions. The Math Libraries approach is to establish performance baselines, refactor and revise algorithms and data structures for new architectures, and research new algorithms for next-generation predictive science. The SDKs (software development kits) concept is the key vehicle for ECP. The math libraries teams are delivering high-quality mathematical libraries that provide scalable, robust algorithms, interoperate with the ECP software stack, and can readily be used in combination with ECP applications.

Discussion

Lethin noted there is an obvious relationship between the dependence database and the compiler dependence analysis. The database implies formalizing those relationships in the software elements; tools exist to formally describe dependencies. He asked if there has been any exploration of using an automated tool to configure systems, etc. **Diachin** said there has been no exploration in the compiler space and the tools used by compilers, rather the dependence database is on the project management side. Project management tools can formally schedule activities and trigger dates; the analysis began here. The goal is to understand when the STs need to be ready and if there is a timing factor that needs to be tracked (i.e., dictated by the delivery of the machines or when the applications need them). ECP anticipates using the SPACK tool moving forward. SPACK was developed as part of ECP and is used for software integration and dependencies in terms of version control and optimization.

Dongarra asked about implications if there are no dependencies. For example, a software component, funded by ECP, without any binding to an application. **Diachin** noted that Michael Heroux has de-scoped projects, whose target was applications, that were not being impactful. STs are fulfilling many dependencies with the vendor software stack or with the facility software stack; those are not captured in this database. **Curfman McInnes** added the numerical software community wants long-term stability for math libraries as well. The math libraries are striving to create innovative algorithms but are driven and motivated by partnerships with applications. It is a constant struggle to assure applications of our long-term stability. Anything the DOE could do to help us as a community would be very much valued.

Dunning explained that for an application developer, software technologies and math libraries always require a risk/ reward analysis. The risk is a timely delivery of the functionality needed in the application and long-term support after the project ends. He encouraged ECP to be proactive and state, to the application developers, the long-term viability of the software technologies and the math libraries. Addressing this problem will increase the uptake of these new technologies and libraries. **Kothe** agreed and gave assurances that ECP plans to be more aggressive to ensure E4S is sustained. The dependencies database is a first step to get a handle on the critical products. The Key Performance Parameters-3 metric incentivizes integration and dependencies on the software side. ECP will work closely with ASCR in the post-CD-2 period to ensure the ecosystem is sustained. ECP is committed to regular, formal releases, containerizing, and making uptake easy for both facilities and applications. ECP wants to work with the community to collectively understand the resources and the technologies that need to be sustained.

Levermore asked for an example of an application that utilized the SuperLU speed-up. He requested ECP share a list of those teams using SuperLU speed-ups with the Exascale Transition committee. **Curfmann McInnes** said SuperLU is a great exemplar of having strong partnerships with applications as well as strong partnerships among the community. Some of the high level libraries, such as PETSc, Trillinos, and HYPER interface with SuperLU – this is another example of how the interfaces are functioning in a practical way.

Berzins asked about the characterization of the applications' compute intensity. **Kothe** explained the applications teams wanted to investigate GPU utilization. One team will work with performance tools, in this case NVIDIA, to better understand what is occurring in the GPUs. This year's review focused on utilization of the GPU's in general; next is quantitative performance detail.

FEDERATED IDENTITY MANAGEMENT, Rich Carlson, ASCR

ASCR has a long history of activities necessary to build a federated environment. In the past, DOE has accomplished this by working with individual science domains. The threat, or challenge, now is moving beyond supporting a single science community to supporting a large community of scientists who can use all DOE facilities effectively. The Future Lab Computing

Working Group was created to evaluate what was happening inside the laboratory computing facilities (not the leadership class machines) across the SC complex – buying resources (compute and storage), using resources, and moving forward in the future. Verifying identity was a major issues in making this access effective. From this information a pilot project was created to investigate the use of a federated identity service to access laboratory resources.

In 2019, the Distributed Computing and Data Ecosystem (DCDE) pilot project was established with the goal of having engineering staff at the labs work together to deploy existing technologies and services. The DCDE was not a research project but an engineering challenge built and run by those who run the machines. There are two identity management fundamentals, Authentication and Authorization; identity management separates these two fundamentals. The information required to identify someone is similar across the labs surveyed. The services used included AuthN/AuthZ, InCommon, CILogon +COManage, Globus and auth-ssh, applications and containers, Jupyter notebook, and Parsl workflow. The pilot study highlighted that – (1) there is a learning curve for users (resolved with a template solution), and (2) site administration overhead will require scripted install steps. Federated identity management is a key enabling service to foster science discovery. The DCDE pilot project demonstrated that identity management services are ready for full-scale deployment and, despite issues in policy and trust, there are significant benefits.

Discussion

Gregurick noted that NIH is engaged in something similar. She asked about future uses of identities from commercial platforms, such as Google, to allow people to log in with their Gmail accounts, for example. **Carlson** said identity management is currently a lab-only activity but it will expand with more experience.

Hey was exasperated that this issue is still not resolved. **Carlson** stated although the activity has taken a long time, progress is being made. This activity is deploying the technologies developed over the last 20 years in a production environment. Policy issues are the sticking point.

Helland asked how this project is utilizing DOE's OneID data. **Carlson** explained since OneID and InCommon are compatible, OneID is acting as an identity provider. **Helland** suggested adding a statement in Carlson's presentation that OneID is being utilized.

Berzins asked to what extent this project creates a single point of failure, from a security point-of-view. **Carlson** explained users only have access to tools and systems they have permission to use.

ENGAGEMENT WITH EXPERIMENTAL FACILITIES, Tom Uram, Argonne Leadership Computing Facility

Uram discussed developments at ALCF (Argonne LCF) in programming a national computing infrastructure to address the computing needs of experimental and observational facilities. ALCF supports experimental scientific computing through allocation programs (Director's Discretionary fund, Early Science program, and ALCF Data Science program) and technologies (Balsam workflows, Cobalt scheduler, and Globus Transfer).

Over the course of near real-time processing of light-source workloads with Balsam workflows, the experiment was able to continuously transfer over 23TB of input data from the Advanced Photon Source/ Advanced Light Source, analyzed over 500 datasets, and transferred 179GB of output data from ALCF to the light sources. ALCF won the SC19 Technology

Challenge Demonstration where they simulated ten beamlines operating, pushing the data over the network, processing it at ANL, and pushing the data back to the SC show floor.

The next steps with light sources include deploying XPCS (X-ray Photon Correlation Spectroscopy) production, engaging with more applications and beamlines, maintaining reliability, and meeting computational demands. ASCR support is critical to this venture and recent efforts have shown the national computing infrastructure can be leveraged to meet the needs of the light sources.

Discussion

Hey inquired if the experimental jobs are not parallelized. **Uram** said XPCS is not parallelized, however, parallelization opportunities have been identified in the code. For example, a lot of analyses at the light sources operate on a single frame at a time, parallelization could be across frames. Given a serial code, if the code is run at the ALCF and parallelization opportunities are identified, ALCF could fix the code for the user, could work with them to have them fix it, or could have their computing people think in a parallel sense, modernize their codes, and reflect differently about getting their work done. In the case of XPCS, significant speed-ups are apparent as the code is modified. **Hey** asked if a leadership scale facility is needed since the parallelism will be much smaller. **Uram** shared that the LCLS (Linac Coherent Light Source) expects to have experiments for which they will require 1.5 exaflops of online computing. With similar parallelization approaches, speed-ups may develop, thus requiring fewer resources. However, a significant amount of computing power will still be required.

Hey requested the percentage, amount of time, and number of people it took to human annotate images of mouse brain synapses for training the neural network in the Connectomics experiment. **Uram** admitted annotation is the most constraining bottleneck. Scientists on the Connectomics project spent nine months annotating data by hand. Some things being explored now are utilizing the hand annotated brain tissue of a fly and determining if it is representative, if it be used on brain tissue of other flies, if there are characteristics of electromagnetic imaging and tissue structure that are similar to an octopus, etc. If so, transfer learning may be a possibility. That might be followed by light-weight training on the octopus image to achieve a model with good accuracy. Taking a stack of images, reconstructing them, and handing them off to a neuroscientist to manually trace is a huge bottleneck. Replacing that process with the application of a pre-trained network and getting an initial result is a huge time savings. The neuroscientist can look at the result and indicate what is right or wrong.

Gregurick asked for a comment on future plans with cryoelectromicroscopy where the data sets are quite large and the real-time analysis is a challenge. **Uram** confirmed they are linking facilities. Electron microscopy data fits with the Connectomics program and ~75% of the pipeline has been built.

Lethin asked if it is accurate to argue that the leadership computer systems is a pooled resource that is available when the capacity is needed and can be shared with others. **Uram** said it is interesting to think about a continuum of computing between experimental facilities and the large-scale compute facilities. Immediately one would ask if a facility, needing the resources installed at another facility, could access those resources. In the context of Carlson’s presentation on federated identity management with DOE-wide identification authorization that would be possible. This infrastructure of dedicated resources located at experimental facilities will involve large and small islands; determining how big each of those should be is a strategic and difficult question. For example, Amazon required tremendous computing to meet peaks in their workloads. They faced the question of what to do with the extra computing when it is not being used. Amazon realized if they built such an infrastructure they could enable a lot of people, meaning if they drop to 50% utilization and sell off the 50% unutilized resources, others will benefit; it looks like Amazon World Services. If the DOE compute facilities have some fraction of unutilized computing available to others it can be a huge benefit. **Lethin** mentioned that the true gem of the leadership computing facilities is the existing fabric among the nodes that allows job to run. As the codes are optimized scientific insights can be scaled. **Helland** added that within the contract for Aurora, Argonne has the option to buy up to 4 cabinets. The idea is to get something similar to what is available at the light sources. The scaling problem has been taken into consideration; it is a form of edge computing.

PUBLIC COMMENT AFTER WHICH ASCAC WILL ADJOURN

None.

ASCAC adjourned at 11:40 a.m.

Respectfully submitted,
January 23, 2020
T. Reneau Conner, PhD, PMP, AHIP
Science Writer, ORAU/ORISE