ADVANCED SCIENTIFIC COMPUTING ADVISORY COMMITTEE

MEETING MINUTES

American Geophysical Union
2000 Florida Avenue, NW – Washington, DC

November 1 - 2, 2011

PARTICIPANTS

Advanced Scientific Computing Advisory Committee (ASCAC) members present:
Dr. Roscoe Giles, Chair
Dr. Marsha Berger
Dr. Jackie Chen
Dr. Sharon Glotzer
Dr. Susan Graham
Dr. James Hack
Dr. Anthony Hey
Dr. Thomas Manteuffel
Dr. John Negele (by telephone)
Dr. Vivek Sarkar (by telephone)
Dr. William Tang
Ms. Victoria White

Members of the U.S. Department of Energy (DOE) present for all or part of the meeting:
Daniel Hitchcock, Acting Associate Director, Office of Advanced Scientific Computing Research (ASCR), Office of Science, DOE
Mrs.. Christine Chalk, ASCAC Designated Federal Officer
Ceren Bennett
Laura Biven
Richard Carlson
Lali Chatterjee
Vince Dattoria
Thomas Ndousse Fetter
William Harrod
Barbara Helland
Randall Laviolette,
Steve Lee
Elaine Lessner
John Mandrekas
Lenore Mullin
Lucille Nowell
Karen Pao
Walt Polansky
Sonia Sachs
Yukiko Sekine

Others present for all or part of the meeting:
John Bell, Lawrence Berkeley National Laboratory
Arthur Bland, Oak Ridge National Laboratory
Barbara Chapman, University of Houston
Jim Corones, Krell Institute
Steve Cotter, Lawrence Berkeley National Laboratory
Bruce Hendricks, Sandia National Laboratory

Paul Houland, Argonne National Laboratory
Gary Johnson, Computational Science Solutions
Wesley Jones, National Renewable Energy Laboratory
Arthur Macek, Oak Ridge National Laboratory
Paul Messina, Argonne National Laboratory
Edmund Ng, Oak Ridge National Laboratory
Jeff Nichols, Oak Ridge National Laboratory
Kalyan Perumalla, Oak Ridge National Laboratory
Miriam Quintal, Lewis-Burke Associates LLC
K.J. Roche, Pacific Northwest National Laboratory
Lee Schroeder, TechSource
T.P. Straatsma, Pacific Northwest National Laboratory
Jack Wells, Oak Ridge National Laboratory
Julia White, INCITE
Anil Vullikanti, Virginia Tech University
Taieb Znati, University of Pittsburgh

MEETING MINUTES

Tuesday, November 1, 2011

The meeting of the Advanced Scientific Computing Advisory Committee (ASCAC) was convened at 9:00 a.m. EST by ASCAC Chair Roscoe Giles.

Presentation: View from Washington, Advanced Scientific Computing Research (ASCR)

Daniel Hitchcock provided an update from the ASCR. A copy of the presentation is available at: http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/nov11/Hitchcock2.pdf

Dr. Hitchcock informed the Board that there is no definitive news on the forthcoming budget. Dr. Hitchcock described a recent high-performance computing (HPC) conference that focused on advances being made in European Union that may hold high value for the ASCAC. European nations are examining exascale computing and its hardware presence from semiconductors up to full systems, and extending software via existing and commonly used community codes. Similar activities are underway in China, Japan and the U.S.

The PRESET initiative is advancing with a new center in the works and aggressive hardware plans. A considerable level of work in large data is occurring in Europe from metadata up through bitwaves and storage. Over the next few years, 2.5 billion Euros will be invested if an existing proposal comes to fruition, presenting collaborative opportunities. Coupling between infrastructure developers and IT experts will need improvement. It was also reported that Brazil is seeking to do a full genomes inventory of all of the species in the Amazon.

ASCR's budget request for FY 2012 is $465M and has been subject to House and Senate discussions. The House Committee recommends $427M and the Senate recommends $441M. There is a push for a plan to establish an operational excascale system in this decade. DOE, with the National Nuclear Security Administration (NNSA) is working on a plan for delivery by February 10th to the House and Senate and trying to determine the major risks going forward, the costs to mitigate risks, determine how much extra risk you would have left on table, and build a plan to address that. By mid-November 2011, a draft plan will be at DOE with the Office of Management and Budget (OMB) and Congress to follow.

Within the Office of Science (SC), a new round of RFPs have been established with the Scientific Discovery through AdvancedComputing Program (SciDAC) that will be awarded in mid-FY12. These represent partnerships with other offices and high priorities. The SciDAC and SC applications address things like uncertainty quantification in partnerships. The work is focused on future activities and is a way to link the best physicists and mathematicians over the next three to five years. There is work being done in multi-scale dynamics and tracer transports in the atmosphere or oceans, in cosmic frontier scientific simulations, and in lattice usage and accelerated scientific modeling and simulation. The work is high risk and represents new levels of strategic engagement, but can change the way that people do computing. The maximum value of the funding opportunity announcement across five institutes is $127.5M over three to five years. Dr. Hitchcock gave an update on American Recover and Reinvestment Act (ARRA) support and noted that interim reports from the SciDAC-e Energy Frontier Research Centers (EFRC) were reviewed by ASCR and Basic Energy Sciences (BES) in mid-September. Project reports from the ARRA-supported postdoctoral researchers were also reviewed by ASCR and SC project managers. There may be interest in keeping these individuals on a permanent basis.

The Board learned that the Magellan Project has a no-cost extension until December 31st and that disposition plans for the use of hardware and a final report are being prepared. Magellan has been useful to DOE science communities, but also demonstrated some gaps. It turns out that clouds were not relayed for mid-concurrency computing use and that because nodes are in different places elicit solvers performed badly. However, it solved problems for those with flex capacity needs; their software runs very well. Users were able to replicate environments and have done some work to help others download their environment. Open administration tools are still somewhat fragile; users can virtualize their own environment when they go to Magellan, but have to build their own environment which means that one has be his own administrator. For mid-range work, concurrency workloads may not be a great solution. The disposition plans call for some elements to end up in knowledge space with some going into test beds and into the National Energy Research Scientific Computing Center (NERSC). Next steps are being discussed with Argonne National Laboratory (ANL).

Advanced Network Technologies and Services is one of two topics identified for Small Business Innovation Research (SBIR) in FY12. The second is to increase the adoption of high power density (HPD) modeling and simulation in advanced manufacturing and engineering industries, and will try to transfer research results into buyable products. The approach will try to incentivize people to build new types of service businesses where a challenge is presented and businesses step forward with solutions. The effectiveness of this approach will be known following the review of 74 proposals by January 2012. An approach for the next year is focusing on accelerator facilities and determining how to get best the intellectual property.

A number of workshops have been held since July 2011. The Programming Challenges Workshop on July 27th – 29th (Workshop report: http://science.energy.gov/~/media/ascr/pdf/program-documents/docs/ProgrammingChallengesWorkshopReport.pdf) addressed strategies to deal with data challenges coming from large facilities. In particular, the system at Lawrence Berkeley National Laboratory (LBNL) has a high data rate presenting questions about how to move, store and share data. Workshops are demonstrating trends such as the increasing use of input/output (I/O), threads**,** and dealing with data bottlenecks. Tools are needed to manage decreasing performance due to complex hardware. Poster sessions at meetings has helped generate collaboration more effectively than the standard meeting and presentation approach.

Of note, the DOE HPC Best Practices Workshop on September 26th and 27th brought together those who run big systems to discuss problems such as file systems and archives, the need for more spindles and tape drives, serialization, and data retrieval. Data management may be the topic for this workshop in 2012 as file abstraction needs to scale to the future.

In 2011, the NERSC played a key role in accelerating the expansion of the universe as Saul Perlmutter of LBNL was awarded the 2011 Nobel Prize in Physics along with two others for their discovery. ASCR's accomplishments in 2011 also include recognition of Witold Nazarewicz of the University of Tennessee, a SciDAC Principal Investigator, as the winner of the 2012 Tom Bonner Prize in Nuclear Physics. LBNL Associate Laboratory Director for Computing Sciences Kathy Yelick was been appointed to the Computer Science and Telecommunications Board (CSTB) of the National Academies. Two ASCR Early Career Research Program Awardees in 2010, Grigory Bronevetsky of LBNL and Christiane Jablonowski of the University of Michigan, were selected as Presidential Early Career Award for Scientists and Engineers (PECASE) winners

Facilities also made progress in 2011, as four ASCR sites completed annual operation assessments by September 30th. The Oak Ridge National Laboratory (ORNL) showed leadership in computing undergoing an update from Jaguar to Titan, aimed at achieving somewhere near 10 to 20 petaflops and serving as a hybrid system with 2.6 petaflops for its operation alone.

Research highlights included work done by Patrick Chiang at Oregon State University that shared new results on off-chip communications that counter energy challenges and lead to better off-chip efficiencies. Flexible display technologies have also advanced through work with organic semiconductors to learn how and which materials to work with and look for. ASCR supported new simulations of cellulosic ethanol being done at ORNL; the research has demonstrated that plants have evolved defense mechanisms to prevent breakdown by enzymes. However, new methods have identified enzymes to breakdown sawgrass into alcohol.

Industry benefitted from ASCR results in 2011. Specifically, the Boeing Corporation has used the BlueGene at the Argonne Leadership Computing Facility at ANL to reduce aircraft landing gear and landing noise. This work will contribute to the design of safe and quiet technologies.

Dr. Hitchcock shared with the Board that multi-scale simulation and nanoscale simulation capabilities were also developed over 2011.

The Energy Sciences Network (ESnet), a is a high-speed computer network that serves DOE scientists and their collaborators. It made its first transcontinental 100G network deployment on October 13th. On October 25th, ESnet demonstrated the world's first long distance 40Gbps RDMA Data Transfer. The network can get 96 percent of bandwidth delivered to data.

Discussion

Prompted by Dr. Giles' question about communicating ASCR accomplishments, Dr. Hitchcock explained that ASCR is driven by the prioritization of activities and then the development of strategies designed to mitigate identified risks. One approach to this is thinking about how to deliver value on DOE missions; that can mean looking at facilities to ensure spending efficiencies. In response to Dr. Manteuffel's interest in the upcoming budget and interest in contingency planning, Dr. Hitchcock expressed that budget numbers may be below ASCR's House and Senate requests, hence compromises will have to be made and contingency plans put into place to avoid eliminating existing work and starting something new. Even minimal changes to budget request language can change ASCR's level of flexibility. Still, at this point in the budget process, it is too premature to say what ASCR will do with speculative language.

Dr. Hitchcock elaborated on the ASCR and NNSA integrated exascale initiative in response to a question from Dr.Chen, explaining that there is a collaborative effort to identify the highest risk and develop a plan that is motivated by mitigating risks. The two organizations have many strategic similarities and are working together to co-design the right approach.

Reflecting on his presentation, Dr. Hitchcock responded to Dr. Chen's interest in collaborative opportunities in Europe by noting that ASCR is identifying a number of opportunities and ones that fit within existing science and technology agreements and agreements for funding and collaboration at a European Union-level. ASCR may pick three or four opportunities that it can manage successfully. Currently, a tools project is in the proposal phase and there may be work in grid infrastructure. Exploration is focused on things that can deliver the most value in a short timeframe and the desire to demonstrate big wins and support future efforts and investments.

Dr. Hitchcock was asked by Dr. Negele about the House language that specifies a target date for an exascale platform and how ASCR is approaching problems that arise when organizations fail to meet a committed date. Although DOE is undergoing intense discussions on this issue, a great answer has yet to emerge. The current budget picture is unclear through 2018, so DOE is aiming toward a contingency plan to deliver what is right for science. Delivering something that has a lot of flops and no memory is something to which ASCR is not eager to commit.

In response to Dr. Tang's interest in ASCR's prioritization processes, Dr. Hitchcock explained that the ASCR continues to prioritize the development of an exascale machine while getting to scientific objectives. He addressed Dr. Mantueffel's interest in this prioritizing exascale work versus other scientific priorities noting that work in exascale depends on solving power, energy and mini-core silicon issues, but will lead to understanding power management in computers and mini-core issues and things where synchronization is too challenging to go forward. The exascale goal is to provide 1,000 times the capabilities for science. This involves balancing between floating point operations and integer operations and delivering a balanced system to actually do science. Exascale seems like the right target to deliver a higher capability system for science.

Presentation: 2011 Committee of Visitors for Advanced Networking Research Program

Taieb Znati shared findings and recommendations for the Committee of Visitors (COV) and introduced the COV members: Victor Frost, Greg Monaco, George Rouskas, Peter Steenkiste and Vicky White. A copy of the presentation can be found at: http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/nov11/Znati.pdf.

The charge to the COV covers operations of the Advanced Networking Research Program (ANRP) from FY08 through FY10 and focused on the efficacy and quality of the processes used to solicit, review, and recommend proposals; document proposal actions; and monitor active projects and programs. The COV also assessed how the award process has been affected by the boundaries defined by DOE missions and available funding, and the national and international standing of the program with regard to other computer science research programs that are focused on high performance networking tools and middleware for science.

The Next Generation Networks Program (NGNS) process used to generate solicitations was well-managed and administered. Despite having a small budget, impacts have been significant and improved many technologies useful to network labs.

ANRP decisions were based on solid data but some of the data was not available or archived. It was recommended that the efforts to automate archiving for solicitation, review, decision making, and post-award management, and archiving data for analysis be continued. There is also a need to provide greater visibility into the award process, budget and scope reduction, and the tracing of progress for the next COV. Data should be placed somewhere to enable assessment.

Communication and collaboration was part of the COV's assessment, and it recommended that a broader base of HP networking research be built to gain insight on building the next network and data center protocols. Active notifications could be used to give workshop and funding alerts, and alerts based on a predetermined schedule for FOAs and proposal due dates. The COV noted that more should be done to engage young investigators in DOE research and to help them submit DOE-relevant proposals. Efforts should continue to coordinate networking research and development efforts with other agencies and at national and international levels.

The proposal process used a secure website with solicitation information. COV members read randomly selected sets of proposals from different programs and discussed this at a planning meeting on October 10th and 11th. The NGNS used an effective solicitation development process and workshops to broaden participation and understanding; workshop participation should be increased and including other communities could help.

The COV observed that the role of the letter of intent (LOI) needs clarification and that across solicitations there is no consistency about intent and if the requirement has been enforced. NSGS should also be clear about the expectation of the use of ESnet and other DOE networking infrastructure to ensure fair and appropriate reviews.

The review process follows DOE peer review process standards; lacking, however, is a central repository for reviewers from which to create diverse and knowledgeable panels able to more effectively respond. Also lacking is a panel summary that explains pros and cons to guide program managers, inform COV members of the basis for decisions, and a review analysis for highly ranked yet declined proposals. It was thought that DOE's mission was not adequately communicated; a rejection letter could include rationale to inform future proposal development.

The COV noted that NGNS participated in the Early Career Program but did not fund any proposals in this program. It encouraged outreach to young investigators and communication of the DOE mission and the NGNS, and support for long-term research in addition to short-term work to provide opportunity to engage and attract young investigators.

While awards monitoring processes uses effective mechanisms, it was recommended that PIs address reviewers' concerns prior to an award to the satisfaction of the program managers. NGNS should also formalize and document the negotiation of awards, particularly in cases of budget reduction. All annual progress reports should be available online for analysis and review, and the communication between managers and investigators should be documented.

The COV determined that NGNS awards are impactful in enabling various useful technologies at the level of network control and measurement and have lead to the type of virtualization breakthroughs and mechanisms to manage the resources in a network in a more efficient way. Many projects advanced the state-of-the-art. Going forward, NGNS should still establish clear strategic goals regarding future funding allocations between long-term fundamental research, near-term research, and development and test-bed support.

Discussion

Dr. Berger voiced concern that the recommendations specify some work that will fall on staff who lack sufficient support, also noting that some things may make the review easier but may not contribute to better outcomes in award process. Dr. Znati recognized this but also highlighted the importance of work in areas like the categorization of information and reviewer diversity. He expressed that it is hard to collect data. Ms. White added that the COV recommends more facilitation via the IT system; existing tools could enable better communication between PIs and PMs and may not be so onerous. Dr. Berger suggested that a balance is needed between things that are built-in and things that would be time-consuming. Dr. Znati responded that a deeper analysis of these aspects could be achieved with more information and better collection tools. He clarified his concern that a balance be achieved between the sophistication of holding information and the work needed to get information to inform decision-making.

Dr. Giles asked the Board about the phrasing of the recommendations. Dr. Graham suggested that the abbreviation "CS" be clarified to distinguish between computer science and computational science. Dr. Hack noted that strong wording could clarify how to develop the

LOI rather than simply what is done with it, adding that this is not a screening criteria but just a way to convey the proposal's purpose. Dr. Znati explained that the LOI can convey the meaning of proposals prior to receipt and to enable review. Another purpose is to filter things that are out of scope. An aspect that was unclear was that the letter sounded like a requirement for the second purpose but the policy was not enforced. He urged the need to clarify the intent of letter and then enforce the policy and a desire to go back and make that recommendation stronger.

Board action

Dr. Giles moved the acceptance of the report. The Board approved the report with none opposed or abstaining.

Board business: New Charge – COV for Computer Science

Dr. Hitchcock presented a new charge that requests examination of a basic research program in computer science that is supported by the ASCR. Dr. Giles has a copy that will be posted for Board members and he will identify a chairperson to lead the charge. The charge is available at: http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/nov11/ASCAC_Oct11_CS_Charge.pdf

Presentation: ASCR-BES Data Workshop

On behalf of Peter Nugent, Walt Polansky of the ASCR described the ASCR-BES Data Workshop held on October 24th – 25th, 2011, in Bethesda, Maryland. Information on the workshop is available at: https://www.orau.gov/dataworkshop2011. Dr. Polansky's presentation is available at: http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/nov11/Polansky.pdf

The workshop organizers were given a charge with five components related to data and communication pathways, tools and methodologies, opportunities that could enhance experimentation, future research areas, and the foundation for information exchanges and collaborations among ASCR-BES research and facilities communities. A draft report from the workshop will be available on December 9th with a final report on January 23rd, 2012.

Bridging these communities required common understanding of each environment. There are 16 BES scientific user facilities, and representatives helped convey each site's capabilities.

BES facilities such as the Linac Coherent Light Source, Spallation Neutron Source and the National Synchroton Light Source can produce terabytes from single beam lines. User communities are pushing toward time-resolved and topographic studies and there is a ground-swell of interest in special research opportunities and doing analysis 'on the fly'. There is a broad spectrum of BES data needs and levels of interest using applied math, computer science and algorithms to address BES challenges. One consideration is the amount of science per dollar that can be produced and increasing the impact of the data.

The larger context explored was to address data activities within DOE and other agencies and the breadth of digital data work. The workshop included participants from the National Science

Foundation and the National Institute of Standards and Technology, national laboratories, academia, and the international community. Discussions highlighted interagency groups and efforts such as the Office of Science and Technology Policy's materials genome activity.

Speakers from academia elaborated on the necessity for layout strategies prior to doing experiments in a data intensive environment. Break-out sessions pointed out differences in languages being spoken and were accounted for with descriptive content. Workflow processes at facilities was discussed, as there is movement from the beginning of data acquisition through viewing data to changes in configuration, proposing new experiments to reduce and view data, and then determining what to do next. The process can take months and the community voiced the need for sampling approaches during the process to shorten the time lag to reconfiguring if needed. There is a desire to short-circuit the workflow, processing and visualization process.

A discussion of theory and algorithms pointed to the need for interdisciplinary engagement, and helped identify near-term and long-term activities. The community also discussed bringing Ab initio theory into experiments in real-time.

One breakout on enabling transformative science illuminated the risk in detector and source advancements and the potential story of data that will occur and overwhelm current analysis pipelines. Thinking on data strategy and tools are needed to achieve successful upgrade paths.

Discussants brought up the integration of theory and analysis. There is interest in common data formats, theory guidance for experimental design, and recognition of ASCR's investment in visualization and analysis tools, and determining how tools are being used and the extent to which they can be used.

The movement of analysis closer to experiment and moving real-time (in-situ) streaming analysis closer to the beam line was proposed. Experiments have a certain period of time allocated at facilities and users want to maximize the ability to produce scientific data in that period of time. Participants also urged the matching of data management access and capabilities with advances in detectors and sources. This is supported by the need to remove bottlenecks and apply existing data transport and mobility toolsets, and the need for integrated teams to address these problems.

The workshop featured a poster from the LBNL light source community and laid out examples for today and tomorrow. It pointed out that one way to categorize the data rate in terms of acquisition is to look at beam brightness and multiple by the number of beam lines. The community noted that they are at the terabyte per hour in terms of data collection and as they upgrade they will be at the petabyte and are facing large challenges in terms of the number of bits. They realize the need to stay competitive and the value of ASCR in further collaborations.

Discussion

Dr. Mantueffel asked about the length of time involved in the data process, doing the simulation, and if the two could be simultaneous. Dr. Polansky agreed that this could be done but the overlap can vary widely. For instance, the neutron community may have a different view than photon community. He has observed that ASCR tools appear to be separate from experiments.

There could be close coordination going on but there does not appear to be a coordinated effort to use HPC to design experiments and analyze data. There are pockets in communities that rely on HPC but Dr. Polansky does not know the extent to which it is embedded.

Dr. Graham wondered if there was discussion of having a more refined notion of what data needs to be preserved as a lot of data is generated over the process. This is a problem that needs to be addressed, said Dr. Polansky, and there was some discussion. This group was more committed to developing mutual understanding of each other's community.

Dr. Chapman pointed out 'on the fly' data analysis and the analysis of tools. The interim and final reports will say more about existing tools, according to Dr. Polansky, and will speak to how tools are being used and the options. The BES scientific user community in general is knowledgeable about SciDAC and SciDAC supported things like the open science community.

Dr. Giles asked about high energy physics as a model as there is a lot of data correction that is done. Dr.. Polansky replied that the model was presented at the workshop, but the extent to which it is disseminated remains to be seen. Another view was presented by the telescope,astrophysicists who shared a similar, clear and compelling model. An advantage in physics is that there is a central focal point on what they are doing..Dr. Eliane Lessner of BES added that users are aware of the problem. Users are very different from one another; they are part of a science digital working group and are getting all of the parameters possible. She noted that there are helpful ASCR tools, but the community is diverse and not very unified.

Dr. Tang commented on the necessity to make sure that SciDAC is well known and that the SC makes sure there is a connection.

Public comment:

None

Presentation: Review of DOE Computational Science Graduate Fellowship

Tom Manteuffel reported on the charge to the ASCAC to review the DOE Computational Science Graduate Fellowship. The presentation is available at: http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/nov11/Manteuffel.pdf

A subcommittee was asked to examine the Computational Science Graduate Fellowship (CSGF) compared to other educational activities, and the quality and breadth of the program over the past decade. The group distilled the charges down to the projected need for trained computational scientists in the DOE labs with a focus on diversity, and for continued U.S. leadership in CS.

The review committee drew information for the ASCR and reports from federal and academic sources that informed the review and informed the subcommittee of the project needs in CS. It was concluded that the need for well-trained computational scientists in government labs and in industry will continue over the next decade. It is recommended that DOE and the SC continue to view stimulation of this workforce as important to its mission.

Program effectiveness and impact were based on the effectiveness of the education process and the program impacts. The CSGF is unique in that it requires a program of study that is viewed as a contract, a practicum that places a student at a DOE lab for at least one summer, and that encourages additional practicum. Participants join the CSGF Annual Conference and have an alumni outreach program to encourage further engagement and mentorship of current Fellows.

Between 2001 and 2006, 96 percent of the Fellows completed their degrees and 101 of 102 worked in technical positions. At some point, 27 were DOE lab employees. While some have started their own companies, 42 percent are in academia with 24 percent in industry. The subcommittee determined that the CSGF has impacted the national CS infrastructure. Many alumni retain contact with labs as they pursue their careers. The subcommittee concluded that CSGF should receive additional funding and be placed on a path to double over the next five years. The CSGF budget has grown from $5M in 2001 to $7.8M in 2010.

In examining the program’s quality and breadth, the selection process, and the management of fellowships, the subcommittee looked at applicants and awardees by field. GPA and GRE scores are high among applicants, and the selection process is well executed with outreach conducted properly. The program saw an increase in applicants, uses a triage method for application, has a rigorous selection process, and has a thorough flowchart process that leverages accomplished computational scientists to review applications. In 2011, the program selected 17 Fellows.

The subcommittee noted that selection criteria were not clearly documented although some terms are defined in the application. The subcommittee presented questions that could improve the selection process and was informed by the CSGF that criteria are kept vague to allow the program to grown organically and relies on the reviewers and selection committee. The program has an annual discussion to set priorities. Krell, the institution that administers the CSGF, noted that the Fellowship must be application driven. The current application excludes enabling disciplines and fields that are important to CS. Krell is aware of this and advocates for a new program to address this need. In FY10 and FY11, ASCR proposed fellowship programs in applied math and HPC at $2M but continuing resolutions made this impossible to support. The subcommittee concluded that the program should be expanded to include additional disciplines.

An examination of data on women and minority applicants and awardees was conducted, identifying that the majority of participants are male and non-minority. Outreach efforts strive to connect with women and minority through conferences, professional societies, email and postal mail. The subcommittee concluded through voluntary input and data that efforts have increased applicants from these two groups and that the Krell Institute should continue such activities.

The subcommittee examined other federal agency educational programs and found little data on CS emphasis. It concluded that the CSGF is unique in its focus, plan of study, and practicum component, and it produces interdisciplinary scientists. It recommended that DOE continue its stewardship of the CSGF.

Discussion

Dr. Graham noted the low applicant acceptance rate in 2011 compared with 2002. It was suggested that comments be included about the consequences of such a low success rate and the strength of the applicants. This would seem to support the recommendation that the program grow; however, more justification is needed to defend the recommendation that CSGF be doubled. Even if the number of applicants doubles, that will only result in a shift from 17 awardees to approximately 35; a strong quantitative reason needs to be offered.

Dr. Graham suggested that the overall report may benefit from an introductory paragraph that explains that the Fellowship provides complete support for a specific number of years and that the Krell Institute currently administers the program which supports 17 new awardees.

Dr. Hack asked about the term “enabling sciences” and the recommendation to expand the program to cover this aspect. Dr. Manteuffel noted that if a proposal comes forward without being tied to a specific discipline then it is “not competitive,” per reviewers. He commented that the program is constrained by the proposal written back in 2008 that brought in Krell and the limits the program. Dr. Glotzer added that an awardee’s plan of study may not change and that a small-scale enabling mathematics and sciences program could be added. The practicum, outreach activities and meetings, and the focus may not differ, and similar results would be achieved, added Dr. Manteuffel. In response to a question from Dr. Glotzer about doctoral students being tied to a particular discipline, Dr. Manteuffel noted that a participant could be in one program but have a different focus and particular plan of study focused on their discipline. Dr. Giles clarified this, adding that a student could get a PhD in a discipline but they would have to focus the impact of their work on some well-defined focus area.

Jim Corones of the Krell Institute responded to Dr. Chen’s question about students’ awareness of opportunities at different laboratories, explaining that each lab has an identified practicum person. These individuals are proactive and reach out at the annual meeting to describe what is happening at labs. At the same time, different labs have individual information strategies.

Dr. Negele expressed that there are many gifted students who are non-U.S. citizens and are not green card holders, encouraging the Board to think of ideas to stimulate and support non-U.S. citizens and entice their interest early in their careers. Dr. Sarkar served on the subcommittee and noted that while the Fellowship has an impact, it is recognized that it is likely benefitting the U.S. and other countries. It would be feasible to include data on the number of non-U.S. citizen awardees and how many remain in the U.S., commented Dr. Sarkar, indicating that this may be in the DOE’s favor. He indicated that this information could be added to the report.

Dr. Sarkar commented that the program is focused on generating high-quality students for the laboratories, but some schools have master’s degree programs in CS and these students could be of value. Dr. Giles added that a large number or significant fraction of CSGF students are already post-masters. Dr. Sarkar responded that there is a need for computationally-trained students at al levels. Dr. Manteuffel added that this comment could be added to the report.

The Board visited each recommendation. Dr. Giles proposed rewording recommendation one to strengthen its impact. Specifically, it was urged that ASCR be added to the recommendation that the SC continue to view stimulation of the CS workforce as important to the its mission, and that "stimulation" be replaced with "development." This might also argue for a similar program to enhance scientists in enabling disciplines and fields. The recommendation should also reflect that the CSGF is important for the Nation and its mission. The committee found this to be acceptable.

Ms. White commented that recommendation two to double the program should be reworded to reflect the CS context versus its current wording that reflects a broad and non-specific context. Dr. Giles also urged that the metric for doubling be explained in the report. Comments from Drs. Glotzer, Graham and Giles are demonstrated in the following rewording: In light of the effectiveness and impact of this program and in the context of the growing projected need in computational science, the Subcommittee has concluded that funding for this program is not only well spent, but that additional funding is needed and should be provided.

Dr. Giles suggested that recommendation three that speaks to enabling disciplines and fields be reworded to enable flexibility and that exascale might be included to show a parallel between the focus on enabling a range of CS applications and that it is a unifying theme that calls for an interdisciplinary view. Dr. Mantueffel responded that this can be added to the text but may not belong in the recommendation wording.

Dr. Graham commented that the fourth recommendation on the participation of women and minorities reflects the need to develop recommendations that can stand on their own and are written out in the report.

Dr. Graham added that recommendation five should name and include DOE and ASCR.

Public comment

Buddy Bland commented that ORNL has had many CSGF Fellows. He considers the recommendation to double the program hard to do. He urged that the subcommittee consider concrete recommendations and ways to accomplish them. Bland offered that the program could expand to include industry and consider what they do to get high-quality PhD computational scientists into their workforces. DOE might also consider practicum for industry.

Board action

Dr. Giles moved the acceptance of the report, as modified. The motion was seconded by Dr. Graham and the Board approved the report with none opposed or abstaining. A revised draft report will be circulated.

Presentation: SciDAC-e: Updates, Highlights and Lessons Learned

Karen Pao presented an update on SciDAC-e. The presentation is available at:

http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/nov11/Pao.pdf

Dr. Pao explained that SciDAC-e was a one-time stimulus to establish a computation foundation to advance DOE's mission. It consisted of three components: 1) Applied mathematics research grants at $8M, 2) the support of 30 new postdoctoral students at $10M, and 3) supplemental awards to ASCR SciDAC Centers and Institutes of $10.8M. The expected impact is the development of algorithms to simulate the performance of electrical grids.

The ASCR's Applied Mathematics Program (APM) Research Grants have funded seven projects since 2009. Highlights include the development of tools for grids to simulate and reconfigure power systems, extensions of the Kalman filter to do data simulation and look at data in electrical grids and use that to model the dynamics of power grids, and the determination of a method for optimally allocating simulation budgets.

The support of postdoctoral students is part of a petascale initiative. Eleven students are at the Argonne Leadership Computing Facility (ALCF), 10 are at the Oak Ridge Leadership Computing Facility (OLCF), and nine are at NERSC. Project reports were reviewed by ASCR, Fusion Energy Sciences and BES in October 2011 and comments were shared with the three facility leads. Highlights include the optimization of performance of a fusion code called GTS and demonstration of weak-scaling by Robert Preissl at NERSC. Dr. Preissl's paper is one of two finalists for best paper at the International Conference for High Performance Computing, Networking, Storage and Analysis. NERSC's Jihan Kim improved CPU time needed to screen a dataset of 5,000,000 materials from several years to a few weeks. At ALCF, Aaron Knoll developed volume rendering techniques to visualize materials according to physical quantities. A postdoctoral researcher at OLCF examined nanoribbons and discovered a boron nitrid nanoribbon that could be used as a nanoconductor and could be used for graphene nanoelectronics. And, OLCF's Josh Hursey proposed run-trough stablization that allows a message passing interface (MPI) application to continue even if one or more MPI processes fail; this proposal may be included in the next MPI standard.

The supplemental awards component of SciDAC-e supports collaborations with EFRCs and institutes to support BES EFRCs and develop an HPC capability relevant to EFRC goals. Funding for SciDAC PIs has been extended through August 2014 to all for tool development. Interim reports were received in August 2011 and were reviewed. Some projects are still gathering data, need to examine more challenging problems, and two have to submit new letters of commitment due to EFRC leadership changes.

The use of the SciDAC infrastructure consisting of experience, personnel, breadth of technical expertise, and organizations allows the rapid transfer from ASCR to collaborative application scientists. There are application scientists open to working with ASCR scientists and SciDAC has helped them build working relationships. Consequently, ASCR researchers have learned to speak the language in the other fields and can deploy their capabilities more rapidly.

An additional lesson learned from this investment is that it is important to encourage students in CS to gain a broad set of skills and experiences. The corollary is that SciDAC engages them and keeps them engaged. The individual sites have also strongly supported the researchers and helped mentor them.

Discussion

Dr. Giles asked about the path forward for the postdoctoral researchers. Dr. Pao noted that they are currently finishing two-year terms and the limit at the laboratories is three years. It is not clear yet if these researchers will be hired on and converted to regular staff members.

Presentation: Workshop on Analysis, Simulation and Optimization of Complex Systems

John Bell presented an overview of the workshop held September 13 – 14, 2011. The presentation is available at:
http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/nov11/Bell.pdf.

Workshop details are available at http://www.orau.gov/mathworkshop2011/. The workshop was held at the request of ASCR program managers and was motivated by the need to identify research in applied mathematics for DOE to meet its mission requirements.

The APM has been in place since the 1950s and is funded at approximately $40M per year. Its core strengths are critical and core areas of research remain very active. Multi-scale mathematics and optimization for complex systems, mathematics for distributed interconnected systems, mathematics for analysis of petascale data, and joint math / computer science institutes are recent thrust areas reflected in solicitations for FY09.

The workshop sought to identify areas of applied math research to address challenges within understanding complex systems. Discussants dealt with key aspects of complex systems, various types of research needs, and the types of problems that the APM needs to address.

Research needs in four mathematical themes were identified and were ubiquitous across all areas that were examined:

1. Building and using hierarchies or collections of models
2. Modeling systems characterized by diverse phenomena
3. Systems that are inherently chaotic, stochastic, and/or uncertain
4. Integration of data and observations with modeling and simulation

The attendees examined computational challenges in these themes associated with multi-mode architecture but were not asked to address exascale.

Each theme was addressed by attendees who were mostly from optimization and multi-scale mathematics communities, yet the workshop focused less on big data and that perspective. The second theme on modeling systems characterized by diverse phenomena brought out discussion of examples such as elements in power grids. Coupling strategies is an area where elements such as synchronicity arise. The third theme that examined characteristics of systems spoke to efficient techniques for capturing rare events. For instance, users and operators do not want nuclear reactors to have problems or for the power grid to go down – they just want them to work. Dr. Bell described the fourth theme as having to do with how data is used to constrain system dynamics and how to solve inversed problems with sparse and noisy data. The theme touches on uncovering ways to use an assimilation model and data to make a prediction.

The greatest potential impact from these themes and discussions will come from synergistic uses of these methodologies together. This is evidenced in examples and what the mathematics community could do based on discussions at the workshop. One example is environmental remediation and how to bridge the gap from pore scale to field in order to define geochemical rates and use data to constrain system behavior and assess risks associated with remediation alternatives. Another example is increasing the reliability of power. A question is how one can develop a predictive model, use real time data to improve model capability, develop control algorithms, and develop strategies to operate the grid in the presence of variable generation of wind and solar power? Additional examples include catalyzing fuel generation by coupling physical models and bridging temporal and spatial scales, and managing nuclear reactor rods to optimize their usefulness and dispersion within a reactor.

The workshop pointed out that there are a huge number of DOE problems with solutions that require significant efforts in applied math. Research can make various aspects of models and modeling fit together. This can lead to mathematical challenges that could also help design new numbers of algorithms and software that exploits next generation parallel architectures.

Discussion

Dr. Bell responded to Dr. Giles' question about how to continue this by proposing that recommendations will evolve from the report that point to investment areas to maximize applied mathematical spending.

Dr. Mantueffel noted that the discussion looked at multi-core architectures and that the work seemed to be platform independent and almost counter to what is happening in the exascale environment as we move toward greater complexity. He asked what the new platform for exascale will look like or if there should be another kind of platform. Dr. Bell noted that there is a lot of discussion of multi-core systems and issues, and that the machine may become the tool to solve DOE problems as efficiently as possible. However, he believes that a missing piece is not just doing better simulations but having enough required computing power to solve problems. The other challenge is building a mathematics infrastructure to solve more complex problems.

Dr. Pao of ASCR commented that the workshop intentionally left out large data and analysis as an integrated endeavor as this is not something that can be addressed in this workshop. It left out exascale computing as it will enable solving complex systems problems. An aim is to unite these two groups to determine if there is synergy in building a more innovative approach in algorithm space to address these problems; exascale by itself without algorithmic advances is not helpful.

Dr. Tang voiced an interest in time to solution. Advanced computers will all have low memory core systems. He asked Dr. Bell what the most optimal would be, and that dealing with people in larger communities in these domains will mean convincing them to use different algorithmic approaches that are properly cross-benchmarked. Prioritization and moving forward presents choices. Dr. Bell commented that the mathematics community needs to address low-core math algorithms, noting that a base program continues to exist and people at the labs are trying to be responsive to that. What is needed is to work in lower-memory environments and convince people that different formulations are better, and thinking about alternative formulations.

Dr. Chen commented that exascale and mathematics are not mutually exclusive. The backdrop will be working with exascale in mind and thinking about new algorithms in an exascale context. Dr. Bell responded that in thinking about new algorithms, most people in the applied mathematics community are already doing multi-core stuff. New approaches will invite them to work with what we have on the machines that we have.

Public comment

None

Presentation: Titan: A new leadership computer for science

Dr. Arthur "Buddy" Bland of the ORCL described work on Titan that began just after ORCL accepted Jaguar in January 2009. His presentation is available at: http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/nov11/Bland.pdf

Titan is aimed at leadership computing to increase computational ability by 20 to 40 petaflops. New resources are needed to avoid gaps between science needs in scientific programs and available resources. Titan grew out of two projects at ALCF and OLCF. Another motivation for Titan is oversubscription on the Innovative and Novel Computational Impact on Theory and Experiment Program (INCITE) where the request in 2011 is at 5,743 hours versus the 1,662 hours allocated. The requests have increased while resources have leveled out, and in 2012, requests are down a little as requestors are going elsewhere due to oversubscription.

Titan is currently called "OLCF-3" and design work started in April 2009. ORNL will upgrade Jaguar from 2.3 petaflops to between 10 and 20, will be built on Cray's newest XK6 computer blades, and will be operational in 2013.

Within Jaguar, the old node boards are being replaced with new node boards. It will go from 12 to 16 cores on the AMD Operton side. There will be an increase in the number of GPUs compared with the current system in order to upgrade faster. Memory will be increased per node from DDR2 to DDR3, and Gemini interconnectors will be used in the new blade.

Titan will reuse the Jaguar cabinets, cooling features, and other components. The O/S will not change too much but will use new Gemini interconnectors and support new features such as globally addressable memory and advanced synchronization features. The new accelerated node design will use NVIDIA multi-core accelerators and disk bandwidth will be increased.

The upgrade will be phased from XT5 to XK6 and essentially split Jaguar in two; the half that contains cabinets will be upgraded and users can use the other half of system. Eventually, users will migrate to the new half and the other 104 cabinets will be upgraded. By December, the two will be combined to move up to 3.3 petaflops. An acceptance test will be performed.

The integration of a GPU system is useful for several reasons. Single-thread performance takes more power and multi-core architectures are a first response to the power issue. However, conventional processing architecture is optimized for single-thread performance and not energy efficiency; the processor has a lot to do beyond delivering flops and doing arithmetic. I t can be more efficient to use simpler processors, exploit vector/SIMD parallelism, and have a slower clock rate. You also have to deal with Amdahl's law. Overall, one has to be careful about saving energy. This can be addressed by using heterogeneous cores and an architecture through fast serial threads, explicitly managed memory hierarchy, and micro-architecture to exploit parallelism at all levels of code. NVIDIA Fermi has made GPUs feasible for HPC. The real problem is that programmability is still a barrier to adoption but Cray is focusing on tools to make adoption easier.

There are application programmability issues for Titan. The biggest problem is not getting applications to run well and include node level concurrency, data locality, and programming models and how to use threads of execution. The key is code portability as codes will need to run on other machine architectures. ORCL is conducting an application readiness review to determine if good performance can be achieved and if it can do this in a code portable way.

Six applications were chosen to enable the development of Titan. S3D will look at combustion code for direct numerical simulation. This works on science problems for auto-ignition and the effort is now moving to more complex fuels. There is a need to study complex fuels, and so far, the needed application will run four times faster on Titan versus Jaguar. The second application is WL-LSMS and will present a physics code using WL Monte-Carlo and first principle density function theory. Titan will compute thermodynamic properties of magnetic materials and it is anticipated that there will be a 25 times increase in speed versus the CPU core on Jaguar. A third application is CAM-HOMME that will examine new atmospheric chemistry to make reliable, long-term climate predictions. Titan shows a two times improvement in processing speed. The fourth application is PFLOTRAN – reactive groundwater code written in F90 used to simulate CO_2 sequestration in subsurface geologic formations that contain saline aquifers. ORCL has found a four times speedup of host+Fermi processors and expects the code to give good results. The fifth application is LAMMPS. This is a parallel particle simulator that will help with biomass recalcitrance and biofuels and turn cellulosic material into biofuels. Titan is running this up to five times faster using all 16 cores. A new MSM algorithm has been written and may be a scalable algorithm. The sixth application is Denovo and solving radiant transport problems for nuclear reactor design. The new GPU-aware sweeper runs two times faster than on CPUs.

The Titan programming environment is focused on ease of use and application portability. There will be a full-fledged programming environment for hybrid CPU / GPU nodes and a hardware agnostic programming model that is portable. Titan will leverage emerging existing software capabilities. To build this environment, OLCF has been working with component developers to address compilers, debuggers, performance analysis tools, and mathematics libraries.

Titan's training program seeks to help users achieve break-through science by developing codes to expose hierarchical parallelism, tools, codes and techniques. Courses have begun.

Discussion

Dr. Chapman asked about the team that is enabling early porting and how that will be preserved and carried forward. A team of individuals from Cray and NVIDIA are restructuring code to expose more parallelism. Lessons learned are emerging and there is work to get algorithms that have been paralleled and share these with people. The training team takes these lessons and shares them with others. There will be “bring your code” workshops that move from examining code and to finding ways to make it better. To get to exascale, ORCL has to learn how to expose more parallelism in the codes. More GPUs in all 200 cabinets would result in 10m parallelism.

Dr. Graham asked about how to compare the performance of Jaguar versus Titan. ORCL noted great baselines at the beginning of Titan and experiments will be done by running the same code on the CPU part and running it both ways to see performance levels. This is the job of the performance analysis team. Some codes will have better performance than others.

Dr. Bland mentioned that the longer term goal for user applications is that they will use the directive space requirements. This was in response to Dr. Glotzer’s question about code development and if it is being done using CUDA or open CL or if it depends on the project. Within the applications that Dr. Bland mentioned, some use CUDA, some use directive space, and some are mixed. The real goal is to write code using parallelism that works across directives. Dr. Bland explained that you can get a card with C2090 and use that same chip. The four-on-a-blade product is a packaging element and there are still four separate nodes. He explained that the new Keplers will use exactly the same chip on those boards. This takes full advantage of economies of scale and will use the same product that NVIDIA develops for its core business.

Dr. Bland noted that the cooling system will be same as before in response to Dr. Giles’ question about any changes to the macro architecture. A possible exception based on the final version of Kepler is that ORCL may need new fans to support the power system. Eventually, there will be one Operton and one NVIDIA which is hotter but is much faster. Dr. Bland estimates that use of all 200 cabinets will put things in the 10 megawatt range versus seven. ORCL uses about five megawatts for its typical and current workload.

In response to a question from Dr. Giles about browning out the power or if there is modeling of that, Dr. Bland noted that AMD guarantees that if the system is not using all of the cores effectively then they can speed up activities in other areas. ORCL needs sufficient resiliency as it is running many chips.

Dr. Tang asked about comparisons to the new Chinese system, Tianhe-1. Dr. Bland views Titan as more integrated as Cray has developed interconnectivity to the board and all the way through. Fundamentally, the Tianhe-1 uses similar chips, but Titan will have one-to-one correlation as this is a transitional type machine with people learning to use the hybrid parallelism.

Public comment

None

Overall public comment

None

Board business

Dr. Giles welcomed Dr. Glotzer as a new Board member

The meeting was adjourned at 4:35 p.m.

Wednesday, November 2, 2011

The Board Chair Dr. Giles started the session at 8:31 a.m. EST.

Presentation: Early Career Research: Reversible Software Execution Systems

Kaylan Perumalla of the ORNL presented efforts by a team of collaborators to address large scale concurrency and inherent software challenges such as synchronization, tolerance and debugging. The long-term view is that there is a need to explore new paradigms to meet these challenges. The presentation is available at: http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/nov11/Perumalla.pdf

Dr. Perumalla explained that most simulations are forward-only in that they think in one direction and are built to execute forward only. This means that components must always be correct. This is the basis for constraints in large-scale execution as it requires all components to be coordinated continuously. Dr. Perumalla and his team propose enabling reversibility at a software-level and at scale, and are exploring ways to advance the paradigm of reversible execution for efficient large-scale concurrency.

Traditional check-pointing requires large state memory size, forward execution is slowed due to the use of copying overheads, and there is reliance on memory that increases energy costs. In the reversible software approach, there is a reduced state memory size and overheads, and reliance computation is more energy efficient. One application is the simulation of elastic collisions reversibly. The reversible solution has generated new reverse code and the results will soon be published. There is related work, but very little exists that aids solving reversibility.

The spectrum of reversibility moves from general purpose to specialized use and from coarse-grained to fine-grained execution units. This approach addresses fully reversibility in an asynchronous mode.

The approach is hard to realize in practice yet most existing simulations are reversible. The challenges are determining how to come up with reversible models of physical systems and how to automatically make existing code reversible. The team is looking at possible automation as well as programmer-assisted approaches and will likely end up with a hybrid of these. They are also working with a reverse compiler that generates two forms of code: modified code and a reverse code. The existing infrastructure can be reused for compilation but it is difficult to automate completely. The reversal methodology for source-to-source provides upper bounds so optimizations are needed to get to a higher level.

The team will look at model-based approaches and reversibility of the Fibonacci number. To work in reverse of this, it will generate a big log but will also generate a reverse code that can be obtained and maintained. Reversible versions of commonly-used libraries are needed; at a high-level, once a matrix is factored, the check-pointing approach will save the whole matrix.

Additional work is needed to develop efficient runtime engines for reversible execution at scale. Once local reversibility is achieved, a runtime process will be conducted. Computation times are pushed backward which has the potential for relaxed synchronization. Causal errors are detected at runtime and intra- and inter-processor reversal can be done. Highlights from reversible runtime in feasibility include achieving high efficiency through reversal asynchronous execution. The team has been able to scale well up to the full Jaguar 216k core scale. The number of dependency violations is fairly low and they are seeing up to a 60 percent improvement over synchronous execution. Another example is a reversible execution of a reaction-diffusion process model to obtain high efficiency at scale and achieve a rate of asynchronous execution at a rate that is two times higher than synchronous runtime. This leads to a very fast execution rate due to a relaxed approach.

Questions remaining include determining the memory lower bounds for reversibility, the trade-offs between reversal, recomputation and memory costs, and the need to develop backward-compatible interfaces for reversibility. For components, research is looking at automation techniques, the runtime approach, the theory, and experimentation.

Theoretical aspects point to challenges in ensuring phase space coverage while being reversible. Methods are not always reversible and have uncovered this problem with strange sine reversals in elastic collision. There are some approaches that can be taken immediately but are not optimal with respect to memory. Reversible root-finding is one of the challenges that may not be feasible through the reversibility method, but it is being investigated using the Newton method. Reversible Turing machines are also being explored by executing a reversible core over a reversible platform.

When the team comes up with more runtime schemes, they will have to have a platform to experiment with these. They have developed a new platform that allows for an MPI application to be tested over this test-bed that is run in parallel. It can be scaled to 10^7 virtual MPI ranks. This is state-of-the-art today as very few tools exist to support this experimentation.

This is viewed as the seed-corn of HPC in the sense that part of this is being used to grow the next generation crop. The approach can take real MPI codes and this helps in co-design approaches. One challenge is scale and simulating real MPI code to work in millions of ranks to support collectives. Dr. Perumalla sees this eventually being used in DOE mini-applications that will make this experimentation more efficient.

The vision for this is to have integrated reversible software working efficiently at scale. The near-term plan is to share software releases and have a "reversible challenge series" to stimulate work and research in this area.

Discussion

Dr. Giles noted the connection to theoretical reversibility work and pointed out that all computation data must be saved. That can be a serious problem due to the complexity of the machine and he wondered how this work relates to that. Dr. Perumalla expressed that if one sees the whole computer as a state machine, no one imagines a specific number needed to reverse that. If one wants to reverse practical code, they have to work in a stable state. He is trying to bridge the gap between the two. A lot of the codes can take a much smaller amount of state, and this work is transforming the data to a control flow problem rather than data transformation.

Dr. Graham brought up modeling at runtime and wondered what would be done with the results if there was much faster computation. She asked how much of this depends on saving a fair amount of the communication. Dr. Perumalla noted that the good performance is not so surprising as his team is only getting it at the highest scale. He also noted that the amount of overhead needed for reversal can be equal to or more than forward. The thing that helps is that there is some concurrency in the model and when that plays out dynamically the hit is for violating dependencies and this occurs only when there are those violations. There are multiple factors at play at a larger scale. In computation, you can imagine processor speed; if the speed gets very high, then there is essentially no charge for corrective computation.

Dr. Berger noted that it seems like overheads would have to be dependent on something very specific and compared this with a transaction-based approach. Dr. Perumalla mentioned database transaction as another instance where corrections are made. This is conceptually different; databases do not have a computational element. The dependencies are logged and recovered from the log. He commented that in the grand scheme and range of applications, we may have to be able to use some of that.

Presentation: Improving computational effectiveness metric – FY11 results

Ken Roche of Pacific Northwest National Laboratory described FY11 software effectiveness conducted by four software application groups with an annual goal to improve computational science capabilities. The presentation is available at:
http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/nov11/Roche.pdf

Work in this area has been inspired by Toring machines and finite resources. Machines have to be programmed and there is a level of determinism being used today. The process of evolving languages brings new thought process and the generation of the language of problems plus time to generate the language of the result. This leads to accepting or rejecting the result

In computational effectiveness, the metric is looked at through simulating the problem in less time but modifying it with different levels of machinery and over reductions. A second view is simulating a larger problem in the same time and using a weak-scaling model and scaling up the number of processing elements. It is important to look at what the machine sees when the problem is executed. With strong scaling, the tie goes down. In weak-scaling, you might increase the number of processing elements.

Processing of the metric is culminated in a report that will go to ASCR. This process is essential describes the problem and looks at available resources today. The report contains details on how to do computation, how software was built, the software environment, and other elements.

In 2011, two target platforms were Jaguar and the GPU test-bed at Dirac. The Dirac system contains nodes that have different GPUs on them. The one with a homogeneous set was around 36 nodes. The examination looks at how to make good use of multi-core and hybrid computer nodes, and might also look at hiding use of a GPU.

Some of the techniques used in processing may have to be rewritten. A rewrite can avoid non-temporal writes, loop unrolling, vector shuffling, vectorization, and reordering. With stack alignment, it might be better if alignment makes better use of the cache. Function inlining and prefetching are things that programmers can use to direct computers. Dr. Roche explained some challenges for the semiconductor industry per the fabrication and demonstration of devices.

OMEN is software that enables the study of the next generation nanoelectronic devices where quantum mechanical effects such as energy quantization and electron tunneling play a very important role. The OMEN tool is similar to CAD for people in the chip industry. There is a GUI interface where you can run studies. A system at Purdue University allows you to feed in questions and the approach can model many types of structures in 3-D and 2-D. This work looks at quantum transport in the tens to thousands range. OMEN takes advantage of modern software capabilities and looks at double-gate ultra-thin body field effects. It is predicted that there will be a 22nm device forthcoming. The calculations they do are self-consistent and solve for the short integer equations. As there are combinations to account for, this has to be solved thousands of times. This can be parallelized over voltage and then there is 10-d spatial domain decomposition; programming must start at the bottom and work back. Dr. Roche explained that one of OMEN's problem formations was rewritten as they were working with a larger cell that was working well; it was recast in terms of a primitive unit cell. The same results were reproduced in Q2 but by Q4 had replicated the unit cell and connections above and below the cell. This achieved more than a factor of two speed up showing that scaling to this new change actually works.

NEMO5 addresses optimized ballistic quantum transport and looked at computations that were different strains. This led to straining that influenced charge carriers and the need to measure each charge strain. When working with a large number or atoms, one must deal with the size of a quantum dot. The number of degrees of freedom is 20 atoms and the experimenters increased the spatial geometry around the quantum dots. This work recompiled PETSc and improved matrix preconditioning which made a huge difference. The physics domain decomposition required coupling to neighboring partitions and neighbor-neighbor couplings. Improvements were substantial and achieved 3.5 times faster results; these are solutions that could be scaled. The dramatic increase of atoms was used in Q4 and it was found to be three times faster.

The LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator) study looked at user CUDA and the GPU package and two different runs for their systems. A function evaluation led to an interaction for these problems and the need to evaluate for a specific number of atoms and problems. One problem addressed was a copper cluster problem that used embedded atomic interaction and looked at atoms before adding an impurity. The second was an atomic “melt“ problem. A comparison of the development of the two was conducted to view GPUs as they emerge. In one case, one MPI process and GPU are used. In the other case, the research looked at exposed GPU. Due to the nature of the inprocessor cores on the host, you can only use one MPI process on one node in +CUDA. Working with 1 million atoms demonstrated a dramatic jump in time; however, one should be able to use eight processing elements. Using the CUDA enhanced version with one process versus MPI with eight processes, you still win. Hardware limited pushing this out to 30 nodes and 30 processers. This pointed out that the GPU has to be kept busy. Minimizing the amount of work means you end up losing out as you amortize. Next, the project will look at the fraction of work needed to go to GPU for arbitrary problems. When part of the computation is not done, there is a reduction of time when you use the algorithm designed for LAMMPS. The user can decide how much of the GPU to use and have it do something else. The group did a strong scaling study and was successful in reducing times and achieved success in weak-scaling. This means that there are applications where work can be moved to the GPU.

ORIRIS looked at a problem space with particle beams that propagate a plasma near the speed-of-light. When one boosts electrons, the ions pull the electrons back. This creates a wakefield. There is a desire to study this process wherein electrons are displaced. The computation moves from current disruption, solving Maxwell’s questions, then uses fields and creates a force to look at atoms displacement. When this system is pulsed it creates an imbalance. Balancing the particle load is hard. Within the laser wakefield scenarios, this was used as a base case to understand how to do this in other scenarios. Much of the code was vectorized, particulate loaded in the vector unit, and fields were interpolated. The grids were upgraded based on understanding how to store the components. This was done through transpose operation and is extremely efficient. Enhancements included dynamic load balancing and an SMP version of the major distributed kernels to test load balance and achieve proper alignment of processing elements with actual geometry for the problem. The researchers conducted full systems runs on these problems looking at particle injection in laser wakefield. In studying interpolation schemes, ways were found to justify the study -- better interpolation led to better physics.

eSTOMP is a subsurface transport over multiple phases. Bacteria were used to catalyze metal in the way that water plumes can be brought back to above useable standards. This uses more than 2 million grid cells and various problems were fixed to conduct the study over several hundred days. Two major enhancements came from fixing memory issues by modifying temporary allocation for distributed arrays. A problem was fixed in Q2 where the project used over-prescribed nodes where only 52,000 were active. In I/O operations, many features were changed. The project fixed the ASCII front and did a binary dump. They achieved more than a factor of two in improving speed by fixing code. The researchers also wrote a code to study their problem and software that looks for best way to run this. When done with the problems, this will reduce time and the set of parameters to produce I/O results 54 times faster and 10 times faster that the best binary version.

There are four FY12 application nominees and a desire to do something to impact novel product development. QMC PACK is a desirable challenge. Drekar::CFD looks at damages to cladding due to turbulent fluid flow. NIMROD looks at sparse inverse iteration and could move forward with new hardware development and new software technologies. The Materials Project at MIT (with LBNL) uses VASP which may be a concern; this projects studying machine learning techniques to mine chemical bonding knowledge. The challenge is that VASP is not free.

Discussion

Regarding OMEN, Dr. Graham pointed out that gains can be due to algorithms yet sometimes improvements are really technology gains. A compelling thing about algorithmic change is that it can exist when moving to a new approach. She asked what bounds were used last year. Dr. Roche responded that he tends to avoid enhancements that are not portable. This year, the project adjusted to new hardware features and enhancement was difficult to achieve. Most of the portability is kernel-based and this project showed end-to-end results. There is an enhancement delivered from hardware developments. That is being examined and adjustments are being made.

Dr. Mantueffel asked if there was a team that brought physics into the calculations. These efforts work directly with people who work in applications and they relied on physicists, according to Dr. Roche. The studies drew from collaborations that relied on first defining the problem successfully and identifying interesting science problems.

Dr. Giles noted that due to OMB benchmarks, an annual report and metrics are necessities. He expressed concern that calendar-driven reporting fails to adequately capture advances that have longer-reaching impacts and lead to scientific improvements. The studies look at power consumption between a host and GPU and can look at end-to-end results and do a comparative only analysis of MPI vs MPI-only runs, said Dr. Roche. There could be another study to show power gains. Dr. Giles replied that it is hard to answer this in a report as it goes to OMB, but one can voice an interest in seeing this work develop for scientific purposes. Dr. Roche hopes that this is not a single endeavor for these groups and that they see benefits in continuing. The problems tackled can be dissected further and something can be done in specific areas as benchmarking has now been done.

Dr. Tang recognized that examination of the application codes showed some demonstrated capability. The ante has been upped in looking at GPU hybrid clusters and similar things. In selecting a suite of codes, he asked if there is interest in making optimal use of GPU platforms already. Dr. Roche noted that the current territory is novel and there are groups that use the hardware today. Many groups do not know how to use the hardware effectively, however, and some are underusing host nodes but starting to feel comfortable with hardware use. The selection process has to be careful.

Dr. Chen asked if any of the problems were compared with different programming models such as CUDA versus open MPI using directives. There are GPU packages in LAMMPS, said Dr.

Roche, with an example where CUDA is used and there is only CL. There is concern with trying to hit a certain metric.

Public comment

An audience member asked about using SSE instruction and what kind of abstract ion machine would allow one to generalize that to map it. Dr. Roche commented that one could probably write code to take advantage of the memory structure.

Presentation: Advanced Networking Initiative

Steve Cotter of the Energy Sciences Network (ESnet) described the Advanced Networking Initiative (ANI). The presentation is available at:
http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/nov11/Cotter.pdf

The ANI is funded through ASCR and has grown by nearly 50 percent since 2008. ANI supports more than 200 universities and has traffic growth comparable to commercial industry. The network is built to handle the needs of different science applications. ANI handles a small number of large flows versus commercial industry that must be equipped to handle many more small flows. Utilization of the ANI backbone is highly variable and driven by the type of flow which depends on peer and partner networks.

Since its beginning, ANI has seen an explosion of data in disciplines such as genomics, high energy physics and light sources. Large experiments drive large data rates, and traffic growth has grown consistently by 10 times over an average of every 48 months. Right now, the flow is expected to be around 100 petabytes / month of data in 2015.

The network typology in 2009 was represented by a 10 gigabit link between various sites around the U.S. In some places, gigabit use levels varied, but as various data uses increase, more links will be added. In preparing for the 2011 – 2015 timeframe, ANI had to develop a prototype to demonstrate that growth was possible. This was based on recognition that U.S. firms were investing less in R&D versus their competitors and that the U.S. economy slowdown might result in fewer qualified personnel, one-third of whom who work in telecommunications are in R&D. ANI worked across four sites initially with a stretch goal of working with other sites. It developed a network research test-bed facility to do large-scale testing and acquired dark fiber IRU to avoid limitations to testing technologies. An advisory committee was set up to look at test-bed proposals, and a research infrastructure stood up for dark fiber projects.

Early challenges included not having a 100 Gbps standard at the time. This made it difficult to understand where technology would be at the time and what equipment would cost when it came to market. It was challenging to understand what ANI would be able to offer the community. Discussions with the major carriers pointed out that none had plans for a 100 Gbps network at that time. Consolidation in the telelcom market and little infrastructure pointed out that there would be a challenge with acquiring the dark fiber infrastructure.

In 2009, optical equipment was purchased and a lab set up with OpenFlow switches fitted with measurement devices to set up a tabletop test-bed right away. Separate to this, ANI bought dark fiber on Long Island to get to Brookhaven so that it could use tabletop equipment deployed in Long Island. This gave researchers their own wavelength to control separate from other traffic. ANI worked to get a message to industry that there would be a 100 Gbps opportunity.

In 2010, technology became advanced enough to issue an RFP. There were many extension requests from manufacturers as they finalized their equipment and pricing details. There were also some vendors that provided routers for testing and it was found that the routers were far behind the optical makers. By the end of 2010 and after two rounds of research proposals, equipment was prepared for redeployment in phase two.

After negotiations were complete, a footprint network was built with the four initial sites consisting of backbone between major cities. The approach selected allowed access to 50 percent of the backbone and in metropolitan networks, there is full access.

The 100 Gbps network used Internet2 stimulus funds to build a broader national footprint. Level3 Communication and Indiana University will work with LBNL and Internet2 to manage equipment and infrastructures. ANI will purchase its own transponder and will own all of the equipment to meet anyone's request. There is a move to get all of the sites connected. Work is underway in the Bay Area, Chicago, Oak Ridge and Long Island.

Four locations will be operational before the end of December 2011 and the rest of the network will be built by July 2012. Work will continue to add sites. By July 2012, there will be six to nine months of experience to support moving the production network onto new infrastructure.

This represents the only science network facility in the world with production-class owned infrastructure for cost-effectiveness and scalability and an infrastructure capable of supporting disruptive national-scale network research. There will ultimately be three network options that will be available to commercial entities and universities and not just R&E researchers.

Early accomplishments have included developing the award-winning MAVEN application to show real-time network energy consumption. In October 2011, ANI built the world's first transcontinental 100 Gbps link, the world's first 40 Gbps RDMA over a long haul network based on a public-private partnership, and demonstrated end-to-end circuit service to show OpenFlow on campus infrastructure and the ability to automatically connect with another end point without including a network administrator.

The new network will be fully instrumented to create an energy-aware network to create an ecosystem where tools and business models will incorporate traffic proportionality into their consumption.

By end of 2011, the network may support up to 30 projects using the testbed including scaling TCP well beyond 10 Gbps (see http://www.es.net/RandD/advanced-networking-initiative/current-testbed-research/)

Network projects will be demonstrated at the 2011 Super Computing Convention including linking from New York to NRSC up to Seattle and down to the conference showfloor and connecting from Chicago down to ORNL. Eight demonstrations will use 100 Gbps links at the conference. One example is LBNL showing the difference between 10 Gbps and 100 Gbps.

Discussion

Dr. Cotter noted that transitioning to the new network will occur in July 2012 with many of the same locations, so there will be little disruption. Other universities and sites can apply to use the test-bed. Some NSF-funded projects and universities are using it and commercial partners like Alcatel and Lucent have used it, too.

Public comment

None

Presentation: Early Career Research: Algorithmic Foundation of Diffusion on Complex Networks

Anil Vullikanti described challenges to social systems and complex networking data. The presentation is available at:
http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/nov11/Vulikanti.pdf

The structure for graphed systems is very amorphous and irregular. Recently, there has been excitement about rallying people via social networks, but scientists have used diffusion processes to understand underlying processes. There are networks of nodes that each have states that change based on their neighbors. This leads to a dynamic system and supports phenomena such as epidemics, packet flows, power grids, cascading failures, and fads and social conventions. Looking at questions of interest, data can be seen in terms of graphic dynamical systems. A first goal is to determine the appearance of the networks and characterize the dynamic properties for the structure, and then look at techniques to control dynamics.

While there are challenges such as dealing with the underlying complex dynamics, a first step is modeling and the analysis of complex networks and very large graphs. Historically, early studies of complex networks looked at very simple models designed to capture degree distribution. This moved to understanding social evolution models that were useful but couldn't measure second order properties. This led to the HOT model and ways to optimize objectives with random evolution. Current research looks at first principle approaches and modeling social networks to deal with getting data from very large complex networks. This requires HPC tools.

One needs to go beyond degree distribution to lead to random graph models to preserve non-local properties. When making random edge swaps, one can see changes as networks change along with changes in the dynamics of networks. Good models are needed so that as there is input one can get very different results if the structures of the applications do not catch them. Results include developing a "first principles" approach for social and infrastructure networks and a need to focus on computing problems of very large networks.

A general goal and basic problem is to find whether there is an embedding of models within smaller nodes. Looking at one or more embeddings of the small subgraph in a complex model. the applications for this model are diverse. This type of graph analysis works with problems that arise in many areas beyond just social networks. From a social side, however, one can determine a mapping of subgraphs and connectivity between subgraphs. Applications include systems and networking, and examples include identifying memory leaks.

A summary of results in this area includes randomized algorithms for counting and approximating the number of subgraphs. This leads to an approximation of property and eventually, understanding of complexity and scale to graphs with large number (500 million) of edges. This can be run in heterogeneous environments. Results include achieving a broad class of relational subgraph queries. This can handle certain types of grammar based-queries. One focus is on tree and tree-like queries. None of the sequential versions move beyond 100,000 node graphs at this point.

One difficulty with graphing is determining ways to partition work and identify smaller subgraphs and then put them together. There can be overlaps in networks, but this can be avoided by keeping track of extra information on the sub-graphs to understand and avoid overlaps. One strategy is color-coding to color the nodes and then count only colorful embeddings that indicate where overlap has occurred. This leads to the development of a program to count the number of colorful embeddings. One can write to detect node neighbors and identify relationships based on building a dynamic analysis program. Tables for these dynamic programs become very large.

Mapreduce/Hadoop is an open source for mapping relationships that is used. This gives ways to do text analysis. Google uses this. This can scale down to parts but then also to trees. One of the first things is to take a tree and partition it into sub-trees. This can help determine ways to map nodes and one can show that in a dynamic program.

Work so far demonstrates different types of graphs (in an experimental setting). One can search from different types of templates to identify local clusters. Generally, graph development takes about 30 minutes for the larger graphs and this is the first timetable to produce computations for graphs this size. The process creates huge-sized intermediate files.

The second part of Dr. Vullikanti's work examines dynamics and control. Within a graph, there are nodes with different states. Depending on the specification of the node and orders, you can get different systems. These are relevant for different applications. Examining different models can help one understand different dynamic properties. Interventions can be undertaken to control dynamics and deal with specific applications such as malware. Dr. Vullikanti's research is analyzing dynamical properties, exposing efficient algorithms and scalable simulation tools for computing dynamical properties, and seeking to control and optimize dynamics.

The third part of this research activity examines efficient simulation tools for computing. A focus is on malware spread in communication networks and efficient computational tools. In malware, specifically, Dr. Vullikanti is looking at two dimensions such as time and space and at how different worms can spread and develop, and conducting modeling and simulation of malware and other worms. In this case, space is not an issue but there are other dimensions such as the study of human diseases that have a much smaller timescale.

Dr. Vullikanti is investigating many types of approaches but a motivating question, particular to the malware example, is how to use an approach that captures worm characteristics reasonably well but scales to very large graphs. His work is leading to an approach entitled "Epicure" that has real features that can scale to large networks. It looks at a stochastic model to simplify the details of the malware problem and comes up with a model that is based on a probabilistic timed transition system. The sample results show that mobility matters and has a significant impact on Malware movement. It shows that centralized dissemination is more effective than local.

To summarize, graphic dynamical systems help capture a rich framework of problems in diffusion. This has lead to challenging algorithmic problems and computational challenges. Dr. Vullikanti's work demonstrated that the graphs resulting from these scenarios are very large, labeled and dynamic, and are very irregular and hard to partition to parallel algorithms and are difficult to compute.

Discussion

Dr. Giles brought up weighted graphs and wondered how that can change the picture.
Dr. Vullikanti responded that weight is an important aspect in examples such as malware spread. Long or short contacts can have different impacts.

Dr. Tang notes that in the physical sciences, there are complex diffusive processes and there are questions to be solved. He proposed that for validation, that would be a nice test-bed for some of the complex algorithms that Dr. Vullikanti presented. He responded that some physicists have found that these work in some settings but it may not be obvious what would work. Dr. Tang clarified that in the context of HPC and convincing a wide user-base, one could take a complex problem in disparate science and just test it. Dr. Vullikanti agreed that one specific way to do this is in percolation although there may be concerns about threshold phenomena. He agreed that this could be explored further and because these are slightly different mathematically, it could be challenging to adapt these techniques to those processes.

Overall public comment

Jack Wells of ORNL commented on the presentation by Dr. Roche. In the presentation, there was a concern expressed that the exercise would improve commercial codes for others.
Mr. Wells does not know the details of the recommendations offered to the ASCR, but pointed out that one reference was made to an example of commercial code and that was the whole content as it pertains to commercial code.

Adjourn

Dr. Giles adjourned the ASCAC meeting at 11:54 a.m. EST.