ADVANCED SCIENTIFIC COMPUTING ADVISORY COMMITTEE
OFFICE OF SCIENCE

American Geophysical Union – 2000 Florida Avenue, NW

Washington, DC
Tuesday, March 22, 2011 – 9:00 am to 5:15 pm

Agenda Tuesday, March 22, 2011

Time	Topic	Page	Speaker
9:00	OPENING REMARKS FROM THE COMMITTEE CHAIR		Roscoe Giles, ASCAC
9:05	VIEW FROM WASHINGTON		William Brinkman, Director of the Office of Science
9:50	NEW CHARGES		Dan Hitchcock, ASCR
10:05	BREAK		
10:20	VIEW FROM GERMANTOWN		Daniel Hitchcock, ASCR
11:10	THE ROLE OF HPC IN THE OFFICE OF ENERGY EFFICIENCY AND RENEWABLE ENERGY		Henry Kelly, EERE
12:00	COMMITTEE LUNCH		
1:45	GETTING READY FOR HYBRID, MULTI-CORE COMPUTING		Richard Murphy, Sandia National Laboratories
2:30	MAGELLAN – AN EXPERIMENT IN CLOUND COMPUTING FOR SCIENCE		Susan Coghlan, Argonne National Laboratory
3:30	BREAK		
3:45	INCITE PROCESS AND AWARDS		Julia White, Oak Ridge National Laboratory
4:30	SUMMARY OF NATIONAL ACADEMY PANEL ON HPC		Kathy Yelick
5:00	PUBLIC COMMENT		
5:15	ADJOURN FOR THE DAY		

ADVANCED SCIENTIFIC COMPUTING ADVISORY COMMITTEE
2010-2011

Agenda Wednesday, March 23, 2011

Time	Topic	Page	Speaker
8:30	UPDATE ON EXASCALE RESEARCH		Karen Pao, ASCR
9:15	IPv6		Kevin Oberman, ESnet
10:00	BREAK		
10:15	UPDATE ON NETWORKING RESEARCH		Rich Carlson, ASCR
10:45	UPDATE ON NETWORKING COV		ASCAC
11:00	UPDATE ON ASCR RECOVERY ACT PROJECTS		Walt Polansky, ASCR
11:45	COMMITTEE DISCUSSION		
12:00	ADJOURN MEETING		

TUESDAY, MARCH 22, 2011

Dr. Roscoe C. Giles, Chairman, was presiding.

WELCOME

Dr. Roscoe Giles welcomed committee members including Dr. John Negele (by telephone) and opened the first of two days of the Advanced Scientific Computing Advisory Committee (ASCAC). He indicated that his goals for the meeting included staffing and organizing sub-committees into groups to undertake the three charges that would soon be in place. He noted that ASCR (Advanced Science Computing Research) reports would occur the following day and this was due to scheduling issues, rather than a matter of priority. Finally he took the opportunity to welcome Dr. William F. Brinkman to the committee meeting.

ROLL CALL

CHAIR - Dr. Roscoe C. Giles
Dr. John Negele (by phone)
Dr. Jack J. Dongarra
Dr. Susan L. Graham
Dr. William M. Tang
Dr. Thomas A. Manteuffel
Dr. Jackie Chen
Dr. Vivek Sarkar
Dr. James J. Hack
Ms. Victoria White

VIEW FROM WASHINGTON

Dr. William F. Brinkman, *Director of the Office of Science*

- Discussed the FY 2012 budget and the delays for the FY 2011 budget year.

- Noted that the FY 2012 budget would be a positive budget from their point of view considering all the Exascale goals that they had.

- Stated that the total swing in their budget was potentially $1.5 billion so there was that unknown factor and that it was a challenging time to manage an organization.

- Discussed the Office of Science:

    - Office of Science had close to 100 Nobel Laureates, 22 in the last decade alone.

    - Has 45 percent of the physical sciences research in the federal government with a higher percentage in some fields. Noted that with particle physics it was approximately 80 percent.

    - It supported 27,000 people.

    - It had a large number of user facilities with 26,000 users.

- Discussed the February State of the Union address by President Barack Obama in which he said that it was our generation's 'sputnik' moment. Suggested that the world was going through a change and would be a different place from a competiveness and innovation point of view and that the United States needed to respond to that challenge in order to stay ahead from an economic point of view. Noted that the President said that the government would invest in biomedical research, information technology and clean energy technology. Stated that for the Department of Energy, clean energy technology was of great importance and the area to focus on.

- Emphasized that The Office of Science was in a good position to respond to the President's demand having an arsenal of basic research, major scientific user facilities, national laboratories and university researchers.

- Stated that The Office of Science had critical technologies that addressed national needs including the Bio-Energy Research Centers, the Energy Frontier Research Centers (EFRC) and the New Energy Innovation Hub, the newest being the Joint Center for Artificial Photosynthesis.

- Described how the issue of clean energy would be addressed from the point of view of science. Noted that there were major changes occurring in two important areas that addressed clean energy:

Materials by Design:

- Noted that computing capabilities enabled changes being made to materials and materials could be designed in a different way, by simulating them for example.

- Described an example of a material science issue completed with a numerical simulation which simulated crack propagation in pure nickel.

Bio-Systems by Design:

- Noted they could generate as many DNA sequences as they wanted to and they had huge databases of DNA sequences which they as yet had to determine how to use most effectively. Discussing the databases he noted that the DOE (Department of Energy) and the NIH (National Institutes of Health) had created a database of approximately 40,000 proteins over the last 10 to 15 years. He noted the issue in biology today was how the proteins functioned and how the database information defined the function.

- Gave a second example concerning systems biology with a microbe called *Cyanothece* with an interesting property which was described in additional detail.

- Noted that the department was focused on plants and microbes, not human biology which was the focus of NIH.

- Noted that they were interested in microbes as they played two important roles in the creation of bio-fuels. The other role was in the climate world and the issue of how microbes absorbed CO2 in soil was seen as an important part of trying to model climate. Microbes and plants played an important role in the same way.

High Power Computing, Modeling and Simulation:

- Discussed two computers, the Cray at Oak Ridge Institute for Science and Education (ORISE) and the BlueGene at the Argonne Leadership Computing Facility (ALCF).

- Advised that they would be upgraded to 10 to 20 petaflops by the end of 2012 resulting in additional capability.

- Noted how the Office of Science was responding to the goal for clean energy. Increases were allotted to Advanced Science and Computing with an increase of 21 percent; Basic Energy Sciences with an increase of 24 percent; and Biological and Environmental Research with an increase of 22 percent. It was noted that the other three categories were kept constant except for nuclear physics which received a small increase as new facilities were being built.

- Stated that they were closing down the Holyfield Accelerator at Oak Ridge so there was available funding for the FRIB (Facility for Rare Isotope Beams) located at Michigan State University.

- Described a new hub created called Fuels from Sunlight Energy Innovation Hub. The process was that instead of using the biological approach to photosynthesis that a chemical catalytic approach to biosynthesis would generate a process that absorbed the sun and created fuels at a greater efficiency of ten times that of natural biological efficiencies. Noted that it was very much a research hub as opposed to applied studies.

- Described a second hub called Batteries and Energy Storage Energy Innovation Hub. Noted that this hub would be an across-the-board hub reaching out to the energy efficiency people with large battery programs. This would involve a mix of people and would try to determine the important battery issues and major questions.

Advanced Scientific Computing Research:

- Noted some challenges and highlights:
  - Uncertainty quantification for high-speed computers.
  - Co-design centers.
  - Excellent progress made with next generation scientific applications, noted as a real strength.
  - Challenged by hardware issues involved with Exascale.
  - Exascale computing lead to many challenges in terms of power and error rate considering the fact that a billion micro-processors would be in operation.
  - Noted the importance of the getting the two 10 petaflop machines in.
  - Discussed the high performance computing SmartTruck/DOE Partnership regarding an aerodynamic study regarding trucks. It concerned a numerical simulation that was done at Oak Ridge National Laboratory (ORNL) and helped to make an advance by reducing resistance to airflow around long-haul trucks. The simulation showed that modifications which added boxes underneath the trailer could decrease the resistance by 12 percent and increase gas mileage by 6 percent resulting in a saving of 1.5 billion gallons of diesel fuel.

Basic Energy Sciences:

- Noted their strong interest in hubs.
- Stated that they were trying to enhance their computational materials in chemistry work and design and nanoscience research.
- Noted that one of the exciting things that they had done during the last year or two had been the free electron laser at SLAC (Stanford Linear Accelerator Center). Discussed in detail the new process using x-ray protein nano-crystallography. Noted the importance that not only had the free electron laser been built but also that the technique worked. Noted that it was a big boon to the community and would make a huge difference.
- Discussed progress made at Argonne National Laboratory (ANL) when in 2000 work was begun on manganese-based materials for cathodes. Noted that over a ten-year period it had gone from

basic research on manganese lithium, manganese oxide compounds all the way to the volt and Argonne licensed General Motors and LG to make the batteries. Noted it was a great example of a start with basic research.

- Discussed the Energy Frontier Research Centers. Involved taking an electron-microscope and built and tiny battery inside the microscope. Described how the battery inside it worked.

- Discussed work done at EFRC research center at Los Alamos National Laboratory (LANL) in which they tried to find materials that could take an enormous number of displacements per atom in a highly irradiated environment like a fusion machine or a high-temperature reactor for example. Researchers had been simulating use of finely-grained materials for this purpose and it was elaborated on further.

Biological and Environmental Research:

- Noted that though it had become a thorny issue, research into climate was still essential. Referenced a comprehensive study of the Arctic, and noted the significance of disastrous potentialities including release of methane. Stated that climate simulation work was ongoing in a positive manner.

- Noted that there were three biofuel centers at Lawrence Berkeley National Laboratory (LBNL) at Oak Ridge National Laboratory and one at the University of Wisconsin-Madison.

    - Discussed the research at ORNL where they have modified switchgrass genes, reducing the amount of lignin by 30 percent. The result meant that it was easier to extract the cellulose. The issue with lignin is that is a tightly-bound molecule on to the cellulose and holds the cellulose and makes it hard to process the cellulose which is then converted to sugars. Noted that it increased by 30 to 40 percent the amount of sugar extracted from switchgrass.

    - Noted that they were able to define a process whereby they got isobutanol out of the switchgrass.

    - Noted that JBEI (Joint BioEnergy Institute) in California built a microbe scaffold in which they processed sugars and cellulose and obtained diesel directly. Noted that they were successful at completing the process without damage to the microbes.

    - Noted that the Great Lakes Bioenergy Research Center (GLBRC) was looking at marginal crops. Could land that was only good for marginal crops be used to cultivate biofuel crops? Mentioned that in three years that they have existed they had 66 patent applications. It was a large community but on the industrial side it was totally dominated by ethanol and the challenge was to initiate change to break from that.

- Discussed two major uncertainties in all climate modeling; aerosols and clouds. Asked if they could do a better job at characterizing clouds and noted what the researchers who ran the

atmospheric radiation measurement systems were working on. Noted they were not able to make three-dimensional projections of clouds so a three-dimensional image of water vapor density could be obtained to understand better what is going on in cloud dynamics.

Fusion Energy Sciences:

- Discussed ITER (International Thermonuclear Experimental Reactor) and noted that although there were initial problems with the change of management over the last year and a half that had changed the situation dramatically. Stated that they believed the project would go forward. Discussed and showed the construction site and project and noted that it represented enormous challenges in science and in the construction of it.
    - Noted that the United States was responsible for the central solenoid which was made out of niobium tin. Noted that the magnet made out of niobium tin was more than twice as big than had ever been made. Discussed the properties and therefore difficulties of working with niobium tin.
    - Discussed the machine and the challenges from a scientific point of view. Discussed the possible problems associated with the tokomaks, plasma and disruptions considering the amount of 500 megajoules of kinetic energy in the plasma. Stated that would be equivalent to 250 sticks of dynamite.
    - Noted that one thing that has been solved was edge-localized modes. Explained that as the plasma evolved the edge of the plasma would get sharper and sharper and become unstable. Noted that several years ago DIII-D out of General Atomics showed that you could put a coil around the tokomak at the right distances and it stopped these instabilities. Noted that it was controversial for several years but now it had been observed in Europe in other tokomaks and they were now confident about it.

Nuclear Physics:

- Noted it was an interesting field and was divided up into three sub-fields. Stated that it had people interested in quark-gluon plasmas based at Brookhaven National Laboratory (BNL), people interested in the quark structure of matter and protons and neutrons, and finally people interested in conventional nuclear physics, rare isotopes and cross- sections.
- Noted there were three areas of reference represented by RHIC (Relativistic Heavy Ion Collider), ATLAS (Argonne Tandem Linac Accelerator System) and CEBAF (Continuous Electron Beam Accelerator). Noted that CEBAF was at the Thomas Jefferson National Accelerator Facility (TJNAF). Stated that the building FRIB (Facility for Rare Isotope Beams) was not yet built. Noted that they have proposed closing the Holyfield reactor.
- Noted the discovery of element II7 by several laboratories, including Oakridge, Dublin, and Russia.

- Discussed the fact that several years ago Congress dumped the isotope program on the nuclear physics group rather than a nuclear energy organization. Noted that they had run a series of workshops and tried to understand the isotope needs of the medical community and the research community. Stated that they had done a great job. Related several occasions when there were shortages of isotopes. Noted several problems one of them being Helium 3 which has been in short supply. Noted that they have taken steps to divide the use of Helium 3 among certain parties. Said that it depended on figuring out how to get the tritium from the heavy water reactors in Canada down into the United States.

- Expanded on hearings held concerning shortages related to isotopes. Mentioned Molybdenum 99 as well as the Helium 3. Noted that now the United States used 60 percent of the world's Molybdenum 99 but produced none of it. Stated that it was a big medical business and ought to be produced by private enterprise. Stated that this issue was being given attention.

High Energy Physics:

- Noted that there were some exciting things going on in two different areas.

- Noted that several years ago they launched a satellite called Fermi and it was a big gamma ray detector which was very directional. Observations had been made in last couple of years. Discussed pulsars and nova explosions which emitted gamma rays. Stated that it was not thought that nova explosions were that powerful. Noted that the latest thing was the observation of a charged-particle plume extending out from the top and bottom of the Milky Way that was 50,000 light years long.

- Noted that the other was the LHC (Large Hadron Collider) which was coming along and was starting to define what was happening in various parts of the energy spectrum. Based on the fact that this was running well a decision was made that they would not continue to run the Tevatron past 2011. Discussed why they had taken this position and the differences in the two types of machines.

- Advised that they wanted to do a long baseline neutrino experiment. He outlined some details about what research had found over the past few decades since early experiments in the 1970s and what research still needed to be done because of the large body of unknowns.

- Reviewed the various STEM (Science, Technology, Engineering, Mathematics, Education) programs that were important.

    - Started a fellowship program in 2010 but had not received authorization to do it in 2011. Hoped to restart it in 2012.

    - Noted they educated a huge number of graduate students through their research grants but believed it was important to have general fellowships in the scientific community because they were attractive to students contemplating graduate school.

  - Described the Science Bowl noting the presence of the First Lady Michelle Obama and Mr. Steven Chu presenting the awards.

- Discussed the America COMPETES Reauthorization Act of 2010 (America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science). Mentioned the open literature question and stated that they had to decide what it was going to do on the question. Asked the committee to decide if the DOE should pay the publication charges or exactly what should they do as far as the open literature issue was concerned.

COMMITTEE DISCUSSION

Dr. Giles noted that for the FY12 budget there was a large percentage increase in workforce, although the numbers were small. Dr. Brinkman responded that indeed the numbers were small and they were trying to get the graduate fellowship program back.

Dr. Giles also asked about certain areas of science receiving strong public interest or scrutiny like the climate issue and whether it affected adversely the number of people that would enthusiastically enter the field of science. Dr. Brinkman responded it was a major concern regarding education. He noted that he had several concerns. He said that they had been living off foreign graduate students and it was standing at about 50 percent in almost all the areas of science. He noted that the majority were coming from India and China. He explained the reason for the concern was that historically in the universities the top third of the class used to come to the United States but were now heading not to graduate school but directly into business opportunities in their own countries. This meant that the United States was not getting the best students and he felt the situation was probably similar in China. As a result of the situation Dr. Brinkman felt that they would need to turn to American students again and therefore felt that the graduate fellowship programs were of great importance and should be in place.

Dr. Thomas Manteuffel asked about the cut to Fusion Energy Sciences (FES) and remarked he had heard the fusion simulation project was on hold. Dr. Brinkman responded that one of the issues was selling the fusion program which was challenging and there was a large financial drain with ITER. He noted that the problem was finding enough money to keep the program going.

Dr. William M. Tang asked about the education of young scientists. He stated that in China they were very aggressive in looking forward and one of things they did well was providing attractive long term career tracks for people, and sustaining longer term programs. They were also recruiting them back from the United States. He said that academic tracks in 2008 had suffered given the downturn in the economy affecting the private universities. Dr. Brinkman responded that he also considered it very worrisome. He also encouraged committee members that if they had any contacts on the Hill that they should encourage them to fund science and technology in a more positive way.

Dr. Vivek Sarkar echoed comments about India, noting that undergraduates could get a job with a global company with a salary twice what their parent could make on retirement. He also said that the funding was plentiful in the universities, that faculty members were not constrained by funding only by research

staff but they were looking for top PhD talent. He added the IITs (Indian Institutes of Technology) are multiplying so they needed faculty. Dr. Brinkman agreed saying it was a worry and that these people would be competitive. He also said that American faculty in some universities were taking reduced salaries. For example he said California was paying faculty salaries only and all other university costs were from federal grants.

Dr. John Negele (by telephone) commented regarding the graduate student issue. He noted that the applications from foreign and domestic students applying to MIT (Massachusetts Institute of Technology) and other top schools showed a worrisome difference with the background and level of achievement of domestic and foreign applicants. In his opinion that problem was increasing and indicated a need to redouble efforts in public education in science and technology in order to have a strong workforce and good graduate students coming from the U.S.

NEW CHARGES

Dr. Daniel A. Hitchcock, *Acting Associate Director*, *Immediate Office of the Associate Director*

First Charge:

- Would discuss COMPETES first and displayed the charge. Noted that all the advisory committees were getting the same charge. Noted that the ASCR charge was for the research communities primarily supported by ASCR which is mostly mathematics, computer science and networking. Computational science and in the other disciplines were mostly covered by other advisory committees.
- Facilities are also involved because the facilities have data policies that have to be involved.
- It is on a short time scale because they need results back as a result of Congressional deadlines. This is a chance to inform what happens in the policies for the long term. It is preferable to act sensibly rather than having random things happen that might prove difficult to live with.
- Some things are already in place in ASCR for software but the publications are different and how publications are replicated is an issue that needs thought.
- The report has to be delivered back before the next meeting which means it will be necessary to have an open teleconference meeting for comments as they have had in the past.
- There is great importance on getting this process right as it will affect the community for a long time, and they are issues that have implications in law.
- The charge involves people working with SIAM (Society for Industrial and Applied Mathematics), with IEEE (Institute of Electrical and Electronic Engineers) and ACM (Association for Computing Machinery) and it is important to be aligned with what they are doing so we don't change things that work already in order to move toward public access.

- Want to ensure that the facilities don't inherit a large intellectual property and publication burden.
- A volunteer was encouraged to help with this process and they would provide assistance from the office to help with logistics.

COMMITTEE DISCUSSION

Regarding the first charge Dr. Giles asked what the downstream scenario was beyond this report. He noted the individual offices make their reports. They are then synthesized in the Office of Science. Dr. Brinkman responded that they would be synthesized. He noted that the COMPETES Act mandated that the OSTP (Office of Science & Technology Policy) form an inter-agency group that would monitor and guide the government on the issue. Dr. Brinkman noted that they wanted to clarify their position so when they were in a position to work with the inter-agency group they would know exactly what they wanted.

Dr. Giles asked about the fact that some of the artifacts they publish are software, and whether that also went across other offices. Dr. Hitchcock noted that it was primarily an ASCR issue. Dr. Hitchcock emphasized again the importance of determining what they wanted before it got to the level of the science advisor and they meld together things from all research agencies.

Dr. Giles asked about whether they would interact with other advisory committees. Dr. Hitchcock responded that bearing in mind their deadline trying to accomplish a consolidated view with other advisory committees might create problems. He also said that the Office of Science would do the integration and within the Director for Programs based on input from the advisory committees.

Second Charge:

- He noted that the other charge is an advisory committee review on the Computational Science Graduate Fellowship Program running since 1991.It received a major review in early 2000s by a committee not part of the advisory committee. He noted that over two decades changes have taken place and it is time for an advisory committee review to:
    - Look at the Computational Science Graduate Fellowship and ask if it is still optimal now; is this the right thing for us to continue to do?
    - What are its successes? What are things that we should maybe think about changing?
    - When set up originally it was explicitly designed for computational science and to not include computer science and applied mathematics specifically. Is this now the right choice going forward or should it be broadened?

- There is a strategy with the practicum that has been effective but it should be looked at now to determine if it is the right thing going forward and whether the administration is done in a way that we think is scalable and builds us the future workforce we need.
- Like a Committee of Visitors (COV) but it is actually a more broad-based program review.
- A lot of people at ASCAC (Advanced Scientific Computing Advisory Committee) are also involved with the fellowship and they need to find someone who is very interested in education issues and the STEM (science, technology, engineering, and mathematics) issues but who is not on the board that selects the fellows.
- Important that this fellowship designed to be as effective at addressing the challenges of the next decade as it was for challenges of the past 20 years.
- Hoping that a Chair can be located who is enthusiastic about the issue and to review this.

Third Charge:

- This is a charge that concerns networking. Gwendolyn Huntoon, the Director at the Pittsburgh Supercomputing Center (PSC) has been networking at the PSC for a long time and has agreed to Chair the COV. An ASCAC member was still need on the COV to report back to ASCAC. So a Chair has been identified and Gwendolyn Huntoon was also a major force in the implementation of the regional optical networks. He noted it was a great opportunity for someone from ASCAC to understand their networking programs.

Regarding the third charge Dr. Giles confirmed that ASCAC as a group would take advantage of the opportunity to serve on the networking committee.

PUBLIC COMMENT

A member of the public made a comment about the first charge concerning peer review and the journals and access. He recommended a model that might be of use in that area that had extreme coverage in mathematics, computer science and it was the archive system. He noted that people updated their final published version on the archive and it had been run since the early 90s. He indicated it might be a useful model for open science.

Dr. Brinkman responded that he used that model and that was one of the issues under discussion whether that was an appropriate model and how would it evolve going forward. He also commented that there was the issue of not wanting to drive publishers out of business. Peer-reviewed, published articles were of great importance to the university communities. For that reason they did not want to give up on reviewed journals. Dr. Brinkman also noted that he was not sure how many different communities used the archive.

BREAK

The advisory committee stopped for a ten minute break.

Dr. Giles noted that many of the presentations and documents were available on the ASCR meeting website.

VIEW FROM GERMANTOWN

Dr. Daniel Hitchcock, *Acting Associate Director, Immediate Office of the Associate Director, ASCR*

- Noted several challenges:
  - Delivering today's petascale science for across the Office of Science.
  - Figuring out what is to be done to get the next factor of 1,000 and developing that in partnership with U.S. industry and all the applied mathematics and computer science that goes along with it.
- Discussed the 2012 target budget:
  - Want to increase work on uncertainty quantification.
  - Establish co-design centers and have them up and running.
  - Investment to U.S. and address critical technology challenges in hardware architecture and critical technologies needed for Exascale.
  - Looking to install a 10 petaflop IBM BlueGene/Q at ALCF and a hybrid multi-core system at ORNL.
- Discussed the ASCR budget overview.
  - Exascale is all over the ASCR program because the challenges of multi-core and CMOS are all over scientific computing. He said that whether you are doing an Exascale system or a 1 petaflop system in the next 10 to 15 years multi-core was coming like a freight train. It is in applied mathematics, in computer science, in the computational partnerships, in the facilities. It is worse for some programs like BES (Basic Energy Sciences) the LCLS (Linac Coherent Light Source) which is the electron laser is BES's first petabyte a year facility. Next generation light source which is on the drawing board has 10,000 times the data rate. He elaborated on that statement.
  - Have a variety of budget numbers and something between that and the 2008 budget number which is $380 million is probably where they will end up and that makes detailed planning challenging.

- He talked about new staff, Dr. William Harrod joined from DARPA (Defense Advanced Research Projects Agency). He ran studies that conditioned what Exascale is going to be and what the constraints of Exascale multi-core are. Before that he was at SGI and Cray Inc. and has broad experience.

- Also mentioned that Dr. Steven Lee had joined ASCR in the SciDAC Division (Scientific Discovery through Advanced Computing). He is involved in helping with the re-competition of the SciDAC institute.

- Also hired Dr. Bennett who came from NIST. She is a nano-scientist and improves interaction with the material science world and is also involved in SciDAC.

- Discussed phase change in computing. He explained there was a concurrency shock coming because of multi-core. Explained that there were investments on both sides of the shot, investments like SciDAC involved in delivering petascale on today's machines. On the other hand there is co-design, X-Stack, advanced architectures on the other side. Noted that SciDAC was a decade old and changing but there was still a dichotomy as science has to be continued but prepare for the future at the same time.

- Noted that they were re-competing SciDAC. In the past it had a dozen or so centers and institutes that no one could differentiate between clearly and a wide variety of partnerships with the applications. Noted that this made it difficult for scientists to know where to go if they needed a problem solved. Stated that they want to consolidate them into a smaller number of centers that were more focused on delivering the software needed by scientists. Advised that they were also in discussion with all other programs about what the partnerships for the next round of SciDAC would look like and the general discussion had been that they would do things as opposed to just partnerships in broad areas and concentrate on focused partnerships in areas with a combination of our and their expertise.

- Noted that the DOE funding opportunity concerning SciDAC was out now and the applications are due May 2nd and letters of intent are not required but strongly encouraged by March 30th.

- Discussed massive challenges to Exascale. Listed some of the main points such as system power, memory, concurrency, processor issues, programming model, algorithms, I/O (input/output) bandwidth, reliability and resiliency and bisection bandwidth and discussed in more detail. Noted that if something is not done the National Academy will make the same report and this impacts everything. Noted this will impacted the computer-intensive area a lot but impacted the data-intensive computing more substantially because it relied on data movement.

- Discussed the ASCR Exascale research kick-off meeting and stated that they had started basic research in certain areas. Had a kick-off meeting and brought together all of the PIs (Principal Investigators) from the solicitations and advanced architectures and the co-design centers that were currently under consideration to discuss what the interactions between the projects were.

Explained that before you started things that came in on proposals you had to first decide how to make that into a program. Noted that a meeting was held March 7th to 11th in San Diego with people from NNSA (National Nuclear Security Administration). Described in more detail the processes of the sessions. Concluded that they had data about the interactions between the projects and the gaps in the portfolio and they were now in the process of analyzing that.

- Commented on the NNSA Exascale meeting that was ongoing in San Francisco which was a parallel one but from their point of view with the weapons program. Stated that they were trying to have as much collaboration as possible to ensure that it was one strategy going forward for critical issues.

- Noted there was a workshop on Terabit networks. Topics covered included identifying the challenges, discussing participants and the major technical areas of discussion. One point raised was that it had a lot of interesting features one of which is that once you have nailed down the optical end points so they are not spoof-able then you have a secure channel.

- Reviewed ASCR interactions with applied programs because computational engineering is important for them and gave several examples of this. Discussed how important it was to understand how the electric grid worked and to see the stability limits on the control system were. Gave an example of the minute and a half that the control people during the Northeast power blackout had to avoid cascading failure.

- Discussed the SciDAC 2011 conference at the Brown Palace Hotel in Denver, Colorado, July 10th to 14th. Dr. Patricia Dehmer, who was the acting associate director for ASCR was going to be the keynote speaker.

- Discussed INCITE (Innovative & Novel Computational Impact on Theory and Experiment). Noted that Dr. Julia White would discuss INCITE in more detail.

- Noted that Dr. Juan Meza was named one of the top 200 influential Hispanics in technology this year. Noted that Dr. John Birge was recognized who is in the School of Business.

- Discussed the work of Dr. Patrick Chiang, an early career PI in 2010 concerning sustainable silicon: energy efficient VLSI interconnects. Described the technical details of the graphics on the slide.

- Discussed a slide concerning applied mathematics and the topology for statistical modeling of petascale data. It concerned a way of looking at the features in a simulation to tell what was going on so everything does not have to be visualized. Noted that it was a topological method not an averaging method so it was more sensitive to features.

- Discussed another slide concerning applied mathematics and optimal derivatives of noisy numerical simulations. Reviewed the technical graphs and concepts in detail.

- Discussed work done by Barton P. Miller concerning MRNet (Multicast Reduction Network). Stated that the slide concerned an infrastructure of highly scalable and reliable tools and middleware. This was a way of hooking together these large systems so faster transitive information is obtained. Noted as important if data is being collected from large systems. Discussed more details about the system and how it is currently being used.

- Discussed the work of Dr. J. Chen concerning feature tracking visualization work in situ. Detailed the technical details of the work.

- Discussed APDEC, a SciDAC center and the work of Dr. Phillip Colella. Noted that it was the tool that was used to model water resources for California for the San Francisco Bay and Delta. Gave an example saying in the Bay if they drew too much water out then the salt water flowed back into the water table and once it was in there it did not come back out again.

- Discussed stress corrosion cracking and particularly discussed nickel. Noted that some people said 3 percent is a low estimate of what corrosion costs the U.S. every year and so understanding corrosion was important not only for nuclear issues but also bridges that might be under stress.

- Discussed earthquake simulations being done in the Southern California. Used to study effects for seismologists and also with questions about how medical supplies could be transported into the area if roads were damaged.

- Reviewed the staff structure of ASCR and contact information.

COMMITTEE DISCUSSION

Dr. Tang asked about the Exascale co-design activities and asked what the timeline was. Dr. Hitchcock responded that they anticipated having the reviews in May to decide which one was the final portfolio. He noted also that there were widely varying budget scenarios dependent on what the budget ended up being for 2012. He concluded by saying that the reviews would be done in May and hopefully budget information would be available by then.

Dr. Giles noted they were seeing the impact of the large scale simulations and he wondered what Dr. Hitchcock's feeling was now about the state of uncertainty quantification across those. Dr. Hitchcock responded that uncertainty quantification was in its early stages. He gave several examples to illustrate that.

Dr. Jackie Chen asked him to comment about V&V (Validation and Verification) for some of the Office of Science Exascale co-design activities. Dr. Hitchcock responded that the ideas of UQ (Uncertainty Quantification) and V&V are critically important. He noted that in terms of the traditional software validation that would be done for a real-time system he noted they did not have much investment in

this area at the moment. He noted that they had done a lot of thinking about how in an uncertainty quantification world how you would actually link up with experimental devices and so some of our partners have explored doing models of the diagnostics and he went on to elaborate further.

Dr. Chen asked whether with computers getting larger and being able to capture more realism that if there was a greater intersection between experiments and computation, and therefore the validation part should be considered carefully. Dr. Hitchcock said absolutely.

Dr. Tang mentioned it was a good comment about the synthetic diagnostics implementation and this was a cost-effective way to proceed. He noted another point that with regard to the topic about more interactions between the modeling activities and experiments he felt that that impacted Dr. Giles' question about uncertainty quantification in that if you can match variations in the experiments versus the models' sensitivity tests that would gain you a lot of mileage. Dr. Hitchcock agreed it was important and also it was important to use the models when the experiments go wrong. He gave an example and discussed it further.

Dr. Jack Dongarra asked about the budget slide. He noted that the budget had $124 million for Exascale. Dr. Hitchcock agreed this was correct but that it was between ASCR and NNSA and some of the things that ASCR started already last year like the efforts in X-Stack and advanced architecture were included in the $90 million that was in the ASCR budget. Dr. Dongarra asked if there was anything in the base program which was now being reclassified as Exascale. Dr. Hitchcock said that some of the uncertainty quantification work, the increases in that were clearly Exascale and he elaborated on that further.

Dr. Giles noted that public comment would be done at the end of the section. He welcomed Dr. Henry Kelly from EERE (Energy Efficiency & Renewable Energy).

THE ROLE OF HPC IN THE OFFICE OF ENERGY EFFICIENCY AND RENEWABLE ENERGY

Dr. Henry Kelly, *US Department of Energy*

- Starting by commenting that EERE were an applied research program and noted there were a lot of intriguing intersections between information technology and what EERE does.
- Noted that he intended to focus on the design and operation of equipment and the operation of complex systems which might involve sensors and other activities. Stated that he would be going through their programs one by one and discussing the connections.

Computation in Buildings:

- Buildings consume 72 percent of all electricity so this area is a core part of the entire U.S. electricity sector. Discussed the difficulties in modeling buildings with many different factors to consider such complex activities, and air flows.

- Working on tools to design buildings and tools to operate them once they are in place. Interested in systems that learn how a building operates and adjust timing of heating and cooling.

- They have found that modeling has not connected itself to real-world data. New models have been validated by comparing the results of other models and they are working to correct this.

- What they want the operational software to do is detect problems when they exist, come up with a hypothesis of what went wrong and then call for help. These small problems can translate into a very large energy problem if not addressed.

- They have been looking at how building systems operate together and the main thing they have been doing is to set up a new hub started in Philadelphia being run by Pennsylvania State University and they have the whole navy yard. Its focus is going to be buildings as systems, building computational tools. They will be coming up with the variable mesh approach and then validating it against buildings that are in the navy yard.

Computational Manufacturing:

- About one third of U.S. energy goes into manufacturing and most of the goals of the programs have to do with driving down the cost of making products.

- Have to try to find ways to improve efficiency of industrial processes and finding materials that have a lower embedded energy. Noted that there was an intersection with the inter-agency program focused on materials by design.

- Another theme is the design of systems for manufacturing ability. There is a break between people who design complex things and then someone has to make them.

- Importance of simulation tools was discussed as they allow designers to update models and listen to feedback.

Transportation:

- Transportation is another big thrust and they are working together with the Office of Science on advanced combustion simulation. Thought that they could get 30 to 40 percent more efficiency out of internal combustion engines by better design and more exquisite controls during the combustion process which was a very complicated problem involving many factors.

- Moving forward toward electrification and so they have a parallel effort on designing advanced batteries. More details were given about the research and goals regarding batteries.

- Also working out in the field on the flow of vehicles on the roads and they are able to track traffic in real time.

- Next generation of studying traffic was working with predictions regarding congestion.

Computation in Wind:

- On the supply side they do all of the renewable energy technologies and this includes the wind. They are working on computational tools for the design of wind blades and the interaction between blades and towers.

- Simulation tools developed in San Diego have been central to increasing the efficiency and lifetime of blades.

- Offshore wind farms are finding new problems with the surface phenomena of the anchors and other devices. Some problems are currents, surface effects such as waves and towers and wind stresses. They are doing simulation testing right now and will be doing testing in the Netherlands as they have a large tank not available in the U.S.

- They are looking at novel ways of capturing hydrokinetic energy from waves, tides and other things. Noted that it was a long shot at the moment because there are ideas but they do not have good handle on the economics. So far what they have looked at is not promising but again this is an area that they want to study in a simulated environment as field experiments are expensive.

Computation for Biomass and Solar:

- There is a significant program in biomass, a place where the simulation of biological pathways, gene expressions and other phenomena are playing a large role. It is now possible to take an e-coli that eats cellulose and produces small amounts of drop-in fuels like isobutene and the efficiencies are very low. It is materials by design problem but in a biological setting.

- They deal with the operation of the electric utility grid which has become more complex. Going from a countable number of generating devices to one where you would have a very complex network of things attached to the grid and examples were shown of wind data and solar and wind data from California. Issues were discussed in more detail regarding a complex grid system.

Computation for Education and Training:

- Computation tools can be used for education in that they have to train people to operate in new environments, how do you build the buildings and how do you retrofit them, how do you hook up smart meters and simulation-based training lends itself to this.

- They have built an open-source version of the powerful defense-based training tools and they are being used online, some free and some at community colleges. Some examples were given of the areas covered in this type of training. Performance-based testing is used for students in this type of training.

COMMITTEE DISCUSSION

Dr. Susan Graham asked if they were doing any economic modeling of what is important and where the choices might be because Dr. Kelly mentioned "what we can afford to pay for". Dr. Kelly responded yes and said that he found they were drowning in incredibly sophisticated economic modeling that do 2,000 coupled equations and produce mystery results that cannot be explained. He noted that in the last year he had tried to move slightly away from things that are sophisticated macro-economic models and take simple bonded like models that will simulate the electricity grid or find out what we can afford to pay for storage. He also said one of the dilemmas is the input data is unknown within huge ranges. He said his goal was to get the first significant figure correct and then proceed from there. He noted that everything that they were doing was trying to drive down the price and gave examples.

Dr. Graham asked where he expected these projects to be done, in the national labs or in the universities or industry. Dr. Kelly responded they were trying to rebalance what they were doing. In the case of labs trying to define what would be the unique value added of the laboratories such as large machinery, a team of people, focus on a project for a significant period of time is important. He also added they were opening up to competition which would include labs and universities. So in summary he noted that they had an even split between labs, universities and industry.

Dr. Manteuffel asked where the EERE sat in the Office of Science and in relation to NREL (National Renewable Energy Laboratory) and other things. Dr. Kelly responded there were three undersecretaries, one for energy that EERE reports to and under energy they have nuclear fossil, EERE and then waste management. Dr. Manteuffel asked how their community could interact with the EERE and how could they help. Dr. Kelly responded that they were trying to strengthen a lot of the ties. He noted that there were parts of the EERE that had worked closely with the Office of Science for years. He elaborated on some of the areas where they hoped to have more interactions in the future.

Dr. Giles asked about the buildings and mentioned the large multi-scale systems that they were involved in designing. He noted that it was quite an art combining information from multiple scales. So he said with the buildings the systems were actually engineered systems and if they became difficult to simulate in their existing form that they could be changed to an easier form to simulate. He asked if that was also happening in the smart grid and some of the other areas like for the cars and traffic. Dr. Kelly responded that they were beginning to look at it and in vehicles for example what they had done was to try to get to a larger community was to put together an external advisory group and this would include at least one computational scientist. He noted that with traffic the job of the Department of Transportation was safety and congestion and not energy. The EERE had more sophisticated computational tools to bring to that relationship. He discussed more details with ongoing challenges in buildings.

PUBLIC COMMENT

Mr. Douglas Ray, Pacific Northwest National Laboratory

Asked about improving engine efficiency as one of the most cost effective approaches to improving vehicle efficiency and he asked if he could comment on others like materials, improved aerodynamics, decreased rolling resistance, improved after-treatment technologies. Is improving combustion the best way to do that, the most effective or just the one we know how to do now? Dr. Kelly responded that the portfolio of things that they were doing on vehicles was investing in all of the things that Mr. Ray mentioned. He noted that their challenge was to match their research in places where the payoff would be highest. One of the hardest problems was the combustion problem and if a 30 to 40 percent improvement in the fuel economy of the engine was achieved that would be important. He elaborated on this in more detail.

COMMITTEE LUNCH

The advisory committee recessed for one hour for lunch.

GETTING READY FOR HYBRID, MULTI-CORE COMPUTING

Dr. Richard C. Murphy, *Sandia National Laboratories*

- Began with itemizing where they were now and what the future looked like.
    - Looking at a 125-500MW investment in an Exascale computer in the 2018 timeframe.
    - Combined with the application trends which are increasingly toward unstructured science codes and in some sense new data analytics applications which may be even less structured than the science codes.
    - The result is that scaling today is petascale design targets. Noted that it was not a feasible path to get into a reasonable Exascale machine at a reasonable power budget.
- Wanted to talk about over the course of the presentation one example of how they were approaching the problem at Sandia National Laboratories (SNL).
- Noted that he led one of four UHPC (Ubiquitous High Performance Computing) projects and described the technical details.
- Noted that CPUs (Central Processing Unit) and GPUs (Graphics Processing Unit) were likely to converge over the next decade.
- Gave a proper definition of hybrid multicore. Noted that the conventional definition of this, the processor-oriented definition, is the hybrid says I've got some combination of CPUs and GPUs on the same device and multicore is saying I can get lots of these on a chip for free.

- Gave a more data-centric definition of hybrid multicore as putting the functionality where it could achieve the desired result with minimum number of picojoules. Stated that multicore was a trend toward having to rely on parallelism to achieve performance.
- Noted what was required for transition:
    - Support for data movement combined with heterogeneous compute elements.
    - Advances in programming models, dynamic runtime systems, resource management and underlying implementation technology.
- Noted that thinking about the Exascale problem correctly would impact computing technology in all scales.
- Noted that he needed to define a DOE centric research path forward to get to Exascale.
    - Noted there have been pessimistic discussions of what is required by 2015 or 2018 to reach Exascale.
    - The Exascale report projections were optimistic so they probably have a harder problem than originally thought.
    - The application base may be changing so they cannot plan for the application set that they have now but the application set in 2018.
    - Key restraint is one of data movement.
    - Noted that they should be application focused not power budget focused.
- Looked at the historical perspective:
    - HTMT (High Technology Multithreaded) was an early petaflop study in the mid 90s supported by NSF (National Science Foundation), DARPA and NSA (National Security Agency).
    - One of eight sponsored petascale design points generated over a short period of time.
    - Reviewed how they were able to get to petascale a decade later and reviewed the power envelopes.
- Discussed HTMT perspectives and noted that many of the things that they claimed they needed were explored in the early concepts in HTMT, so multithreading, message-driven computation or light-weight active messages and a distributed global shared memory.
- Noted deficiencies in the machines of today:

  - Dynamic adaptive resource management.
  - Smart memory operations
  - Intelligent data movement.
  - High bandwidth low latency interconnects for short messages.

- X-caliber perspective used the above core concepts as a notional design target in thinking about HTMT, focusing everyone as they co design, so they can understand tradeoffs in future.
- Discussed how applications are changing in detail as per chart of spatial versus temporal locality.
- Discussed DOE physics applications in detail as per graphic, comparing percentages of different types of instructions, this instruction mix was different from industry requirements. This showed that the majority of instructions were used to do memory references, and these support instructions could be made more efficient.
- Discussed co-design in detail as per graphic, which displayed how applications can inform the architects and those doing the run time software?
- Reviewed how the X-caliber team thought about co-design. Gave a definition of a computation model also known as an execution model. Said that it described five elements of a machine:
  - Concurrency
  - Coordination
  - Movement of work and data
  - How things are named
  - Introspection or how you look at the system as the program is running.
- Noted that their execution model for UHPC was Parallax. Described each element as it applied to Parallax and discussed in detail.
- Noted that looking at it compared to today's model which he called stylized communicating sequential processors and was harder to define.
- Discussed a comparison between Parallax and Today's Dominant Model. Reviewed each element of the two systems.
- Reviewed system balance for petascale racks and system balance and discussed in more technical numerical detail.

- Noted the DARPA challenged problems from a UHPC perspective. Discussed five problems, graph, stream, decision support, shock physics and molecular dynamics.
- Discussed the enabling technologies including advanced packaging, 3D integration and optics.
- Discussed the node heterogeneous architecture which would look like hybrid CPU or GPU technology surrounded by advance memory cubes.
- Discussed the memory systems in details as per graphic.
- Discussed thread coalescing observation and noted that there was high temporal locality mode of operation and low temporal locality mode of operation. Described in more technical detail.
- Discussing sprinting and noted the observation that one could wire up more capability in one of the systems to processor, memory system and network. Reviewed each area in more detail.
- Discussed one of the side benefits of doing this had been that they are working closely with industry. Stated that it led to a closer engagement between Intel and SNL on this topic. Advised that they would have a number of combined activities that would arise from UHPC that would target at 2016. Reviewed the list.
- Stated the initial and preliminary findings from UHPC.:
    - Look at the execution and programming model change
    - Note the execution model acceptance/adoption challenge.
    - Co-design is the key to efficiency.
    - Fundamental technology advances have to occur to support Exascale.
    - Begin the disruptive research.
- Summarized his recommendations and conclusions:
    - Continue to quantify what happens if we don't do anything. That would change over the life of the program and there should be a clear definition of what the roadmap is.
    - Create an application-driven DOE execution model.
    - Proffer DOE Exascale design points to complement industry design points. Described this area in detail providing a preliminary list.

COMMITTEE DISCUSSION

Dr. Tang commented that given that the presentation emphasized the fact that the new generation of codes moving to the Exascale would have to take into account efficient data movement, and some of these same challenges were faced by LANL with Roadrunner, what were the lessons learned from that in terms of the successful application suite on Roadrunner.

Dr. Murphy responded that Roadrunner was an interesting design point. He noted that if you spoke with the father of Roadrunner, Peter Hofstee he would say the one thing that he wanted to do differently was have a beefier scaler processor. He gave additional technical details on that. Dr. Tang stated that if the successful app suite for Roadrunner the active applications domain focused on co-design these days would come to the conclusion that you would probably have to rewrite code, however the app suite could be used as a test bed, and they were not starting from scratch.

Dr. Sarkar thanked him for providing a definition of the execution model and he said he liked the five components. He asked about one of them, introspection because it seemed like it would be useful for tools. He asked what kind of tools did he think it would enable through introspection and he asked if the application programmer would be aware of introspection at all? Dr. Murphy stated the programmer could be aware of introspection. He noted that the approach they had been taking was more that the run time was aware of introspection. He mentioned that that might be an erroneous bias on their part. He thought his point about tools was dead-on and he elaborated.

MAGELLAN – AN EXPERIMENT IN CLOUD COMPUTING FOR SCIENCE

Susan Coghlan, *Argonne National Laboratory*

- Noted that she would be co-presenting with Mr. Shane Canon from LBNL and he would present the second half of the presentation.
- Advised that they would give a progress report on the Magellan Cloud Testbed Project and described the outline of that progress report.

Overview of the Magellan Project:

- Noted that the Magellan Project was a joint project between ALCF and the National Energy Research Scientific Computing Center (NSERC).

Overview of Cloud Computing:

- Funded by the DOE with ARRA (American Recovery and Reinvestment Act) funds to investigate how clouds might be used to support DOE-HPC computing.
- Noted that the primary mission was to look at private cloud computing for the midrange computing space for DOE/SC.

- Deployed two testbeds one at each site to do the investigation and it included different things like deploying different cloud service models and supporting them and running science applications within them. Looked also at different scientific applications to understand how they might fit within the cloud space.

- Magellan started in September 2009. Noted that both sites were in the process of competitive bid procurements in order to upgrade production facilities they piggybacked on that. Deployed core compute cloud systems rapidly. Had hardware at both sites by winter 2009. Outlined progress noting that the Advanced Network Initiative (ANI) projects will be supported despite the initial project period expiring.

- Gave a definition of a cloud from the National Institute of Standards and Technology (NIST). Described its broad definition of what a cloud is and it included:

    - Resource pooling
    - Broad network access, remotely accessible
    - Measured service for transparency
    - Rapid elasticity, you only pay for what you use
    - On demand self service, consumers can provision it themselves.

- Noted within the definition a number of service models and listed three as:

    - Infrastructure as a service
    - Platform as a service – the provider gives you the tools and programming language to write your own applications.
    - Software as a service – the provider give you the application, you provide data.

Overview of Magellan Distributed Testbed:

- Computing - Procurements were won by the preexisting hardware; IBM iDataPlex
- Storage was a mix of disk storage, archival and two classes of flash storage.
- Architected for flexibility and to support research. Similar to high end hardware in HPC clusters. Suitable for scientific applications and included specialized hardware such as GPUs.

Early Findings – Based on Progress to Date:

- Noted that these were the lines of enquiry.

    - When is it cost effective to run DOE HPC science in a cloud?

  - What are the ramifications for data intensive computing? Not far enough in the work to date to present those findings.
  - Other questions would be discussed in detail during the presentation.

- Discussed cloud software stacks and the question considered was, are the open source cloud software stacks ready for DOE HPC science. The characteristics of the existing stacks were discussed:
  - Mature and stable
  - Scalable
  - Depth and breadth in tool availability
  - Integrated I/O
  - High performance
- Evaluated five open source cloud software stacks. All but one was deployed on Magellan itself. Noted the evaluations were done by Magellan team people plus some special users.
- Discussed the evaluation criteria which included:
  - Feature set
  - Stability – most were unstable
  - Infrastructure Scalability
  - Usability – impact on the scientific users
  - Manageability
  - Sustainability – big enough community behind it
  - Did not do a formal performance evaluation except where there were problems
- Reviewed the evaluation results. Discussed the communities and noted there were two Eucalyptus ones. Started out with Eucalyptus because it had the biggest community. It was also unstable and a lot of time was spent on it. She compared the different communities and improvements. Discussed the Open Stack community in detail.
- Discussed early findings and next steps in cloud software stacks. Noted it was moving rapidly, but was not production ready yet. Noted that networking was complicated and challenging to get right. Discussed the next steps and advised that they would be spending some time on their

public cloud that they had at ANL (Argonne National Laboratory) trying to scale it up and implement more distributed. Discussed scalability, performance and features in detail.

- Discussed the question can the DOE cyber security rules be met within the cloud infrastructure. Noted that current cyber security was developed for onsite HPC cluster installations. Reviewed some of the issues regarding security in more detail.

- Reviewed the IaaS (Infrastructure as a Service) cyber security overview and discussed machine definition and management, system instance configuration and management and network authorization and management.

- Discussed early findings and next steps regarding cyber security. Discussed trust issues and network separations were complicated. Noted that fundamental threats were the same but have to put in different controls to manage issues. Discussed next steps for Magellan:

    - Can they try and use the hypervisors to do some new roles in monitoring and auditing and tracking.

    - The type of forensic analysis could be done on virtual m*achines.*

Mr. Shane Canon, Lawrence **Berkeley National Laboratory**

- Noted that Susan Coghlan had concentrated her presentation on the provider side but the next questions are going to be more application-focused.

- Noted one of the questions they were interested in exploring with Magellan is in conjunction with these cloud models was also new programming methods. Noted this was not specifically cloud computing but associated with it. Noted some examples as being Hadoop and Azure. Discussed some technical background.

- Discussed programming models and Hadoop and Pig for bioinformatics. Reviewed and discussed bioinformatics using MapReduce. Discussed researchers and 12 recent developments, construction of an end-to-end pipeline and complex operations that generated parallel executions.

- Discussed some evaluation work that they did on Hadoop primarily benchmarking activities. Noted that they also looked at supporting specific applications to the framework. Discussed many of the technical details.

- Discussed early findings and next steps for programming models. Noted they were useful tools for data-intensive applications because they can hide some complexities. Noted that going forward they would continue to look at some of the scaling aspects of Hadoop and also interesting in looking at some underlying file systems. Stated that they were also looking other applications that could take advantage of all the capabilities and features of the programs.

- Another area of interest was looking at the performance of applications running inside the environments. Stated that some questions included how well would the parallel applications run within the virtualized environments? Also noted another question such as the importance of the interconnects? Noted that they were also interested in finding out if there were some applications better than others.

- Reviewed several performance plots on several graphs in detail.

- Reviewed application performance and application scaling graphs in detail.

- Discussed application performance and early findings and next steps. Started with early findings. Illustrated that the non HPC tuned offerings gave poor performance at modest scales and illustrated the importance of having a good interconnect on the systems. Discussed next steps started to look at the price performance looked like on the offerings compared with the traditional HPC centers. Wanted more statistics and to run more applications. Wanted to gather data from various mid-range clusters both institutional clusters as well as applications run at NSERC.

- Discussed user experience. Wanted to know how difficult it was to port applications to run with cloud environments and how should users manage their data and workflow.

- Looked at user experience and user community. Noted that they had a range of applications running on Magellan. Focused on three examples, MG-RAST, STAR and the Joint Genome Institute:

    - MG-RAST: Deep Soil Analysis. Described the background that they took samples from an experimental facility in the United Kingdom and they were looking at a microbial community around the root structures of plants and how that evolved. There was a need to analyze the data from the samples.

    - STAR: Interested in using Magellan to do real-time analysis of data coming off the detector. Using about 20 instances that they are continually running. Described in more detail.

    - Joint Genome Institute: Had a facilities incident that caused their computing to be unavailable so they had a need for additional computing to keep up with the data coming off from the sequencers. Extended their network into the Magellan system and deployed it coming up in their network.

- Reviewed some of the findings and next steps. Determined that there was a strong need for system administration. Noted that I/O performance remained a challenge. Noted there was no batch system, not a lot of easy ways to move data around. Noted that the projects were successful and they were happy with the results. Stated for next steps that they would continue to recruit new projects to come on to the system. Noted they were looking at other ways to give

customized environments besides virtual machines. Interested in making tools to make it easier for users to deploy clusters on their own.

- Discussed conclusions on cloud potential:
  - Enables rapid prototyping at a larger scale than desktop
  - Support different service levels
  - Support surge computing
  - Support tailored environments
  - Facilities resource pooling
- Discussed the cloud challenges:
  - Open source cloud software stacks are still immature but evolving rapidly
  - Small scale MPI-based applications not suited for more business focused offerings
  - Security challenges
- Discussed the next steps for conclusions:
  - Continue to look at different mid-range applications for suitability
  - Cost analysis
  - Finish performance analysis
  - Support the Advanced Networking Initiative (ANI)
  - Final Magellan Report closer to end of fiscal year

COMMITTEE DISCUSSION

Dr. Graham asked if the cloud existed in an organization rather than out there in the world how much would that help. Mr. Canon responded that you could start to look at that by first examining the definition of a cloud and compare that to what an HPC center already did. He noted they look similar but there are small differences. He thought that it would help in the sense that you would expect that you would do better resource pooling and running at larger scales so better cost structure. He added that you would still get some of the benefits of tailored environments. He noted the one advantage would be that you could tailor that system for the unique needs of for example, HPC or the community that you want to support.

Dr. Sarkar thanked him for the presentation. He asked for his thoughts regarding data storage, noting that the slides viewed the cloud more as a computation resource. He asked for his viewpoint regarding the practicality of moving data back up and down for science projects where there are large volumes of data involved. Mr. Canon responded that the best fits to date for commercial clouds were things that had modest data needs and a lot of computational needs that are de-coupled. Thought that bioinformatics was well suited because even though they had a reputation for large amounts of data, for the chunk of data that they would work on, it would be reasonable. So he noted that was an area that was a good fit but if you were talking about moving a large fraction of data into the cloud that could be expensive because you have to pay for storage.

Ms. Victoria White thanked them both for the presentations. She noted that they have had a long history of working on grids and asked what interaction they had with the grid community and asked if he would like to comment on if there were any lessons learned from that, like security for example. Mr. Canon responded that they had been actively engaged with OSG (Open Science Grid) to use Magellan. He noted that several projects were tied into the grid community as well. He replied that there was a lot of technology within the grid space that can be leveraged for the cloud space because he had talked about the fact that they would still have to deal with data management and workflow management so leveraging some of the condor or grid tools to facilitate that was one value. Mr. Conan and Ms. Coghlan discussed some more technical details regarding the security issue.

Dr. Giles asked about the final report in the Magellan project. He asked if there was some event or evolution of cloud systems that he thought they should be paying attention to and maybe reinvestigate. Mr. Canon clarified and asked so how long would the conclusions be applicable? Mr. Canon noted that the space was evolving quickly. He said some significant changes within the commercial space would be one example. He noted it would also be interesting to see how other technologies outside of cloud would impact them, for example with Exascale. He noted as far as specific triggers it would depend on what the broader market does.

Ms. White noted that their investigations were technology-based but she didn't see any mention of the costs or economic model. She asked if that was something that they would look at. Mr. Canon advised that they would be doing a cost analysis as part of the project. He noted that the way they looked at it was that cloud technology was not a technology it is a business model. So he said one could say are there aspects of the business model that could be applied to HPC centers for example.

BREAK

The advisory committee stopped for a ten minute break.

INCITE PROCESS AND AWARDS

Dr. Julia White, *Oak Ridge National Laboratory*

- Discussed the INCITE (Innovative and Novel Computational Impact on Theory and Experiment):
    - Entering the ninth year of operations of the INCITE program.
    - Entering the third year of the INCITE program being run jointly by the Argonne and Oak Ridge Leadership Computing Facilities (OLCF).
- Discussed the INCITE awards that were made in the calendar year 2011 and the process that was used to award funds. Noted that they would also go over what is in store for the future and what was planned for the next call for proposals which would open in April 13th and run through June 30th.
- Noted that in our last call for proposals they had requests exceeding 5.7 billion processor hours. Stated that she would state today the awards that were selected and process by which the selections were made.
- Explained that INCITE was the Innovative and Novel Computational Impact on Theory and Experiment program. They allocated large awards of time to researchers around the world from academia, laboratories and industry. Looked for projects that could maximize effective use of the leadership computing facilities.
- Described the mission goals. First she noted that when proposals are received and they begin the review process they are looking for projects that can clearly articulate high potential impact science and engineering projects. Added they are looking for projects that are computationally intensive. Also noted that they are looking for campaigns that require tens of millions of CPU hours and that are efficient.
- Discussed some of the projects that were successful:
    - Earthquake simulation which required the leadership class systems given the size of the calculations. She reviewed the technical requirements of the study.
    - Projects that are involved in the production of clean energy. This was a three-year award. It was carried out by the National Energy Technology Laboratory (NETL) and they were looking at coal gasification. She explained rather than burning the coal to generate energy they were looking at thermo-chemical mechanisms to take that biomass and convert it into various cleaner types of energy.
    - Another project was groundwater remediation work. The researchers were looking at a site where there is quite a bit of legacy waste and trying to understand the flow of that legacy waste through the sub-surface in a very heterogeneous type of environment and putting boundaries with greater fidelity on to an understanding of how fast and where that uranium plume is moving.

  - Discussed next-generation energy and propulsion. Discussed industry which was a user of the leadership computing facilities through the INCITE program and talked about General Electric which had received time allocated to them.

- Discussed the LCF allocation programs.
  - The Leadership Computing Facilities allocate a total of 2.7 billion processor hours annually.
  - Divided into three programs.
  - Allocate 60 percent of the allocable hours on the Jaguar and Intrepid systems at Oak Ridge and Argonne Leadership Computing Facilities.
  - Annual call for proposals and award sizes average 26 million processor hours per project.
  - Other program called ALCC (ASCR Leadership Computing Challenge) and they allocate up to 30 percent of the resources on the Oak Ridge and Argonne systems.
  - Third program is called the Director's Discretionary. Proposals for this can be received at any time during the year and are typically awarded hours approximately 1 million processor hours per project. Explained that the Director's Discretionary referred to the director of the leadership computing facility.
- Noted that through the INCITE program a wide range of different types of science topics were considered. Noted that they look at all recommendations received through their peer-review science panels and through computational readiness reviews.
- Noted that the INCITE program had no particular requirements in terms of funding source.
- Noted that on average for the past four years they were about two and half times over-subscribed. In 2011 they were about three and half times over-subscribed in terms of requests versus hours.
- Discussed awards made for the current calendar year. Awarded a total of 1.7 billion processor hours to a total of 57 projects. Stated that of those 32 were new projects and 25 were renewals of multi-year awards. In terms of acceptance rate that was 33 percent of new submittals received some award of time and 89 percent of the renewals continued through the INCITE program.
- Discussed trends in INCITE awards using graphs.
- Discussed the project decreases and average and median increases from 2010 to 2011.
- Discussed demographics and noted that if you looked at the PI's of the 57 projects the breakdown would be 44 percent were led by PIs from DOE institutions, nearly half by PI's from

universities and then smaller allocations to projects led by government but not DOE laboratories and industry.

- Examined the 2011 INCITE project collaborations.
- Reviewed the timeline for the 2011 INCITE review process and awards.
- Discussed the review process and the most important was on which projects had the potential for the highest science or engineering impact. Noted peer review and computational readiness. Noted that several of the things they would consider would be the appropriateness of the proposal method or how written the proposal was and the team qualifications.
- Decisions made in late October, and the committee was composed of the leadership team, the management teams at the leadership computing facilities. Noted that they look at the rankings and recommendations made by the peer-review panels. Stated that they secondly look at the computational readiness. This was elaborated on further.
- Reviewed the 2011 INCITE process achievements and discussed in detail.
- Noted the 2011 INCITE peer-review panels. Brought together 66 reviewers and more than half of which were society or laboratory fellows, department chairs or heads or center and program directors. 15 percent were from Europe or Canada. Reviewers from previous years stood at 50 percent. Discussed the panels and process in further detail.
- Noted that they did a panel questionnaire with a 94 percent response rate. Reviewed each category.
- Discussed the details of how and why they build relationships with other centers and programs.
- Discussed the 2012 INCITE process goals. Noted that last year they designed the proposal form and for 2012 they would design the renewal form. Noted also setting expectations for anticipated feature HPC resources. Reminded that the 2012 call for proposals would run from April 13th through to June 30th this year for allocations of time for the calendar year 2012.

COMMITTEE DISCUSSION

Dr. Graham asked what other resources funded proposals used apart from the leadership resources. Dr. White responded that it ran the gamut from receiving resources from their own university or laboratory resources, NSERC allocations or international allocations. Dr. Graham asked if they had a sense if the projects were getting what they needed. Dr. White confirmed that many did not receive what they requested which might be different from what they perhaps needed.

Dr. Giles asked if they had a sense of the over subscription or the over demand. Dr. White advised yes definitely and you could look at the submittals received and awards of allocations.

A member asked how much of the allocations made the previous year were actually used. Dr. White said that that on the balance they were, with some going over allocation and some not using all of it. Further clarification was sought about what a processor hour was, a core, socket or a node? Dr. White asked one of the members of the committee and she confirmed it was a core hour.

SUMMARY OF NATIONAL ACADEMY PANEL ON HPC

Dr. Kathy Yelick, *National Academy*

- Discussed the report just released on computing performance titled, The Future of Computing Performance: Game Over or Next Level.
- Noted she was on the panel that put the report together, and provided a list of people from the NRC who assisted.
- Noted the agenda for the meeting that was held and described the mix of people who took part in the discussions and panel.
- Discussions the highlights that were discussed. Noted one as was there a need for continued growth in computing performance. Noted that the word Exascale would not appear in the report as it was not about Exascale computing or high performance computing but the problem of computing performance focusing on what was going on in a single chip. Noted that the committee spent a great deal of time of that question of what is computing performance. Speed is a product of a large performance picture and less and less has to do with raw clock speed, which has plateaued.
- Reviewed slide concerning the expectation gap due to an inability to increase clock speed in accord with previous expectations as defined by Moore's law.
- Reviewed slide concerning classic CMOS Dennard Scaling and gave technical details regarding the limitations to CMOS semiconductors.
- Discussed the alternatives to CMOS, and stated its future was not indefinite, and had been estimated at 4 generations.
- Revealed cracks in the virtuous cycle – that innovation could not be driven by consumption because the innovation cost too much for too little appreciable gain by consumers.
- Reviewed the recommendations of the report: investment in algorithms, programming methods, rethinking of canonical thought on software stacks, parallel architectures, power efficiency, open interface standards, tools and methods, and finally computer science education.
- Discussed the highlights of the symposium:

- Discussion on the allocation of power from smart phones to PCs
- Discussions of cloud computing challenges, and possible cloud computing benefits from efficiency.
- Discussions about training computer programmers in parallel algorithms from the beginning, rather than sequentially.
- The NSF's being "grilled" on whether they had a plan for computing performance that was coordinated.
- The over engineering of processors, and momentum of the processing industry which had lead them to be stuck in an improvement trajectory that was difficult to escape.
- Lack of political sympathy for the IT industry due to the fact that it was overshadowed by other industries.
- Change in where promising students were pursuing careers tending toward the IT service industry, rather than writing software.
- Most venture capitalists do not want to risk in the marketplace with "brave new world" ideas, and prefer to test things that are "better, faster, cheaper" in the lab.

COMMITTEE DISCUSSION

Dr. Giles commented on the discouraging final slide. Dr. Yelick said the funding area was the gloomy area, but the problem solving questions brought more optimism, as did the receptivity of industry. Later she specifically referenced those discussing programming models and algorithms as being optimistic.

Dr. Sarkar asked about stakeholders and if there was discussion on the role of national laboratories and universities. She responded that there wasn't a distinction or separate discussion about national labs versus universities in the process. She noted that there was general discussion about the need to have long-term research programs.

Dr. Tang asked Dr. Yelick to remark whether her involvement on SciDAC and co design had changed her perspective going into the National Academy Meetings. Dr. Yelick said that it had, especially with a view to recognizing that a bottom up approach wasn't always as useful, and that it was useful to approach some problems as a mission-focused endeavor which was less insular.

PUBLIC COMMENT

Mike (Member of the public) asked a question about the issue Dr. Yelick had raised about where venture capital was likely to go, and said that this wasn't so different from how things were dealt with by the

scientific community. She responded that one of the presenters at her meeting was making the argument that the government's help was needed, that you could not just rely on industry alone to solve a particular problem.

Dr. Giles asked about the last slide regarding the virtuous cycle, and mentioned that it's easy to get discouraged because it became more and more expensive to get results. That said, she said overall the panel was optimistic. Noted that people earlier in the day had been more optimistic about an ability to solve problems and the question was about the funding for the people to solve the problems.

ADJOURNMENT

The advisory committee adjourned for the day at 5:15 p.m. The board will reconvene tomorrow, Wednesday, March 23rd at 8:30 a.m.

MARCH 23, 2011

UPDATE ON EXASCALE RESEARCH

Dr. Karen Pao, *Program Manager, Applied Math Base Programs, ASCR*

- Gave an update on a meeting regarding ASCR Exascale research that took place on March 7th to 11th in San Diego.
- Advised it was a PI meeting for advanced architecture, X-Stack, scientific management and data management analysis.
- Noted there was an X-Stack planning team and seven co-design proposal teams.
- Discussed the expected outcomes:
    - Awareness within each solicitation community
    - What ASCR members are doing and what areas they can leverage
    - About building collaboration an awareness including awareness across the different communities
    - Identification of gaps and leg groundwork for collaboration and co-operation within NNSA and Exascale road mapping activities
- Noted that all the presentations and briefing materials were on the website.
- Discussed the PI meeting which was facilitated and choreographed to ensure that the PIs there would be thinking about relevant issues and be actively engaged. Reiterated that the goal was to get people to work together, interact, network, build collaborations rather than present. The breakdown of the meeting was:
    - Day 1: ASCR presentations
    - Day 2: three poster sessions
    - Day 3: working groups to identify crosscutting topics
    - Day 4: further discussion of crosscutting topics
    - Day 5: gap analysis
- Reviewed a slide which detailed the new PIs and which provided for them information on ASCR. Advised that the point of the illustration to them was to let them know that the Office of Science first and foremost was interested in enabling scientific discovery.

- Discussed the ASCR mission which was to discover, develop and deploy the computational and networking capabilities that enabled researchers to analyze, model, stimulate and predict complex phenomena important to the Department of Energy.
- Noted that the particular challenge of the program was fulfilling the science potential of emerging multi-core computing systems and other novel "extreme-scale" computing architectures which will make significant modifications to today's tools and techniques.
- Detailed the DOE progress toward Exascale. Detailed proposals processed in Exascale-related topic areas:
    - Applied Math (Uncertainty Quantification) – Received 90 proposals – funded six
    - Computer Science Advanced Architectures
    - Computer Science X-Stack
    - Computer Science – Scientific Data Management and Scientific Analysis
    - Computational Partnerships – Co-design – 21 proposals – all still under review
- Noted there were Exascale coordination meetings with other departments and with other federal departments and agencies.
- Formalized partnership with NNSA
- Discussed why Exascale was not just 1,000 X petascale. Noted that they had to let the reviewers know what they meant by co-design. Noted some of the following points which were discussed with technical details:
    - End of increase in clock speed
    - Increasing performance comes through concurrency
    - Floating port operations
    - I/O to disk slower than it is today
    - Silent error
- Discussed power challenges with Exascale. Discussed energy challenges by industry by 2018 and gave technical details. Noted the issue was trading flops for memory so you could compute more.
- Discussed with Exascale and potential system architectures. Stated that the challenge was what would they do? As they were not going to get as much memory.

- Discussed aspects of co-design. Noted that there were still many things that they did not understand about co-design.
- Discussed in more detail the activities on a day by day basis of the meeting:
    - On Day 2 and 3 there were three poster sessions each one a half a day. There were 28 projects and the format was that each project group would do a five-minute presentation with handouts.
    - Organizers assigned the approximate 200 participants and assigned them to one of eight mixed groups.
    - The working groups were moderated and two PIs from ASCR took notes and reported back to the participants.
    - Homework assignments were due right after the poster sessions and there were two forms that researchers were asked to fill out. One was called the Exascale Research Coverage Assessment form and the Research Collaboration "Dance Cards". The Dance Cards were for collaboration.
    - On Day 3 they started the identification of crosscutting topics and there was some slides to facilitate discussion with questions on them. Then the reporter or scribe filled it in to get it ready for the outbrief. There was further discussion of the topics and seed questions.
    - After the participants put out the topics then they each voted on topics with each person had three red and three blue dots. The blue was used for what was important and red is what they worried about.
    - There was then discussion about the topics and how they were rated among the participants using the color coding.
    - Five crosscutting topics were identified based on researcher input.
        - Programming Models
        - Resilience and Fault Tolerance
        - Data Management and I/O Architecture
        - Memory/Data Movement and Energy Efficiency
        - Application Architecture
    - Noted that there were other topics but not all had a critical mass that could have had intelligent discussions.

- Participants were asked to define the crosscutting topics. Asked to discuss those topics with the same groups and they were given guidelines with specific questions.
- There was further discussion and the participants were given five questions to guide the discussion.
- On the last day they worked on gap analysis.

- Noted the website address for interested parties as: http://Exascaleresearch.labworks.org/ascr2011

COMMITTEE DISCUSSION

Dr. Dongarra noted that he had heard good things about the meeting. He noted that it was good to see that the crosscutting issues were the same as the crosscutting issues being developed by the international Exascale software project. He asked what was next step was? Dr. Pao responded that the next step was to analyze all the data they had collected. She noted that through collaborations they got a matrix of who had identified which collaborations were high priority, medium priority and low priority from each team. She noted they were also looking at the research assessment form and there was a lot of outbrief material that they had not had a chance to go through as yet. She said that ideally there would be a follow-up meeting to include more mathematicians especially people working on UQ.

Dr. Dongarra asked if it was their intention to take those topics and turn them into a program that would provide funding for the community to work towards Exascale. Dr. Pao responded that she did not know.

Dr. Hitchcock discussed the issue of making various research top areas into a program with linkages that could move forward. He mentioned that they would also then be able to identify gaps and then see what could be done in those areas.

Dr. Manteuffel thanked her for the presentation and said it sounded like a very effective and active workshop. He noted that she mentioned the applied mathematics algorithms people were not there. He followed up by saying that the Exascale report said that one of the most important things that would be a challenge would be algorithms. He asked if they were planning something similar for that aspect and would it look like the previous one and would they be combined? Dr. Hitchcock responded by noting that what they looked at were the strongest interactions for the present and it looked like UQ was a slightly weaker interaction than the other parts of the stack. Also he noted that the meeting was beginning to get to 230 – 240 people and the challenges of running the meeting with all the mathematicians there which would make it 400 people would be difficult. He said they might have a set of more focused meetings in the future.

Dr. Graham complimented them on the organization of the meeting. She noted that a huge amount of information was generated and although they would use it to determine how to structure programs that

it could also be used for other purposes. She suggested one of them was the opportunity to distribute the information widely so that people external to ASCR and the DOE might get excited about some of the issues. She also said it might be used to explain to the larger community what Exascale was all about based on the intellectual ideas and excitement and not just thinking of it as a tool, in other words use it as a way to talk to the skeptics. Dr. Hitchcock said that their goal was to have as much of the analysis of the data on the website for the participants and general community as they could.

Dr. Chen asked if the topics that they had selected had enough granularity in what she collected from everybody to distinguish between those important or worrisome topics that are more evolutionary versus more revolutionary and she gave an example of programming models. Dr. Pao responded that they were still looking at the data but there had been discussion about evolutionary versus revolutionary and she elaborated. Dr. Chen said when they organized the statistics of the results it would be interesting to see if there are two camps or uniformly spread.

Someone asked where the 20MW number came from and if it came from the DARPA (Defense Advanced Research Projects Agency) report. Dr. Hitchcock said that the 20MW was a number that was about the maximum that you could conveniently bring into any of their existing facilities. There was some general discussion about the DARPA report and some technical discussion about the 20MW number.

Dr. Giles noted that on one of the slides when she was talking about proposals that applied math had 90 proposals and funded six for $45 million versus three compared to all the other categories. He asked what that meant and what could be done about it. Dr. Pao responded that 90 proposals were more than they expected. It was a lot of proposals looking at complex systems. They want more funding toward UQ.

Ms. White asked where they were with the co-design. Dr. Pao noted that they had 21 proposals requesting a large sum of money and of the 21, 5e teams received some funding for planning exercises and then two additional teams were considered as well. She said the funding picture was not clear at the moment.

PUBLIC COMMENT

Ms. Dona Crawford

Ms. Crawford expressed her concern about the community with regard to Exascale. She listed the concerns as: multiple independent efforts, there was no one organization that owned the whole problem, with various different workshops from different agencies. Noted it was an expensive proposition but provided exciting research and increased national security, but the community wasn't really articulating the argument effectively, in a collective way.

Bill Spotz, Sandia National Laboratory

He noted that the co-design FOA (Funding Opportunity Announcement) was structured around applications. Groups were asked to submit a proposal structure around science applications and the meeting seemed to be prospect so it was application independent and he wondered if that had any implications going forward. Dr. Hitchcock responded that they had a number of efforts that they started in response to different FOAs and the meeting was to understand how they could be managed as an integrated portfolio of things going forward to get the maximum benefit out of all of them. He noted that they needed to make sure that the co-design teams knew what computer science to rely on and the computer scientists know who is relying on them.

IPv6

Kevin Oberman, *ESnet*

- Advised the talk would be a technical presentation primarily. Noted that they would be looking at IPv6 and what it actually was today. Reviewed what he would discuss:
  - Review the issue leading to IPv6
  - Why now
  - The OMB IPv6 mandate
  - IPv6 overview and issues
  - IPv6 and DOE organizations
  - Requirements
  - Networks
  - Hosts
  - Services and
  - Suggested path to compliance
- Discussed the internet and where they were today for context. Noted that the internet had hundreds of millions of computers. Noted that each computer had a unique address and the current protocol that handled it was called IPv4 and was developed in the 1970s. IP addresses are 32 bits long and are represented as four decimal numbers separated by dots.
- Stated that in the 1990s the internet moved from a U.S. government research project to a commercial entity. Noted that because of growth the available addresses would be used up quickly. Engineers developed a system to work while they developed a replacement for IPv4.

Developed classless addressing for efficient use of space and Network Address Translation (NAT). Worked on a replacement for IPv4.

- He stated they were working on IPv6 because the internet was running out of addresses. Noted that all the IPv4 addresses have been assigned to regions, e.g. Asia and North America. Noted projections will run out soon.
- Noted that the OMB had mandated IPv6 for public access for 2012. Described the OMB mandate:
  - Government services support IPv6
  - Not a transition as IPv4 is not going away
  - All public internet services on government systems must be reachable via IPv6 by the end of FY12
  - All internet services on government systems must be reachable by IPv6 by end of FY14
  - DOE services and those of DOE laboratories are covered by the mandate
- Outlined what services would be covered as web, mail and DNS (Domain Name Service). Noted that services publicly available are covered. Noted that services requiring authentication were not covered and also services intended for a close audience. Excluded collaboration web services.
- Noted that the 2014 OMB mandate said by the end of FY14 all network services must be IPv6 capable.
- Indicated what made IPv6 different:
  - Uses longer addresses
  - IPv4 has 4.3 billion addresses
  - IPv6 has 340 sextillion addresses
  - Similar to IPv4
  - IPv4 cannot talk to an IPv6 system
  - Systems may and do run on both protocols
- Discussed the technical details regarding v6 space and if it was sufficient.
- Discussed some other important differences. Stated as most notable it supported much easier re-addressing and relied on ICMPv6 for normal operations.

- Discussed NAT and the fact that it allowed one public IPv4 address to be used by an entire network. Used in home and entire private enterprises. Designed mainly to extend the life of IPv4.
- Discussed what a network would need if it was ready to do IPv6. Noted that the mandate only discussed services. Noted that you would need functioning IPv6 connectivity, service providers needed to support IPv6, critical infrastructure must support IPv6, network management, security and accounting.
- Discussed the service requirements of the IPv6.
- Described the DOE IPv6 taskforce:
  - Chaired by Samara Moore
  - Participation by all DOE programs
  - Developing the DOE response for input to OMB
  - Developing scope for FY2012
  - Defining training requirements for both procurement and technical personnel
  - Developing status display to allow easy check on IPv6 compliance
- Discussed what would be required to happen before starting to implement IPv6.
- Outlined the general steps for support of IPv6 services.
- Noted areas likely to require attention, for example hosts, security gear, limits and staff training.
- Outlined some problems that might occur such as bugs, IPv6 tools not fully tested, tools performing poorly, operations differences, provider issues, reliability and connectivity.
- Summarized some of the main points describing the system and as discussed during the presentation.

COMMITTEE DISCUSSION

Dr. Giles asked about ESnet . He asked if he expected in the transitional year some set of disruptions of the other services on ESnet? Mr. Oberman stated there would be no disruption on ESnet because ESnet had been running IPv6 as a full production service for the better part of a decade.

Dr. Graham said the OMB mandate was a U.S. national mandate but IPv6 was an international standard so she asked if other countries were preparing in the same way and what problems were likely to arise for them in their international interactions. Mr. Oberman stated that yes it was an international

standard and most end systems, the hosts, supported these things. Internationally the uptake of IPv6 has been widely touted as being ahead of the U.S. particularly in Asia. He noted the reality was that he had colleagues in Asia and they have stated they are not that much further along. He noted that there may be some issues internationally.

BREAK

The advisory committee stopped for a ten minute break.

UPDATE ON NETWORKING RESEARCH

Mr. Rich Carlson, *ASCR*

- Stated that he would be giving an overview of what the current program was and thoughts on moving into the future.
- Stated the mission was to develop a program to do research, develop test and deploy the advanced technologies and services that are needed for the science communities to do their jobs.
- Stated they have two major thrusts for the program the middleware and networking programs.
- Described the networking program. Stated that it dealt with connecting the machines, the facilities, the computers, the instruments together so they could be used by the science community.
- Described the middleware program. Stated it was focused more on making sure that the people that used the resources could effectively communicate with each other and form partnerships and collaborations that allow them to explore their science base and do the work that they needed to do to perform their science tests.
- Stated they also had a program that integrated with their deployment issues so they talked to their networking community and the network provider that they have ESnet and see how they could get the technologies that they research into their production network.
- Stated that they had a large science community in the U.S. and it was actually a global community. Noted that as per the slide that it showed the people using the facilities and their collaboration middleware program addressed their needs.
- Discussed performance gaps throughout the infrastructure as per an illustrated graph.
- Discussed the FY11 budget and noted their projections and noted that the Next-Gen program was a small slice of the overall ASCR budget. Discussed programs, reserves and overhead in more detail.

- Noted that historically ASCR was involved in networking research for a long time. Discussed the historical perspective and main events from 1989 to 2010.

- Mentioned several highlights, the first was the TCP/IP (Internet Protocol Suite) congestion control algorithm in the late 80s which made the internet possible. Noted that in 2003 the development of GridFTP concerning a way to set up servers.

- Discussed the fact that they have developed software and services at the middleware layer and noted that the Globus community was involved in that for many years.

- Discussed their current portfolio and where Next-Gen was. When breaking down the dollars down it could be seen that the dollars are going to industry, the national laboratories and universities in that order. The people involved followed basically the same breakdown. Noted that long-term R&D (Research and Development) was about 20 percent, short-term issues are about 60 percent, with 20 percent going to activities on test beds and tests.

- Gave a breakdown of their current portfolio, including data movement, advanced network provisioning, network performance monitoring, large scale scientific collaboration and advanced network concepts.

- Highlighted some of the ongoing projects. Described the CEDPS – SaaS (software as a service) Data Management a project in the Globus community. Noted they were making it service rather than a suite of tools. Gave technical details.

- Discussed the open science grid, a large community that is building large-scale collaborations.

- Discussed the climate community. Noted they were addressing their needs and were expanding. Elaborated on their research, and their need for data repositories and access from a global community.

- Discussed the infrastructure with on-demand bandwidth and circuits. Elaborated on the technical details.

- Discussed a new project which they started with regard to advanced diagnostics and analysis. Noted that the PI was attempting to look at data generated with all the flows going through the network and analyze the flows and by doing so determine which one of the links is bad.

- Noted that in addition to their base research program they were involved in the ARRA funding projects so they funded five projects into ARRA to look at networking research activities. Noted they also funded the Magellan project out of their office.

- Noted that a DOE network testbed was going up in the Long Island area. He elaborated on this project.

- Reviewed what they were planning to do in the future including challenges, research activities and core aspects.
- Identified three major challenges:
    - Develop an understanding of how the science communities are using the network and how the network is operating and behaving.
    - Provide the science communities with the technology that they need to access the machines, the computers and their data sources
    - What is the network itself going to do and how are we going to change the infrastructure to allow us to achieve those goals.
- Noted they would do this through two methods including a series of targeted workshops and then a solicitation notice.
- Discussed the terabit network workshop.
- Noted that in 2010 they issued a solicitation and noted the change that they were making in the way that they were doing solicitations. Went into detail regarding the change and solicitations received including budget details.
- Highlighted various applications, both in industry and application of the research they had funded.
- Concluded that their focus is on the end to end performance of large distributed scientific computing environments, an integrated approach combining network, middleware and collaborative activities, interaction with ESnet operation staff, and interaction with the community to identify new challenges.

UPDATE ON NETWORKING COV

Dr. Giles Roscoe - *ASCAC*

- Indicated that he would be talking more about the networking COV later, they are making slow progress, they do have a chair, were working to identify members of the committee to serve. The visit will take place in May or June.
- Noted that Dr. Manteuffel had agreed to Chair the CSGF report and noted that he and Dr. Manteuffel would be working to identify ASCAC folks who were on that committee. Asked if anyone was interested in participating in that to contact Dr. Manteuffel and they would make that happen.

- Noted for the Data Services Report Dr. James Hack would be taking a leadership role in doing the report. Dr. Giles sent out a message to the other advisory committee chairs and he had heard from all of them quickly.

- Stated that they are surveying what the current practices are in their committees. He noted that they were using one mechanism or another but were trying to do a clear report on what the data situation is. He noted that the report and it may go beyond the survey to mention some key issues or concerns for them. He did not want them to be in a situation where there were policies put in place that were later regretted. Noted that it was important for the members of the committee to commit to work with Dr. Hack to make the report happen.

COMMITTEE DISCUSSION

Dr. Giles noted that he was taking the opportunity to discuss Exascale as there were some comments that members wanted to make regarding the earlier presentation.

Dr. Dongarra noted the importance of Exascale and trying to understand what the plan was going forward. He felt that it was important that both sides of DOE needed to get together and he mentioned the ASCR office and the NNSA. He asked what the plan was. He asked what the strategy forward was in terms of working with the other side of the house.

Dr. Hitchcock commented that they had done a lot of planning with the other side of the house. He noted that they were working on a formal MOU (Memorandum of Understanding) between ASCR and the NNSA which is percolating through the DOE approval process accumulating signatures. He thought that it was almost ready for final signature. Dr. Dongarra suggested it might be something that the ASCR office could help with. Dr. Hitchcock noted that if they had budget numbers for 2011 then their planning would be much more transparent.

Dr. Graham asked if Dr. Hitchcock could comment on where DARPA and NSF are in the Exascale issue. Dr. Hitchcock commented that he couldn't say where DARPA was. He noted that DARPA was moving in its own direction and he gave some comments on NSF but indicated it was not clear where they were going regarding the issue.

Dr. Hack discussed the budget and said priorities are determined at every level. He thought what Dr. Dongarra was asking was how it got coordinated. He asked how was ASCR going to determine the priorities when there were other players involved in contributing important pieces to the overall enterprise. Dr. Hitchcock indicated that the plan was that the people who they had the closest collaboration with were the NNSA people who had the most similar end requirements. He elaborated on that concept.

Dr. Chen asked about the Exascale drivers and wondered if there had been more of a prioritization in terms of the drivers based on just pure science or based on competitiveness or based on the environment or based on the environment for security. She asked if he could comment. Dr. Hitchcock

responded that they had a goal going forward, a broad enough set of co-design centers so that they would learn things that would impact the broad spectrum of science across DOE's missions. So he gave for example combinations of CFD plus radiation plus chemistry were one sort of class of multi-physics problems that they had and materials was one sort of class of problems that they had. So they wanted to ensure that when they selected things going forward that they had the most likely and the highest priority of things.

Dr. Hack repeated again that as a committee if there was anything that they could do to help they would like to be able to do this considering its importance.

Mr. Christianson (on the telephone) asked how someone who was currently involved in the SciDAC II research effort could optimally become involved with the new initiatives. He asked what advice he could give to people presently involved in SciDAC II. Dr. Hitchcock responded they were involved in discussions with all the other programs in the Office of Science to figure out what the strategic things that ASCR and those programs should do as part of the SciDAC III applications calls going forward. He hoped that those discussions would be further along but the associate directors had all been busy trying to make decisions while considering budget issues.

Dr. Giles brought up the issue of what the committee could do. He noted that they had looked at Exascale and sponsored the workshops, the initiatives that got the community to begin to engage with it and looking at the outcomes of the workshops there is lots of material to think about. He thought it needed a real sense of mission in overall direction at a high and visible level.

UPDATE ON ASCR RECOVERY ACT PROJECTS

Walt Polansky, *ASCR*

- Advised he would be giving an update on the recovery act activities that the ASCR managed, discuss what the researchers had achieved to date, and give a perspective on the existing and future impact of the research results. Noted the bottom line was he felt that there was a good story and it would get better as the recovery acts projects matured.
- Noted that when they reported ARRA projects they had to report it to a different level of precision than they typically did. Noted all the numbers had to be vetted through the budget office in a separate review process.
- He listed the recovery act activities that they had. Noted he was going to focus on SciDAC-e primarily but noted that each of the five major activities crosses the threshold of what the ASCR mission is.
- Discussed the progress that was made on SciDAC-e in terms of three focus areas, EFRCs, applied math projects, and the post-docs.

- Discussed the next step in terms of SciDAC-e activity is for ASCR and basic energy sciences to conduct a programmatic review.
- Discussed the applied mathematics research activity which they got a head start on because they used a solicitation they had on the street that was issued in FY08. They received about 100 proposals in response to the solicitation, made 15 awards under the base program and when they had an opportunity to fund additional competitive proposals they found seven proposals in the area of smart grids, and mathematics for smart grids. They were able to make the awards very early.
- Discussed the ANI testbed overview and noted there were three phases.
- Emphasized the importance of the information on the website regarding the ANI testbed.
- Discussed a sample project from the testbed with a goal that if you looked at the baseline flow of information across a network and you disrupted that with very large files what would you do. Asked how you would optimize that network. Noted the PI was at the University of Virginia. Described the process in more detail.
- Discussed the smart grid activity. Noted that the project looked at can you do a meaningful simulation of an electrical grid or a major component of the electrical grid and could you detect the onset of what could result in a catastrophic failure and if you could, could you mitigate that. Noted that you would need a sensor network to detect and how would the signals be interpreted. Reviewed the progress results.
- Gave some illustrative examples of SciDAC-e with some technical details. Each example including background and results were discussed in detail.

Dr. Giles indicated that with no further interest in public comment, they would take the opportunity to have further discussion on Exascale.

COMMITTEE DISCUSSION

Dr. Giles asked about the reporting mechanisms. He asked whether the reporting mechanisms were easier or harder than you thought. Mr. Polansky noted that it was easier once he had started to see the signs coming out. He noted that when they first ventured into the new arena decisions were made that they were going to use their standard peer-review process, standard award process to hold the recovery act projects to the same standards as any other proposal. He noted that the added complexity they had with the reporting requirements that came down after the fact. That created some complexity in the office.

Dr. Graham noted that the program has provided a welcome infusion of funds for a lot of very worthy projects but with bounded duration. She referred to one of the opportunities that he mentioned being

the cohort of post-docs who had two-year appointments and given that the money is going away and the current climate was so iffy she asked if anything was being done to ask what happens next with that group of young people. Mr. Polansky noted that she was right it was one-term funding. He noted from his perspective what it created an expectation across the board and that expectation was that it was about improving the economic situation. He was not aware of any plan in place after the two-year term has ended.

ADJOURNMENT

The Advanced Scientific Computing Advisory Committee adjourned for the day at 12:00 p.m.