ASC ASCeNews Quarterly Newsletter

NA-ASC-500-08 Issue 8

September 2008

The New Fiscal Year Cometh…

Editorial by Bob Meisner

For the past two years our budget has been determined by continuing resolution and it seems that we are facing that situation at least for the immediate future. There will be a continuing resolution up to the election and beyond and we are waiting to find out if the level of funding will be at 2008 levels or at the lower of the House and Senate marks. With a $66M difference between the marks, that decision will have a substantial impact on how we will be able to proceed.

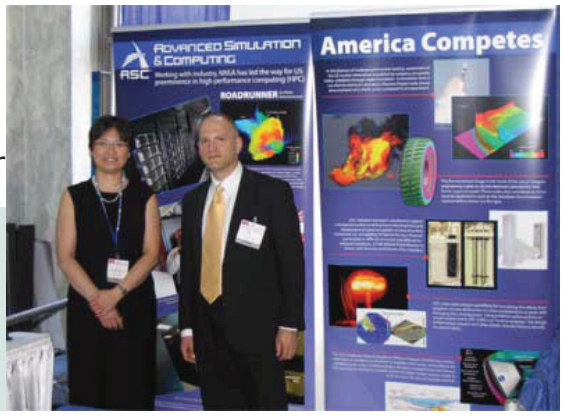
Furthermore, the President will not submit an FY10 budget to allow the incoming administration to have a clean slate for implementing its direction. However, as part of the FY09 budget process there are various activities underway both at NNSA and at the laboratories to inform the transition teams about our priorities and to describe the importance and the range and breadth of our activities. Money is tight in Washington right now and the nation is facing a significant deficit. The competition for resources will affect programs in all agencies and ASC is no exception.

I believe that our work over the past several years to develop and publish increasingly specific direction documents, in concert with the laboratories, which outline our vision and detail our directions, will be of value to us in explaining our activities to new players both in NNSA and in the new administration. These direction documents include our vision statement for a predictive capability, our business plan, a detailed roadmap to predictability by addressing major areas of scientific uncertainty, a Platform Strategy, and, most recently, outlining our code strategy that depends on the special talents across the three laboratories and minimizes unnecessary redundancies. The plans we have developed and our reputation for strong management and excellence in execution should serve us well during these unsettled times. Simulation and computing will continue to be essential to the future of national security sciences including nuclear weapons certification and the grand challenges associated with the broad spectrum of potential threats, and I will do everything that I can to ensure that our program emerges from this transition recognized and healthy.

Third Annual Modeling and Simulation Expo Held on Capitol Hill

On Tuesday July 15, 2008, the ASC Division exhibited for the second time in the third annual Modeling and Simulation (M&S) Expo sponsored by the National Training and Simulation Association (NTSA) and the Congressional M&S Caucus, co-chaired by Congressman J. Randy Forbes and Congressman Soloman P. Ortiz.

ASC's exhibit, entitled "America Competes," showcased the current two fastest computers in the world, the Roadrunner and the BlueGene/L, as well as Red Storm, whose development brought major breakthrough to the high-performance computer (HPC) industry. The exhibit further highlighted the importance of supercomputers, modern high-fidelity physics and engineering codes, and the theory and model



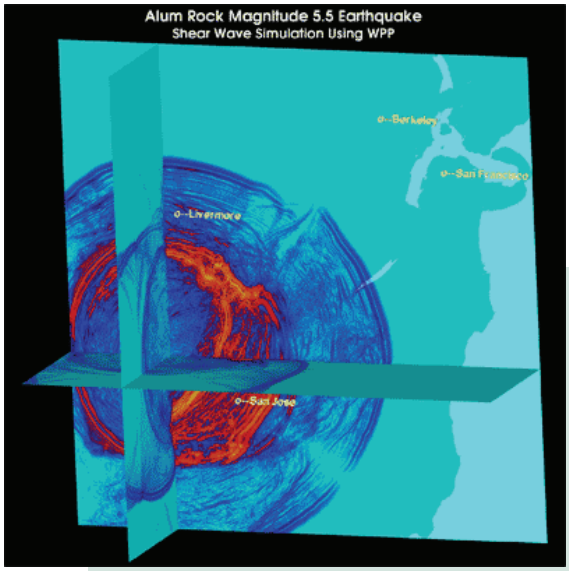
Karen Pao and Ben Bergen from Los Alamos National Laboratory in front of the boothNNSA Creates an Alliance to Pursue Computing at Extreme Scales

development to the core nuclear stockpile stewardship mission, as well as the broader national security mission. The DOE/NNSA/ASC booth was staffed by Karen Pao, Scott Doebling, and Kim Yates, all of NA-121.2, and Ben Bergen, SME on Roadrunner, of Los Alamos National Laboratory.

A handout describing the exhibit panels can be accessed on the following link: http://www.sandia.gov/NNSA/ASC/pdfs/ASC-DC-HandOut08.pdf

Earthquake Simulations Broaden Scientific Understanding

Using the ASC BlueGene/L supercomputer, Lawrence Livermore's modern simulation tool, Wave Propagation Program (WPP), helps seismologists and earthquake engineers understand the ground motion hazard posed by earthquakes and other seismic events. The code computes ground motions for any seismic event in the San Francisco Bay Area region using a 3D geology/seismic wave-speed model of the area provided by the U.S. Geological Survey. Moderate earthquakes, such as the magnitude 5.6 quake that rocked the Calaveras Fault near the town of Alum Rock (east of San Jose) on October 30, 2007, can be used to validate earthquake simulations by testing the 3D model and allowing scientists to constrain uncertainties for large, less frequent earthquakes.



The image shows the damaging shear-wave motions from the Alum Rock earthquake simulated with the WPP code. The image shows shaking intensity (orange being the most intense) from a point of view below the surface with east and west reversed. Note that shaking is most intense in the Santa Clara Valley around San Jose and that the wave fronts spread at different speeds to seismic wave-speed variations.

Multi-Agency Workshop Addresses Issues Facing Network Technologies

The Institute for Advanced Architectures and Algorithms (IAA), which is jointly operated by Sandia National Laboratories (SNL) and Oak Ridge National Laboratory (ORNL), hosted a highly successful workshop on advanced interconnection network technologies in San Jose, California on July 21 and 22. The workshop was organized by Scott Hemmert from Sandia and Jeff Vetter from Oak Ridge National Laboratory and sought to characterize the challenges facing interconnect technologies, looking toward building exascale supercomputers within the next decade.

Two days of discussion focused on prioritizing the challenges in four important areas: network topologies and routing; processor network interfaces; device technologies; and performance prediction and simulation. The workshop provided a broad view of the problem by bringing together experts in the fields of interconnects, system software, applications, and integrated device technologies. The seventy-six attendees were from a wide variety of institutions, spanning academia, industry and government, including program managers from NNSA, National Science Foundation, Defense Advanced Research Projects Agency and the Department of Defense. A final report outlining the top priorities and challenges will be generated from the proceedings of the workshop. Follow-up workshops to delve more deeply into these issues are anticipated.

Project Provides Thirteen Years of Industry-Shaping Deliverables

Over the past 13 years, the ASC File Systems and I/O project has provided products used by the ASC program directly, leveraged and shaped the HPC industry, and built and shaped the HPC file systems and I/O community.

The (former) ASCI Path Forward Program had two projects that were particularly noteworthy. The first Path Forward Global Parallel File System project was instrumental in guiding the design of the Panasas®

File System, one of the most popular global parallel file systems in industry and in use in production at Los Alamos. The second project was the guiding force behind the Lustre® file system, also a top parallel file system industry competitor and in use at Sandia and Lawrence Livermore national laboratories.

The ASCI Level 3 Alliance Program and its descendants have three projects that have produced important results. The partnership with University of California Santa Cruz (UCSC) has provided the world's only open source research object parallel file system. The University of Michigan partnership produced NFSv4 and parallel NFS (pNFS), an IETF standard advancement of the NFS protocol. NFSv4 will be deployed at the three laboratories in the next two years. Additionally, the partnership with Northwestern University produced enhancements to the MPI-IO (ROMIO) parallel I/O library to help with small and unaligned I/O, the hardest I/O pattern for parallel file systems to deal with.

Since 2005, the ASC Program, in partnership with Office of Science, Department of Defense, and NASA, has been coordinating the High End Computing (HEC) Inter-agency Working Group File Systems and I/O (FSIO) technical advisory group, as well as coordinating all HEC government agency investments in the FSIO area. In the last four years this group has assisted with soliciting, awarding, and managing about $20M in FSIO R&D investments in over 25 projects with dozens of universities, national labs, and industry.

ASC's Roadrunner Supercomputer Energy Efficient

In addition to breaking the petaFLOPS record, the Roadrunner supercomputer located in Los Alamos is #3 on the June 2008 Green500 list (http://www.green500.org/lists/2008/06/green500.php).

To illustrate the dramatic improvements in power efficiency, the first vector supercomputer installed at Los Alamos in 1976 needed 115 kW to deliver about 100 megaFLOPS on highly optimized matrix operations hand-coded in assembly language. To match Roadrunner's performance, ten million such machines would require 1,150,000 megaWatts of power — more than the entire electricity-generating capacity of the United States.

Power efficiency of the fastest computers in the world has increased more than 500,000-fold. The Roadrunner system delivers record-breaking 1.026 petaFLOPS while using only 2.345 megaWatts of power, an important advantage since construction of larger facilities would have been costly and time consuming. In addition, significant savings in operating costs are enabled by Roadrunner's power efficiency of 437 megaFLOPS/Watt, more than double that of the #2 machine on the TOP500 list.

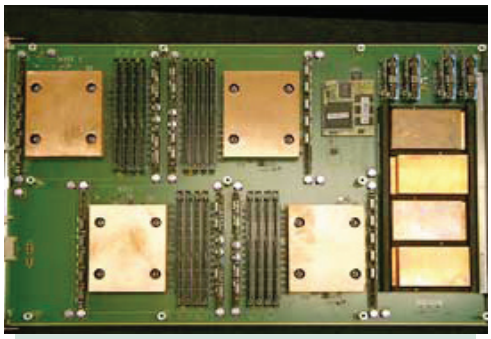
Roadrunner uses power efficient PowerXCell 8i processors in IBM's QS22 Cell Blades, which can reach up to 488 megaFLOPS/Watt. Each Cell chip contains two kinds of processor cores to perform the computational work and the overall coordination. Roadrunner also benefits from IBM's LS21 Opteron processor blades, which are less power efficient but enable less computationally intensive code to run without modifications. Roadrunner combines advantages of three kinds of processor cores in a hybrid system, each of which is programmed in a high-level language such as C or FORTRAN.

The Roadrunner supercomputer

Red Storm Supercomputer Upgraded to 284 teraFLOPS

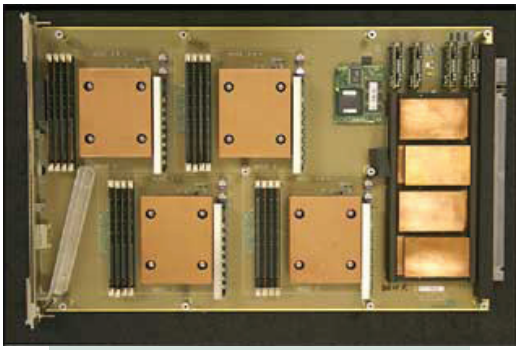
Located at Sandia, Red Storm is undergoing its second large-scale upgrade in four years of active duty. Sixty-five of the 135 compute cabinets are being upgraded from dual-core to quad-core Opteron processors by swapping dual-core modules for newer, quad-core boards. Processing power will increase from a theoretical maximum of 124 teraFLOPS to 284 teraFLOPS across 38,400 cores (12,960 nodes). To hold the computational results, roughly 1.5 petabytes of disk space are being deployed. A corresponding memory upgrade will provide 2 gigabytes of memory per core across the entire machine.

Throughout this upgrade, the system has remained operational with only brief interruptions for testing and emplacing new modules.



Original XT3 Compute Module

Since 2004, Red Storm has increased in performance from 40 to 284 teraFLOPS given an investment of 20% over the original cost of the machine, and with essentially no change in the power requirements. Created by Sandia engineers, the Red Storm design, continues to show its excellent architectural foundation.



New XT4 Compute Module

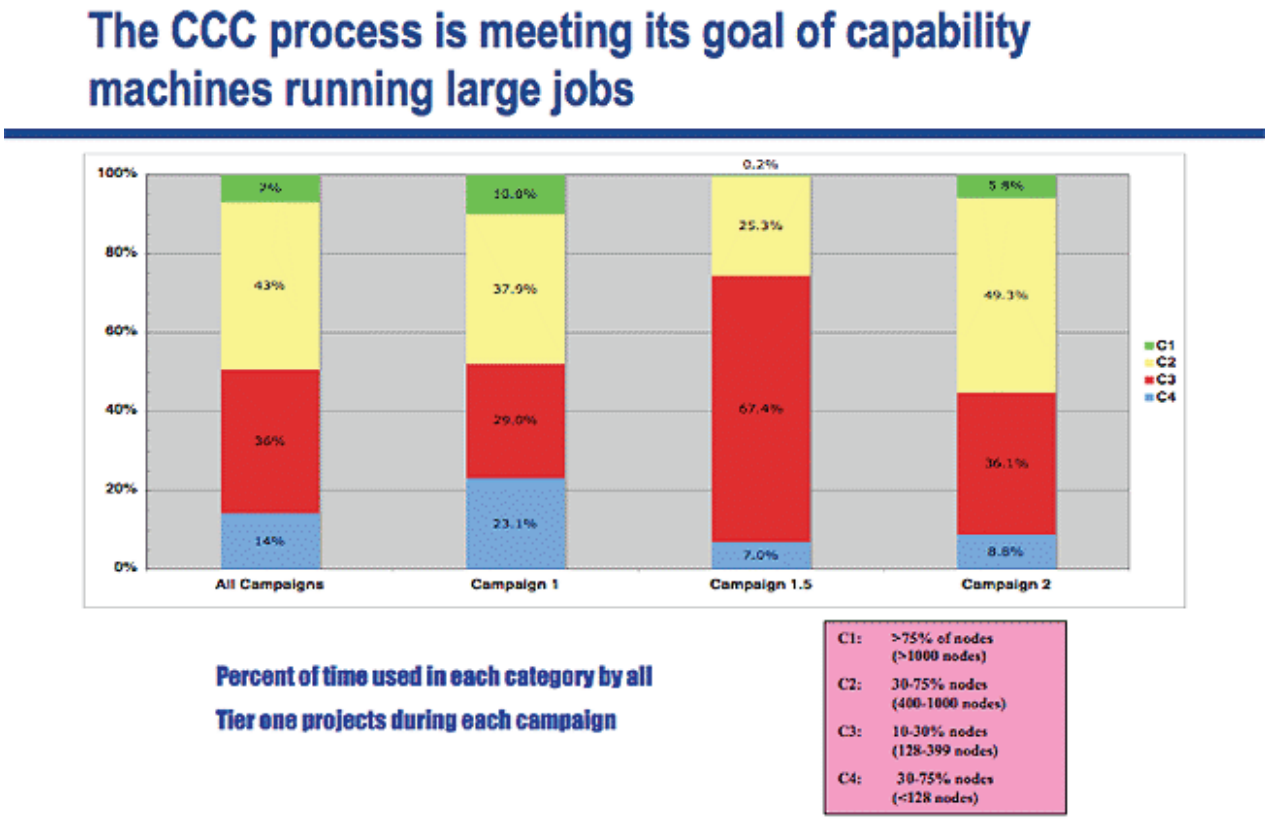
Pictured left and right are the older dual-core and newer quad-core modules, showing that the basic architecture remains the same, while the implementation improves.

Maximizing Use of the ASC's Purple Supercomputer

In the fall of 2005, the ASC Program appointed a team to formulate a governance model for allocating resources and scheduling the stockpile stewardship workload on ASC capability systems. Since October 2006, the ASC Purple capability system at Lawrence Livermore has successfully run as a national user facility using this model. The two primary objectives of the model are to ensure the capability system resources are allocated on a priority-driven basis according to program requirements and to use ASC capability systems for the large capability jobs for which they were designed and procured.

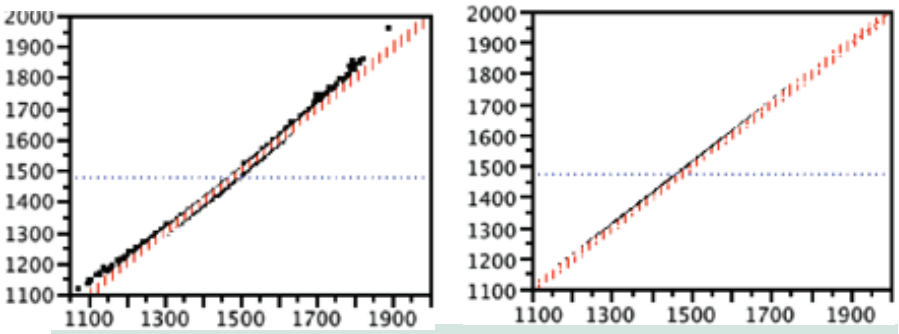
Capability systems' resources are allocated by six-month duration work packages—Capability Computing Campaigns (CCC)—which consist of at least one major calculation needing a significant portion of the machine and related supporting jobs of smaller sizes. Every six months, a new call for proposals is issued. ASC Program scientists submit their proposals, which are vetted through the tri-lab review process. The tri-lab Capability Planning Advisory Committee (CPAC) ranks proposed CCCs according to priorities of the Stockpile Stewardship Program (SSP) and allocations are made following SSP priorities.

Within the constraints of meeting the two primary objectives, this model has effectively maximized use of the Purple machine, both by minimizing idle cycles and by enhancing the probability of productive and useful capability calculations.



Collaborative Campaign Funding Aids Researchers to Develop New Modeling Capabilities

Researchers in Sandia's ASC Verification and Validation Program, leveraging funding from NW Campaign 6 experiments, have developed new failure modeling capabilities within Sandia's SIERRA Mechanics software framework. This new capability permits quasistatic and dynamic modeling of complex failure modes in composite materials, which are used in many weapon and non-weapon applications. This study employed Sandia's DAKOTA software toolkit to generate over 800 SIERRA simulations, which were used to optimize parameters in the composite material failure models to match experimental test data, and, to perform sensitivity studies that identified the most critical parameters out of SIERRA's 37-parameter material failure model.



Each black dot in these figures represents results from one of the 800 sets of parameters chosen using LHS. The vertical red hash marks represent the 95% confidence interval for the model fit.

In addition to providing new capabilities to SIERRA, the results of this study were presented in a paper titled, "Modeling Interlaminar Failure Using Cohesive Zone Models," at the 23rd Technical Conference for the American Society for Composites, (September 9-11, 2008). The results were also documented in a Sandia technical report, "Mode II Fracture—Simulations and Experiments."

Sandia System Software Researchers Receive SC08 Best Paper Nomination

A paper authored by Sandians Kurt Ferreira and Ron Brightwell together with Patrick Bridges from the University of New Mexico has been nominated for the Best Paper and Best Student Paper awards at the upcoming SC08 conference in Austin, Texas http://www.sc08.supercomputing.org/

The paper was one of four such papers selected from a total of 59 accepted papers. The paper entitled "Characterizing Application Sensitivity to OS Interference using Kernel-Level Noise Injection" is a detailed study performed on the ASC Red Storm machine of how operating system activity can impact the performance of parallel applications on very large-scale machines. The paper describes how the Catamount lightweight compute node operating system was extended to allow for creating various kinds of artificial noise to provide more insight into the important characteristics of applications and operating systems that influence performance and scalability.

Roadrunner Technical Seminar Videos Now Streaming on Web

Now available to the high-performance computing community are Roadrunner technical seminars presented by Computer, Computational, and Statistical Sciences (CCS) Division staff at Los Alamos National Laboratory.

These presentations about the first multi-core petaFLOPS speed computer, Roadrunner, were given March–June 2008 at Los Alamos. Using Flash Media Server, the hour-long sessions can now be seen in their entirety, including question and answer sessions, and can be used as tutorials as the adoption of similar systems increases. Visit the website at http://www.lanl.gov/roadrunner/rrseminars.shtml.

Sequoia Request for Proposals Approved, Proposals Accepted

The ASC Sequoia Supercomputer Request for Proposals (RFP) was approved for release by DOE/NNSA in July, and proposals were submitted by industry on August 21, 2008. The contract award should be announced by the end of the year. The RFP specified two machines: Sequoia with a mandatory 20-petaFLOPS peak with target delivery in 2011–2012 and Dawn with a mandatory peak of 0.5 petaFLOPS peak with target delivery in 2009. Bidders were required to run a series of benchmarks and include the results as part of their proposals and make projections of the performance of these benchmarks on the

Dawn and Sequoia platforms.

"Designing scalable benchmarks is hard, but estimating the performance of a machine that is 20–50x faster than BlueGene/L and has not yet been designed poses even greater challenges," said Mark Seager, assistant department head of advanced technologies.

Roadrunner Supercomputer Installed at Los Alamos

The first Roadrunner Phase 3 Connected Unit (CU) was delivered to Los Alamos on July 15, 2008. Shortly after delivery, IBM successfully kicked off a Linpack burn-in run on the first CU in the Metropolis Center for Modeling and Simulation.

Los Alamos will begin acceptance testing this month. In January 2009, the file systems will be connected and scientists will begin system and code stabilization using weapons and open-science code runs. This installation is very exciting because of the potential for advancing science and engineering toward predictive capabilities for the ASC stockpile stewardship mission and other national needs.

See the video online at http://www.sandia.gov/NNSA/ASC/enews/0908/0908-rr-installed.html

Roadrunner Phase 3 System Stabilization Proposals Selected

Proposals have been selected in a competitive process to help with the Roadrunner Phase 3 system stabilization efforts once the machine is accepted at Los Alamos National Laboratory.

To advance science and programming knowledge, LANL has selected open-science and weapons-science projects.

Eighteen connected units of the Roadrunner accelerated petascale system will be available for a select group of users for approximately four months early in 2009 for high-impact science and engineering simulations. This hybrid computing system has a peak performance of 1.026 petaFLOPS with over 80 terabytes of memory. Selected proposals include thermonuclear burn studies, molecular dynamics, turbulence, magnetic reconnection, petascale atomistic bioenergy, viral phylogenetics, parallel-replica dynamics, and supernova light-curves.

For an example of the science, see the movie of a supernova explosion of a 15-solar-mass Star simulation at http://www.sandia.gov/NNSA/ASC/enews/0908/0908-rr-phase3.html. The simulation was carried out with the LANL code SNSPH, a parallel three-dimensional smoothed particle hydrodynamics code.

Work on Roadrunner will extend these simulations by calculating light curves and spectra from full radiation-hydrodynamic models of these explosions. Comparisons with actual light curves and spectra from supernova observations—particularly from upcoming programs like the Large Synoptic Survey Telescope and the Joint Dark Energy Mission—will allow us to more accurately interpret observed phenomena and to further constrain the mechanism responsible for supernovae.

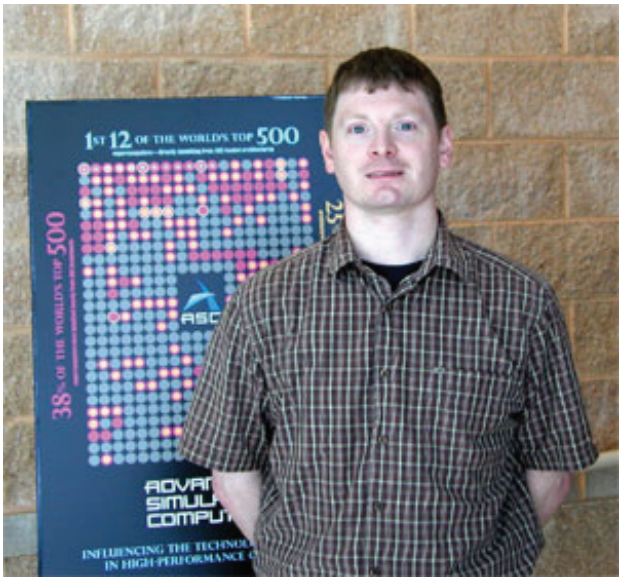
ASC Salutes Matthew L. Leininger

As a deputy for Advanced Technology Projects within the ASC Program at Lawrence Livermore National Laboratory, Matt Leininger works at the center of advances in supercomputing that are enabling predictive scientific simulations.

"It's exciting to see the direct impact advanced supercomputing architectures and predictive models are having on both scientific discovery and real-world problems such as stockpile stewardship," said Matt.

The new 14–20 petaFLOPS supercomputer, Sequoia, is one of the projects on which Matt works. Due for deployment in late 2011, Sequoia will be preceded by a 500-teraFLOPS early development system called Dawn, due for deployment in early 2009.

"Sequoia is a prime example of how LLNL is pushing the frontiers of supercomputing," said Matt. "It's a huge technology leap forward." Sequoia will have on the order of one million processor cores and deliver a 20-time improvement over BlueGene/L for science codes while simultaneously delivering a 24-time improvement over ASC Purple for the integrated design codes. While serving as a stepping stone to petascale com-



puting, Sequoia will also put the ASC Program on the path to reaching exascale computing within the next ten years.

This past year, Matt has also been part of the team developing a new lab collaboration where LLNL computer scientists are working with industry to field a large-scale Linux cluster testbed. This cluster, called Hyperion, is a dedicated development and testing environment for critical enabling of Linux cluster technologies and an evaluation testbed for new hardware and software technologies. The Hyperion partnership will help ensure the success of future multi-petascale platforms, such as Sequoia, and prepare commodity Linux cluster technologies for the petascale era.

The Hyperion partnership involves LLNL, Dell, Intel, Cisco, Sun, Mellanox, QLogic, RedHat Data Direct Networks, LSI, and Supermicro. The first phase of Hyperion is currently being integrated onsite at LLNL and is made up of 576 nodes. The second phase will be delivered in January and will double the size of the Hyperion cluster to 1,152 nodes. Partnership personnel will be using Hyperion to develop and scale new hardware technologies and system software such as the Lustre parallel file system, the ASC Tri-lab Operating System Software (TOSS) based on RedHat Linux, the common parallel programming models such as MPI and OpenMP, the OpenFabrics high performance networking software for InfiniBand, and low latency ethernet.

Prior to becoming a LLNL scientific staff member in 2007, Matt received a PhD in theoretical and computational chemistry from the University of Georgia in 1999, joined Sandia National Laboratories (Livermore) as a post-doctoral candidate, and later became a Sandia staff member in the High Performance Computing and Networking group. Matt led the recent ASC PathForward project, which funded the development of the OpenFabrics InfiniBand software stack now run on 25 percent of the machines on the Top 500, including the ASC tri-lab Linux capacity clusters. Matt's areas of expertise include parallel quantum chemistry methods, application middleware, and high-performance networks and storage. His current interests include scalable system architectures, commodity-based and high-performance computing, and enabling predictive simulation.

Newsletter Points of Contact

Send submittals to:

Denise Sessions—Los Alamos National Laboratory denise@lanl.gov

Andrea Baron—Lawrence Livermore National Laboratory baron1@llnl.gov

Reeta Garber—Sandia National Laboratories ragarbe@sandia.gov

Who's Who

ASC Program Managers—Headquarters

Director, Office of Advanced Simulation and Computing, NA-114

Bob Meisner—bob.meisner@nnsa.doe.gov
Njema Frazier—njema.frazier@nnsa.doe.gov
Thuc Hoang—thuc.hoang@nnsa.doe.gov
Sander Lee—sander.lee@nnsa.doe.gov
Ed Lewis—edgar.lewis@nnsa.doe.gov
Erich Rummel—erich.rummel@nnsa.doe.gov
April Commodore—april.commodore@nnsa.doe.gov
Karen Pao—karen.pao@nnsa.doe.gov
Watti Hill—watti.hill@nnsa.doe.gov

ASC Program Managers—Labs

Los Alamos National Laboratory

John Hopson—jhopson@lanl.gov

Cheryl Wampler—clw@lanl.gov

Lawrence Livermore National Laboratory

Michel McCoy—mccoy2@llnl.gov

Lynn Kissel, Deputy Program Manager—kissel1@llnl.gov

Sandia National Laboratories

James Peery—jspeery@sandia.gov

Art Ratzel—acratze@sandia.gov

Developed and maintained by Sandia National Laboratories for NA.114

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

SAND 2008-6301 W