cug 2009 • atlanta • may 4-7

image courtesy of Sean Ahern (ORNL) • simulation by Fred Jaeger (ORNL) • radio-frequency heating of NSTX fusion plasma

Welcome

COMPUTE THE FUTURE

Dear Friends,

The National Center for Computational Sciences at Oak Ridge National Laboratory (ORNL) and the National Institute for Computational Sciences (NICS) are pleased to host CUG 2009. Much has changed since ORNL was the host in Knoxville for CUG 2004. The Cray XT line has really taken off, and parallelism has skyrocketed. Computational scientists and engineers are scrambling to take advantage of our potent new tools.

But why? Why do we invest in these giant systems with many thousands of processors, cables, and disks? Why do we struggle to engineer the increasingly complex software required to harness their power? What do we expect to do with all this hardware, software, and electricity?

Our theme for CUG 2009 is an answer to these questions...

Compute the Future

This is what many Cray users do. They design the next great superconductor, aircraft, and fusion-energy plant. They predict the efficiency of a biofuel, the track of an approaching hurricane, and the weather patterns that our grandchildren will encounter. They compute to know now, because they should not or must not wait.

We welcome you to the Omni Hotel in Atlanta for CUG 2009. This beautiful hotel and conference facility are in the world-famous CNN Center, overlooking Centennial Olympic Park, and within walking distance of a myriad of restaurants and attractions. We hope the convenience and excitement of Atlanta, the skills and attention of Cray experts, and the shared wisdom of your CUG colleagues will convince you.

Compute the future with us at CUG 2009!

Sincerely,

James B. White III (Trey)
Local Arrangements Chair

Sherry Hempfling
Local Arrangements Coordinator

NATIONAL CENTER FOR COMPUTATIONAL SCIENCES

A UT/ORNL PARTNERSHIP
NATIONAL INSTITUTE FOR COMPUTATIONAL SCIENCES
NICS

Welcome

On Behalf of the CUG Program Committee

Welcome to the 51st Cray User Group meeting, CUG 2009 "Compute the Future." This technical conference provides a unique opportunity for you to exchange problem-solving information and enjoy professional interactions with your Cray Inc. high performance computing system colleagues.

The CUG Program Committee has assembled a diverse and impressive array of detail-packed presentations in General and Parallel Technical Sessions, Tutorials, and Special Interest Group (SIG) meetings. The technical program includes many talks encompassing all aspects of the system from managing the system to taking full advantage of the architecture. A number of talks will highlight the improvements and benefits to the scientific community. Furthermore, Cray Inc. is committed to having many technical experts on site throughout the entire program. In addition, we expect many of Cray's vendor partners to participate in the technical program and to be available for informal discussions.

- General Sessions provide you with the latest corporate and technical information from Cray Inc. executives, as well as topics of general interest to this audience.
- Parallel Technical Sessions give you the opportunity to focus on the specific knowledge domains of the SIGs. The presentations in the technical sessions have been reviewed and selected for you by the Program Committee.
- Tutorials are a great opportunity for you to update your technical skills with the help of selected experts from Cray Inc. and/or other CUG sites.
- SIG Meetings include discussion of the technical issues surrounding Legacy Systems, Applications and Programming Environment, User Support and System Support. These meetings are where we work together to maintain and improve our CUG technical program.

In addition to these prepared presentations and workshops, there are many opportunities for informal discussions.

- Birds of a Feather (BoF) Sessions, the most dynamic aspect of a CUG conference, are scheduled as needed. You are welcome to organize a BoF session. Notices of BoF sessions will be posted on the Message Board.
- The CUG Night Out, the Cray Social, luncheons, and breaks are among the many occasions you will have to exchange information with your colleagues from other CUG sites and to engage the insight and expertise of representatives from Cray Inc. and their technology partners.

Two excellent Keynote Speakers are part of this year's program:

- **James Hack**, Director of the National Center for Computational Sciences, Oak Ridge National Laboratory (ORNL)
- **Thomas Schulthess**, Director, CSCS-Swiss National Supercomputing Centre (CSCS)

This year's program includes a third General Session talk on Jaguar, the most powerful Cray system in the world from:

- **Arthur Bland**, Director of the Leadership Computing Facility project, Oak Ridge National Laboratory (ORNL)

CUG 2009 promises to be a rewarding, educational, and entertaining experience for all attendees. I look forward to seeing you in Atlanta, Georgia, at CUG 2009, Compute the Future.

Nicholas P. Cardo
CUG Vice-President and Program Chair
National Energy Research Scientific Computing Center (NERSC)

CRAY USER GROUP INCORPORATED

Atlanta's Centennial Park next to the Omni Hotel

Program Notes

Contents

Cell Phone and Pager Policy

Please turn off your cell phone and/or pager during conference sessions. There will be ample time to check your messages outside of the conference sessions. Please extend this courtesy to our presenters and your fellow attendees.

Smoking Policy

There is no smoking allowed at the Conference.

Pocket Schedule

The Pocket Schedule provides the schedule and other information in a handy pocket size and replaces the program schedule layout used in past programs.

Special Interest Group (SIG) Meetings

The Special Interest Groups (SIGs) hold interactive sessions that are open to all interested CUG attendees. A separate interactive session will be held again this year for each of the four standing groups. Please check the schedule for day and time of the meetings. These meetings provide a forum to become acquainted with people from other CUG sites who share your interests, to discuss the future direction of the SIG and other important issues, and to hear talks of special interest. Liaisons from Cray Inc. will be available for each SIG to help address questions and to foster communications between the SIG and Cray. You are encouraged to attend any of these open meetings.

This year's program welcomes the addition of the XTreme Special Interest Group (SIG). This group works very closely with Cray, under Non-Disclosure Agreements, to provide valuable input into the Cray XT system development cycle. For this reason, these are "closed door" sessions on Tuesday and Thursday, and attendance is restricted. This group typically focuses on XT system management and reliability, as well as customer requirements and feedback on site issues, especially for large-scale XT systems. For more information please contact the SIG Chair, Ann Baker (ORNL) or Vito Bongiorno, Cray Inc.

Program Suggestions

The Program Committee needs your ideas! We encourage you to share your suggestions on how to improve the conference with the CUG Vice-President, CUG Directors, and/or any of the SIG Chairs. The CUG Conference Program is the net result of the combined efforts of all interested CUG participants. You can make a unique contribution to the next CUG conference with your suggestions. Help establish the future direction of CUG and let us know how to help you make the most of your Cray system!

Changes to the Program

We always anticipate there will be some changes to the Program schedule. Check the Message Board on site at the conference each day for schedule changes. Thank you for your patience and cooperation.

Conference Evaluation

The conference evaluation will be available online starting Wednesday at 8:00 AM until Thursday at 5:00 PM. Please go to www.cug.org/4-organization/eval_survey/. Once you've completed it, please stop by the registration desk. We have a gift for you.

Keynote Speaker Bio's

James J. Hack

James J. Hack directs the National Center for Computational Sciences (NCCS), a leadership computing facility at Oak Ridge National Laboratory supporting transformational science. He identifies major high performance computing needs from scientific and hardware perspectives and develops strategies to meet those needs as machines evolve past the petascale, the ability to carry out a quadrillion calculations per second. An atmospheric scientist, Dr. Hack also leads ORNL's Climate Change Initiative.

After receiving a Ph.D. in atmospheric dynamics from Colorado State University in 1980, Dr. Hack became a research staff member at the IBM Thomas J. Watson Research Center, where he worked on the design and evaluation of high-performance computing architectures. In 1984 he moved to the National Center for Atmospheric Research, a National Science Foundation-sponsored center, where his roles included senior scientist, head of the Climate Modeling Section, and deputy director of the Climate and Global Dynamics Division. He was one of the principal developers of the climate model that ran on NCCS supercomputers to provide more than one-third of the simulation data jointly contributed by the Department of Energy and the National Science Foundation to the most recent assessment report of the United Nations' Intergovernmental Panel on Climate Change, the group that shared the 2007 Nobel peace Prize with Al Gore. He has also held an adjoint professor position at the University of Colorado at Boulder and is author or co-author of 98 scientific or technical publications.

Dr. Hack is a member of the DOE's Advanced Scientific Computing Advisory Committee and the Working Group on Numerical Experimentation, sponsored by the Joint Scientific Committee for the World Climate Research Program and the World Meteorological Organization Committee for Atmospheric Sciences. He also chairs the Lawrence Livermore National Laboratory/ Program for Climate Model Diagnosis and Intercomparison Advisory Committee. He has served on the ORNL Computer Science and Mathematics Division Advisory Committee, a wide variety of NSF high-performance computing review and advisory panels, the U.S. Water Cycle Scientific Steering Group, and was editor for the *Journal of Climate*. He was also a founding member of DOE's Computational Science Graduate Fellowship Program, a highly successful educational initiative in which he continues to participate.

Thomas C. Schulthess

Thomas Schulthess received his PhD in Physics in 1994 from the ETH Zurich. After a postdoctoral fellowship at Lawrence Livermore National Laboratory he joined Oak Ridge National Laboratory (ORNL) in 1997, initially as a postdoctoral fellow in Metals and Ceramics Division. In 1999 he became a research staff member and from 2002 to 2008 served as group leader of the Computational Materials Sciences Group of ORNL's Computer Science and Mathematics Division. Additionally, in 2005, Thomas was appointment at ORNL's Nanoscience Research Center, the Center for Nanophase Materials Sciences, where he was leading the Nanomaterials Theory Institute until 2008. Since October 2008, Thomas holds a chair in computational physics at ETH Zurich and directs the National Supercomputing Center of Switzerland in Manno. Thomas' research interests are in condensed matter, nano-, and materials sciences, as well as the development and use of high performance computing to solve important problems in these fields. He led the team that was awarded the 2008 ACM Gordon Bell Prize for the development of DCA++, an application to study models of high temperature superconductors and that sustained a petaflop/s in November 2008.

For the location of Tutorials, Tracks A-C, General Sessions, Interactive Sessions, BoFs and Lunch, see the Meeting Room Map on the last page.

Monday

1A Tutorial

8:00 **Performance Tuning on the Cray XT System,** *Luiz DeRose and John Levesque, Cray Inc.*

In this tutorial we will present tools and techniques for application performance tuning on the Cray XT system, with focus on multi-core processors. The tutorial will consist of a brief discussion of the Cray XT architecture, focusing on aspects that are important for understanding the performance behavior of applications; followed by a presentation on the Cray performance measurement and analysis tools; and concluding with optimization techniques, including discussions on MPI, numerical libraries and I/O.

1B Tutorial

8:00 **Introduction to Chapel,** *Steve Deitz, Cray Inc.*

Chapel is a new parallel programming language under development at Cray Inc. as part of the DARPA High Productivity Computing Systems (HPCS) program. Chapel has been designed to improve the productivity of parallel programmers working on large-scale supercomputers as well as small-scale, multicore computers and workstations. It aims to vastly improve programmability over current parallel programming models while supporting performance and portability at least as good as today's technologies. In this tutorial, we will present an introduction to Chapel, from context and motivation to a detailed description of Chapel concepts via lecture and example computations. We will also perform a live demonstration of using Chapel on a laptop to compile and run a sample program. Although not part of the tutorial, we will also provide interested audience members with several Chapel exercises in the form of handouts. We'll conclude by giving an overview of future Chapel activities.

2 Opening General Session

11:00 **Welcome,** *David Gigrich, CUG President, The Boeing Company (BOEING)*

Introduction, *James B. White III (Trey), Local Arrangements Chair, Oak Ridge National Laboratory (ORNL)*

Invited Talk: Challenges in Climate Change Research: Computation as an Enabling Technology, *James J. Hack, Director, National Center for Computational Sciences, Oak Ridge National Laboratory (ORNL)*

The problem of quantifying the consequences of climate change over the next few decades is motivated by the increasingly urgent need to adapt to near term trends in climate and to the potential for changes in the frequency and intensity of extreme events. There are significant uncertainties in how climate properties will evolve at regional and local scales, where the background signal of natural variability is large. Consequently, the climate community is now facing major challenges and opportunities in its efforts to rapidly advance the basic science and its application to policy formation. Meeting the challenges in climate change science will require qualitatively different levels of scientific understanding, numerical modeling capabilities, and computational infrastructure than have been historically available to the community. This talk will touch on how climate science will need to develop in the context of the community's scientific capabilities as assessed by the recent IPCC AR4 activity, and where it will need to be in order to accurately predict the coupled chemical, biogeochemical, and physical evolution of the climate system with the fidelity required by policy makers and resource managers.

12:00 Lunch

3A

1:00 **GPFS on a Cray XT,** *Shane Canon, Matt Andrews, William Baird, Greg Butler, Nicholas P. Cardo, and Rei Lee, National Energy Research Scientific Computing Center (NERSC)*

The NERSC Global File System (NGF) is a center-wide production file system at NERSC based on IBM's GPFS. In this paper we will give an overview of GPFS and the NGF architecture. This will include a comparison of features and capabilities between GPFS and Lustre. We will discuss integrating GPFS with a Cray XT system. This configuration relies heavily on Cray DVS. We will describe DVS and discuss NERSC's experience with DVS and the testing process. Performance and scaling for the configuration will be presented. We will conclude with a discussion of future plans for NGF and data initiatives at NERSC.

1:30 **DVS as a Centralized File System in CLE– Installation, Configuring, Usability, Benchmarks, and the Future,** *Jason Temple and Fabio Verzelloni, CSCS-Swiss National Supercomputing Centre (CSCS)*

Using a centralized filesystem to share home directories between systems is essential for High Performance Computing Centers. In this paper, we will discuss the installation and utilization of DVS in the Cray CLE environment, including current worldwide usage, benchmarks, usability and missing or desired features.

2:00 **The Spider Center Wide File System: From Concept to Reality,** *Galen Shipman, David Dillow, Sarp Oral, and Feiyi Wang, Oak Ridge National Laboratory (ORNL) and John Carrier and Nicholas Henke, Cray Inc.*

The Leadership Computing Facility at Oak Ridge National Laboratory has a diverse portfolio of computational resources ranging from a petascale XT4/XT5 simulation system (Jaguar) to numerous other systems supporting

Monday (continued)

visualization and data analytics. In order to support the I/O needs of these systems, Spider, a Lustre-based center wide file system was designed to provide over 240 GB/s of aggregate throughput with over 10 Petabytes of capacity. This paper will detail the overall architecture of the Spider system, challenges in deploying a file system of this scale, and novel solutions to these challenges which offer key insights into file system design in the future.

3B

1:00 **Comparison of Scheduling Policies and Workloads on the NCCS and NICS XT4 Systems at Oak Ridge National Laboratory,** *Troy Baer, National Institute for Computational Sciences (NICS) and Don Maxwell, Oak Ridge National Laboratory (ORNL)*

Oak Ridge National Laboratory (ORNL) is home to two of the largest Cray XT systems in the world: Jaguar, operated by ORNL's National Center for Computational Sciences (NCCS) for the U.S. Department of Energy; and Kraken, operated by the University of Tennessee's National Institute for Computational Sciences (NICS) for the National Science Foundation. These two systems are administered in much the same way, and use the same TORQUE and Moab batch environment software; however, the scheduling policies and workloads on these systems are significantly different due to differences in allocation processes and the resultant user communities. This paper will compare and contrast the scheduling policies and workloads on these two systems.

1:30 **Unifying Heterogeneous Cray Resources and Systems into an Intelligent Single-Scheduled Environment,** *Scott Jackson and Michael Jackson, Cluster Resources*

As Cray systems are expanded and updated with the latest chip sets and technologies (for example, memory and processors), system managers may want to allow users to run jobs across heterogeneous resources to avoid fragmentation. In addition, as next-generation platforms with key differences (such as partition managers like ALPS and CPA) are added, system managers want the ability to submit jobs to the combined system, automatically applying workload to the best-available resources and unifying reporting for managers. This paper will describe how Moab Workload Manager has been integrated with Cray technologies to provide support for running jobs across heterogeneous resources and disparate systems.

2:00 **Select, Place, and Vnodes: Exploiting the PBS Professional Architecture on Cray Systems,** *Bill Nitzberg, Altair Engineering, Inc.*

In 2009, Altair is porting our new PBS Professional workload management and job scheduling architecture to Cray systems. This architecture significantly improves the extensibility, scalability, and reliability of systems by providing two key abstractions: jobs are described in generic "chunks" (e.g., MPI tasks and OpenMP threads), independent of hardware, and resources are described in terms of generic "vnodes" (e.g., blades and/or sockets), independent of jobs. By reducing both jobs and resources to their most basic components, PBS is able to run the right job at the right time in the right place, run it as fast as possible on any hardware, and reduce waste to zero. In this paper, we provide a detailed look at the PBS Professional architecture and how it is mapped onto modern Cray systems.

3C

1:00 **The Cray Programming Environment: Current Status and Future Directions,** *Luiz DeRose, Cray Inc.*

The Cray Programming Environment has been designed to address issues of scale and complexity of high end HPC systems. Its main goal is to hide the complexity of the system, such that applications can achieve the highest possible performance from the hardware. In this paper we will present the recent activities and future directions of the Cray Programming Environment, which consists of state of the art compiler, tools, and libraries, supporting a wide range of programming models.

1:30 **Scaling the MPT Software on the XT5 and Other New Features,** *Mark Pagel, Kim McMahon, and David Knaak, Cray Inc.*

The MPT 3.1 release allowed MPI and SHMEM codes to run on over a hundred and fifty thousand cores and was necessary to help the Cray XT5 at ORNL to achieve over a petaflop in performance on HPL. MPT 3.1 was also used in quickly getting numerous other applications to scale to the full size of the machine. This paper will walk through the latest MPT features including both performance enhancements and functional improvements added over the last year including improvements for MPI_Allgather, MPI_Bcast as well as for MPI-IO collective buffering. New heuristics for better default values for a number of MPI environment variables resulting in less application reruns will also be discussed.

2:00 **Porting GASNet to Portals: Partitioned Global Address Space (PGAS) Language Support for the Cray XT,** *Dan Bonachea and Paul Hargrove, Lawrence Berkeley National Lab and Michael Welcome and Katherine Yelick, National Energy Research Scientific Computing Center (NERSC)*

Partitioned Global Address Space (PGAS) Languages are an emerging alternative to MPI for HPC applications development. The GASNet library from Lawrence Berkeley National Lab and the University of California Berkeley provides the network runtime for multiple implementations of four PGAS Languages. This paper describes our experiences porting GASNet to the Portals network API on the Cray XT series.

4A

3:00 **Performance Evaluation of Chapel's Task Parallel Features,** *Michele Weiland, HPCX Consortium (HPCX) and Thom Haddow, Imperial College London*

Chapel, Cray's new parallel programming language, is specifically designed to provide built-in support for high-level task and data parallelism. This paper investigates the performance of the task parallel features offered by Chapel, using benchmarks such as N-Queens and Strassen's algorithm, on a range of different architectures, including a multi-core Linux system, an SMP cluster and an MPP. We will also be giving a user's view on Chapel's achievements with regards to the goals set by the HPCS programme, namely programmability, robustness, portability and productivity.

3:30 **HPCC STREAM and RA in Chapel: Performance and Potential,** *Steve Deitz, Brad Chamberlain, Samuel Figueroa, and David Iten, Cray Inc.*

Chapel is a new parallel programming language under development at Cray Inc. as part of the DARPA High Productivity Computing Systems (HPCS) program. Recently, great progress has been made on the implementation of Chapel for distributed-memory computers. This paper reports on our latest work on the Chapel language. The paper provides a brief overview of Chapel and then discusses the concept of distributions in Chapel. Perhaps the most promising abstraction in Chapel, the distribution is a mapping of the data in a program to the distributed memory in a computer. Last, the paper presents preliminary results for two benchmarks, HPCC STREAM Triad and HPCC RA, that make use of distributions and other features of the Chapel language. The highlights of this paper include a presentation of performance results (2.78 TB/s on 4096 cores of an XT4), a detailed discussion of the core components of the STREAM and RA benchmarks, a thorough analysis of the performance achieved by the current compiler, and a discussion of future work.

4:00 **An Evaluation of UPC in the Ludwig Application,** *Alan Gray, HPCX Consortium (HPCX)*

As HPC systems continue to increase in size and complexity, Unified Parallel C (UPC), a novel programming language which facilitates parallel programming via intuitive use of global data structures, potentially offers enhanced productivity over traditional message passing techniques. Modern Cray systems such as the X2 component of HECToR, the UK's National High Performance Computing Service, are among the first to fully support UPC; the XT component of HECToR will offer such support with future upgrades. This paper reports on a study to adapt Ludwig, a Lattice Boltzmann application actively used in research, to utilize UPC functionality: comparisons with the original message passing code, in terms of performance (on HECToR) and productivity in general, are presented.

4:30 **Multithreading MADNESS: Implementation and Benefits,** *Scott Thornton and Robert Harrison, University of Tennessee and Oak Ridge National Laboratory*

In preparation for the petaflop Cray XT5, MADNESS was multithreaded using Pthreads and concepts borrowed from the Intel Thread Building Blocks (TBB). MADNESS had been designed for multithreading but Catamount did not support threads that were instead emulated using an event queue. Transitioning to truly concurrent execution greatly simplified the overall implementation but required appropriate use of various mutual exclusion mechanisms (atomic read and increment, spinlock, scalable fair spinlock, mutex, condition variable). Idiosyncrasies of the AMD memory architecture and a non-thread-safe MPI had also to be addressed. We shall discuss the implementation and performance of the multithreaded MADNESS motivating the presentation with reference to an application from solid-state physics that is currently under development.

4B

3:00 **Post-Mortem of the NERSC Franklin XT4 Upgrade to CLE 2.1,** *James Craw, Nicholas P. Cardo, and Helen He, National Energy Research Scientific Computing Center (NERSC) and Janet Lebens, Cray Inc.*

This paper will discuss the lessons learned of the events leading up to the production deployment of CLE 2.1 and the post install issues experienced in upgrading NERSC's XT4 system called Franklin.

3:30 **System Administration Data Under CLE 2.2 and SMW 4.0,** *Jason Schildt, Cray Inc.*

With SMW 4.0 and CLE 2.2, Cray is making significant improvements in how system administrators can access information about jobs, nodes, errors, and health / troubleshooting data. This talk and paper will explain the changes and how administrators can use them to make their lives easier.

4:00 **Slow Nodes Cost Your Users Valuable Resources: Can You Find Them?,** *Ricky Kendall,Don Maxwell, Oak Ridge National Laboratory (ORNL) and Cathy Willis and Jeff Beckleheimer, Cray Inc.*

Many HPC applications have a static load balance which is easy and cheap to implement. When one or a few nodes are not performing properly this makes the whole code slow down to the rate limiting performance of the slowest node. We describe the utilization of a coded called bugget which has been used on Catamount and the Cray Linux Environment to quickly identify these nodes so they can be removed from the user pool until the next appropriate maintenance period.

Monday (continued)

4:30 **Cray Operating System Plans and Status,** *Charlie Carroll, Cray Inc.*

Cray continues to improve and advance its system software. This paper and presentation will review progress over the past year and discuss new and imminent features with an emphasis on increased stability, robustness and performance.

4C

3:00 **Some Issues in the Development of Overset Method in CFD Using Coarray Fortran,** *YikLoon Lee, ERDC MSRC (ERDCMSRC)*

Overset (overlapping) grids are very useful for CFD problems with multiple moving bodies. The parallel implementation of the grid connectivity algorithm with a 2-sided communication model like MPI, however, imposes formidable challenge. One of the reasons is that only a few remote data are known to be needed at any one time in the cell search process. A 1-sided model is well suited for this problem. This paper describes the porting and development of a Navy rotorcraft CFD code using Coarray Fortran on the Cray X1.

3:30 **Effects of Floating-point Non-associativity on Numerical Computations on Massively Multithreaded Systems,** *Oreste Villa, Daniel Chavarria-Miranda, Vidhya Gurumoorthi, Andres Marquez, and Sriram Krishnamoorthy, Pacific Northwest National Laboratory (PNNL)*

Floating-point addition and multiplication are not necessarily associative. When performing those operations over large numbers of operands with different magnitudes, the order in which individual operations are performed can affect the final result. On massively multithreaded systems, when performing parallel reductions, the non-deterministic nature of numerical operation interleaving can lead to non-deterministic numerical results. We have investigated the effect of this problem on the convergence of a conjugate gradient calculation used as part of a power grid analysis application.

4:00 **Optimizing High Resolution Climate Variability Experiments on the Cray XT4 and XT5 Kraken Systems at NICS,** *Richard Loft, John Dennis, Mariana Vertenstein, and Nathan Hearn, National Center for Atmospheric Research, James Kinter, Center for Ocean Land Atmosphere Studies, and Ben Kirtman, University of Miami*

This paper will present XT4 (and hopefully XT5) performance and scaling data for a high resolution (0.5º atmosphere and land surface coupled to 0.1º ocean/sea ice) development version of the Community Climate System Model (CCS) in configurations capable of running efficiently on up to 6380 processors. Technical issues related to tuning the MPI runtime environment, load balancing multiple climate components, and I/O performance will also be discussed.

4:30 **Future-proof Parallelism for Electron-atom Scattering Codes on the Cray XT4,** *Mike Ashworth, Andrew Sunderland, Cliff Noble, and Martin Plummer, HPCX Consortium (HPCX)*

Electron-atom and electron-ion scattering data are essential in the analysis of important physical phenomena in many scientific and technological areas. A suite of parallel programs based on the R-matrix ab initio approach to variational solution of the many-electron Schrodinger equation has been developed and has enabled much accurate scattering data to be produced. However, future calculations will require substantial increases in both the numbers of channels and scattering energies involved in the R-matrix propagations. This paper describes how these huge computational challenges are being addressed by improving the parallel performance of the PRMAT code towards the petascale on the Cray XT4.

5A Interactive Session

5:15–6:00 **Open Discussion with CUG Board**, *Chair, Nicholas P. Cardo, CUG Vice-President, National Energy Research Center (NERSC)*

The CUG Board members are making themselves available for a period of open discussions. Site representatives are encouraged to attend to engage in discussions with the board members as well as other site representatives.

5B Interactive Session

5:15–6:00 **Applications and Programming Environments SIG,** *Chair, Robert Ballance, Sandia National Laboratories (SNLA)*

The Applications and Programming Environment SIG welcomes attendees with a focus on compilers and programming environments. Topics include compilers, scientific libraries, programming environments and the Message Passing Toolkit. SIG business will be conducted followed by open discussions with other attendees as well as representatives from Cray. All attendees are welcome to participate in this meeting.

5C Interactive Session

5:15–6:00 **System Support SIG,** *Chair, Joni Virtanen, CSC - IT Center for Science Ltd. (CSC)*

The System Support SIG welcomes attendees interested in discussing issues around managing and supporting XT systems of all scales. SIG business will be conducted followed by open discussions with other attendees as well as representatives from Cray. All attendees are welcome to participate in this meeting.

Tuesday

6 General Session

8:30 **Introduction,** *Nicholas P. Cardo, Program Chair, National Energy Research Scientific Computing Center (NERSC)*

Cray Corporate Update, *Peter Ungaro, President & Chief Executive Officer, Cray Inc.*

Cray Product Roadmap, *Steve Scott, Chief Technology Officer, Cray Inc.*

Cray Software Update, *Peg Williams, Senior Vice President of Engineering, Cray Inc.*

7 General Session

10:30 **Jaguar: The World's Most Powerful Computer,** *Arthur Bland, Douglas Kothe, Galen Shipman, Ricky Kendall, and James Rogers, Oak Ridge National Laboratory (ORNL)*

The Cray XT system at ORNL is the world's most powerful computer with several applications exceeding one-petaflops performance. This talk will describe the architecture of Jaguar with combined XT4 and XT5 nodes along with an external Lustre file system and login nodes. The talk will also present early results from Jaguar.

11:00 **1 on 100 (or more),** *Peter Ungaro, President & Chief Executive Officer, Cray Inc. (No other Cray personnel and no Cray Vendor Partners please.)*

12:00 Lunch

8A

1:00 **Fast Generation of High-Quality Pseudorandom Numbers and Permutations Using MPI and OpenMP on the Cray XD1,** *Stephen Bique and Robert Rosenberg, Naval Research Laboratory (NRL)*

Random number generators are needed for many HPC applications such as real-time simulations. Users often prefer to write their own pseudorandom number generators. We demonstrate simple techniques to find and implement fast, high-quality customized parallel pseudorandom number generators and permutations. We present results of runs on the Cray XD1.

1:30 **Cray XD1 100X Speedup/FPGA Exceeded: Timing Analysis Yields Further Gains,** *Olaf Storaasli, Oak Ridge National Laboratory (ORNL) and Dave Strenski, Cray Inc.*

Our CUG 2008 paper demonstrated 100X/FPGA speedup on Cray XD1s for very large human DNA sequencing with speedup scalable to 150 FPGAs. This paper examines FPGA time distribution (computations vs. I/O). Testing shows I/O time dominates to such an extent that the actual computation time taken by the FPGAs is almost negligible compared to Opteron computations. The authors demonstrate procedures to significantly reduce I/O, yielding even greater speedup, far exceeding 100X, scalable to 150 FPGAs, for human genome sequencing.

2:00 **Challenges and Solutions of Job Scheduling in Large Scale Environments,** *Christopher Porter, Platform Computing*

Many large computational processes running on large systems are embarrassingly parallel – large parallel jobs consisting of many small tasks. Typically, job schedulers only schedule the parallel job, leaving the application to handle task scheduling. To avoid such complexity, some developers leverage standard parallel programming environments like MPI. This presentation will discuss a solution using Platform LSF and Platform LSF Session Scheduler to handle task level scheduling. As a result, application developers can be more focused on the application itself rather than dealing with system and task scheduling issues.

8B

1:00 **DCA++: Winning the Gordon Bell Prize with Generic and Object-oriented Programming,** *Michael Summers, Oak Ridge National Laboratory (ORNL)*

The 2009 Gordon Bell prize was won by the DCA++ code which was the first code in history to run at a sustained 1.35 petaflop rate. While many GB prize wining codes are written in FORTRAN, the DCA++ code is fully object-oriented and makes heavy use of generic programming. This paper discusses the programming trade-offs which were used to simultaneously achieve both world class performance and also the maintainability and elegance of modern software practice.

1:30 **Applications of the LS3DF Method in CeSe/CdS Core/shell Nano Structures,** *Zhengji Zhao, National Energy Research Scientific Computing Center (NERSC) and Lin-Wang Wang, Lawrence Berkeley National Laboratory*

The Linear Scaling 3 dimensional fragment (LS3DF) code is an O(N) ab initio electronic structure code for large scale nano material simulations. The main idea of this code is a divide-and-conquer method, and the heart of the method is the novel patching scheme that effectively cancels out the artificial boundary effect that exists in all the divide-and-conquer schemes. This method has made the ab intio simulations of the thousands-atom nano systems tractable in terms of the simulation time and yield essentially the same results as the traditional calculation methods. The LS3DF method has won the Gordon Bell Prize in SC 2008 for its algorithmic achievement. We have applied this method to study the internal electric field in the CdS/CdSe core shell nano structure, which has potential applications for the electronic devices and energy conversions.

Tuesday (continued)

2:00 **Benchmarking and Evaluation of the Weather Research and Forecasting (WRF) Model on the Cray XT5,** *Oralee Nudson, Craig Stephenson, Don Morton, Lee Higbie, and Tom Logan, Arctic Region Supercomputing Center (ARSC)*

The Weather Research and Forecasting (WRF) model is utilized extensively at ARSC for operational and research purposes. ARSC has developed an ambitious suite of benchmark cases and, for this work, we will present results of scaling evaluations on the Cray XT5 for distributed (MPI) and hybrid (MPI/OpenMP) modes of computation. Additionally, we will report on our attempts to run a 1km-resolution case study with over one billion grid points.

8C

1:00, 1:30, and 2:00 **XTreme,** *Chair, Ann Baker, Oak Ridge National Laboratory (ORNL)*

This group works very closely with Cray, under Non-Disclosure Agreements, to provide valuable input into the Cray XT system development cycle. For this reason, these are "closed door" sessions.

9A

3:00 **Application of Cray XMT for Power Grid Contingency Selection,** *Yousu Chen, Shuangshuang Jin, Daneil Chavarria, Zhenyu Huang, Pacific Northwest National Laboratory (PNNL)*

Contingency analysis is a key function to assess the impact of various combinations of power system component failures. It involves combinatorial numbers of contingencies which exceed the capability of computing power. Therefore, it is critical to select contingency cases within the constraint of computing power. This paper presents a contingency selection method of applying graph theory (betweenness centrality) to power grid graph to remove low-impact components using Cray XMT machine. The implementation takes the advantage of the graph processing capability of Cray XMT and its programming features. Power grid sparsity is explored to minimize memory requirements. The paper presents the performance scalability of Cray XMT and comparison with other multi-threaded machines.

3:30 **Optimizing Loop-level Parallelism in Cray XMT Applications,** *Michael Ringenburg and Sung-Eun Choi, Cray Inc.*

The Cray XMT, a massively multithreaded shared memory system, exploits fine-grained parallelism to achieve outstanding speed. The system contains a set of software, including a compiler and various performance tools, to help users take advantage of the parallelism opportunities provided by the architecture. Despite the strengths of this software, writing efficient codes for the XMT requires a level of understanding of the architecture and of the capabilities provided by the compiler, as well as a knowledge of how they differ from their more conventional counterparts. This paper will provide a brief overview of the XMT system before diving into the techniques used by the compiler to parallelize user codes. We will then describe how this knowledge can be used to write more efficient and parallelizable codes, and discuss some of the pragmas and language extensions provided by the compiler to assist this process. We will finish with a demonstration, using real application kernels, of how the XMT performance tools can be used to assess how the compiler transformed and parallelized user codes, and to determine what changes may be required to get the codes to run more efficiently.

4:00 **Cray CX1 Overview,** *Ian Miller, Cray Inc.*

This talk will introduce the technology and capability of the Cray CX1 personal supercomputer, including microprocessor advancements, product roadmap, application results and customer testimonials. In addition, there will be discussion on creating SMP machines using Cray CX1 blades and how the CX1 leverages GPGPU computing. Directly following this presentation will be an interactive BOF hosting a live demonstration of applications running on the Cray CX1.

9B

3:00 **Improving CASINO Performance for Models with Large Number of Electrons,** *Lucian Anton, HPCX Consortium (HPCX), Dario Alfe, University College of London, Randolph Hood, Lawrence Livermore National Laboratory, and David Tanqueray, Cray UK Limited*

CASINO is used for quantum Monte Carlo calculations which have at their core algorithms that use sets of independent multidimensional random walkers and which are straightforward to use on parallel computers. However, some computations have reached the limit of the memory resources for models describing more than 1000 electrons because of the large amount of electronic orbital related data that is needed on each core for the computation. Besides that, for models with large number of electrons it is interesting to study whether the swap over one configuration can be done in parallel in order to improve computation speed. We present a comparative study of several ways to solve these problems: i) Second level parallelism for configuration computation, ii) distributed orbital data done with MPI or unix System V shared memory and iii) mixed mode programming (OpenMP and MPI).

3:30 **Using Processor Partitioning to Evaluate the Performance of MPI, OpenMP and Hybrid Parallel Applications on Dual- and Quad-core Cray XT4 Systems,** *Xingfu Wu and Valerie Taylor, Texas A&M University*

Chip multiprocessors (CMP) are widely used for high performance computing. While this presents significant new opportunities, such as on-chip high inter-core bandwidth

and low inter-core latency, it also presents new challenges in the form of inter-core resource conflict and contention. A major challenge to be addressed is how well current parallel programming paradigms, such as MPI, OpenMP and hybrid, exploit the potential offered by such CMP clusters for scientific applications. In this paper, we use processor partitioning as a term about how many cores per node to use for application execution to analyze and compare the performance of MPI, OpenMP and hybrid parallel applications on two dual- and quad-core Cray XT4 systems, Jaguar with quad-core at Oak Ridge National Laboratory (ORNL) and Franklin with dual-core at the DOE National Energy Research Scientific Computing Center (NERSC). We conduct detailed performance experiments to identify the major application characteristics that affect processor partitioning. The experimental results indicate that processor partitioning can have a significant impact on performance of a parallel scientific application as determined by its communication and memory requirements. We also use the STREAM memory benchmarks and Intel-s MPI benchmarks to explore the performance impact of different application characteristics. The results are then utilized to explain the performance results of processor partitioning using NAS Parallel Benchmarks with Multi-Zone. In addition to using these benchmarks, we also use a flagship SciDAC fusion microturbulence code (hybrid MPI/OpenMP): a 3D particle-in-cell application Gyrokinetic Toroidal Code (GTC) in magnetic fusion to analyze and compare the performance of these MPI, OpenMP and hybrid programs on the dual- and quad-core Cray XT4 systems, and study their scalability on up to 8192 cores. Based on the performance of GTC on up to 8192 cores, we use Prophesy system to online generate its performance models to predict its performance on more than 10,000 cores on the two Cray XT4 systems.

4:00 **Communication Characteristics and Hybrid MPI/OpenMP Parallel Programming on Clusters of Multi-core SMP Nodes,** *Rolf Rabenseifner, High Performance Computing Center Stuttgart (HLRS), Georg Hager, Erlangen Regional Computing Center (RRZE), and Gabriele Jost, Texas Advanced Computing Center (TACC)*

Hybrid MPI/OpenMP and pure MPI on clusters of multi-core SMP nodes involve several mismatch problems between the parallel programming models and the hardware architectures. Measurements of communication characteristics between cores on the same socket, on the same SMP node, and between SMP nodes on several platforms (including Cray XT4 and XT5) show that machine topology has a significant impact on performance for all parallelization strategies and that topology awareness should be built into all applications in the future. We describe potentials and challenges of the dominant programming models on hierarchically structured hardware. Case studies with the multizone NAS parallel benchmarks on several platforms demonstrate the opportunities of hybrid programming.

9C

3:00, 3:30, and 4:00 **XTreme,** *Chair, Ann Baker, Oak Ridge National Laboratory (ORNL)*

This group works very closely with Cray, under Non-Disclosure Agreements, to provide valuable input into the Cray XT system development cycle. For this reason, these are "closed door" sessions.

10A Birds of a Feather

4:45 **CX1,** *Ian Miller, Cray Inc.*

This is a continuation of the Cray CX1 Overview session. Details regarding the new CX1 system will be discussed as well as some presentations by Cray. Opportunities for discussions regarding the CX1 will also be available during this BoF.

10B Birds of a Feather

4:45 **CrayPort,** *Mary Johnston, Cray Inc.*

CrayPort has been evolving since its release. This BoF provides an opportunity for attendees to learn more about CrayPort as well as provide valuable feedback to Cray on the current version. All attendees are welcome to participate in this BoF.

10C Birds of a Feather

4:45 **Open for BoF**

Cray Inc. Social

6:00 Omni Hotel Atrium Terrace B

The Cray Social will be held in the Atrium Terrace B in the South Tower overlooking the CNN Center.

Oak Ridge National Laboratory campus.

Wednesday

11 General Session

8:30 **CUG Business: AC Introductions, Election Presentations, Voting,** *Chair: David Gigrich, CUG President, The Boeing Company (BOEING)*

9:00 **Introduction,** *James B. White III (Trey), Local Arrangements Chair, Oak Ridge National Laboratory (ORNL)*

Invited Talk: The DCA++ Story: How New Algorithms, New Computers, and Innovative Software Design Allow Us to Solve Challenging Simulation Problems in High Temperature Superconductivity, *Thomas Schulthess, Director, CSCS-Swiss National Supercomputing Centre (CSCS)*

Staggering computational and algorithmic advances in recent years now make possible systematic Quantum Monte Carlo simulations of high temperature superconductivity in a microscopic model, the two dimensional Hubbard model, with parameters relevant to the cuprate materials. Here we report the algorithmic and computational advances that enable us to study the effect of disorder and nano-scale inhomogeneities on the pair-formation and the superconducting transition temperature necessary to understand real materials. The simulation code is written with a generic and extensible approach and is tuned to perform well at scale. Significant algorithmic improvements have been made to make effective use of current supercomputing architectures. By implementing delayed Monte Carlo updates and a mixed single/double precision mode, we are able to dramatically increase the efficiency of the code. On the Cray XT5 system of the Oak Ridge National Laboratory, for example, we currently run production jobs on up to 150 thousand processors that reach a sustained performance of 1.35 PFlop/s.

9:45 **CUG Business: Election Results,** *James Kasdorf, CUG Secretary, Pittsburgh Supercomputing Center (PITTSCC)*

12A

10:30 **Enhanced Productivity Using the Cray Performance Analysis Toolset,** *Heidi Poxon, Steve Kaufmann, Dean Johnson, Bill Homer, and Luiz DeRose, Cray Inc.*

The purpose of an application performance analysis tool, is to help the user identify whether or not their application is running efficiently on the computing resources available. However, the scale of current and future high end systems, as well as increasing system software and architecture complexity, brings a new set of challenges to today's performance tools. In order to achieve high performance on these petascale computing systems, users need a new infrastructure for performance analysis that can handle the challenges associated with multiple levels of parallelism, hundreds of thousands of computing elements, and novel programming paradigms that result in the collection of massive sets of performance data. In this paper we present the Cray Performance Analysis Toolset, which is set on an evolutionary path to address the application performance analysis challenges associated with these massive computing systems by highlighting relevant data and by bringing Cray optimization knowledge to a wider set of users.

11:00 **Debugging Scalable Applications on the XT,** *Chris Gottbrath, TotalView Technologies*

Debugging at large scale on the Cray XT can involve a combination of interactive and non-interactive debugging; the paper will review subset attach and provide some recommendations for interactive debugging at large scale, and will introduce the TVScript feature of TotalView which provides for non-interactive debugging. Because many users of Cray XT systems are not physically co-located with the HPC centers on which they develop and run their applications, the paper will also cover the new TotalView Remote Display Client, which allows remote scientists and computer scientists to easily create a connection over which they can use TotalView interactively. The paper will conclude with a brief update on two topics presented at the previous two CUG meetings: memory debugging on the Cray XT and Record and Replay Debugging.

11:30 **General Purpose Timing Library (GPTL): A Tool for Characterizing Performance of Parallel and Serial Applications,** *James Rosinski, Oak Ridge National Laboratory (ORNL)*

GPTL is an open source profiling library that reports a variety of performance statistics. Target codes may be parallel via threads and/or MPI. The code regions to be profiled can be hand-specified by the user, or GPTL can define them automatically at function-level granularity if the target application is built with an appropriate compiler flag. Output is presented in a hierarchical fashion that preserves parent-child relationships of the profiled regions. If the PAPI library is available, GPTL utilizes it to gather hardware performance counter data. GPTL built with PAPI support is installed on the XT4 and XT5 machines at ORNL.

12B

10:30 **Petascale I/O Using The Adaptable I/O System,** *Jay Lofstead, Georgia Institute of Technology and Scott Klasky, Karsten Schwan, and Chen Jin, Oak Ridge National Laboratory (ORNL)*

ADIOS, the adaptable I/O system, has demonstrated excellent scalability to 16,000 cores. With the introduction of the XT5 upgrades to Jaguar, new optimizations are required to successfully reach 140,000+ cores. This paper explains the techniques employed and shows the performance levels attained.

11:00 **Administration Experiences with a Petabyte Scale Lustre File System,** *John Walsh, Victor Hazlewood, Troy Baer, and Nathaniel Mendoza, National Institute for Computational Sciences (NICS) and Art Funk, Cray Inc.*

The National Institute for Computational Sciences (NICS), located at the Oak Ridge National Labs, installed a 66K core Cray XT5 in December 2008. A 2.3PB (usable) lustre file system was also configured as part of the XT5 system. NICS would like to present their experiences in configuring, cabling, building, tuning, and administrating such a large file system. Topics to be discussed include, MDS size and speed implications, purge policies, performance tuning, configuration issues, determining the default stripe size and block size, and reliability issues.

11:30 **Cray Scaling Efforts to Reach a PetaFlop,** *Kevin Peterson, Cray Inc.*

This paper describes the activities leading up to reaching sustained PetaFlop performance on the Jaguar XT5 system fielded at the DOE Leadership Computing Facility in Oak Ridge National Laboratories (ORNL). These activities included software development, scaling emulation, system validation, pre-acceptance application tuning, application benchmarking, and acceptance testing. The paper describes what changes were necessary to the Cray software stack to execute efficiently over 100K cores. Changes to Cray System Management (CSM), Cray Linux Environment (CLE) and Cray Programming Environment (CPE) are described, as well as methodologies used to test the software prior to execution on the target machine.

12C

10:30 **Red Storm/XT4: A Superior Architecture for Scalability,** *Mahesh Rajan, Douglas Doerfler, and Courtenay Vaughan, Sandia National Laboratories (SNLA)*

The benefits of the XT4 node and interconnect architecture on scalability of applications are analyzed with the help of a micro benchmarks, mini-applications and production applications. Performance comparisons to a large infini-band cluster with multi-socket nodes incorporating a similar AMD processor brings out the importance of architectural balance in the design of HPC systems. This paper attempts to quantify application performance improvements in terms of simple balance ratios.

11:00 **Catamount N-Way Performance on XT5,** *Ron Brightwell, Sue Kelly, and Jeff Crow, Sandia National Laboratories (SNLA)*

This paper provides a performance evaluation of the Catamount N-Way (CNW) operating system on a dual-socket quad-core XT5 platform. CNW provides several operating system-level enhancements for multi-core processors, including the SMARTMAP technology for single-copy MPI messages and the ability to easily choose between small and large memory pages. Our evaluation will include an analysis of the performance of important micro-benchmarks and applications.

11:30 **Access to External Resources Using Service-Node Proxies,** *Ronald Oldfield, Andrew Wilson, Craig Ulmer, Todd Kordenbrock, and Alan Scott, Sandia National Laboratories (SNLA)*

Partitioning massively parallel supercomputers into service nodes running a full-fledged OS and compute nodes running a lightweight kernel has many well-known advantages but renders it difficult to access externally located resources such as high-performance databases that may only communicate via TCP. We describe an implementation of a proxy service that allows service nodes to act as a relay for SQL requests issued by processes running on the compute nodes. This implementation allows us to move toward using HPC systems for scalable informatics on large data sets that simply cannot be processed on smaller machines.

12:00 Lunch

13A

1:00 **Performance Characteristics of the Lustre File System on the Cray XT5 with Regard to Application I/O Patterns,** *Lonnie Crosby, National Institute for Computational Sciences (NICS)*

As the size and complexity of supercomputing platforms increases, additional attention needs to be given to the performance challenges presented by application I/O. This paper will present the performance characteristics of the Lustre file system utilized on the Cray XT5 systems and illuminate the challenges presented by applications that utilize tens of thousands of parallel processes.

1:30 **Shared File MPI-IO Performance on the Lustre File System,** *Katie Antypas, John Shalf, Shane Cannon, and Andrew Uselton, National Energy Research Scientific Computing Center (NERSC)*

As we enter the petascale computing era, the need for high performing shared file I/O libraries becomes more urgent. From simpler file management and data analysis to the portability and longevity of data, parallel shared file I/O libraries increase scientific productivity and the potential for sharing simulation data. However, the low performance of MPI-IO on the Lustre file system makes it difficult for scientists to justify switching from a one-file-per processor I/O model, despite all the benefits of shared files for post-processing, data sharing and portability. This paper will discuss some of the reasons and possible solutions for low MPI-IO performance on the Lustre file system and the implications for higher level self-describing parallel I/O libraries such as HDF5 and Parallel-NetCDF.

Wednesday (continued)

2:00 **Practical Examples for Efficient I/O on Cray XT Systems,** *Jeff Larkin, Cray Inc.*

The purpose of this paper is to provide practical examples on how to perform efficient I/O at scale on Cray XT systems. Although this paper will provide some data from recognized benchmarks, it will focus primarily on providing actual code excerpts as specific examples of how to write efficient I/O into an HPC application. This will include explanations of what the example code does and why it is necessary or useful. This paper is intended to educate by example how to perform application checkpointing in an efficient manner at scale.

13B

1:00 **CCE,** *Nathan Wichman, Cray Inc.*

In 2008, Cray released its first compiler targeted at the X86 instruction set and over the last several months code developers have begun to test its capabilities. This paper will briefly review the life of the Cray compiler, how to use it, and its current capabilities. We will then present performance numbers, for both some standard benchmarks as well as real applications.

1:30 **The Accelerator Programming Model for Multicore,** *Brent Leback, Douglas Miles, Michael Wolfe, and Steven Nakamoto, The Portland Group*

PGI has developed a kernels programming model for accelerators, such as GPUs, where a kernel roughly corresponds to a set of compute-intensive parallel loops with rectangular limits. We have designed directives for C and Fortran programs to implement this model, similar in design to the well-known and widely used OpenMP directives. We are currently implementing the directives and programming model in the PGI C and Fortran compilers to target x64+NVIDIA platforms. This paper explores the possibility of implementing the kernels programming model on multicore x64 CPUs as a vehicle for creating algorithms that efficiently exploit SIMD/vector and multicore parallelism on processors that are increasingly memory-bandwidth constrained.

2:00 **Cray Programming Environment's Implementation of Dynamic Libraries,** *Geir Johansen and Barb Mauzy, Cray Inc.*

This paper will describe the Cray Programming Environment's implementation of dynamic libraries, which will be supported in a future Cray XT CLE release. The Cray implementation will provide flexibility to users to allow specific library versions to be chosen at both link and run time. The implementation will provide support for running executables that were built using the Cray Programming Environment on other platforms. Also, the paper will discuss how executables built with software other than the Cray Programming Environment may be able to be run on the Cray XT.

13C

1:00 **Understanding Aprun Use Patterns,** *Hwa-Chun (Wendy) Lin, National Energy Research Scientific Computing Center (NERSC)*

On the Cray XT, aprun is the command to launch an application to a set of compute nodes reserved through the Application Level Placement Scheduler (ALPS). At the National Energy Research Scientific Computing Center (NERSC), interactive aprun is disabled. That is, invocations of aprun have to go through the batch system. Batch scripts can and often do contain several apruns which either use subsets of the reserved nodes in parallel, or use all reserved nodes in consecutive apruns. In order to better understand how NERSC users run on the XT, it is necessary to associate aprun information with jobs. It is surprisingly more challenging than it sounds. In this paper, we describe those challenges and how we solved them to produce daily per-job reports for completed apruns. We also describe additional uses of the data, e.g., adjusting charging policy accordingly or associating node failures with jobs/users, and plans for enhancements.

1:30 **Coping at the User-Level with Limitations of the Cray Message Passing Toolkit at Scale: How Not to Spend Your Summer Vacation,** *Forrest Hoffman and Richard Mills, Oak Ridge National Laboratory (ORNL)*

As the number of processor cores available in Cray XT series computers has rapidly grown, users have increasingly encountered instances where an MPI code that has previously worked for years unexpectedly fails at high core counts ("at scale") due to resource limitations being exceeded within the MPI implementation. Here, we examine several examples drawn from user experiences and discuss strategies for working around these difficulties at the user level.

2:00 **Experiences and Challenges Scaling PFLOTRAN, a PETSc-based Code for Implicit Solution of Subsurface Reactive Flow Problems–Towards the Petascale on the Cray XT5,** *Richard Mills, Oak Ridge National Laboratory (ORNL), Glenn Hammond, Pacific Northwest National Laboratory (PNNL), Peter Lichtner, Los Alamos National Laboratory, and Barry Smith, Argonne National Laboratory*

We will describe our initial experiences running PFLOTRAN (a code for simulation of coupled hydro-thermal-chemical processes in variably saturated, non-isothermal, porous media) on the petaflop incarnation of Jaguar, the Cray XT5 at Oak Ridge National Laboratory. PFLOTRAN utilizes fully implicit time-stepping and is built on top of the Portable, Extensible Toolkit for Scientific Computation (PETSc). We discuss the hurdles to "at scale" performance with PFLOTRAN and make some observations in general about implicit simulation codes on the XT5.

14A

3:00 **Fault and Performance Monitoring Using NAGIOS/Cacti,** *Thomas Davis and David Skinner, National Energy Research Scientific Computing Center (NERSC)*

This paper will examine the issues in monitoring a large (9k+ node) installation using NAGIOS and Cacti. Thermal, Fault, and Performance, display, and notification will be examined.

3:30 **Scalable Tool Infrastructure for the Cray XT Using Tree-Based Overlay Networks,** *Philip C. Roth and Jeffrey S. Vetter, Oak Ridge National Laboratory (ORNL)*

Performance, debugging, and administration tools are critical for the effective use of parallel computing platforms, but traditional tools have failed to overcome several problems that limit their scalability, such as communication between a large number of tool processes and the management and processing of the volume of data generated on a large number of compute nodes. A tree-based overlay network has proven effective for overcoming these challenges. In this paper, we present our experiences in bringing our MRNet tree-based overlay network infrastructure to the Cray XT platform, including a description and preliminary performance evaluation of proof-of-concept tools that use MRNet on the Cray XT.

14B

3:00 **XT5 Performance Testing,** *Lee Higbie, Arctic Region Supercomputing Center (ARSC)*

A project to evaluate the compilers on ARSC's XT5 revealed some hazards of performance testing. In this study we compared the performance of "highly optimized" code to "default optimization." Some results on optimization were surprising, and we discuss those results and obstacles we faced in gathering our results.

3:30 **Early Evaluation of the Cray XT5,** *Patrick Worley, Oak Ridge National Laboratory (ORNL)*

We present preliminary performance data for the Cray XT5, comparing with data from the Cray XT4 and the IBM BG/P. The focus is on single node computational benchmarks, basic MPI performance both within and between nodes, and on impact of topology and contention on MPI performance. Example application performance, at scale, will be used to illuminate how the subsystem performance impacts whole system performance.

14C

3:00 **Optimization and Auto Tuning of 3 Dimensional Fast Fourier Transform on Cray XT4,** *Manisha Gajbe and Richard Vuduc, Georgia Institute of Technology and Andrew Canning, Lin-Wang Wang, John Shalf, and Harvey Wasserman, National Energy Research Scientific Computing Center (NERSC)*

In this work we show optimization, performance modeling and auto tuning, of 3 Dimensional Fast Fourier Transform on Cray XT4 (Franklin) system. At the core of many real-world scientific and engineering applications is the necessity for computing 3D FFT. FFT is a very commonly used numerical technique in computational physics, engineering, chemistry, geosciences, and other areas of high performance computing. The FFT has many properties useful in both engineering and scientific computing applications (example include molecular dynamics, material science, Fluid Dynamics etc.). The problem with a parallel FFT is that the computational work involved is O(NlogN) while the amount of communication is O(N). This means that for small values of N (64 x 64 x 64 3D FFTs), the communication costs rapidly overwhelmed the parallel computation savings. A distributed 3D FFT represents a considerable challenge for the communications infrastructure of a parallel machine because of the all-to-all nature of the distributed transposes required, and it stresses aspects of the machine that complement those addressed by other benchmark kernels, such as Linpack, that solves system of linear equations, Ax = b. In addition, by the examination of the key characteristics of an application kernel, an analytical performance model is formed. The performance model can be a very useful tool for predicting the performance of many scientific applications that use 3 dimensional Fast Fourier Transforms. This model can be used to explore the achievable performance on future systems with increasing computation and communication performance.

3:30 **Hierarchical Auto-Tuning of a Hybrid Lattice-Boltzmann Computation on the XT4 and XT5,** *Samuel Williams, Jonathan Carter, Leonid Oliker, John Shalf, and Katherine Yelick, National Energy Research Scientific Computing Center (NERSC)*

We apply auto-tuning to a hybrid MPI-pthreads lattice-Boltzmann computation running on the Cray XT4 at National Energy Research Scientific Computing Center (NERSC) and the XT5 at Oak Ridge National Laboratory. Previous work showed that multicore-specific auto-tuning can significantly improve the performance of lattice-Boltzmann magnetohydrodynamics (LBMHD)—by a factor of 4x when running on dual- and quad-core Opteron dual-socket SMPs. We extend these studies to the distributed memory arena via a hybrid MPI/pthreads implementation. In addition to conventional auto-tuning at the local SMP node, we tune at the message-passing level to determine the optimal aspect ratio as well as the correct balance between MPI tasks and threads per MPI task. Our study presents a detailed performance analysis when moving along isocurves of constant hardware usage: fixed total memory, total cores, and total nodes. Overall, our work points to approaches for improving intra- and inter-node efficiency on large-scale multicore systems for demanding scientific applications.

Wednesday (continued)

15A Birds of a Feather

4:15 **External Services,** *Jim Harrell, Cray Inc.*

The Cray XT architecture separated service from compute. This was a well worn path used by MPP systems in the past but important for scaling. Now a number of customers and Cray are working on making the service nodes even more separate from the compute nodes by using more standard servers that are outside the High Speed Network. This BOF will offer an opportunity for Cray and customers to share experiences and directions for External Service Nodes.

15B Interactive Session

4:15 **Legacy Systems SIG,** *Chair, James Kasdorf, Pittsburgh Supercomputing Center (PITTSCC)*

The Legacy System SIG welcomes attendees still utilizing Legacy systems such as XD1 and X1. This meeting provides an opportunity for SIG business followed by open discussions with other attendees as well as representatives from Cray. All attendees are welcome to participate in this meeting.

15C Interactive Session

4:15 **User Support SIG,** *Chair, James Glidewell, The Boeing Company (BOEING)*

The User Support SIG welcomes attendees interested in discussion issues with regard to supporting their user community. SIG business will be conducted followed by open discussions with other attendees as well as representatives from Cray. All attendees are welcome to participate in this meeting.

CUG Night Out

5:30 **Buses depart from the Omni Hotel starting at 5:30 p.m. The last bus leaves at 6:30 p.m. See page 22 for details.**

Thursday

16A

8:30 **Automated Lustre Failover on the Cray XT,** *Nicholas Henke, Cray Inc.*

The ever increasing scale of current XT machine requires increased stability from the critical system services. Lustre failover is being deployed to allow continued operation in the face of some component failures. This paper will discuss the current automation framework and the internal Lustre mechanisms involved during failover. We will also discuss the impact of failover on a system and the future enhancements that will improve Lustre failover.

9:00 **The Lustre Monitoring Tool,** *Andrew Uselton, National Energy Research Scientific Computing Center (NERSC)*

NERSC has deployed The Lustre Monitoring Tool (LMT) on the "Franklin" Cray XT, and several months of data are now on file. The data may be used in real time for monitoring system operations and system testing, as well as for incident investigation and historical review. This paper introduces LMT and then presents several examples of insights gained though the use of LMT as well as anomalous behavior that would have otherwise gone unrecognized.

9:30 **Exploring Mass Storage Concepts to Support Exascale Architectures,** *David Fellinger, DataDirect Networks, Inc.*

There are many challenges that must be faced in creating an architecture of compute clusters that can perform at multiple petaflops and beyond to exascale. One of these is certainly the design of an I/O and storage infrastructure that can maintain a balance between processing and data migration. Developing a traditional workflow including checkpoints for simulations will mean an unprecedented increment in I/O bandwidth over existing technologies and designing a storage system that can contain and deliver the product of the computation in an organized file system will be a key factor in enabling these large clusters. Potential enabling technologies will be discussed which could achieve dynamic scaling in both I/O bandwidth and data distribution. Concepts will also be presented that could allow the distribution elements to act upon the data to simplify postprocessing requirements and scientific collaboration. These technologies, if affected, would help to establish a much closer tie between computation and storage resources decreasing the latency to data assimilation and analysis.

16B

8:30 **XT9? Integrating and Operating a Conjoined XT4+XT5 System,** *Don Maxwell, Josh Lothian, Richard Ray, Jason Hill, and David Dillow, Oak Ridge National Laboratory (ORNL) and Cathy Willis and Jeff Becklehimer, Cray Inc.*

The National Center for Computational Sciences at Oak Ridge National Laboratory recently acquired a Cray XT5 capable of more than a petaflop in sustained performance. The existing Cray XT4 has been connected to the XT5 to increase the system's computing power to a peak of 1.64 petaflops. The design and implementation of the conjoined system will be discussed. Topics will include networks, Lustre, the Cray software stack, and scheduling with Moab and TORQUE.

9:00 **Investigating Real Power Usage on Red Storm,** *James Laros, Kevin Pedretti, Sue Kelly, John Vandyke, and Courtenay Vaughan, Sandia National Laboratories (SNLA) and Mark Swan, Cray Inc.*

This paper will describe the instrumentation of the Red Storm Reliability Availability and Serviceability (RAS) system to enable collection of power draw and instantaneous voltage

measurements on a per-socket basis. Additionally, we will outline modifications to the Catamount Light Weight Kernel Operating System which have realized significant power savings during idle periods. We will also discuss what we call Application Power Signatures and future plans for research in this area.

9:30 **Deploying Large Scale XT Systems at ORNL,** *Jim Rogers, Oak Ridge National Laboratory (ORNL) and Bob Hoehn and Doug Kelley, Cray Inc.*

Jaguar, the DOE leadership system at the ORNL Leadership Computing Facility, and Kraken, the NSF petascale system managed by UT-ORNL's National Institute for Computational Sciences, are two of the largest systems in the world. These systems rely on the Cray ECOphlex™ technology to routinely remove more than 10MW of heat. This presentation will describe the installation of these two systems, the unique packaging of the XT5, and how this system design can reduce power and cooling requirements, reduce energy consumption, and lower operating costs.

16C

8:30 **Band Parallelism in CASTEP: Scaling to More Than 1000 Cores,** *Mike Ashworth, Phil Hasnip, Keith Refson, and Martin Plummer, HPCX Consortium (HPCX)*

CASTEP is the UK's premier quantum mechanical materials modeling code. We describe how the parallelism is implemented using a 3-layer hierarchical scheme to handle the heterogeneously structured dataset used to describe wave functions. An additional layer of data distribution over quantum states (bands) has enhanced the scaling by a factor of 8, allowing many important scientific calculations to efficiently use thousands of cores on the HECToR XT4 service.

9:00 **Exploiting Extreme Processor Counts on the Cray XT4 with High-Resolution Seismic Wave Propagation Experiments,** *Mike Ashworth and David Emerson, HPCX Consortium (HPCX) and Mario Chavez and Eduardo Cabrera, UNAM, Mexico*

We present simulation results from a parallel 3D seismic wave propagation code that uses finite differences on a staggered grid with 2nd order operators in time and 4th order in space. We describe optimizations and developments to the code for the exploitation of extreme processor counts. The ultra high resolution that we are able to achieve enables simulations with unprecedented accuracy as demonstrated by comparisons with seismographic observations from the Sichuan earthquake in May 2008.

9:30 **Parallel Visualization and Analysis with ParaView on the Cray XT4,** *John Patchett, Dave Pugmire, Sean Ahern, and Daniel Jamison, Oak Ridge National Laboratory (ORNL) and James Ahrens, Los Alamos National Laboratory*

Scientific data sets produced by modern supercomputers like ORNL's Cray XT4, Jaguar, can be extremely large, making visualization and analysis more difficult as moving large resultant data to dedicated analysis systems can be prohibitively expensive. We share our continuing work of integrating a parallel visualization system, ParaView, on ORNL's Jaguar system and our efforts to enable extreme scale interactive data visualization and analysis. We will discuss porting challenges and present performance numbers.

17A

10:30 **User and Performance Impacts by Franklin Upgrades,** *Yun (Helen) He, National Energy Research Scientific Computing Center (NERSC)*

The NERSC flagship computer named “Franklin,” the 9660 compute nodes Cray XT4, has gone through two major upgrades (quad core upgrade and OS 2.1 upgrade) during the last year. In this paper, we will discuss the various aspects of the user impacts such as user access, user environment, and user issues etc. The performance impacts on the kernel benchmarks and selected application benchmarks will also be presented.

11:00 **Performance Evaluation for Petascale Quantum Simulation Tools,** *Stanimire Tomov and Shirley Moore, UTK, Jerzy Bernholc, Jack Dongarra, and Heike Jagode, Oak Ridge National Laboratory (ORNL) and Wenchang Lu, NCSU*

This paper describes our efforts to establish a performance evaluation "methodology" to be used in a much wider multidisciplinary project on the development of a set of open source petascale quantum simulation tools for nanotechnology applications. The tools to be developed will be based on existing real-space multigrid (RMG) method. In this work we take a reference set of these tools and evaluate their performance using state-of-the-art performance evaluation libraries and tools including PAPI, TAU, KOJAK/Scalasca, and Vampir. The goal is to develop an in-depth understanding of their performance on Teraflop leadership platforms, and moreover identify possible bottlenecks and give suggestions for their removal. The measurements are being done on ORNL's Cray XT4 system (Jaguar) based on quad-core 2.1 GHz AMD Opteron processors. Profiling is being used to identify possible performance bottlenecks and tracing is being used to try to determine the exact locations and causes of those bottlenecks. The results so far indicate that the methodology followed can be used to easily produce and analyze performance data, and that this ability has the potential to aid our overall efforts on developing efficient quantum simulation tools for petascale systems.

11:30 **Parallel Performance Optimization and Behavior Analysis Tools: A Comparative Evaluation on the Cray XT Architecture,** *Timothy Stitt and Jean-Guillaume Piccinali, CSCS–Swiss National Supercomputing Centre (CSCS).*

In the field of high-performance computing (HPC) optimal runtime performance is generally the most desirable trait of an executing application. In this paper we present a comparative evaluation across a set of community parallel performance and behavior analysis tools which have been ported to the Cray XT architecture. Using the vendor-supplied CrayPat and Apprentice tools as our benchmark utilities, we evaluate the remaining tools using a set of objective and subjective measurements

ranging from overhead and portability characteristics, through feature comparisons to quality of documentation and "ease-of-use". We hope such a study will provide an invaluable reference for Cray XT users wishing to identify the best analysis tools which meet their specific requirements.

Thursday (continued)

17B

10:30 **Characteristics of Barcelona Floating Point Execution,** *Ben Bales and Richard Barrett, Oak Ridge National Laboratory (ORNL)*

In almost all modern scientific applications, developers achieve the greatest performance gains by tuning algorithms, communication topographies, and memory access patterns. For the most part, instruction level optimization is left to compilers. With increasingly varied and complicated architectures, it has become extraordinarily unclear what benefits these low level code changes can even bring, and, due to time and complexity constraints, many projects can not find out. In this paper we explore the gains of this last mile effort for code executing on an AMD Barcelona processor, leaving readers able to decide if investment in advanced optimization techniques make sense for their codes.

11:00 **Performance of Variant Memory Configurations for Cray XT Systems,** *Wayne Joubert, Oak Ridge National Laboratory (ORNL)*

In late 2009 NICS will upgrade its 8352 socket Cray XT5 from Barcelona (4 cores/socket) processors to Istanbul (6 cores/socket) processors, taking it from a 615 TF machine to nearly 1 PF. To balance the machine and keep 1 GB of memory per core, NICS is interested in reconfiguring the XT5 from its current mix of 8- and 16-GB nodes to a uniform 12 GB/node. This talk examines alternative memory configurations for attaining this, such as balancing the DIMM count between sockets of a node vs. unbalanced configurations. Results of experiments with these configurations are presented, and conclusions are discussed.

11:30 **Solution of Mixed-Integer Programming Problems on the XT5,** *Michael Schultze, Rebecca Hartman-Baker, Richard Middleton, Michael Hilliard, and Ingrid Busch, Oak Ridge National Laboratory (ORNL)*

In this paper, we describe our experience with solving difficult mixed-integer linear programming problems (MILPs) on the petaflop Cray XT5 system at the National Center for Computational Sciences at Oak Ridge National Laboratory. We describe the algorithmic, software, and hardware needs for solving MILPs and present the results of using PICO, an open-source, parallel, mixed-integer linear programming solver developed at Sandia National Laboratories, to solve canonical MILPs as well as problems of interest arising from the logistics and supply chain management field.

17C

10:30, 11:00, and 11:30 **XTreme,** *Chair, Ann Baker, Oak Ridge National Laboratory (ORNL)*

This group works very closely with Cray, under Non-Disclosure Agreements, to provide valuable input into the Cray XT system development cycle. For this reason, these are "closed door" sessions.

12:00 Lunch

18A

1:00 **DOE IAA: Scalable Algorithms for Petascale Systems with Multicore Architectures,** *Ron Brightwell and Mike Heroux, Sandia National Laboratories (SNLA) and Al Geist and George Fann, Oak Ridge National Laboratory (ORNL)*

The DOE Institute for Advanced Architecture and Algorithms (IAA) was established in 2008 to facilitate the co-design of architectures and applications in order to create synergy in their respective evolutions. Today's largest systems already have a serious gap between the peak capabilities of the hardware and the performance realized by high performance computing applications. The DOE petascale systems that will be in use for the next 3-5 years have already been designed so there is little chance to influence their hardware evolution in the near term. However, there is considerable potential to affect software design in order to better exploit the hardware performance features already present in these systems. In this paper, we describe the initial IAA project, which focuses on closing the "application-architecture performance gap" by developing architecture-aware algorithms and the supporting runtime features needed by these algorithms to solve general sparse linear systems common in many key DOE applications.

1:30 **Gemini Software Development Using Simulation,** *Kevin Peterson, Cray Inc.*

This paper describes Cray's pre-hardware development environment, the activities, and the testing for the Gemini software stack. Gemini is the next generation high-speed network for the Cray XT-series. The simulation environment is based on AMD's SimNow coupled with a Gemini device model that can be aggregated to form multi-node systems. Both operating system and programming environment software components have been developed within this environment. The simulated batch environment, regression test suite, and development progress are also described.

2:00 **Integrating Grid Services into the Cray XT4 Environment,** *Shreyas Cholia and Hwa-Chun Wendy Lin, National Energy Research Scientific Computing Center (NERSC)*

The 38640 core Cray XT4 "Franklin" system at NERSC is a massively parallel resource available to Department of Energy researchers that also provides on-demand grid computing to the Open Science Grid. The integration of grid services on Franklin presented various challenges, including fundamental differences between the interactive and compute nodes, a stripped down compute-node operating system without dynamic library support, a shared-root environment and idiosyncratic application launching. In our work, we describe how we resolved these challenges on a running, general-purpose production system to provide on-demand compute, storage, accounting and monitoring services through generic grid interfaces that mask the underlying system-specific details for the end user.

18B

1:00 **Simulating Population Genetics on the XT5,**

Michael D. Vose, National Institute for Computational Sciences (NICS),

We describe our experience developing custom C code for simulating evolution and speciation dynamics using Kraken, the Cray XT5 system at the National Institute for Computational Sciences. The problem's underlying quadratic complexity was problematic, and the numerical instabilities we faced would either compromise or else severely complicate large-population simulations. We present lessons learned from the computational challenges encountered, and describe how we have dealt with them within the constraints presented by hardware.

1:30 **Computing Atomic Nuclei on the Cray XT5,** *Hai Ah Nam and David Dean, Oak Ridge National Laboratory (ORNL) and Pieter Maris and James Vary, Iowa State University*

Understanding the structure of atomic nuclei is crucial for answering many fundamental questions such as the origin of elements in the universe. The computational challenge for large-scale nuclear structure calculations, particularly ab initio no-core shell model calculations, stems from the diagonalization of the Hamiltonian matrix, with dimensions in the billions, which can be stored in memory, on disk, or recomputed it on-the-fly. Here, we discuss the issues of scaling MFDn, a nuclear shell model application, including the I/O demands on the CrayXT5.

2:00 **A Generalized Framework for Auto-tuning Stencil Computations,** *Shoaib Kamil, University of California, Berkeley, Cy Chan, Massachusetts Institute of Technology, John Shalf, National Energy Research Scientific Computing Center (NERSC), and Leonid Oliker and Sam Williams, Lawrence Berkeley National Laboratory*

This work introduces a generalized framework for automatically tuning stencil computations to achieve optimal performance on a broad range of multicore architectures. Stencil (nearest-neighbor) based kernels constitute the core of many important scientific applications involving block-structured grids. Auto-tuning systems search over optimizations strategies to find the combination of tunable parameters that maximize computational efficiency for a given algorithmic kernel. Although the auto-tuning strategy has been successfully applied to libraries, generalized stencil kernels are not amenable to packaging as libraries. We introduce a generalized stencil auto-tuning framework that takes a straightforward Fortran77 expression of a stencil kernel and automatically generates code tuned implementations of the kernel in Fortran, C, or CUDA to achieve performance portability across a diverse computer architectures that range from conventional AMD multicore processors to the latest NVidia GTX280 GPUs.

18C

1:00, 1:30, and 2:00 **XTreme,** *Chair, Ann Baker, Oak Ridge National Laboratory (ORNL)*

This group works very closely with Cray, under Non-Disclosure Agreements, to provide valuable input into the Cray XT system development cycle. For this reason, these are "closed door" sessions.

19 General Session

2:45 **LAC Appreciation**

3:00 **Next CUG–Edinburgh,** *Alan Gray, EPCC*

3:15 **Adjourn**

Local Arrangements

How to Contact Us

Until May 1, 2009, and after the conference

James B. White III (Trey) (ORNL)
Local Arrangements Chair
Oak Ridge National Laboratory
1 Bethel Valley Road
Oak Ridge, TN 37831-6008 USA
(1-865) 241-2103 Fax: (1-865) 241-2850
trey@ornl.gov

Sherry Hempfling (ORNL)
Local Arrangements Coordinator
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, TN 37831-6008 USA
(1-865) 574-5162 Fax: (1-865) 241-2850
hgv@ornl.gov

During the conference, May 7–10, 2009

CUG 2009 Conference
Omni Hotel
100 CNN Center
Atlanta, Georgia 30303 USA
(1-404) 659-0000 Fax: (1-404) 525-5050

Conference Registration

CUG Office
3224 Southwood Drive
Philomath, Oregon 97370 USA
(1-541) 929-3776 Fax: (1-509) 272-7239
bobwinget@toolboxcm.com

Attendance and Registration

Who May Attend?

CUG meetings are not open sessions, according to the organization's corporate bylaws and guidelines. Admittance to any meeting session requires the appropriate identification badge. Generally speaking, employees of a CUG Member site (usually a facility or company using a Cray Inc. computer and identified by its CUG site code), and users of computing services provided by a CUG member site, may attend a CUG meeting as a registered Installation Participant from that CUG Member site.

In addition, CUG bylaws specify that a "CUG Visitor" is "Any individual who is not an Installation Representative and who is invited to attend a specific function of the Corporation ... [with] the prior approval of the Board of Directors." This generally applies to an occasional invited speaker who brings important perspectives to our meeting in support of a strong technical program. Finally, CUG bylaws state that, "Meetings shall not be used for marketing or other commercial purposes."

Conference Office and Registration

Registration for pre-registered, paid attendees is in the North Tower, M4 Level directly opposite the North Tower escalators. Registration for Presenters, SIG Chairs, new and unpaid registrants is in the Conference Office directly behind the registration desk. Badges and registration materials are available:

Sunday.....3:00 p.m.–5:00 p.m.
Monday–Wednesday.....7:30–a.m.–5:30 p.m.
Thursday.....7:30 a.m.–3:00 p.m.

All attendees must wear badges during CUG Conference activities.

Conference Registration Fees

Your registration fee includes

- Admission to all program sessions, meetings, and tutorials
- Continental breakfast, morning and afternoon breaks, and lunch Monday through Thursday
- CUG Night Out including dinner that evening

Proceedings

All conference attendees will have access to the conference Proceedings on our secure web server and will receive them on CD-ROM distributed approximately two months after the conference.

Cancellations

Conference registration cancellations must be received by the CUG Office before April 22, 2009. All registration fees will be refunded (less a $25 handling fee) if the cancellation was received by this date.

Conference Hotel Room Rates

The conference room rate is $141 USD (plus tax) for single and $161 USD (plus tax) for double rooms. The deadline for reservations at the conference rate is Wednesday, April 15, 2009, 5 PM EDT. After April 15 rooms are available on a space and rate availability basis.

Special Assistance

For special assistance during the conference come to the Conference Office.

Dining Services

Breakfast, Lunch, and Refreshments

The conference registration fee includes continental breakfast, lunch, and snacks for the duration of the conference. Breakfast and lunch will be served Monday through Thursday in the Grand Ballroom (D1) from 7:00 am until sessions begin. Refreshments will be provided in the Grand Ballroom Lobby during the morning and afternoon breaks.

Dinner on Wednesday

Dinner Wednesday evening is provided at the CUG Night Out for all attendees who have registered for the full conference. Tickets for attendees' guests are $55 and are available in the Conference Office. This special event is organized by the Local Arrangements Committee.

Dining Out

For dining out, everything from fast food to premium steak and seafood is available in the CNN Center, with more choices within walking distance. Visit the hotel concierge for assistance. There are also a number of useful links from the CUG 2009 web pages.

On-Site Facilities

Business Services

The hotel provides business support services. CUG presenters rarely provide paper copies of their presentations, but, if you wish to do so, please bring copies with you. The hotel business office can make limited copies for a charge.

Messages

Call the hotel at (1-404) 659-0000 to leave telephone messages for attendees and ask that the message be either delivered to your room or to the conference office during the conference (May 4–7).

Faxes

A fax machine is available in the hotel business office. Incoming fax messages may be sent to the CUG Conference Office eFax™ at (1-509) 272-7239. They will arrive as PDFs via E-mail to the CUG Office computer. We will post notices of incoming fax messages on the conference message board.

Wireless Access, E-mail, and Power

There is wireless internet access throughout the conference facility for computers equipped with an 802.11b/g-compatible card. A limited number of wire connections (ethernet) will be available in the E-mail room, and some guest rooms will be available with wired ethernet. The ethernet uses RJ-45 modular jacks.

Each conference meeting room will have power strips servicing the first few rows of seating. If you want to plug in, sit near the front!

The Omni hotel lobby and guest rooms have high-speed internet access. Guest room access is available for about $10 per day or for free for Omni "Select Guest" members. Membership is free; go to www.omnihotels.com/Home/FindAHotel/AtlantaCNNCenter/HotelHighlights/WirelessInternetAccess.aspx for more information.

E-mail and Personal Computers

An E-mail room will be equipped with a limited number of personal computers. SSH will be available on all systems. There will be a laser printer on the LAN as well. One of the personal computers and the printer are designated for the use of speakers for last minute changes to presentations. The E-mail room will not be available after 4:00 p.m. on Thursday.

E-mail Room hours are:

Monday–Wednesday..... 7:00 a.m.–9:00 p.m.

Thursday..... 7:00 a.m.–4:00 p.m.

Speaker Preparation

We invite speakers to use the personal computers and the printer in the E-mail room for last minute changes. We will provide video projectors in all technical and general session rooms. Please find time before sessions and during breaks to test your laptop connection to the video projectors. Please come to the Conference Office if you need assistance.

Photocopying

A copy machine for making a limited number of copies is available in the Conference Office. If you plan to distribute copies of your presentation, please bring sufficient copies with you. You may also have copies made at the hotel business office for a charge.

Voltage

The United States operates on a 120V system. Those traveling from outside the US may need an adaptor. Check the Voltage Valet at voltagevalet.com/idx.html for information on requirements and to find adapters.

The CNN Center Atrium.

Local Arrangements

Travel Information

Atlanta Airport

The Hartsfield-Jackson Atlanta International Airport (ATL) is about 12 miles from the conference hotel.

A taxi to the airport should cost about $30 one way, plus tip. Additional passengers should add about $2 per person.

Airport Shuttle Services

There is no free Omni Hotel shuttle service, but the Atlanta Link at www.theatlantalink.com provides shared-ride shuttles to the airport for about $17 per person, or $30 for a round trip.

Parking

The only parking available at the Omni is indoor valet parking, currently at $30 per night. For other nearby parking, ask the hotel concierge for assistance.

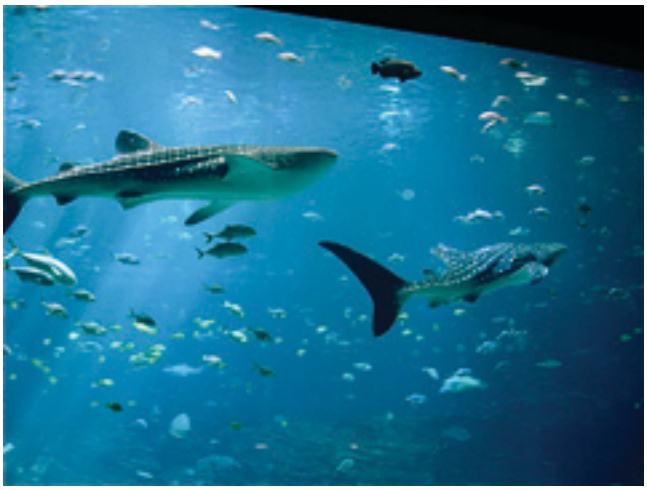
Many restaurants and attractions are within easy walking distance, even within the CNN Center itself, so a car is not necessary.

Ground Transportation

MARTA, the Atlanta subway, has stops at the airport and at CNN Center. One-way trips cost just $2.25.

Rental cars are available from:

Avis	(404) 530-2725
Budget Car Rental	(404) 530-3000
Dollar Rent A Car	(866) 434-2226
Enterprise	(404) 763-5220
Hertz	(404) 530-2925
National/Alamo	(404) 530-2800
Thrifty Car Rental	(770) 996-2350



The Georgia Aquarium is walking distance from the Omni Hotel.

Currency

Foreign currency can be exchanged at banks near the hotel. At the Atlanta airport Currency Exchange Travelex has two currency exchange locations: in the Atrium–main terminal and Concourse "E" near gate 26. Also look for mobile carts at Concourse "E". The Omni Hotel also has ATM, check cashing, and currency exchange services. Go to the Universal Currency Converter™ web site located at www.xe.net/ucc/ for help with foreign exchange rates.

Weather

May is a beautiful time of year in Atlanta. Temperatures will be warm mid 70's F (mid 20's C).

Tourist Information

Tourist information is available from the CUG 2009 Local Arrangments pages on cug.org. Brochures and additional information are available at the hotel.

Social Events

Cray Inc. Social

Cray Inc. will host a social on Tuesday evening in the Omni Hotel Atrium Terrace B. All conference participants and their guests are invited to attend.

CUG Night Out

The Local Arrangements Committee invites all attendees to enjoy the CUG Night Out at Dave & Buster's, Marietta, GA. This last minute change in venue was beyond our control, so please watch and listen for updates at the conference.

Dave & Busters is an entertainment and dining venue with lots of fun activities. From the classics to the latest in cutting-edge action games, there are hundreds of games to play in their Million Dollar Midway. And hundreds of prizes to take home, too. Cash in your tickets at the Winner's Circle and take some of the fun with you.

You will receive 2 drink tickets and a "Power Play Card," worth $15.00, but tickets are half price on Wednesday's, so you will really be getting a card worth $30.00. Dinner is a special Italian Feast Buffet.

Buses depart from the Omni starting at 5:30 p.m. The last bus leaves at 6:30 p.m.

We will arrive at Dave & Busters between 6:30 and 7:30 p.m. (It will take about an hour to get there due to traffic at that time of day.) Buses depart from Dave & Buster's starting at 9:30. The last bus leaves at 10:00 p.m. We will arrive back at the Omni around 10:00 to 10:30.

Contacts

CUG Board of Directors

President
David Gigrich
The Boeing Company
P.O. Box 3707
MS 7R-47
Seattle, WA 98124-2207 USA
(1-425) 865-6342 Fax: (1-425) 865-6844
david.j.gigrich@boeing.com

Vice-President
Nicholas P. Cardo
National Energy Research Scientific Computing Center
1 Cyclotron Rd, Bldg 943r0256
Berkeley, CA 94720 USA
(1-510) 486-4765 Fax: (1-510) 486-4316
cardo@nersc.gov

Treasurer
Mark Fahey
National Institute for Computational Sciences
1 Bethel Valley Road
PO Box 2008, MS 6008
Oak Ridge, TN 37831-6008 USA
(1-865) 574-5622 Fax: (1-865) 241-2850
faheymr@ornl.gov

Secretary
James Kasdorf
Pittsburgh Supercomputing Center
300 South Craig Street #304
Pittsburgh, PA 15213 USA
(1-412) 268-6802 Fax: (1-412) 268-1490
kasdorf@psc.edu

Director-at-Large
Mike Ashworth
HPCX Consortium
STFC Daresbury Laboratory
Warrington, Cheshire WA4 4AD United Kingdom
(44-01925) 603178 Fax: (44-01925) 603634
mike.ashworth@stfc.ac.uk

Director-at-Large
Liam Forbes
Arctic Region Supercomputing Center
P.O. Box 756020
909 Koyukuk Drive, Suite 105
Fairbanks, AK 99775 USA
(1-907) 450-8618
lforbes@arsc.edu

Director-at-Large
Ladina Gilly
CSCS–Swiss National Supercomputing Centre
Galleria 2
Via Cantonale
Manno 6928 Switzerland
(41-91) 610 82 67 Fax: (41-91) 610 83 08
lgilly@cscs.ch

Director-at-Large
Claire Hepwood
Atomic Weapons Establishment
E2.1, AWE Aldermaston
Reading RG7 4PR United Kingdom
(44-118) 9826886
claire.hepwood@awe.co.uk

Conference Chair
James B. White III (Trey)
Oak Ridge National Laboratory
1 Bethel Valley Road
Oak Ridge, TN 37831-6008 USA
(1-865) 241-2103 Fax: (1-865) 241-2850
trey@ornl.gov

CRAY USER GROUP INCORPORATED

Contacts

Special Interest Groups

Special Interest Groups (SIGs) are organized by subject areas rather than by platform. The Applications and Programming Environment SIG deals naturally with Applications and Programming Environments issues. The Systems Support SIG covers topics related to Operations, Systems & Integration, and Operating Systems. The User Support SIG deals with issues related to User Services. The Legacy Systems SIG provides a forum for discussions relating to systems no longer sold but still supported by Cray. This year, a new XTreme SIG works very closely with Cray, under Non-Disclosure Agreements, to provide valuable input into the Cray XT system development cycle.

Applications and Programming Environments

Chair: Robert Ballance (SNLA)
Deputy Chair: Rolf Rabenseifner (HLRS)

Cray Inc. SIG Liaisons
Applications: Jef Dawson
Programming Environments: Luiz DeRose

Systems Support

Chair: Joni Virtanen (CSC)
Deputy Chair: Jim Glidewell (BOEING)

Cray Inc. SIG Liaisons
Systems & Integration: Jim Harrell
Operating Systems: Charlie Carroll
Operations: Charlie Clark

User Support

Chair: Jim Glidewell (BOEING)
Deputy Chair: Yun (Helen) He (NERSC)

Cray Inc. SIG Liaison
User Services: Nola Van Vugt

XTreme Systems

Chair: Ann Baker (ORNL)

Cray Inc. SIG Liaison
XTreme Systems: Vito Bongiorno

Legacy Systems

Chair: James Kasdorf (PITTSC)
Deputy Chair: Open

Cray Inc. SIG Liaisons
TBD

SIG Chairs and Cray Inc. SIG Liaisons

Ann Baker
Oak Ridge National Laboratories
Bldg. 5600, MS 6016
Oak Ridge, TN 37831-6016 USA
(1-865) 574-6324
bakerae@ornl.gov

Robert Ballance
Sandia National Laboratories
P.O. Box 5800, MS 0807
Albuquerque, NM 87185-0807 USA
(1-505) 284-9361
raballa@sandia.gov

Charlie Carroll
Cray Inc.
1340 Mendota Heights Road
Mendota Heights, MN 55120 USA
(1-651) 605-8987 Fax: (1-651) 605-9001
charliec@cray.com

Charlie Clark
Cray Inc.
1050 Lowater Road
P.O. Box 6000
Chippewa Falls, WI 54729 USA
(1-715) 726-4020 Fax: (1-715) 726-4739
cmc@cray.com

Jef Dawson
Cray Inc.
1340 Mendota Heights Rd.
Mendota Heights, MN 55120 USA
(1-651) 605-9036 Fax: (1-651) 605-9123
jef@cray.com

Luiz DeRose
Cray Inc.
1340 Mendota Heights Road
Mendota Heights, MN 55120 USA
(1-651) 605-8925 Fax: (1-651) 605-9001
ldr@cray.com

Contacts

Jim Glidewell
The Boeing Company
PO Box 3707, MC 7M-RM
Seattle, WA 98124-2207 USA
(1-425) 865-1516 Fax: (1-425) 865-1768
james.glidewell@boeing.com

Jim Harrell
Cray Inc.
1340 Mendota Heights Rd.
Mendota Heights, MN 55120 USA
(1-651) 605-9131 Fax: (1-651) 605-9001
ejh@cray.com

Yun (Helen) He
National Energy Research Scientific Computing Center
Lawrence Berkeley National Laboratory
1 Cyclotron Road, MS 943
Berkeley, CA 94720 USA
(1-510) 486-5180 Fax: (1-410) 486-4316
yhe@lbl.gov

James Kasdorf
Pittsburgh Supercomputing Center
300 South Craig Street #304
Pittsburgh, PA 15213 USA
(1-412) 268-6802 Fax: (1-412) 268-1490
kasdorf@psc.edu

Rolf Rabenseifner
High Performance Computing Center Stuttgart
Nobelstr. 19
D-70550 Stuttgart Germany
(49-7) 685 65530 Fax: (49-7) 685 65832
rabenseifner@hlrs.de

Joni Virtanen
CSC–IT Center for Science Ltd.
Keilaranta 14
P.O. Box 405
FI-2101 Espoo Finland
(358-9) 4570009 Fax: (358-9) 4572302
Joni.Virtanen@csc.fi

Nola Van Vugt
Cray Inc.
411 First Avenue S., Suite 600
Seattle, WA 98104 USA
(1-206) 701-2239 Fax: (1-206) 701-2500
nolav@cray.com

Cray Inc. Representatives to CUG

Vito Bongiorno, Business Liaison
Cray Inc.
1340 Mendota Heights Road
Mendota Heights, MN 55120 USA
(1-651) 605-8935 Fax: (1-651) 605-9123
vb@cray.com

Ellen Roder, Program Liaison
Cray Inc.
1340 Mendota Heights Road
Mendota Heights, MN 55120 USA
(1-651) 605-9058 Fax: (1-651) 605-9001
ellen@cray.com

Steve Scott, Technical Liaison
Cray Inc.
1050 Lowater Road
P.O. Box 6000
Chippewa Falls, WI 54729 USA
(1-715) 726-4739 Fax: (1-715) 726-4739
sscott@cray.com

Linda Yetzer, Arrangements Liaison
Cray Inc.
1340 Mendota Heights Road
Mendota Heights, MN 55120 USA
(1-651) 605-8961 Fax: (1-651) 605-9123
yetzerll@cray.com

CUG Office

Bob and Karen Winget
Toolbox Conference Management
3224 Southwood Drive
Philomath, OR 97370 USA
(1-541) 929-3776 Fax: (1-509) 272-7239
bobwinget@toolboxcm.com kwinget@fpes.com

Oak Ridge National Laboratories pond.

CUG 2010

The Cray User Group invites you to its annual technical conference, CUG 2010, in Edinburgh, Scotland, in May 2010. EPCC–the supercomputing centre at the University of Edinburgh–is pleased to host this event. The exact dates and the venue of the conference will be announced during CUG 2009.

EPCC, (www.epcc.ed.ac.uk), a leading European centre of expertise in advanced research, technology transfer and the provision of supercomputer services to academia and business, is proud to host HECToR: a Cray XT5h hybrid system based on XT4 and X2 components, which acts as the UK's main national supercomputing service and is used to facilitate world leading research in a wide range of disciplines. In the summer of 2009, an upgrade will increase the peak performance of the XT4 component to 208 TFlops and HECToR will become one of Europe's top supercomputers.

In order to take advantage of the lower fee we encourage you to register early. You will find all the necessary information about the venue, the hotel, plus Edinburgh and its surroundings using the links from the cug.org web site. Registration will be open by December 1, 2009.

Be sure to attend the Closing Session on Thursday at 3:00 pm when we will present information about our plans for CUG 2010.

We look forward to welcoming you to Edinburgh, one of the most spectacular and beautiful cities in the world.

Sincerely,

Professor Arthur Trew
EPCC Director

Dr Alan Gray
Local Arrangements Chair

Call for Papers

You and your colleagues from CUG sites around the world hold the key to sustaining this conference as a source of information and technical interchange. Please begin now to consider presenting at next year's conference in Edinburgh. A year passes very quickly and our speaking requests have been growing consistently of late. Please share your experiences and help us work cooperatively toward solutions to problems encountered in the exciting venue of High Performance Computing.

As with this conference, Technical Presentations are invited in the following categories:

Applications and Programming Environments
- Programming Environment
- Compilers
- Libraries
- Tools
- 3rd Party Applications
- User Code Optimization

Systems Support
- Operations
- Environmental Monitoring
- Facilities and Site Preparation
- Tuning and OS Optimizations
- System Operations
- Architecture
- Mass Storage
- Networking

User Support
- Account Administration
- Consulting
- Documentation
- Training

Legacy Systems
- Cray X1 Series
- Cray XD

Take note of the presentations given this year, and discuss possible contributions with your colleagues. Seek advice from the Special Interest Group Chair or Deputy Chair. Work on your idea and prepare a draft to discuss at your home organization. Look for the official Call For Papers reminders you will receive in the fall and be ready to submit your abstract to www.cug.org following the instructions for publications listed therein. Help us make CUG in 2010 the best yet. Hope to see you in Edinburgh!

CUG 2009 Sponsors

The LAC would like to thank all of the people who have assisted in making this conference possible. We would especially like to thank our sponsors for their support.

Cray Inc. and ***AMD*** for sponsoring our CUG Night Out

DataDirect Networks and ***LSI*** for their sponsorship of the lunches

Altair, ClusterResources, PGI, Platform, and ***Totalview*** for their sponsorship of the coffee breaks

Finally, I would like to thank ***Cray Inc.*** for their assistance and support throughout the entire planning process as well as the CUG Board and CUG Office for their guidance and patience.

James B. White III (Trey), Local Arrangements Chair

Diamond

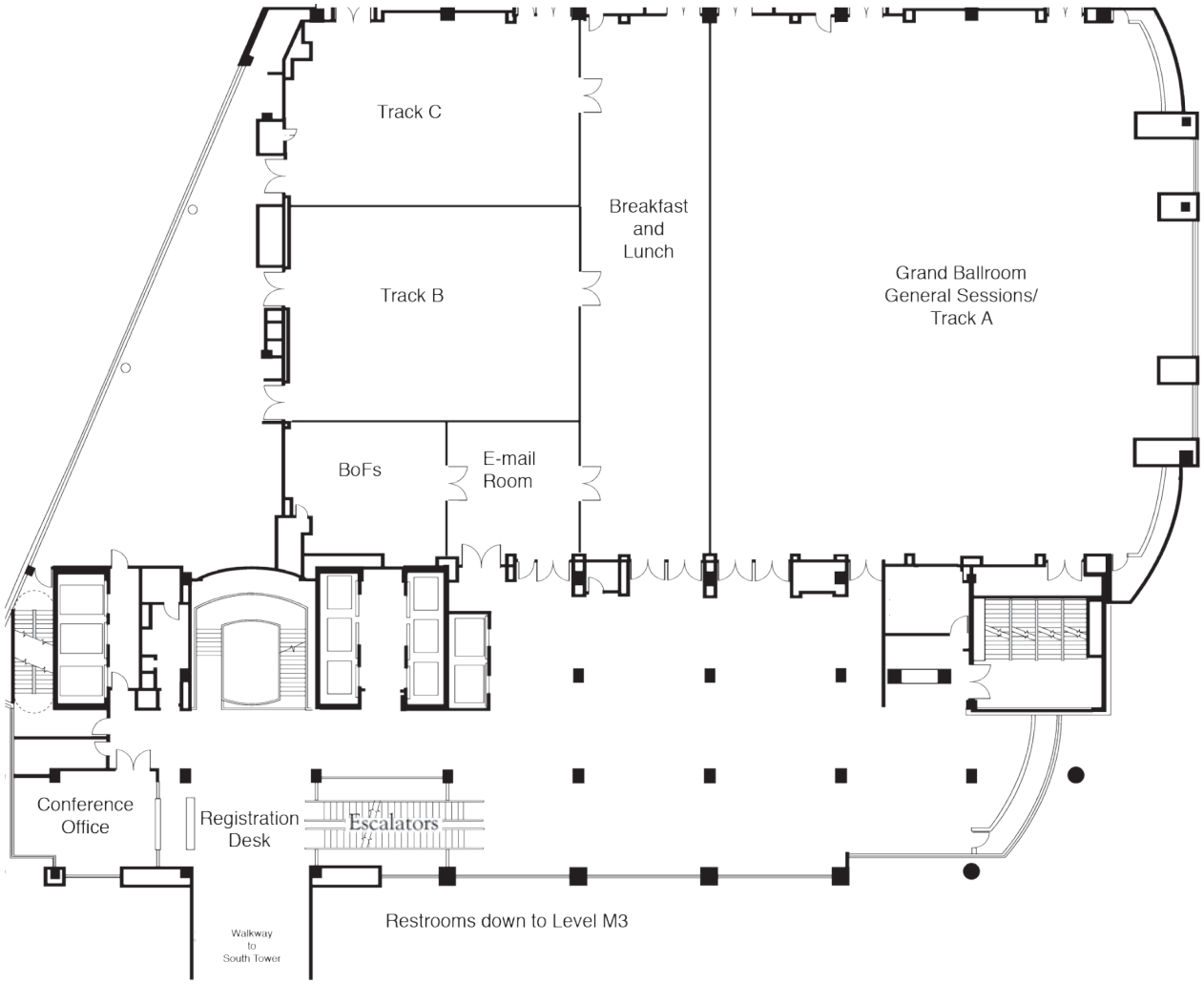
Platinum

Gold

Silver

Meeting Room Map

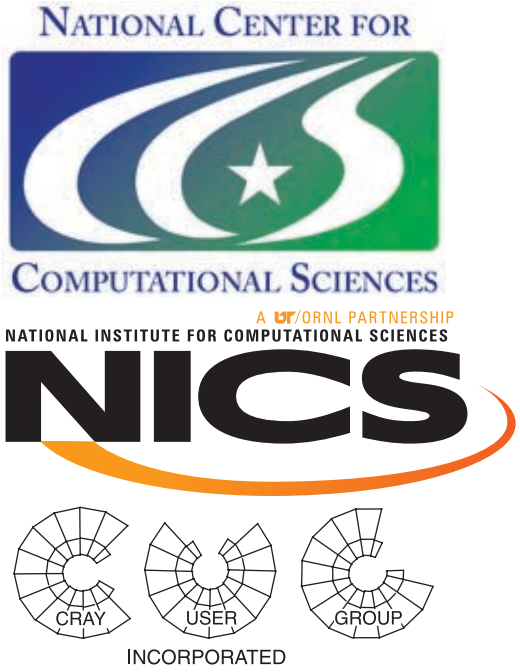
Omni Hotel North Tower
M4/Grand Ballroom Level



CUG 2009 Meeting Rooms

Event/Facility	Room
General Sessions	Grand Ballroom D2 and E
Track A	Grand Ballroom D2 and E
Track B	Grand Ballroom B
Track C	Grand Ballroom C
Lunch	Grand Ballroom D1
Conference Registration	Grand Ballroom Lobby
Refreshments	Grand Ballroom Lobby
E-mail & Speaker Prep	Grand Ballroom A (right)
BoFs	Grand Ballroom A (left)
Conference Office	M4 Behind Registration Desk
Cray Social	Omni Hotel Atrium Terrace B, South Tower

Online and printed program design and production by Toolbox Conference Management (toolboxcm.com)


NATIONAL CENTER FOR
COMPUTATIONAL SCIENCES
A UT/ORNL PARTNERSHIP
NATIONAL INSTITUTE FOR COMPUTATIONAL SCIENCES
NICS
CRAY
USER
GROUP
INCORPORATED