The premier international conference on high performance computing, networking, storage and analysis

Reno-Sparks Convention Center
Reno, Nevada November 10-16, 2007

http://sc07.supercomputing.org
<table>
<thead>
<tr>
<th>Page</th>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Chair’s Welcome</td>
</tr>
<tr>
<td>7</td>
<td>What’s New at SC07</td>
</tr>
<tr>
<td>9</td>
<td>General Information</td>
</tr>
<tr>
<td>10</td>
<td>Registration Desk/Conference Store</td>
</tr>
<tr>
<td>10</td>
<td>Registration Levels</td>
</tr>
<tr>
<td>11</td>
<td>Student Volunteers</td>
</tr>
<tr>
<td>12</td>
<td>Pass Access</td>
</tr>
<tr>
<td>13</td>
<td>Hours of Operation</td>
</tr>
<tr>
<td>13</td>
<td>Social Events</td>
</tr>
<tr>
<td>14</td>
<td>Facilities</td>
</tr>
<tr>
<td>15</td>
<td>Conference Hotels</td>
</tr>
<tr>
<td>17</td>
<td>Schedule/Maps</td>
</tr>
<tr>
<td>33</td>
<td>Tutorials</td>
</tr>
<tr>
<td>47</td>
<td>Invited Speakers</td>
</tr>
<tr>
<td>49</td>
<td>Invited Speakers</td>
</tr>
<tr>
<td>53</td>
<td>Masterworks</td>
</tr>
<tr>
<td>67</td>
<td>Papers</td>
</tr>
<tr>
<td>95</td>
<td>New &amp; Noteworthy</td>
</tr>
<tr>
<td>95</td>
<td>Posters</td>
</tr>
<tr>
<td>113</td>
<td>ACM Student Research Competition</td>
</tr>
<tr>
<td>115</td>
<td>Doctoral Research Showcase</td>
</tr>
<tr>
<td>121</td>
<td>Birds-of-A-Feather Sessions</td>
</tr>
<tr>
<td>Page</td>
<td>Section</td>
</tr>
<tr>
<td>------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>141</td>
<td>Awards &amp; Challenges</td>
</tr>
<tr>
<td>142</td>
<td>Gordon Bell Prize</td>
</tr>
<tr>
<td>144</td>
<td>Seymour Cray Award</td>
</tr>
<tr>
<td>145</td>
<td>Sidney Fernbach Award</td>
</tr>
<tr>
<td>146</td>
<td>Cluster Challenge</td>
</tr>
<tr>
<td>148</td>
<td>Analytics Challenge</td>
</tr>
<tr>
<td>149</td>
<td>Storage Challenge</td>
</tr>
<tr>
<td>151</td>
<td>Bandwidth Challenge</td>
</tr>
<tr>
<td>155</td>
<td>Panels &amp; Workshops</td>
</tr>
<tr>
<td>155</td>
<td>Panels</td>
</tr>
<tr>
<td>159</td>
<td>Workshops</td>
</tr>
<tr>
<td>165</td>
<td>Disruptive Technologies</td>
</tr>
<tr>
<td>167</td>
<td>Education Program</td>
</tr>
<tr>
<td>175</td>
<td>Broader Engagement</td>
</tr>
<tr>
<td>179</td>
<td>Exhibitor Events</td>
</tr>
<tr>
<td>180</td>
<td>Disruptive Technologies</td>
</tr>
<tr>
<td>182</td>
<td>Exhibitor Forum</td>
</tr>
<tr>
<td>201</td>
<td>SCinet</td>
</tr>
<tr>
<td>209</td>
<td>Acknowledgments</td>
</tr>
</tbody>
</table>
Welcome to Reno, Nevada, the host of the 20th annual international conference. SC07 represents the premier forum for the exchange of information and highlights on the most innovative developments in high performance computing (HPC), storage, networking and analysis. The conference is sponsored by the Association for Computing Machinery (ACM) and the IEEE Computer Society. The SC07 Conference Committee has worked very hard to prepare this conference for your scientific exchange, education, and enjoyment. A complete listing of the various offerings over this week is laid out for you in the pages of this program. There is so much to choose from this week that we feel confident you will find your days very full of activities.

SC07 continues in the SC conference tradition of bringing together scientists and engineers, researchers, educators and managers from the world’s leading computing and network installations and companies. It showcases innovative developments that are sparking new ideas and new industries. This week is the culmination of three years of planning by an outstanding committee of your peers. The result is the best SC conference ever! First and foremost, an extraordinary technical program is in store for you. The research portions of the technical program this year reflect ever-higher levels of selectivity, while the invited presentations feature an exceptional lineup of speakers. You will be treated to a selection of 54 refereed papers, a series of excellent plenary sessions, 11 independently planned workshops and 25 tutorials. Seven panel sessions have been selected with an eye toward spanning many of today’s “hot” HPC topics. We have teamed with the U.S. Council on Competitiveness to bring you a set of Masterworks sessions which include industry speakers from global corporations who will discuss complex, real-world problems that demand high performance computing solutions. A new feature at SC07 is the Doctoral Research Showcase. This session provides a venue for 12 Ph.D. students who will be graduating in the next 12 months to present a short summary of their research. Looking further ahead than most forecasts commonly do, two panel sessions and five exhibits on Disruptive Technologies will suggest and display those innovations most likely to overturn today’s dominant HPC technologies in just a few years. A record number of Birds-of-a-Feather submissions will provide you with highlights on a wide variety of technology and software topics. The Storage, Analytics and Bandwidth Challenges have been reshaped, and you
will find fierce competition in all of these areas. An exciting new challenge—the Cluster Challenge—has been added and will showcase the computational power that is easily in reach to anyone today. In this challenge, teams of undergraduate students will assemble small clusters on the Exhibit floor and run benchmarks and applications selected by industry and HPC veterans. To round out the technical program, in the hallway in front of the ballroom, we will have 39 regular posters and six student posters on display. Come to our poster reception on Tuesday evening from 5:15 p.m. - 7:00 p.m. to discuss the research displays with presenters in a casual setting.

The Industry and Research Exhibits in the Reno-Sparks Convention Center occupy over 120,000 square feet of net space—more than any previous SC conference—270 companies and organizations are on-site to offer the most up-to-date information on products and research results. You will have plenty of space in the exhibit halls to network with the exhibitors and your colleagues. In the exhibit hall, please take time to visit the students working on the Cluster Challenge and experience the Disruptive Technologies exhibits that are in Exhibit Hall 1. A highlight of the Exhibits program is the Exhibitor Forum. These talks showcase the latest advances by our Industry Exhibitors, including new products and upgrades, recent research and development and future plans and roadmaps. As in previous years, SCinet has leveraged new advances in network technology to bring you a unique infrastructure for research, demonstrations and general networking.

The Education Program is the start of a new three-year initiative. This year’s program started in the summer with nine week-long summer workshops held at colleges and universities in nine states. Four were held at a minority-serving institutions and one was held at a women’s college. At SC07, there will be 110 participants from 50 colleges, universities and high schools in 21 states, as well as Puerto Rico and Santo Domingo, who will have four days of discussions and hands-on applications of computational science resources, tools, and methodologies to enhance the preparation of the next generation of scientists, technologists, engineers and mathematicians. The Learning and Physical Challenges Education Program is combined with the general Education Program this year. It will engage a dozen educators in ensuring that all students are able to fully participate in the use of computational science tools to enhance learning.

SC07 is committed to broadening the engagement of individuals from groups that have traditionally been under-represented in high performance computing. To do this, we are kicking off a multi-year Broader Engagement (BE) program. Our Broader Engagement initiative provides a number of opportunities and activities from groups that have traditionally been under represented in high performance computing, such as African Americans, Hispanics, Indigenous People, and
women. One of the BE activities that will be useful to everyone is the information kiosk which Broader Engagement will staff with volunteers who can answer questions. They will provide guidance to first-time SC attendees, to support them in determining which conference events and resources best fit their needs and interests. Finally, there is plenty of time to exchange ideas, celebrate past successes and plan for the future with colleagues and friends, both old and new. So, welcome to SC07 and Reno and please enjoy all of the activities we have lined up for you during this exciting week. The SC07 Committee hopes you will find this to be a great learning experience!

Becky Verastegui
SC07 GENERAL CHAIR

WELCOME!
While many components of the SC conference series remain constant from year to year, the organizing committee tries to introduce new ideas along with new perspectives. Four new offerings this year are Doctoral Research Showcase, Cluster Challenge, Broader Engagement, and HPC Ph.D. Fellowship Program. Brief descriptions of these are presented below.

**Doctoral Research Showcase**
This session provides a venue for Ph.D. students who will be graduating in the next 12 months to present a short summary of their research. From the 48 proposals received, we selected 12 students to make presentations in two different sessions. To accommodate as many people as possible, presentations are limited to 15 minutes each. We encourage you to attend this session to see what the next generation of HPC researchers is up to.

**Cluster Challenge**
Did you know that a small cluster today (less than 1/2 rack) would top the Top500.org website from just ten years ago? The computational power that is easily within reach today significantly surpasses that available only to the national labs from that time. This challenge showcases the significance of this and highlights how accessible clusters are to anyone today. In this challenge, teams of undergraduate students will assemble a small cluster on the exhibit floor and run benchmarks & applications selected by industry and HPC veterans. The challenge will be judged on the speed of benchmarks and the throughput of application runs over the first three days of the conference.

**Broader Engagement Initiative**
The Broader Engagement Initiative is committed to broadening the engagement of individuals from groups that have traditionally been underrepresented in the field of high performance computing and networking. The Initiative is providing grants to support participation in the technical program, encouraging technical program submissions, and fostering networking through both a formal mentoring program and informal contact at SC07.

**High Performance Computing Ph.D. Fellowship Program**
The Association of Computing Machinery (ACM), IEEE-Computer Society and the SC Conference Series are pleased to announce the first year of the High Performance Computing Ph.D. Fellowship Program. The program honors exceptional Ph.D. students throughout the world focusing on research in High Performance Computing, Networking, Storage and Analysis and also supports our long-standing commitment to workforce diversity and encourages nominations of women, minorities and all who contribute to diversity. Students must be nominated by a full-time faculty member at a Ph.D. granting, accredited institution. More information is available at: www.sigarch.org/HPC_Fellowships.html.
**SC07** is the premier international conference on high performance computing, networking, storage, and analysis. Following the traditions set with the first SC conference in 1988, SC07 is expanding the frontiers of high performance computing by introducing a new cluster challenge that combines the computation, networking, storage, and analysis elements into a single competition.

SC07 will host an exceptional technical program, tutorials, workshops, an expanded exhibits area, an exciting education program and many other activities. The SC conference series has grown to include scientists, researchers, software developers, network engineers, policy makers, corporate managers, CIOs and IT administrators from universities, industry and government from all over the world.

Attendees become immersed in the latest state-of-the-field developments in technology, applications, vendor products, research results, national policy and national/ international initiatives. SC07 is the one place attendees can see tomorrow’s technology being used to solve world-class challenge problems today. The SC07 conference ties the program components together and demonstrates how high performance computing, networking, storage and analysis touch all disciplines to enhance people’s ability to understand information and lead to new understanding, promote interdisciplinary projects, affect the educational process through the use of computers in modeling and simulation in the classroom, and solve heretofore unsolvable problems in nanoscience, biotechnology, climate research, astrophysics, chemistry, fusion research, drug research, homeland defense, nuclear technologies and many other fields.

SC07 provides attendees the opportunity to speak with industry and academic research leaders through the technical program, industry exhibits, research exhibits and Birds-of-a-Feather sessions. The SC07 Education Program is designed to work with undergraduate faculty, administrators and college students to integrate computational science and the high performance computing and communications technologies highlighted through the SC Conference into the preparation of future scientists, technologists, engineers, mathematicians, and teachers.

We are initiating a new Broader Engagement program with the desired goal of increasing the participation of people from underrepresented groups in the SC conference.

We look forward to your participation.
Registration Desk/Conference Store

The registration desk and conference store are located in the convention center lobby.

Registration Desk & Conference Store Hours

<table>
<thead>
<tr>
<th>Day</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saturday, Nov 10</td>
<td>1:00 p.m. - 6:00 p.m.</td>
</tr>
<tr>
<td>Sunday, Nov 11</td>
<td>7:30 a.m. - 6:00 p.m.</td>
</tr>
<tr>
<td>Monday, Nov 12</td>
<td>7:30 a.m. - 9:00 p.m.</td>
</tr>
<tr>
<td>Tuesday, Nov 13</td>
<td>7:30 a.m. - 6:00 p.m.</td>
</tr>
<tr>
<td>Wednesday, Nov 14</td>
<td>7:30 a.m. - 6:00 p.m.</td>
</tr>
<tr>
<td>Thursday, Nov 15</td>
<td>7:30 a.m. - 5:00 p.m.</td>
</tr>
<tr>
<td>Friday, Nov 16</td>
<td>8:00 a.m. - 11:00 a.m.</td>
</tr>
</tbody>
</table>

Registration Levels

Tutorials

Full-day and half-day tutorials are offered on Sunday and Monday, November 11-12. Tutorials are not included in the Technical Program registration and require separate registration and fees. Attendees may choose a one-day or two-day passport, allowing them to move freely between tutorials on the selected day(s). Tutorial notes and luncheons are provided for each registered tutorial attendee.

Tutorial Notes

This year each registered tutorial attendee will receive a copy of all the tutorial notes on a computer-readable medium; no hardcopy notes will be distributed or available. Some of the tutorials will have hands-on components. For these, attendees must bring their own laptops with SSH software installed. For hands-on tutorials, the rooms will have wired network drops, Ethernet cables, SCinet wireless, and power drops, but there will be no computer support available. Please arrive early, as there may be tutorial-specific software to install on your laptop.

Technical Program

Technical Program registration provides access to plenary talks, posters, panels, BOFs, papers, exhibits, challenges, awards, Masterworks, the new Doctoral Showcase, and workshops.

Exhibitor

Exhibitor registration provides attendees whose companies have booths with access to the Exhibit floor and limited technical program events.

Exhibits Only

Exhibits Only registration provides access to the exhibit floor for all three days of the exhibition (but not the Monday Night Gala Opening). Children under age 12 are not permitted on the floor except during Family Hours (4-6 p.m., Wednesday, November 14), when they must be accompanied by a family member who is a registered conference attendee.

Education Program

Education Program registration provides access to all events except Tutorials and the show floor.
Proceedings
Attendees registered for the Technical Program will receive one copy of the SC07 proceedings on a USB flash drive.

Lost Badge
There is a $40 processing fee to replace lost badges.

Member, Retired Member and Student Registration Discounts
To qualify for discounted registration rates, present your current IEEE, IEEE Computer Society, ACM, or ACM SIGARCH membership number or a copy of a valid full-time student identification card when registering. You may complete the IEEE Computer Society and/or ACM membership application provided in your conference bag and return it to the Special Assistance desk in the registration area to receive the member discounted registration rate.

Student Volunteers
Undergraduate and graduate student volunteers assist with the administration of the conference and receive, in exchange, free conference registration, housing for out-of-town volunteers and most meals.

Student volunteers have the opportunity to see and discuss the latest high-performance networking and computing technology and meet leading researchers from around the world while contributing to the success of this annual event. Conference attendees are encouraged to share information about the Student Volunteers program with their colleagues and to encourage their students to apply for future conferences.
**Pass Access**
Each registration category provides access to a different set of conference activities, as summarized below.

<table>
<thead>
<tr>
<th>Type of Event</th>
<th>Tutorials (on day of passport)</th>
<th>Technical Program</th>
<th>Exhibitor Only</th>
<th>Exhibits</th>
<th>Education Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education Program Meals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>•</td>
</tr>
<tr>
<td>All Tutorial Sessions</td>
<td>•</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tutorial Lunch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exhibitor Party</td>
<td></td>
<td></td>
<td>•</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monday Exhibits Gala Opening</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Tuesday Keynote</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Tuesday Poster Reception</td>
<td></td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wednesday Plenary (Cray &amp; Fernbach Awards)</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td></td>
</tr>
<tr>
<td>Thursday Plenary (Invited Speakers)</td>
<td></td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thursday Night Reception</td>
<td></td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birds-of-a-Feather</td>
<td></td>
<td>•</td>
<td>•</td>
<td>•</td>
<td>•</td>
</tr>
<tr>
<td>Challenge Presentations</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exhibitor Forum</td>
<td></td>
<td>•</td>
<td>•</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exhibit Floor</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masterworks</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panels (Friday Only)</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panels (Except Friday)</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Papers</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posters</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCinet Access</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Workshops</td>
<td></td>
<td>•</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Hours of Operation**

**Exhibit Floor Hours**

<table>
<thead>
<tr>
<th>Date</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tues., Nov. 13</td>
<td>10:00 a.m. - 6:00 p.m.</td>
</tr>
<tr>
<td>Weds., Nov. 14</td>
<td>10:00 a.m. - 6:00 p.m.</td>
</tr>
<tr>
<td>Thurs., Nov. 15</td>
<td>10:00 a.m. - 4:00 p.m.</td>
</tr>
</tbody>
</table>

**Media Room**

**Location: Room M2 (Mezzanine)**

Media representatives or industry analysts should visit the Media Room for on-site registration. The SC07 Media Room provides resources to media representatives for writing, researching and filing their stories, or interviewing conference participants and exhibitors. The Media Room is also available to exhibitors who wish to provide materials to, or arrange interviews with, media representatives and industry analysts covering the conference.

**Media Room Hours:**

<table>
<thead>
<tr>
<th>Date</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sun., Nov. 11</td>
<td>1:00 p.m. - 4:00 p.m.</td>
</tr>
<tr>
<td>Mon., Nov. 12</td>
<td>9:00 a.m. - 6:00 p.m.</td>
</tr>
<tr>
<td>Tues., Nov. 13</td>
<td>8:00 a.m. - 6:00 p.m.</td>
</tr>
<tr>
<td>Weds., Nov. 14</td>
<td>9:00 a.m. - 6:00 p.m.</td>
</tr>
<tr>
<td>Thurs., Nov. 15</td>
<td>9:00 a.m. - 4:00 p.m.</td>
</tr>
</tbody>
</table>

**Social Events**

**Exhibitor Party**

On Sunday, November 11, 7 p.m. - 10 p.m., SC07 will host an exhibitor party for registered exhibitors and education program participants. Please join us for an entertaining evening at the National Bowling Stadium. Busing will start at 6:30 p.m. from the convention center and will continue throughout the evening. The National Bowling Stadium, located at 300 N. Center Street, is the only facility of its kind in the world dedicated to the sport of bowling.

**Gala Opening Reception**

On Monday, November 12, SC07 will host its annual Grand Opening in the Exhibits Hall from 7:00 p.m. to 9:00 p.m. This event is open to all Technical Program and Education Program registrants.

**Conference Reception**

The social event for SC07 technical program attendees will be a special, one-time performance of the Blue Man Group at the Grand Sierra Resort on Thursday, November 15. Blue Man Group is best known for their wildly popular theatrical shows and concerts which combine music, comedy and multimedia theatrics to produce a totally unique form of entertainment.

The Blue Man Group performance will be held in the Grand Theatre from 7:30 p.m. to 8:00 p.m. There will also be a reception in the Grand Sierra Ballroom, starting at 6:00 p.m., with cocktails and snacks,
followed by a buffet dinner back in the ballroom after the performance. Entry into the Ballroom will require a Technical Program badge. Additional buses will transport attendees from the convention center to the Grand Sierra Resort starting at 5:45 p.m., as well as regular hotel shuttles.

Entry into the Grand Sierra Theatre will require a separate wristband. Attendees can visit the Blue Man Group table at the convention center starting at 3 p.m. or the Grand Sierra Ballroom starting at 6:00 p.m. on Thursday, the day of the performance, and show a Technical Program badge to obtain a wristband. Tickets for non-technical program attendees will be available for purchase at the SC07 Conference Store for $100. Due to limited seating for the performance wristbands will be required for entrance into the theatre and doors will shut promptly at 7:25 p.m. The performance will also be available for viewing on two screens in the Grand Sierra Ballroom, with food and beverage service continuing (no wristband required).

After the performance, bus transportation from the Grand Sierra to other conference hotels will be available until 10 p.m.

Facilities

Conference Survey Gift
Technical Program attendees are encouraged to complete the conference survey and turn it in at the SC08 booth in the lobby of the convention center before noon on Friday, November 16, and receive a gift. Survey responses are used in planning future SC conferences. Members of the SC08 committee will be available in the booth to offer information and discuss next year’s conference in Austin, Texas.

Coat and Bag Check
There are self-service locations for hanging coats within the technical program rooms and in the lobby. In addition, there is a coat and bag check on the premises.

First Aid/Emergency Medical Team
The Reno-Sparks Convention Center provides an on-site first aid facility staffed with emergency medical professionals. In the event of a medical emergency, attendees are requested to contact the Reno-Sparks Convention Center immediately by dialing Guest Services from any house phone located in the facility. In addition, all uniformed security personnel are available to assist you in any emergency.

Wheelchair Accessibility
The Reno-Sparks Convention Center complies with ADA requirements and is wheelchair accessible. The center provides a limited number of complimentary wheelchairs on a first come, first served basis.
Conference Hotels

The following hotels have been designated as SC07 conference hotels.

**Grand Sierra Resort Casino**
2500 E. Second Street
775.789.2000
Wireless available in lobby; in room for $10.99 per day
3.5 miles from the convention center

**Circus Circus Reno Hotel Casino**
500 N. Sierra Street
775.329.0711
Internet access included in rate
3.5 miles from the Convention Center and is connected via a skyway to the Silver Legacy.

**Atlantis Casino Resort Spa**
3800 S. Virginia St.
775.825.4700
High speed Internet access: $9.99 per 24 hours
Across the street from the Reno-Sparks Convention Center

**Peppermill Reno Hotel Casino**
2707 S. Virginia Street
775.826.2121
Internet access $10.99 per day
Approximately 6 blocks from the Reno-Sparks Convention Center

**Silver Legacy Resort Casino**
407 N. Virginia Street
775.325.7401
Internet access included in rate
3.5 miles from the Reno-Sparks Convention Center

**John Ascuaga's Nugget Hotel Casino**
1100 Nugget Ave.
775.356.3300
Internet access $11.99 per day
Less than 12 minutes from the Reno-Sparks Convention Center
<table>
<thead>
<tr>
<th>Event Type</th>
<th>Time</th>
<th>Location</th>
<th>Session/Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tutorial</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A5</td>
<td>S01: Parallel Computing 101</td>
</tr>
<tr>
<td>Tutorial</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A4</td>
<td>S02: Parallel I/O in Practice</td>
</tr>
<tr>
<td>Tutorial</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A10</td>
<td>S03: Application Development Using Eclipse and the Parallel Tools Platform</td>
</tr>
<tr>
<td>Tutorial</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A2</td>
<td>S04: A Tutorial Introduction to High Performance Analytics and Workflow on Grids</td>
</tr>
<tr>
<td>Tutorial</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A20</td>
<td>S05: High Performance Computing on GPUs with CUDA</td>
</tr>
<tr>
<td>Tutorial</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A3</td>
<td>S06: Introduction to Globus</td>
</tr>
<tr>
<td>Tutorial</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A7</td>
<td>S07: Introduction to Scientific Workflow Management and the Kepler System</td>
</tr>
<tr>
<td>Tutorial</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A11</td>
<td>S08: Introductory Babel for Massive Supercomputing Software Integration</td>
</tr>
<tr>
<td>Tutorial</td>
<td>8:30 a.m. - 12:00 p.m.</td>
<td>A1</td>
<td>S09: HPC Challenge (HPCC) Benchmark Suite</td>
</tr>
<tr>
<td>Tutorial</td>
<td>8:30 a.m. - 12:00 p.m.</td>
<td>A6</td>
<td>S10: Hybrid MPI and OpenMP Parallel Programming</td>
</tr>
<tr>
<td>Tutorial</td>
<td>1:30 p.m. - 5:00 p.m.</td>
<td>A1</td>
<td>S11: Principles and Practice of Experimental Performance Measurement and Analysis of Parallel Applications</td>
</tr>
<tr>
<td>Tutorial</td>
<td>1:30 p.m. - 5:00 p.m.</td>
<td>A6</td>
<td>S12: Programming using the Partitioned Global Address Space (PGAS) Model</td>
</tr>
<tr>
<td>Workshop</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A8</td>
<td>First International Workshop on High Performance Reconfigurable Computing Technology and Applications</td>
</tr>
<tr>
<td>Workshop</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A9</td>
<td>Workshop on Performance Analysis and Optimization of High-End Computing Systems</td>
</tr>
<tr>
<td>Workshop</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>C2</td>
<td>Workshop on Grid Computing Portals and Science Gateways (GCE 2007)</td>
</tr>
<tr>
<td>Workshop</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>C3</td>
<td>Manycore and Multicore Computing: Architectures, Applications and Directions</td>
</tr>
<tr>
<td>Workshop</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>Atlantis Hotel-Ballroom A</td>
<td>High Performance Computing in China: Solution Approaches to Impediments for High Performance Computing</td>
</tr>
<tr>
<td>Workshop</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>Atlantis Hotel-Ballroom B</td>
<td>Petascale Data Storage Workshop</td>
</tr>
<tr>
<td>Event Type</td>
<td>Time</td>
<td>Location</td>
<td>Session/Title</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------</td>
<td>----------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Tutorial</strong></td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A3</td>
<td>M01: Advanced MPI</td>
</tr>
<tr>
<td><strong>Tutorial</strong></td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A7</td>
<td>M02: Debugging Parallel Application Memory Bugs with TotalView</td>
</tr>
<tr>
<td><strong>Tutorial</strong></td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A4</td>
<td>M03: Application Supercomputing Concepts</td>
</tr>
<tr>
<td><strong>Tutorial</strong></td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A20</td>
<td>M04: Reconfigurable Supercomputing</td>
</tr>
<tr>
<td><strong>Tutorial</strong></td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A11</td>
<td>M05: Component Software for High-Performance Computing: Using the Common Component Architecture</td>
</tr>
<tr>
<td><strong>Tutorial</strong></td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A10</td>
<td>M06: A Practical Approach to Performance Analysis and Modeling of Large-scale Systems</td>
</tr>
<tr>
<td><strong>Tutorial</strong></td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A6</td>
<td>M07: Designing High-End Computing Systems with InfiniBand and 10-Gigabit Ethernet iWARP</td>
</tr>
<tr>
<td><strong>Tutorial</strong></td>
<td>8:30 a.m. - 12:00 p.m.</td>
<td>A2</td>
<td>M08: Large Scale Visualization with ParaView 3</td>
</tr>
<tr>
<td><strong>Tutorial</strong></td>
<td>8:30 a.m. - 12:00 p.m.</td>
<td>A5</td>
<td>M09: Introduction to OpenMP</td>
</tr>
<tr>
<td><strong>Tutorial</strong></td>
<td>8:30 a.m. - 12:00 p.m.</td>
<td>A1</td>
<td>M10: Configuring and Deploying GridFTP for Managing Data Movement in Grid/HPC Environments</td>
</tr>
<tr>
<td><strong>Tutorial</strong></td>
<td>1:30 p.m. - 5:00 p.m.</td>
<td>A2</td>
<td>M11: Clustered and Parallel Storage System Technologies</td>
</tr>
<tr>
<td><strong>Tutorial</strong></td>
<td>1:30 p.m. - 5:00 p.m.</td>
<td>A5</td>
<td>M12: Advanced Topics in OpenMP</td>
</tr>
<tr>
<td><strong>Tutorial</strong></td>
<td>1:30 p.m. - 5:00 p.m.</td>
<td>A1</td>
<td>M13: Parallel Programming Using the Global Arrays Toolkit</td>
</tr>
<tr>
<td><strong>Workshop</strong></td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>C2</td>
<td>Workshop on Grid Computing Portals and Science Gateways (GCE 2007)</td>
</tr>
<tr>
<td><strong>Workshop</strong></td>
<td>8:30 a.m. - 12:00 p.m.</td>
<td>C3</td>
<td>Manycore and Multicore Computing: Architectures, Applications And Directions</td>
</tr>
<tr>
<td><strong>Workshop</strong></td>
<td>1:30 p.m. - 5:00 p.m.</td>
<td>C3</td>
<td>Parallel Computing with MATLAB</td>
</tr>
<tr>
<td><strong>Workshop</strong></td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>Atlantis Hotel-Ballroom A</td>
<td>Ultra-Scale Visualization</td>
</tr>
<tr>
<td><strong>Workshop</strong></td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>Atlantis Hotel-Ballroom B</td>
<td>2nd International Workshop on Virtualization Technologies in Distributed Computing</td>
</tr>
</tbody>
</table>
| **Disruptive Technologies** | 7:00 p.m. - 9:00 p.m. | Exhibit Hall 1B | Disruptive Technologies Exhibit:  
  • High Performance Optical Connectivity Based on CMOS Photonics Technology  
  • Optical Printed Circuit Board Technology and 200+200Gbps Transceiver  
  • NRAM as a Disruptive Technology  
  • Superconducting Quantum Computing System  
  • Disruptive Technology for Many-core Chip System |
<table>
<thead>
<tr>
<th>Event Type</th>
<th>Time</th>
<th>Location</th>
<th>Session/Title</th>
</tr>
</thead>
</table>
| Keynote    | 8:30 a.m. - 10:00 a.m. | Ballroom | Neil Gershenfeld  
Programming Bits and Atoms |
| Disruptive Technologies | 10:00 a.m. - 6:00 p.m. | Exhibit Hall 1B | Disruptive Technologies Exhibit  
• High Performance Optical Connectivity Based on CMOS Photonics Technology  
• Optical Printed Circuit Board Technology and 200+200Gbps Transceiver  
• NRAM as a Disruptive Technology  
• Superconducting Quantum Computing System  
• Disruptive Technology for Many-core Chip System Software and Logic Co-verification |
| Papers     | 10:30 a.m. - 12:00 p.m. | A1 / A6 | Computational Biology  
• A Preliminary Investigation of a Neocortex Model Implementation on the Cray XD1  
• Anatomy of a Cortical Simulator  
• Large-scale Maximum Likelihood-based Phylogenetic Analysis on the IBM BlueGene/L |
| Papers     | 10:30 a.m. - 12:00 p.m. | A2 / A5 | Network Switching and Routing  
• Age-Based Packet Arbitration in Large-Radix k-ary n-cubes  
• Performance Adaptive Power-aware Reconfigurable Optical Interconnects for HPC Systems  
• Evaluating Network Information Models on Resource Efficiency and Application Performance in Lambda-Grids |
| Papers     | 10:30 a.m. - 12:00 p.m. | A3 / A4 | System Performance  
• Using MPI File Caching to Improve Parallel Write Performance for Large-scale Scientific Applications  
• Virtual Machine Aware Communication Libraries for High Performance Computing  
• Investigation of Leading HPC I/O Performance using a Scientific-application-derived Benchmark |
| Masterworks | 10:30 a.m. - 12:00 p.m. | C1 / C2 / C3 | High Performance Racing  
• High Performance Computing: Shaping the Future of Formula One  
• HPC at Chrysler: Bringing NASCAR from the Race Track to Your Driveway |
| Challenge  | 10:30 a.m. - 12:00 p.m. | A10 / A11 | Analytics Challenge Finalists  
• Angle: Detecting Anomalies and Emergent Behavior from Distributed Data in Near Real Time  
• Cognitive Methodology-based Data Analysis System for Large-scale Data |
| Exhibitor Forum | 10:30 a.m. - 12:00 p.m. | A7 | Petascale  
• Production-ready Petascale Computing  
• Innovation beyond Imagination: The Road to PetaFLOPS Computing  
• Beyond Standards: A Look at Innovative HPC Solutions |
## Tuesday, November 13

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Time</th>
<th>Location</th>
<th>Session/Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exhibitor Forum</td>
<td>10:30 a.m. - 12:00 p.m.</td>
<td>A8</td>
<td>Petascale&lt;br&gt;• Scalable, Congestion-free, Low Latency 10 Gigabit Ethernet Fabrics for Compute and Storage Clusters&lt;br&gt;• Dynamic Ethernet Lightpaths: On-demand 10GbE and GbE Connections for Research Networks&lt;br&gt;• Design and Technology for Supercomputers and Grids: Growth of 10 GbE</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m.</td>
<td>A1 / A6</td>
<td>Federal Plan for Advanced Networking Research and Development</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m.</td>
<td>A10 / A11</td>
<td>Parallel File Systems BOF</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m.</td>
<td>A2 / A5</td>
<td>Adapting Legacy Software to Hybrid Multithreaded Systems</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m.</td>
<td>A20</td>
<td>Converged Fabrics: Opportunities and Challenges</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m.</td>
<td>A3 / A4</td>
<td>TORQUE Resource Manager and Moab: New Capabilities and Roadmap Forum</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m.</td>
<td>A2 / A5</td>
<td>NSF Cyber Enabled Discovery (CDI) : Challenges for the Scientific Community</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m.</td>
<td>A8</td>
<td>Multi-core Support in Resource Managers and Job Schedulers</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m.</td>
<td>A9</td>
<td>Grid Operating Systems Community Meeting</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m.</td>
<td>C1 / C2 / C3</td>
<td>The 2007 HPC Challenge Awards</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m.</td>
<td>D4</td>
<td>Career Paths: High Performance Computing Education and Training</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m.</td>
<td>D5</td>
<td>Cyberinfrastructure and Society: Creating Outreach Programs for the Public</td>
</tr>
<tr>
<td>Papers</td>
<td>1:30 p.m. - 3:00 p.m.</td>
<td>A1 / A6</td>
<td>Grid Scheduling&lt;br&gt;• Automatic Resource Specification Generation for Resource Selection&lt;br&gt;• Performance and Cost Optimization for Multiple Large-scale Grid Workflow Applications&lt;br&gt;• Inter-operating Grids through Delegated MatchMaking</td>
</tr>
<tr>
<td>Papers</td>
<td>1:30 p.m. - 3:00 p.m.</td>
<td>A2 / A5</td>
<td>Security and Fault Tolerance&lt;br&gt;• Automatic Software Interference Detection in Parallel Applications&lt;br&gt;• DMTTracker: Finding Bugs in Large-scale Parallel Programs by Detecting Anomaly in Data Movements&lt;br&gt;• Scalable Security for Petascale Parallel File Systems</td>
</tr>
<tr>
<td>Papers</td>
<td>1:30 p.m. - 3:00 p.m.</td>
<td>A3 / A4</td>
<td>System Architecture&lt;br&gt;• The Cray BlackWidow: A Highly Scalable Vector Multiprocessor</td>
</tr>
<tr>
<td>Event Type</td>
<td>Time</td>
<td>Location</td>
<td>Session/Title</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------</td>
<td>----------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Masterworks   | 1:30 p.m. - 3:00 p.m. | C1 / C2 / C3 | • GRAPE-DR: 2-Pflops Massively-Parallel Computer with 512-Core, 512-Gflops Processor Chips for Scientific Computing  
• A Case for Low-complexity MP Architectures |
| Masterworks   | 1:30 p.m. - 3:00 p.m. | C1 / C2 / C3 | Biofuels/Alternative Energy  
• Child’s First Words, Terrible Teens and My Boring Parents: Prediction in the Complex World of Crop Genetic Improvement  
• Diverse Energy Sources for Better Driving |
| Exhibitor Forum | 1:30 p.m. - 3:00 p.m. | A7       | • Fujitsu’s Solutions and Vision for High Performance Computing  
• NEC’s HPC Strategy: Consistency and Innovation  
• Cray Advances Adaptive Supercomputing Vision |
| Exhibitor Forum | 1:30 p.m. - 3:00 p.m. | A8       | Tools - Scheduling and Debugging  
• Node-level Scheduling Optimization for Multi-core Processors  
• What’s In Store: How Organizations Cope as they Mature From Clusters to Adaptive Computing  
• Debugging for PetaScale |
| Challenge     | 1:30 p.m. - 3:00 p.m. | A10 / A11 | Storage Challenge Finalists  
• ParaMEDIC: A Parallel Meta-data Environment for Distributed I/O and Computing  
• ZeSt: The Maximum Reliable TBytes/sec/$ for Petascale Systems  
• Astronomical Data Analysis with Commodity Components  
• Grid-oriented Storage: Parallel Streaming Data Access to Accelerate Distributed Bioinformatics Data Mining |
| Papers        | 3:30 p.m. - 5:00 p.m. | A1 / A6  | Microarchitecture  
• Variable Latency Caches for Nanoscale Processor  
• Data Access History Cache and Associated Data Prefetching Mechanisms  
• Scaling Performance of Interior-Point Method on Large-Scale Chip |
| Papers        | 3:30 p.m. - 5:00 p.m. | A2 / A5  | PDE Applications  
• Data Exploration of Turbulence Simulations using a Database Cluster  
• Parallel Hierarchical Visualization of Large Time-varying 3D Vector Fields  
• Low-Constant Parallel Algorithms for Finite Element Simulations using Linear Octrees |
| Masterworks   | 3:30 p.m. - 5:00 p.m. | C1 / C2 / C3 | Toward Perfect Product Design  
• A Grand Challenge: MultiCore and Industrial Modelling and Simulation  
• HPC in the Kitchen and Laundry: Optimizing Everyday Appliances for Customer Satisfaction and Market Share |
## Tuesday, November 13

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Time</th>
<th>Location</th>
<th>Session/Title</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Challenge</strong></td>
<td>3:30 p.m. - 5:00 p.m.</td>
<td>A10 / A11</td>
<td><strong>Bandwidth Challenge Finalists</strong>&lt;br&gt;• iWarp-based Remote Interactive Scientific Visualization&lt;br&gt;• Streaming Uncompressed 4k Video&lt;br&gt;• Distributed Data Processing over Wide Area Networks&lt;br&gt;• Phoebus&lt;br&gt;• A Virtual Earth TV Set via Real-time Data Transfer from a Supercomputer&lt;br&gt;• Using the Data Capacitor for Remote Data Collection, Analysis, and Visualization&lt;br&gt;• TeraGrid Data Movement with GPFS-WAN and Parallel NFS</td>
</tr>
<tr>
<td><strong>Exhibitor Forum</strong></td>
<td>3:30 p.m. - 5:00 p.m.</td>
<td>A7</td>
<td><strong>Network Management</strong>&lt;br&gt;• The Benefits of Multi-protocol Networking Architecture&lt;br&gt;• Managing Network and Storage Infrastructure as an Application Resource&lt;br&gt;• Application of Direct Execution Technology for Analysis of Packets in a High Speed Application Layer Scenario</td>
</tr>
<tr>
<td><strong>Exhibitor Forum</strong></td>
<td>3:30 p.m. - 5:00 p.m.</td>
<td>A8</td>
<td><strong>Networking - Infiniband</strong>&lt;br&gt;• High-performance Ethernet and Fibre Channel Connectivity for InfiniBand Clusters&lt;br&gt;• Checklist for Selecting and Deploying Scalable Clusters with InfiniBand Fabrics&lt;br&gt;• Creating and Debugging Low Latency, High Bandwidth InfiniBand Fabrics</td>
</tr>
<tr>
<td><strong>Panel</strong></td>
<td>3:30 p.m. - 5:00 p.m.</td>
<td>A3 / A4</td>
<td><strong>How to Get a Better Job in Computing—and Keep It!</strong></td>
</tr>
<tr>
<td><strong>Posters</strong></td>
<td>5:15 pm. - 7:00 p.m.</td>
<td>Ballroom Lobby</td>
<td><strong>Poster Reception</strong>&lt;br&gt;• Posters&lt;br&gt;• ACM Student Research Competition</td>
</tr>
<tr>
<td><strong>BOF</strong></td>
<td>5:30 p.m. - 7:00 p.m.</td>
<td>A1 / A6</td>
<td><strong>Open Standards for Reconfigurable Computing in a Hybrid Computing Environment</strong></td>
</tr>
<tr>
<td><strong>BOF</strong></td>
<td>5:30 p.m. - 7:00 p.m.</td>
<td>A10 / A11</td>
<td><strong>Petascale System Interconnect Project</strong></td>
</tr>
<tr>
<td><strong>BOF</strong></td>
<td>5:30 p.m. - 7:00 p.m.</td>
<td>A2 / A5</td>
<td><strong>Developing Applications for Petascale Computers</strong></td>
</tr>
<tr>
<td><strong>BOF</strong></td>
<td>5:30 p.m. - 7:00 p.m.</td>
<td>A20</td>
<td><strong>Open Standards for Accelerated Computing</strong></td>
</tr>
<tr>
<td><strong>BOF</strong></td>
<td>5:30 p.m. - 7:00 p.m.</td>
<td>A3 / A4</td>
<td><strong>Reliability of High-Speed Networks in Large-Scale Clusters</strong></td>
</tr>
<tr>
<td><strong>BOF</strong></td>
<td>5:30 p.m. - 7:00 p.m.</td>
<td>A8</td>
<td><strong>Globus and Community: Today’s Cyberinfrastructure</strong></td>
</tr>
<tr>
<td><strong>BOF</strong></td>
<td>5:30 p.m. - 7:00 p.m.</td>
<td>A9</td>
<td><strong>Parallel Program Development Tools Users BOF</strong></td>
</tr>
<tr>
<td><strong>BOF</strong></td>
<td>5:30 p.m. - 7:00 p.m.</td>
<td>C1 / C2 / C3</td>
<td><strong>TOP500 Supercomputers</strong></td>
</tr>
<tr>
<td><strong>BOF</strong></td>
<td>5:30 p.m. - 7:00 p.m.</td>
<td>D4</td>
<td>• Scaling I/O Capability in Parallel File Systems</td>
</tr>
<tr>
<td><strong>BOF</strong></td>
<td>5:30 p.m. - 7:00 p.m.</td>
<td>D5</td>
<td>• HPC Centers</td>
</tr>
<tr>
<td>Event Type</td>
<td>Time</td>
<td>Location</td>
<td>Session/Title</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Invited Speakers</td>
<td>8:30 a.m. - 10:00 a.m.</td>
<td>C4 / D1 / D2 / D3</td>
<td>Plenary</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• The American Competitiveness Initiative: Role of High End Computation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Cosmology’s Present and Future Computational Challenges</td>
</tr>
<tr>
<td>Disruptive Technologies</td>
<td>10:00 a.m. - 6:00 p.m.</td>
<td>Exhibit Hall 1B</td>
<td>Disruptive Technologies Exhibit</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• High Performance Optical Connectivity Based on CMOS Photonics Technology</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Optical Printed Circuit Board Technology and 200+200Gbps Transceiver</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• NRAM as a Disruptive Technology</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Superconducting Quantum Computing System</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Disruptive Technology for Many-core Chip System</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Software and Logic Co-verification</td>
</tr>
<tr>
<td>Gordon Bell Prize</td>
<td>10:30 a.m. - 12:00 p.m.</td>
<td>A3 / A4</td>
<td>Gordon Bell Prize Finalists</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• A 281 Tflops Calculation for X-ray Protein Structure Analysis with the Special-Purpose Computer MDGRAPE-3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• First-Principles Calculations of Large-Scale Semiconductor Systems on the Earth Simulator</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Extending Stability Beyond CPU-Millennium: Micron-Scale Atomistic Simulation of Kelvin-Helmholtz Instability</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• WRF Nature Run</td>
</tr>
<tr>
<td>Papers</td>
<td>10:30 a.m. - 12:00 p.m.</td>
<td>A1 / A6</td>
<td>File Systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Noncontiguous Locking Techniques for Parallel File Systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Integrating Parallel File Systems with Object-based Storage Devices</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Evaluation of Active Storage Strategies for the Lustre Parallel File System</td>
</tr>
<tr>
<td>Papers</td>
<td>10:30 a.m. - 12:00 p.m.</td>
<td>A2 / A5</td>
<td>Performance Tools and Methods</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• The Ghost in the Machine: Observing the Effects of Kernel Operation on Parallel Application Performance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• P’nMPI Tools: A Whole Lot Greater than the Sum of Their Parts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Multi-threading and One-sided Communication in Parallel LU Factorization</td>
</tr>
<tr>
<td>ACM Student Research</td>
<td>10:30 a.m. - 12:00 p.m.</td>
<td>A10 / A11</td>
<td>Finalists</td>
</tr>
<tr>
<td>Competition</td>
<td></td>
<td></td>
<td>• A Dynamic Programming Approach to Kd-Tree Based Data Distribution</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Storing and Massive Scale-free Graphs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• GrenchMark: A Framework for Testing Large-Scale Distributed Computing Systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Performance Analysis and Optimization of Large-scale Scientific Applications on Clusters with CMPs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Evolving the GPU-based Cluster</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Obtaining High Performance via Lower-Precision FPGA Floating Point Units</td>
</tr>
<tr>
<td>Event Type</td>
<td>Time</td>
<td>Location</td>
<td>Session/Title</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------</td>
<td>----------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Masterworks</td>
<td>10:30 a.m. - 12:00 p.m</td>
<td>C1 / C2 / C3</td>
<td>Industry INCITE Winners</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Rendering at the Speed of Shrek</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Turbo Charging Gas Turbine Engine Development: How Pervasive Supercomputing is Helping Pratt &amp; Whitney Develop a “Greener” Jet Engine</td>
</tr>
<tr>
<td>Exhibitor Forum</td>
<td>10:30 a.m. - 12:00 p.m</td>
<td>A7</td>
<td>Data Center</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Innovation for the Data Center of Tomorrow</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Combined Effect of Passive and Low-power Active Equalization: Extend the Life of High-speed Copper Interconnects</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• From the 380V DC Bus to sub-1V Processors: Efficient Power Conversion Solutions</td>
</tr>
<tr>
<td>Exhibitor Forum</td>
<td>10:30 a.m. - 12:00 p.m</td>
<td>A8</td>
<td>Storage - Analytics and Reliability</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• The Power of Streaming Analytic Appliances in Supercomputing Environments</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Advances in Reliability and Data Integrity for High Performance Storage Solutions</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Cyber Crimes Center: A Distributed Heterogeneous Mass Storage Federation for Digital Forensics</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m</td>
<td>A1 / A6</td>
<td>Rocks Clusters</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m</td>
<td>A10 / A11</td>
<td>Bayesian Network Awareness</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m</td>
<td>A2 / A5</td>
<td>OSCAR Community Meeting</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m</td>
<td>A20</td>
<td>Fortran</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m</td>
<td>A3 / A4</td>
<td>Coordinated Fault Tolerance in High-end Computing Environments</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m</td>
<td>A7</td>
<td>TeraGrid Operations and Plans in Oak Ridge: User Community Interaction</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m</td>
<td>A8</td>
<td>Open MPI State of the Union</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m</td>
<td>A9</td>
<td>Deploying HPC for Interactive Simulation</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m</td>
<td>C4 / D1 / D2 / D3</td>
<td>Federal Activities Impacting Long Term HEC Strategies</td>
</tr>
<tr>
<td>BOF</td>
<td>12:15 p.m. - 1:15 p.m</td>
<td>D9</td>
<td>Cyberinfrastructure in Education</td>
</tr>
<tr>
<td>Invited Speakers</td>
<td>1:30 p.m. - 3:00 p.m</td>
<td>C4 / D1 / D2 / D3</td>
<td>Cray and Fernbach Awards Lectures</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Kenneth E. Batcher (Kent State University), Seymour Cray Award Winner</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• David E. Keyes (Columbia University), Sidney Fernbach Award Winner</td>
</tr>
<tr>
<td>Exhibitor Forum</td>
<td>1:30 p.m. - 3:00 p.m</td>
<td>A7</td>
<td>Storage and HPC - Innovations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Lustre for the Rest of Us: A Fast, Fully Redundant, Fully Configured Lustre Appliance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• A Paradigm Shift in the Storage Industry: From Monolithic Boxes to Clustered Architectures</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Intel HPC Innovations for the Mainstream and High End</td>
</tr>
</tbody>
</table>
### Wednesday, November 14

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Time</th>
<th>Location</th>
<th>Session/Title</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exhibitor Forum</strong></td>
<td>1:30 p.m. - 3:00 p.m.</td>
<td>A8</td>
<td>Multicore Technologies</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Data Streaming Compilers for Multi-core CPUs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• A New Application Debugging Framework for the Multi-core Age</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Multicore: All Eyes on Bottlenecks</td>
</tr>
<tr>
<td><strong>Papers</strong></td>
<td>3:30 p.m. - 5:00 p.m.</td>
<td>A1 / A6</td>
<td>Grid Management</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Workstation Capacity Tuning using Reinforcement Learning</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Anomaly Detection and Diagnosis in Grid Environments</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• User-friendly and Reliable Grid Computing Based on Imperfect Middleware</td>
</tr>
<tr>
<td><strong>Papers</strong></td>
<td>3:30 p.m. - 5:00 p.m.</td>
<td>A2 / A5</td>
<td>Network Interfaces</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Analyzing the Impact of Supporting Out-of-order Communication on In-order Performance with jWARP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Evaluating NIC Hardware Requirements to Achieve High Message Rate PGA Support on Multi-Core Processors</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• High-performance Ethernet-based Communications for Future Multi-core Processors</td>
</tr>
<tr>
<td><strong>Masterworks</strong></td>
<td>3:30 p.m. - 5:00 p.m.</td>
<td>C1 / C2 / C3</td>
<td>HPC in Entertainment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• HPC Comes to Hollywood: Birth of a New Industry</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• A Tiger by the Tail: HPC Designs High Performance Golf Clubs</td>
</tr>
<tr>
<td><strong>Disruptive Technologies</strong></td>
<td>3:30 p.m. - 5:00 p.m.</td>
<td>A3 / A4</td>
<td>Panel Session</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Interconnects</td>
</tr>
<tr>
<td><strong>Doctoral Research Showcase I</strong></td>
<td>3:30 p.m. - 5:00 p.m.</td>
<td>A10 / A11</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Improving Power-Performance Efficiency in High-End Computing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Qualitative Performance Analysis for Large-Scale Scientific Workflows</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Parallel Performance Wizard: An Infrastructure and Tool for Analysis of Parallel Application Performance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Compiler Techniques for Efficient Communication in Multiprocessor Systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• On Economics and the User-Scheduler Relationship in HPC and Grid Systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Reliability for Scalable Tree-based Overlay Networks</td>
</tr>
<tr>
<td><strong>Exhibitor Forum</strong></td>
<td>3:30 p.m. - 5:00 p.m.</td>
<td>A7</td>
<td>Distributed Computation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Cost-effective Grid with EnginFrame</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Deploying a Geographically Distributed Infrastructure to Enable Global Data Sharing and Continuity of Operations</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• High Performance Java and .NET Applications for Financial and Other Compute-intensive Industries</td>
</tr>
<tr>
<td>Event Type</td>
<td>Time</td>
<td>Location</td>
<td>Session/Title</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------</td>
<td>----------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Exhibitor Forum | 3:30 p.m. - 5:00 p.m. | A8       | FPGAs - Applications  
• Accelerating Key Recovery and Mapping of MD Simulations using FPGAs  
• 60x Faster NCBI BLAST on the Mitrion Virtual Processor |
<p>| BOF        | 5:30 p.m. - 7:00 p.m. | A1 / A6  | Power, Cooling and Energy Consumption for Petascale and Beyond |
| BOF        | 5:30 p.m. - 7:00 p.m. | A10 / A11| Unleashing the Power of the Cell Broadband Engine Processor for HPC                |
| BOF        | 5:30 p.m. - 7:00 p.m. | A2 / A5  | PVFS: A Parallel File System for Petascale Computing                           |
| BOF        | 5:30 p.m. - 7:00 p.m. | A20      | FAST-OS (Forum to Address Scalable Technology for runtime and Operating Systems) |
| BOF        | 5:30 p.m. - 7:00 p.m. | A3 / A4  | Evaluating Petascale Infrastructure: Benchmarks, Models, and Applications       |
| BOF        | 5:30 p.m. - 7:00 p.m. | A7       | Large Data Handling                                                           |
| BOF        | 5:30 p.m. - 7:00 p.m. | A8       | Adaptive Routing in InfiniBand                                                  |
| BOF        | 5:30 p.m. - 7:00 p.m. | A9       | Eclipse Parallel Tools Platform                                                 |
| BOF        | 5:30 p.m. - 7:00 p.m. | C1 / C2 / C3 | Introduction to Ranger: The First NSF &quot;Track 2&quot; Petascale System             |
| BOF        | 5:30 p.m. - 7:00 p.m. | C4 / D1 / D2 / D3 | Parallel Network File System (pNFS)                                               |</p>
<table>
<thead>
<tr>
<th>Event Type</th>
<th>Time</th>
<th>Location</th>
<th>Session/Title</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Invited Speakers</strong></td>
<td>8:30 a.m. - 10:00 a.m.</td>
<td>C4 / D1 / D2 / D3</td>
<td>Plenary&lt;br&gt;• HPC in Academia and Industry - Synergy at Work&lt;br&gt;• Toward Millisecond-scale Molecular Dynamics Simulations of Proteins</td>
</tr>
<tr>
<td><strong>Disruptive Technologies</strong></td>
<td>10:00 a.m. - 4:00 p.m.</td>
<td>Exhibit Hall 1B</td>
<td>Disruptive Technologies Exhibit&lt;br&gt;• High Performance Optical Connectivity Based on CMOS Photonics Technology&lt;br&gt;• Optical Printed Circuit Board Technology and 200+200Gbps Transceiver&lt;br&gt;• NRAM as a Disruptive Technology&lt;br&gt;• Superconducting Quantum Computing System&lt;br&gt;• Disruptive Technology for Many-core Chip System Software and Logic Co-verification</td>
</tr>
<tr>
<td><strong>Papers</strong></td>
<td>10:30 a.m. - 12:00 p.m.</td>
<td>A1 / A6</td>
<td>Benchmarking&lt;br&gt;• Optimization of Sparse Matrix-vector Multiplication on Emerging Multicore Platforms&lt;br&gt;• Cray XT4: An Early Evaluation for Petascale Scientific Simulation&lt;br&gt;• An Adaptive Mesh Refinement Benchmark for Modern Parallel Programming Languages</td>
</tr>
<tr>
<td><strong>Papers</strong></td>
<td>10:30 a.m. - 12:00 p.m.</td>
<td>A2 / A5</td>
<td>Grid Performance&lt;br&gt;• Exploring Event Correlation for Failure Prediction in Coalitions of Clusters&lt;br&gt;• Advanced Data Flow Support for Scientific Grid Workflow Applications&lt;br&gt;• Falkon: Fast and Light-weight task execution framework</td>
</tr>
<tr>
<td><strong>Masterworks</strong></td>
<td>10:30 a.m. - 12:00 p.m.</td>
<td>C1 / C2 / C3</td>
<td>Large Scale Data Analysis and Visualization&lt;br&gt;• High Performance Computing: Pragmatism in the Corporate World&lt;br&gt;• From the Earth to the Stars: Supercomputing at the Hayden Planetarium</td>
</tr>
<tr>
<td><strong>Panel</strong></td>
<td>10:30 a.m. - 12:00 p.m.</td>
<td>A3 / A4</td>
<td>Fifty Years of Fortran</td>
</tr>
<tr>
<td><strong>Doctoral Research Showcase II</strong></td>
<td>10:30 a.m. - 12:00 p.m.</td>
<td>A10 / A11</td>
<td>• Efficiently Solving Large-scale Graph Problems on High Performance Computing Systems&lt;br&gt;• High Performance Non-rigid Registration for Image-Guided Neurosurgery&lt;br&gt;• Statistical and Pattern Recognition Techniques Applied to Algorithm Selection for Solving Linear Systems&lt;br&gt;• Migratable and Re-parallelizable OpenMP Programs&lt;br&gt;• Runtime Coupling Support for Component Scientific Simulations&lt;br&gt;• Adaptive Fault Management for High Performance Computing</td>
</tr>
<tr>
<td><strong>Exhibitor Forum</strong></td>
<td>10:30 a.m. - 12:00 p.m.</td>
<td>A7</td>
<td>Portable Clusters&lt;br&gt;• Clusters&lt;br&gt;• Maximum Performance Cluster Design&lt;br&gt;• Measured Performance of CIFS (SMB) as a Global File System for a Moderate Sized Cluster</td>
</tr>
</tbody>
</table>
### Thursday, November 15

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Time</th>
<th>Location</th>
<th>Session/Title</th>
</tr>
</thead>
</table>
| Exhibitor Forum    | 10:30 a.m. - 12:00 p.m. | A8       | **FPGA Programming**  
• Accelerating HPC Applications using C-to-FPGA Techniques  
• FPGA for High Performance Computing Applications |
| BOF                | 12:15 p.m. - 1:15 p.m. | A1 / A6  | Advancements in Distributed Rendering Open New Visualization Capabilities                                                               |
| BOF                | 12:15 p.m. - 1:15 p.m. | A10 / A11| OpenMP 3.0: Tasks Rule!                                                                                                                        |
| BOF                | 12:15 p.m. - 1:15 p.m. | A2 / A5  | Supercomputers or Grids: That is the Question!                                                                                               |
| BOF                | 12:15 p.m. - 1:15 p.m. | A20      | Partitioned Global Address Space (PGAS) Programming Languages                                                                                 |
| BOF                | 12:15 p.m. - 1:15 p.m. | A3 / A4  | MPICH2: A High-Performance Open-Source MPI Implementation                                                                                     |
| BOF                | 12:15 p.m. - 1:15 p.m. | A7       | Parallel Debugging and Correctness Checking                                                                                                |
| BOF                | 12:15 p.m. - 1:15 p.m. | A8       | Meeting the Feature Needs of the LSF User Community                                                                                  |
| Papers             | 1:30 p.m. - 3:00 p.m.  | A1 / A6  | **Storage, File Systems, and GPU Hashing**  
• RobustStore: A Distributed Storage Architecture with Robust and High Performance  
• A User-level Secure Grid File System  
• Efficient Gather and Scatter Operations on Graphics Processors |
| Papers             | 1:30 p.m. - 3:00 p.m.  | A2 / A5  | **Modeling in Action**  
• A Genetic Algorithms Approach to Modeling the Performance of Memory-bound Computations  
• Performance under Failure of High-end Computing  
• Bounding Energy Consumption in Large-scale MPI Programs |
| Masterworks        | 1:30 p.m. - 3:00 p.m.  | C1 / C2 / C3 | **Supply Chain Optimization**  
• From the Molecule to the Pump: Global Energy Supply Chain Optimization with HPC for Maximum Energy Security  
• High Performance Computing in a 24x7 Operational Environment |
| Panels             | 1:30 p.m. - 3:00 p.m.  | A3 / A4  | Progress in Quantum Computing                                                                                                                |
| Awards             | 1:30 p.m. - 3:00 p.m.  | A10 / A11| SC07 Conference Awards                                                                                                                        |
| Exhibitor Forum    | 1:30 p.m. - 3:00 p.m.  | A7       | **Programming Tools**  
• New Technologies in Mathematica  
• Supercomputing Engine for Mathematica  
• A Unified Development Platform for Cell, GPU, and CPU Programming with RapidMind |

---

Event Type: Time: Location: Session/Title
Thursday, November 15

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Time</th>
<th>Location</th>
<th>Session/Title</th>
</tr>
</thead>
</table>
| Exhibitor Forum    | 1:30 p.m. - 3:00 p.m. | A8       | Networking - Performance  
• iWARP and Beyond: Performance Enhancements to Ethernet  
• Scaling I/O Performance with Storage Aggregation Gateways  
• Testing Application Performance over a WAN: Using Network Emulation to Find the Weakest Link |
| Papers             | 3:30 p.m. - 5:00 p.m. | A1 / A6  | Performance Optimization  
• Application Development on Hybrid Systems  
• Multi-level Tiling: M for the Price of One  
• Implementation and Performance Analysis of Non-blocking Collective Operations for MPI |
| Papers             | 3:30 p.m. - 5:00 p.m. | A2 / A5  | Scheduling  
• Efficient Operating System Scheduling for Performance-asymmetric Multi-core Architectures  
• A Job Scheduling Framework for Large Computing Farms  
• Optimizing Center Performance through Coordinated Data Staging, Scheduling and Recovery |
| Masterworks        | 3:30 p.m. - 5:00 p.m. | C1 / C2 / C3 | CTO Roundtable: Strategic Decisions in Industrial HPC |
| Panels             | 3:30 p.m. - 5:00 p.m. | A3 / A4  | Supercomputer Operating System Kernels: A Weighty Issue |
| Disruptive         | 3:30 p.m. - 5:00 p.m. | A10 / A11| Panel Session  
• Memory Systems |

Friday, November 16

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Time</th>
<th>Location</th>
<th>Session/Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel</td>
<td>8:30 a.m. - 5:00 p.m.</td>
<td>A3 / A4</td>
<td>Third International Workshop on High Performance Computing for Nano-science and Nanotechnology (HPCNano07)</td>
</tr>
<tr>
<td>Panel</td>
<td>8:30 a.m. - 10:00 a.m.</td>
<td>A1 / A6</td>
<td>Is There an HEC Energy Crisis?</td>
</tr>
<tr>
<td>Panel</td>
<td>8:30 a.m. - 10:00 a.m.</td>
<td>A2 / A5</td>
<td>Supercomputing on FPGAs, GPUs, Cell and Other Exotic Architectures: Challenges and Opportunities</td>
</tr>
<tr>
<td>Panel</td>
<td>10:30 a.m. - 12:00 p.m.</td>
<td>A1 / A6</td>
<td>Return of HPC Survivor - Outwit, Outlast, Outcompute</td>
</tr>
</tbody>
</table>
Reno-Sparks Convention Center
Continuing in the tradition of previous SC conferences, the SC07 Tutorials program comprises a key component of the conference’s technical program and is a major reason why many people attend the conference. Attendees of the Tutorials sessions have the opportunity to learn from, and interact directly with, internationally recognized experts on a broad range of topics related to high performance computing, networking, storage, and analytics. This year, 10 half-day and 15 full-day tutorials are offered with a range of introductory to advanced material. There are eight new tutorials scheduled, and 17 that have been presented at previous SC conferences.

**Topics include:**

- programming and construction of parallel, distributed, cluster, and grid systems
- high performance and parallel I/O
- computer security
- program development, performance analysis, and optimization tools and techniques
- new technologies and architectures
- visualization
- collaborative technologies and tools
- algorithms and numerical methods
- application-specific topics

Some of the tutorials (S04, S07, S08, M02, and M13) will have hands-on components. For these, attendees must bring their own laptops with SSH software installed. The tutorial rooms will have wired network drops, Ethernet cables, SCinet wireless, and power drops, but there will be no computer support available. Please arrive early, as there may be tutorial-specific software to install on your laptop.

Each registered tutorial attendee will receive a copy of all the tutorial notes on a computer-readable medium only; there will be no hardcopy notes available.

---

**Reno Fact**

In 1868, the Central Pacific Railroad, building tracks across the west to connect with the Union Pacific, built from the east to form the first transcontinental railroad. Myron Lake, realizing what a rail connection would mean for business, deeded land to the Central Pacific in exchange for its promise to build a depot at Lake’s Crossing. Once the railroad station was established, the town of Reno officially came into being. The new town was named in honor of Major General Jesse L. Reno (an Anglicization of the French family name Renault), a Union officer killed in the American Civil War.
Sunday, Nov. 11

Full-Day Tutorials

8:30 a.m. - 5:00 p.m.

**S01: Parallel Computing 101**

**Room: A5**

*Quentin Stout, Christiane Jablonowski (University of Michigan)*

This tutorial provides a comprehensive overview of parallel computing, emphasizing those aspects most relevant to the user. It is suitable for new users, managers, students and anyone seeking an overview of parallel computing. It discusses software and hardware, with an emphasis on standards, portability, and systems that are widely available. The tutorial surveys basic parallel computing concepts, using examples selected from large-scale engineering and scientific problems. These examples illustrate using MPI on distributed memory systems, Globus on the grid, OpenMP on shared memory systems, and MPI+OpenMP on hybrid systems. It discusses numerous parallelization approaches and software engineering and performance improvement aspects, including the use of state-of-the-art tools. The tutorial helps attendees make intelligent decisions by covering the primary options that are available, explaining how they are used and what they are most suitable for. Extensive pointers to the literature and web-based resources are provided to facilitate follow-up studies.

**S02: Parallel I/O in Practice**

**Room: A4**

*Robert B. Ross, Rajeev Thakur (Argonne National Laboratory); William Loewe (Lawrence Livermore National Laboratory); Robert Latham (Argonne National Laboratory)*

I/O on HPC systems is a black art. This tutorial sheds light on the state-of-the-art in parallel I/O and provides the knowledge necessary for attendees to best leverage I/O resources available to them. We cover the entire I/O software stack from parallel file systems at the lowest layer, to intermediate layers (such as MPI-IO), and finally high-level I/O libraries (such as HDF-5). We emphasize ways to use these interfaces that result in high performance, and benchmarks on real systems are used throughout to show real-world results. This tutorial first discusses parallel file systems (PFSs). We cover general concepts and examine three examples: GPFS, Lustre, and PVFS. Next we examine the upper layers of the I/O stack, covering POSIX I/O, MPI-IO, Parallel netCDF, and HDF5. We discuss interface features, show code examples, and describe how application calls translate into PFS operations. Finally we discuss I/O best practice.
S03: Application Development Using Eclipse and the Parallel Tools Platform
Room: A10
Beth R. Tibbitts, Greg Watson (IBM Research); Craig Rasmussen (Los Alamos National Laboratory)

The Eclipse Parallel Tools Platform (PTP) is an open-source Eclipse Foundation project (http://eclipse.org/ptp) for parallel application development and debugging, and a research base for integrating future parallel tools. Eclipse offers features expected from a commercial quality integrated development environment (IDE): a syntax-highlighting editor, a source-level debugger, revision control (including CVS and Subversion), code refactoring, and support for multiple languages, including C, C++, and Fortran. PTP allows parallel application developers to use Eclipse as a portable parallel IDE across a wide range of parallel systems. PTP also includes a scalable parallel debugger, and tools for development of parallel programs (including MPI and OpenMP). This tutorial will introduce participants to the Eclipse platform and provide hands-on experience in developing and debugging parallel applications using Eclipse and PTP with C, Fortran, and MPI. Integration with performance tools, resource management tools, and remote machine support will also be covered. Students should bring a laptop.

S04: A Tutorial Introduction to High Performance Analytics and Workflow on Grids
Room: A2
Robert Grossman (University of Illinois at Chicago), Michael Wilde (Argonne National Laboratory), Michal Sabala (University of Illinois at Chicago)

In this introductory tutorial, we will: (1) give an introduction to data mining and high performance analytics; (2) give an introduction to workflow; and (3) show through two extended case studies how data mining and workflow can be integrated so that large data sets can be analyzed and processed efficiently. We describe three common patterns when using grids for building statistical and data mining models. The first pattern is to use grids to parallelize parameter search. The second pattern is to use grids to build ensembles or collections of statistical models. The third pattern is to use grids to handle the pipeline processing of data that is usually necessary when preparing data for building analytic models. The tutorial includes a hands-on laboratory.

S05: High Performance Computing on GPUs with CUDA
Room: A20
Massimiliano Fatica, David P. Luebke, Ian A. Buck (NVIDIA); John D. Owens (University of California, Davis); Mark J. Harris (NVIDIA); John E. Stone, James C. Phillips (University of Illinois); Bernard Deschizeaux (CGGVeritas)

NVIDIA’s Compute Unified Driver Architecture (CUDA) platform is a co-designed hardware and software stack that
expands the GPU beyond a graphics processor to a general-purpose parallel coprocessor with tremendous computational horsepower, and makes that horsepower accessible in a familiar environment—the C programming language. Scientists throughout industry and academia are already using CUDA to achieve dramatic speedups on production codes. In this tutorial NVIDIA engineers will partner with academic and industrial researchers to present CUDA and discuss its advanced use for science and engineering domains. In the morning, we will introduce CUDA programming and the execution and memory models at its heart, motivate the use of CUDA with many brief examples from different HPC domains, and discuss fundamental algorithmic building blocks in CUDA. In the afternoon, we will discuss advanced issues such as optimization and “tips & tricks,” and include real-world case studies from domain scientists using CUDA.

S06: Introduction to Globus
Room: A3
Jennifer M. Schopf (Argonne National Laboratory); Ben Clifford (University of Chicago); Ravi Madduri, Lee Liming (Argonne National Laboratory)

Globus is developing fundamental technologies needed to build Grids—persistent environments that enable software applications to integrate instruments, displays, computational and information. The first half of this tutorial is presented lecture-style and is open to attendees of all levels. It provides an overview to the basic concepts behind Globus through a series of use cases that commonly occur in distributed systems use, including security, data management, execution management, and information services. We encourage questions and direction from participants. The second half provides hands-on experience with these concepts for anyone who has a laptop, walking through a series of exercises that instruct users in the basic use of Globus—how to run a job, how to transfer a file, how to work with Globus security, how to track the progress of your jobs and services, and how to use some higher-level data tools such as replicas. Exercises are pitched for a variety of experience levels.

S07: Introduction to Scientific Workflow Management and the Kepler System
Room: A7
Ilkay Altintas (San Diego Supercomputer Center), Mladen Vouk (North Carolina State University), Scott Klasky (Oak Ridge National Laboratory), Norbert Podhorszki (University of California, Davis)

A scientific workflow is the process of combining data and processes into a configurable, structured set of steps that implement semi-automated computational solutions of a scientific problem. Scientific workflow systems provide graphical user interfaces to combine different technologies along with efficient methods for using them, and thus increase the efficiency of the scientists. This tutorial provides an introduction to scientific workflow construction and management (Part I) and includes a detailed hands-on (Part II) using the Kepler system. It is intended for an audience with a computational science background. It will cover principles and foundations of scientific work-
Tutorials
flows, Kepler environment installation, workflow construction out of the available Kepler library components, and workflow execution management that uses Kepler facilities to provide process and data monitoring and provenance information, as well as high speed data movement solutions. This tutorial also incorporates hands-on exercises and application examples from different scientific disciplines.

S08: Introductory Babel for Massive Supercomputing Software Integration
Room: A11
Gary Kunkerts, Thomas G. W. Epperly (Lawrence Livermore National Laboratory)

Babel is an award winning software integration technology for computational science and engineering. Babel is being used in a broad range of applications from big component systems such as CCA, to application specific codes in chemistry, fusion, accelerator beam dynamics, radio astronomy, cell biology, subsurface transport, material science. Babel is also being used in Math and CS infrastructure projects such as solvers, meshing, and performance analysis. This tutorial will show how Babel works and why it is so effective for supercomputing applications. We will cover Babel’s traditional strength, unparalleled language interoperability, as well as its major new feature: Remote Method Invocation (RMI). Babel RMI adds asynchronous, interrupting, one-sided communication primitives to the traditional SPMD model. Participants will learn how RMI and MPI can be combined to bring data-, thread-, and task-parallelism to bear on a single simulation. Laptops are required for hands-on exercises in the afternoon.

Sunday, Nov. 11
Half-Day Tutorials
8:30 a.m. - 12:00 p.m.

S09: HPC Challenge (HPCC) Benchmark Suite
Room: A1
Jack Dongarra, Piotr Luszczek (University of Tennessee); Allan Snavely (San Diego Supercomputer Center)

In 2003, DARPA’s High Productivity Computing Systems released the HPCC suite. It examines the performance of HPC architectures using kernels with various memory access patterns of well understood computational kernels. Consequently, HPCC results bound the performance of real applications as a function of memory access characteristics and define performance boundaries of HPC architectures. The suite was intended to augment the TOP500 list and by now the results are publicly available for 6 out of 10 of the world’s fastest supercomputers. Implementations exist in most of the major high-end programming languages and environments, accompanied by countless optimization efforts. The increased publicity enjoyed by HPCC doesn’t necessarily translate into deeper understanding of the performance metrics that HPCC measures. And so this tutorial will introduce attendees to HPCC, provide tools to examine
differences in HPC architectures, and give hands-on training that will hopefully lead to better understanding of parallel environments.

**S10: Hybrid MPI and OpenMP Parallel Programming**

**Room: A6**

Rolf Rabenseifner (High Performance Computing Center Stuttgart), Georg Hager (University of Erlangen-Nuremberg), Gabriele Jost (Oracle Corporation), Rainer Keller (High Performance Computing Center Stuttgart)

Most HPC systems are clusters of shared memory nodes. Such systems can be PC clusters with dual or quad boards, but also "constellation" type systems with large SMP nodes. Parallel programming must combine the distributed memory parallelization on the node inter-connect with the shared memory parallelization inside of each node. This tutorial analyzes the strength and weakness of several parallel programming models on clusters of SMP nodes. Various hybrid MPI+OpenMP programming models are compared with pure MPI. Benchmark results of several platforms are presented. The thread-safety quality of several existing MPI libraries is also discussed. Case studies will be provided to demonstrate various aspect of hybrid MPI/OpenMP programming. Another option is the use of distributed virtual shared-memory technologies. This tutorial analyzes strategies to overcome typical drawbacks of easily usable programming schemes on clusters of SMP nodes. Details are available at www.hlrs.de/people/rabenseifner/publ/SC2007-tutorial.html.

**Sunday, Nov. 11**

**Half-Day Tutorials**

**1:30 p.m. - 5:00 p.m.**

**S11: Principles and Practice of Experimental Performance Measurement and Analysis of Parallel Applications**

**Room: A1**

Luiz DeRose (Cray Inc.), Bernd Mohr (Research Centre Juelich)

In this tutorial we will introduce the principles of experimental performance instrumentation, measurement, and analysis of HPC applications, with an overview of the major issues, techniques, and resources in performance tools development, as well as an overview of the performance measurement tools available from vendors and research groups. In addition, we will discuss cutting-edge issues, such as performance analysis on large scale multi-core systems, techniques for performance analysis on the emerging multi-threaded architectures, and automatic performance analysis. Our goals are threefold: first, we will provide background information on methods and techniques for performance measurement and analysis, including practical tricks and tips, so that you can exploit available tools effectively and efficiently. Second, you will learn about simple portable techniques for measuring the performance of your parallel application. Finally, we will discuss open problems in the area for students, researchers, and users interested in working in the field of performance analysis.
The Partitioned Global Address Space (PGAS) programming model provides ease-of-use through a global shared address space while emphasizing performance through locality awareness. Over the past several years, the PGAS model has been gaining rising attention. A number of PGAS languages are now ubiquitous, such as UPC which runs on most high-performance computers. The DARPA HPCS program has also resulted in new promising PGAS languages, such as X10. In this tutorial we will discuss the fundamentals of parallel programming models and will focus on the concepts and issues associated with the PGAS model. We will follow with an in-depth introduction of two PGAS languages, UPC and X10. We will start with basic concepts, syntax and semantics and will include a range of issues from data distribution and locality exploitation to advanced topics such as synchronization, memory consistency and performance optimizations. Application examples will be also shared.
**M02: Debugging Parallel Application Memory Bugs with TotalView**  
**Room: A7**  
Christopher Gottbrath, Josh Carlson, Edward Hinkel (TotalView Technologies)

This is a full-day debugging tutorial focusing on ways to solve memory bugs in parallel and distributed applications. Memory problems such as heap memory leaks and array bounds violations can be challenging to track down in serial environments, and the challenges can be even more vexing when involving an MPI program running in a cluster environment. This tutorial will include an introduction to debugging parallel applications with TotalView—which will cover basic operations as well as introduce important concepts such as subset attach and parallel process control. Tutorial participants will learn how to use TotalView to debug memory problems in both serial and parallel applications. In the process, they will learn to take advantage of heap error reporting, pointer annotation, and various report views through an interactive debugging designed to track down memory problems.

**M03: Application Supercomputing Concepts**  
**Room: A4**  
Alice Koniges (Lawrence Livermore National Laboratory); William Gropp, Ewing Lusk (Argonne National Laboratory); David Eder (Lawrence Livermore National Laboratory)

This tutorial provides an overview of supercomputing application development from a practical perspective. We describe current and upcoming architectures, terminology, parallel languages, and development tools giving attendees a sense of what works and why in application development. Aimed primarily at those interested in large-scale applications, we provide information that will both help the attendees in their own application design as well providing a springboard to better understand the myriad of supercomputing information presented at SC07. The architecture overview examines current TOP500-type systems and surveys upcoming designs like Cell. Parallel programming concepts (MPI, OpenMP, HPF, UPC, CAF) are reviewed and compared. We use tools and MPI performance techniques to show how to evaluate and improve application performance. We also consider application team development techniques and tools for verification and validation. We end with a series of terascale applications including recent Gordon Bell winners and discuss specific challenges and performance issues.

**M04: Reconfigurable Supercomputing**  
**Room: A20**  
Tarek El-Ghazawi (George Washington University), Duncan Buell (University of South Carolina), Kris Gaj (George Mason University), Volodymyr Kindratenko (University of Illinois at Urbana-Champaign)

The synergistic advances in high-performance computing and reconfigurable computing, based on field programmable gate arrays (FPGAs), has resulted in hybrid parallel systems of microprocessors and FPGAs. Such systems support both fine-grain and coarse-grain parallelism, and can dynamically tune
their architecture to fit various applications. Programming these systems can be quite challenging as programming of FPGA devices can involve hardware design. This tutorial will introduce the field of reconfigurable supercomputing and its advances in systems, programming, applications and tools. Reconfigurable system developments at SRC, Cray, SGI, and other vendors will be highlighted, and case studies including full application developments will be presented. The tutorial will include scalability studies for real-life applications over entire HPRC systems and will reveal the tremendous promise held by this class of architectures in performance, power and cost improvements. Challenges that remain will be also discussed.

**M05: Component Software for High-Performance Computing: Using the Common Component Architecture**

Room: A11

David E. Bernholdt (Oak Ridge National Laboratory), Rob Armstrong (Sandia National Laboratories), Gary Kunfert (Lawrence Livermore National Laboratory), Boyana Norris (Argonne National Laboratory)

This tutorial will introduce participants to the Common Component Architecture (CCA). Component-based approaches increase software developer productivity by helping to manage the complexity of large-scale applications and facilitating the reuse and interoperability of code. The CCA was designed specifically for high-performance scientific computing. It supports language-neutral component-based application development for both parallel and distributed computing without penalizing the underlying performance, and with a minimal cost to incorporate existing code into the component environment. The CCA environment is also well suited to the creation of domain-specific application frameworks, in which the rich domain-specific computational infrastructure is cast as components. Using lectures and hands-on exercises, we will cover the concepts of components, the CCA’s particular approach to components, and their use in scientific applications. Participants should bring a computer with network connectivity to connect to the tutorial server. The CCA tools and example software will also be available for download.

**M06: A Practical Approach to Performance Analysis and Modeling of Large-scale Systems**

Room: A10

Darren J. Kerbyson, Adolfoy Hoisie (Los Alamos National Laboratory)

This tutorial presents a practical approach to the performance modeling of large-scale, scientific applications on high performance systems. The defining characteristic involves the description of a proven modeling approach, developed at Los Alamos, of full-blown scientific codes, that has been validated on systems containing 1,000’s of processors. We show how models are constructed and demonstrate how they are used to predict, explain, diagnose, and engineer application performance in existing or future codes and/or systems. Notably, our approach does not require the use of specific tools but rather is applicable across commonly used
environments. Moreover, since our performance models are parametric in terms of machine and application characteristics, they imbue the user with the ability to “experiment ahead” with different system configurations or algorithms/coding strategies. Both will be demonstrated in studies emphasizing the application of these modeling techniques including: verifying system performance, comparison of large-scale systems, and examination of possible future systems.

M07: Designing High-End Computing Systems with InfiniBand and 10-Gigabit Ethernet iWARP
Room: A6
Dhabaleswar K. (DK) Panda, Sayantan Sur (Ohio State University); Pavan Balaji (Argonne National Laboratory)

InfiniBand Architecture (IBA) and 10-Gigabit Ethernet iWARP technologies are generating a lot of excitement towards building next generation High-End Computing (HEC) systems. This tutorial will provide an in-depth look at these emerging technologies and examine their suitability for prime-time HEC. It will start with a brief overview of IBA, 10-Gigabit Ethernet iWARP and their architectural features. An overview of the emerging OpenFabrics stack which encapsulates both IBA and iWARP in a unified manner will be presented. IBA and iWARP hardware/software solutions and the market trends will be highlighted. Challenges in designing different kinds of systems using these standards on multi-core platforms for performance, scalability, portability and reliability will be covered. Specifically, case studies and experiences in designing HPC clusters (with MPI-1 and MPI-2 programming models), parallel file systems, networked file systems (NFS), storage protocols, multi-tier datacenters, and virtualization schemes will be presented together with the associated performance numbers and comparisons.

Monday, Nov. 12

Half-Day Tutorials

8:30 a.m. - 12:00 p.m.

M08: Large Scale Visualization with ParaView 3
Room: A2
Kenneth Moreland, John Greenfield (Sandia National Laboratories)

ParaView is a powerful open-source turnkey application for analyzing and visualizing large data sets in parallel. ParaView is regularly used by Sandia National Laboratories analysts to visualize simulations run on the Red Storm and ASC Purple supercomputers, which are currently ranked as the second and fourth fastest supercomputers, respectively, on www.top500.org. Designed to be configurable, extendible, and scalable, ParaView is built upon the Visualization Toolkit (VTK) to allow rapid deployment of visualization components. This tutorial presents the architecture of ParaView and the fundamentals of parallel visualization. Attendees will learn the basics of using ParaView for scientific visualization with highlights on the new features in ParaView.
3. The tutorial features detailed guidance in visualizing the massive simulations run on today’s supercomputers. Attendees are encouraged to bring laptops to install ParaView and follow along with the demonstrations.

**M09: Introduction to OpenMP**  
*Room: A5*  
*Timothy G. Mattson (Intel Corporation)*

OpenMP is an API (Application Programming Interface) for writing parallel programs. Its compiler directives and library routines are the de facto standard for writing parallel applications for shared memory computers. As multi-core processors move into the mainstream, the need for multi-threaded applications will grow; and OpenMP is usually the most straightforward way to write such programs. With recent advances in OpenMP technology, its reach is expanding to distributed memory systems (e.g. clusters) as well. In this tutorial, we will provide a comprehensive introduction to OpenMP. By dedicating a half-day tutorial to the API itself, we will be able to cover every construct within the language and show how OpenMP is used to program shared memory multiprocessor computers, multi-core CPUs and clusters.

**M10: Configuring and Deploying GridFTP for Managing Data Movement in Grid/HPC Environments**  
*Room: A1*  
*Dan Fraser, John Bresnahan, Rajkumar Kettimuthu, Michael Link (Argonne National Laboratory / University of Chicago)*

One of the foundational issues in HPC computing is the ability to move large (multi Gigabyte, and even Terabyte) data files between sites. Simple file transfer mechanisms such as FTP and SCP are not sufficient either from the reliability or the performance perspective. GridFTP is the most widely used Open Source production quality data movers available today. Capabilities include: parallel streams for optimal performance over TCP; striped data streams for moving large files; multiple security options (SSH, GSI); extensibility for integrating with specialized file systems; and extensibility for using different HPC transfer protocols such as UDT. In this tutorial, we will quickly walk through the steps required for setting up GridFTP on Linux laptops. Then we will explore the advanced capabilities of GridFTP such as striping, and a set of best practices for obtaining maximal performance with GridFTP. With an Internet connection, users will perform transfers to see the advantages.
Monday, Nov. 12

Half-Day Tutorials

1:30 p.m. - 5:00 p.m.

M11: Clustered and Parallel Storage System Technologies
Room: A2
Marc Unangst, Brent Welch (Panasas, Inc.)

To meet the demands of increasingly data-hungry cluster applications, cluster-based parallel storage technologies are now capable of delivering performance scaling from 10s to 100s of GB/sec. This tutorial will examine current state-of-the-art high performance file systems and the underlying technologies employed to deliver scalable performance across a range of scientific and industrial applications. The tutorial has two main sections. The first section describes the architecture of clustered, parallel storage systems and then compares several open-source and commercial systems based on this framework. Of particular interest are the newly emerging Object Storage Device (OSD) and Parallel NFS (pNFS) standards. The second half of the tutorial is about performance, including what benchmarking tools are available, how to use them to correctly evaluate a storage system, and how to optimize application I/O patterns to exploit the strengths and weaknesses of clustered, parallel storage systems.

M12: Advanced Topics in OpenMP
Room: A5
Larry Meadows, Timothy G. Mattson (Intel Corporation)

OpenMP is a well known API for parallel programming. Most OpenMP implementations are designed for shared-memory homogeneous systems with flat (SMP) and complex (NUMA) memory architectures. Recently, however, OpenMP has branched out to address more diverse systems such as clusters of SMP systems (with no hardware shared memory), FPGAs, and processors built from multiple smaller cores (GPUs, the Cell architecture, and multi-core CPUs). In this tutorial, we will explore OpenMP and how it works with a diverse range of parallel systems. We will start by discussing advanced OpenMP topics required to understand how OpenMP interacts with computer systems: (1) OpenMP’s relaxed consistency memory model, and (2) Compiler transformations implied by OpenMP constructs. Then we will explore the use of OpenMP in novel settings: (1) OpenMP on NUMA systems; (2) OpenMP on clusters; (3) FPGA; and (4) OpenMP on manycore systems (quadcore, GPUs, Cell, etc.)
M13: Parallel Programming Using the Global Arrays Toolkit

Room: A1

Jarek Nieplocha, Bruce J. Palmer, Manojkumar Krishnam (Pacific Northwest National Laboratory); P. Sadayappan (Ohio State University)

This tutorial provides an overview of the Global Arrays (GA) programming toolkit and describes its capabilities, performance, and the use of GA in high performance computing applications. It will also be compared with other Global Address Space models such as UPC and Co-Array Fortran. GA was created to provide programmers with an interface that allows them to distribute data while maintaining the global index space and programming syntax similar to that in serial programs. The goal of GA is to free the programmer from the low-level management of communication and allow them to deal with their problems in the same index space in which they were originally formulated. At the same time, the compatibility of GA with MPI enables the programmer to take advantage of the existing MPI software when appropriate. The variety of existing GA applications attests to the attractiveness of using higher level abstractions to write parallel code.
Invited Speakers

Plenary

This year's invited speakers include world leaders in high performance computing systems and represent a cross section of the growing HPC community.

Tuesday, Nov. 13
**Keynote Address**
*Programming Bits and Atoms*
*Neil Gershenfeld (Massachusetts Institute of Technology)*
8:30 a.m. - 10:00 a.m.  Room: Ballroom

Wednesday, Nov. 14
**The American Competitiveness Initiative: Role of High End Computation**
*Raymond Orbach (Under Secretary for Science, Department of Energy)*
8:30 a.m. - 9:00 a.m.  Room: C4 / D1 / D3

Wednesday, Nov. 14
**Cosmology's Present and Future Computational Challenges**
*George Smoot (University of California, Berkeley / Lawrence Berkeley National Laboratory)*
9:15 a.m. - 10:00 a.m.  Room: C4 / D1 / D3

Thursday, Nov. 15
**HPC in Academia and Industry - Synergy at Work**
*Michael Resch (High Performance Computing Center Stuttgart)*
8:30 a.m. - 9:00 a.m.  Room: C4 / D1 / D3

Thursday, Nov. 15
**Toward Millisecond-scale Molecular Dynamics Simulations of Proteins**
*David E. Shaw (D. E. Shaw Research / Columbia University)*
9:15 a.m. - 10:00 a.m.  Room: C4 / D1 / D3
Tuesday, Nov. 13

Keynote Address
Room: Ballroom

8:30 a.m. - 10:00 a.m.

Programming Bits and Atoms
Professor Neil Gershenfeld (Massachusetts Institute of Technology)

Today’s high-performance computers are many orders of magnitude more powerful than their ancestors, but they still share implicit assumptions, including: binary information is physically represented by systems with two states; circuits are constructed from components on cards in cases; and programs manipulate descriptions of things rather than the things themselves. While these might appear to be tautological statements, I will explore the consequences of relaxing them, including, respectively, analog logic circuits, conformal computing architectures, and digital fabrication processes. Breaking down these boundaries between bits and atoms can help improve not just the performance but also the relevance of advanced computing for some of the greatest emerging challenges and opportunities in developed and developing countries.

Biography: Prof. Neil Gershenfeld is the Director of MIT’s Center for Bits and Atoms. His unique laboratory is breaking down boundaries between the digital and physical worlds, from creating molecular quantum computers to virtuosic musical instruments. Technology from his lab has been seen and used in settings including New York’s Museum of Modern Art and rural Indian villages, the White House and the World Economic Forum, inner-city community centers and automobile safety systems, Las Vegas shows and Sami herds. He is the author of numerous technical publications, patents, and books including Fab, When Things Start To Think, The Nature of Mathematical Modeling, and The Physics of Information Technology. He has been featured in media such as The New York Times, The Economist, and the McNeil/Lehrer News Hour, and has been selected as a CNN/Time/Fortune Principal Voice and as one of the top 100 public intellectuals. Dr. Gershenfeld has a BA in Physics with High Honors and an honorary Doctor of Science from Swarthmore College, a Ph.D. from Cornell University, and was a Junior Fellow of the Harvard University Society of Fellows and a member of the research staff at Bell Labs.
Wednesday, Nov. 14

Invited Speakers
Room: C4 / D1 / D2 / D3

Session Chair: Fred Johnson (Department of Energy)

8:30 a.m. - 9:15 a.m.

The American Competitiveness Initiative: Role of High End Computation

Dr. Raymond Orbach (Under Secretary for Science, Department of Energy)

With two Leadership Computing Facilities and NERSC—each providing more than 100 teraflops performance, the powerful ESnet dedicated science network, and the innovative SciDAC and INCITE programs for tackling the biggest scientific challenges, the Office of Science is delivering computational science breakthroughs today and leading the way to tomorrow’s scientific discoveries. Over the past seven years, we have launched programs to develop tools to create increasingly detailed simulations, extract the science from massive datasets and support the computing and networking demands of large-scale experimental facilities, thereby changing the very fabric of scientific research. Our combination of facilities, applications expertise, applied mathematics and computer science research is transforming science in areas such as climate research, nanotechnology, astrophysics, energy and the environment throughout the US research community and in partnerships spanning the globe. US industry is substantially reducing R&D costs and shortening time to market. High end computation is transforming basic scientific research and U.S. global competitiveness.

Biography: Raymond Lee Orbach was sworn in by Secretary Samuel W. Bodman as the Department of Energy’s first Under Secretary for Science on June 1, 2006. President Bush nominated Dr. Orbach for the new position, created by the Energy Policy Act of 2005, on December 13, 2005, and he was unanimously confirmed by the U.S. Senate on May 26, 2006.

Secretary Bodman has tasked Dr. Orbach with the department’s implementation of the President’s American Competitiveness Initiative, will help drive continued U.S. economic growth.

The Secretary also has charged Dr. Orbach with leading the Department’s efforts to transfer technologies from DOE national laboratories and facilities to the global marketplace, naming the Under Secretary for Science as the Department’s Technology Transfer Coordinator, in accordance with the Energy Policy Act, and as chair of the DOE Technology Transfer Policy Board, responsible for coordinating and implementing policies for the Department’s technology transfer activities.

Dr. Orbach continues to serve as the 14th Director of the Office of Science at the Department of Energy, a position he has held since the Senate confirmed him and he was sworn in in March 2002. In this capacity, Dr. Orbach manages an organization that is the third largest Federal sponsor of basic research in the United States, the primary
supporter of the physical sciences in the U.S., and one of the premier science organizations in the world.

9:15 a.m. - 10:00 a.m.

**Cosmology’s Present and Future Computational Challenges**

*Professor George F. Smoot (University of California, Berkeley / Lawrence Berkeley National Laboratory)*

The current status of cosmology is reviewed, along with current observations and anticipated forthcoming data sets. The challenges of computing in three different areas are outlined and contrasted with the also challenging simulation efforts. The scale and range of the simulations by themselves are daunting and promise to tax even the largest computing systems even with the most optimum algorithms.

**Biography:** Professor George Smoot was co-awarded the 2006 Nobel Prize in Physics “for discovery of the blackbody form and anisotropy of the cosmic microwave background radiation.” Smoot received Bachelor degrees (1966) in Mathematics and Physics and a Ph.D. (1970) in Physics from MIT. Smoot has been at the University of California Berkeley and the Lawrence Berkeley National Laboratory since 1970. In April 1992, George Smoot made the announcement that the team he led had detected the long sought variations in the early Universe that had been observed by the COBE DMR. NASA’s COBE (Cosmic Background Explorer) satellite mapped the intensity of the radiation from the early Big Bang and found variations so small they had to be the seeds on which gravity worked to grow the galaxies, clusters of galaxies, and clusters of clusters that are observed in the universe today. These variations are also relics of creation. Professor Smoot is an author of more than 200 science papers and is also co-author (with Keay Davidson) of the popularized scientific book *Wrinkles in Time* (Harper, 1994) that elucidates cosmology and the COBE discovery. Another essay entitled *My Einstein Susenders* appears in *My Einstein: Essays by Twenty-four of the World’s Leading Thinkers on the Man, His Work, and His Legacy* (Ed. John Brockman, Pantheon, 2006).

---

**Tuesday, Nov. 15**

**Invited Speakers**

**Room: C4 / D1 / D2 / D3**

*Session Chair: Wilfred Pinfold (Intel Corporation)*

8:30 a.m. - 9:15 a.m.

**HPC in Academia and Industry - Synergy at Work**

*Professor Dr.-Ing. Michael Resch (High Performance Computing Center Stuttgart)*

Typically HPC in academia and industry are two separate worlds. This is mainly motivated by the fact that the two are different both in aim and organizational frameworks. However, with growing costs for HPC hardware and system operation discussions about merging academic and industrial usage of HPC systems has intensified. The merger is
Further supported by the fact that many HPC systems are built from standard components such that architectures in academia and industry are no longer as different as they were 15 years ago. The talk presents an example for a working collaboration of academia and industry in HPC sharing resources and costs. Potential synergistic effects and benefits are discussed. Potential risks and challenges are presented and put into perspective. Some lessons learned will be presented and an outlook for the future collaboration of academia and industry in HPC is given.

**Biography:** Michael Resch is the director of the Höchstleistungsrechenzentrum Stuttgart (HLRS) / High Performance Computing Center Stuttgart of the University of Stuttgart/Germany and the director of the Department for High Performance Computing at the University of Stuttgart since 2003 holding a full professorship for High Performance Computing. He has a 20-year record in high performance computing. In 1999 he led the group that received the NSF Award for High Performance Distributed Computing at SC’99 and was a member of the group that received the SC2003 HPC Challenge Award in 2003. Michael Resch led the team that in 1997 for the first time in the history of high performance computing linked two supercomputers in Europe and the US to solve a single great challenge problem. He initiated the first European Grid computing project METODIS in 1998 and has since led a number of European Grid Projects. Michael Resch is a member of the steering board of the German Grid initiative D-Grid, the PI of the German Engineering Grid Computing Project InGrid and a Co-PI of the German Financial Grid Computing Project FinGrid. He is the chairman of the Technology and Business Council of T-Systems StR, a member of the HPC Customer Advisory Board of Microsoft, and a member of the advisory board of Triangle Venture Capital. He is the chairman of the Scientific Advisory Board of the Swiss Center for Scientific Computing (CSCS) and the chairman of the NEC User Group (NUG). He holds an MSc (Dipl.-Ing.) degree in Technical Mathematics from the Technical University of Graz/Austria and a PhD (Dr.-Ing.) in Engineering from the University of Stuttgart/Germany. In 2002 he held an Assistant Professorship at the Department of Computer Science of the University of Houston, Texas.

9:15 a.m. - 10:00 a.m.

**Toward Millisecond-scale Molecular Dynamics Simulations of Proteins**

*Dr. David E. Shaw (D. E. Shaw Research / Columbia University)*

The ability to perform long, accurate molecular dynamics (MD) simulations involving proteins and other biological macromolecules could in principle lead to important scientific advances and provide a powerful new tool for drug discovery. A wide range of biologically interesting phenomena, however, occurs over time scales on the order of a millisecond—about three orders of magnitude beyond the duration of the longest current MD simulations. Our research group is currently building a specialized, massively parallel machine called Anton which, when completed in late 2008, should be capable of...
executing millisecond-scale classical MD simulations of one or more proteins at an atomic level of detail. We have also recently completed a parallel MD package called Desmond, which uses novel algorithms and numerical techniques to achieve unprecedented simulation speed on an ordinary computational cluster. This talk will provide an overview of our work on parallel algorithms and machine architectures for high-speed MD simulation, and will describe research conducted recently within our lab in which lengthy Desmond simulations helped elucidate the dynamics and functional mechanisms of two biologically important proteins. These computational studies yielded testable predictions which have subsequently been validated through laboratory experiments.

**Biography:** David E. Shaw serves as chief scientist of D. E. Shaw Research, LLC, and as a senior research fellow at the Center for Computational Biology and Bioinformatics at Columbia University. He received his Ph.D. from Stanford University in 1980 and served on the faculty of the Computer Science Department at Columbia University until 1986 when he turned his attention to the emerging field of computational finance. In 1988, he founded the D. E. Shaw group, an investment and technology development firm that now has approximately 1,200 employees and $30 billion in aggregate investment capital. In 2001, Dr. Shaw returned to full-time, hands-on scientific research and now leads a research group in the field of computational biochemistry. His lab is currently involved in the design of novel algorithms and machine architectures for high-speed molecular dynamics simulations of proteins and other biological macromolecules, and in the application of such simulations to basic scientific research in structural biology and biochemistry and to the process of computational drug design. In 1994, President Clinton appointed Dr. Shaw to the President’s Council of Advisors on Science and Technology, in which capacity he served as chairman of the Panel on Educational Technology. He has since testified before the National Science Board and several Congressional committees on various topics related to science and technology policy. Dr. Shaw is a fellow of the American Academy of Arts and Sciences and was elected to the board of directors of the American Association for the Advancement of Science in 2000.
This year, Masterworks collaborated with the Council on Competitiveness to identify industry speakers from global corporations who will discuss complex, real-world problems that demand high performance computing solutions. Masterworks will tap into the expanding use of HPC to drive innovation and productivity, a critical part of the Council’s work for the past three years. Presentations will highlight novel and innovative ways of applying advanced computing, communications and storage technologies to achieve the breakthroughs needed to capture and ensure a competitive advantage. (www.compete.org/hpc/annual_hpc.asp).

Tuesday, Nov. 13

High Performance Racing
Room: C1 / C2 / C3
Session Chair: Thom Dunning (National Center for Supercomputing Applications)

10:30 a.m. - 11:15 p.m.

High Performance Computing: Shaping the Future of Formula One
Torbjörn Larsson (BMW)

As the "Pinnacle of Motorsports," Formula One has become a fiercely competitive game that is highly technology driven. Racing teams invest huge amounts of money each year in aerodynamic research and development. For decades, physical testing using wind tunnel scale models has served as the primary instrument for improving race car aerodynamic efficiency. A few teams even utilize two wind tunnels to further increase the development pace. The near future will see a paradigm shift—emerging new technologies will move the focus in Formula One R&D from physical testing to computational simulations. With the recent launch of Albert2, a state-of-the-art supercomputer tailored for large-scale CFD applications, BMW Sauber F1 Team clearly underlines a strong commitment to simulation technology. And rather than pursuing a second wind tunnel, the team management has decided to take a "pioneering approach" focused around CFD and high performance computing.

Biography: Torbjörn Larsson heads the CFD group at BMW Sauber F1 Team. He joined the team in April 2000 (Former Sauber-Petronas F1). Torbjörn has close to 20 years of experience in CFD, aerodynamics and HPC. He is the author of numerous scientific papers and has delivered several keynote speeches at various international CFD and HPC events. Former positions include Aero Lead Engineer at the GM Tech Centre in Michigan, US and Research Scientist at Saab Aerospace in Sweden. Larsson holds an MSc degree in Vehicle Engineering from the Royal Institute of Technology, Stockholm, Sweden.
11:15 a.m. - 12:00 p.m.

**HPC at Chrysler: Bringing NASCAR from the Race Track to Your Driveway**
*John Picklo (Chrysler LLC)*

Chrysler has employed high performance computing in the vehicle design process since 1986. The HPC tools have evolved over the years and HPC has become tightly integrated into the formal vehicle design and analysis processes. Chrysler has been involved with NASCAR Craftsman Truck Series racing since 1995 and returned to NASCAR Nextel Cup racing in 2002. Collaboration on product development with NASCAR and Dodge racing teams includes the use of HPC resources on a regular basis. Passenger vehicle and race vehicle design activities have common issues. The lessons learned in one discipline frequently apply to the other. This presentation will include an overview of Chrysler’s use of HPC for NASCAR and other motorsports activities, as well as demonstrate how those efforts are applied to help improve passenger vehicle designs.

**Biography:** John Picklo is Manager, High Performance Computing for Chrysler Group. He has been with Chrysler for 14 years working with Information Technology for Product Development. Since 1998, he has been responsible for High Performance Computing, supporting advanced engineering and vehicle simulation. John has a Bachelor’s degree in Economics from Oakland University and a Master’s in Business Administration from the University of Detroit/Mercy.

Tuesday, Nov. 13

**Biofuels/Alternative Energy**
*Room: C1 / C2 / C3*
*Session Chair: Dona Crawford (Lawrence Livermore National Laboratory)*

**1:30 p.m. - 2:15 p.m.**

**Child’s First Words, Terrible Teens and My Boring Parents: Prediction in the Complex World of Crop Genetic Improvement**
*Mark Cooper (Pioneer Hi-Bred International)*

We will consider three case studies in the co-evolution of high performance computing infrastructure and the emergence of industrial scale hybrid maize breeding strategies. The case studies are selected to represent different stages of the innovation to application continuum that operates within the Pioneer crop genetics research community: (1) Hybrid Characterization and Advancement to commercial release (maturity and my boring parents), (2) Molecular Breeding (adolescence and the terrible-teens), and (3) In Silico Breeding (protective attention and your child’s first words).

**Biography:** Mark Cooper is a Research Fellow in Complex Trait Genetics & Molecular Breeding for Pioneer Hi-Bred International, Inc., a DuPont business that is the world’s leading developer and supplier of advanced plant genetics to farmers worldwide. For the duration of his professional career, Mark Cooper has been involved in genetics, with an emphasis on quantitative
genetics and its applications in plant breeding. This has involved research on aspects of theory, computer modeling, field and laboratory-based experimental work and the conduct of applied plant breeding programs. Cooper was an adjunct professor in Quantitative Genetics and Plant Breeding at the University of Queensland, supervising PhD students and Postdoctoral research projects.

2:15 p.m. - 3:00 p.m.

Diverse Energy Sources for Better Driving
Sharan Kalwani (General Motors)

Sharan Kalwani will discuss HPC’s mission-critical role in harnessing and using different energy sources appropriate for the automotive industry and look at a variety of other contributions HPC is making to the needs of the transport industry.

Biography: Sharan is currently the HPC architect and is tasked with continually prospecting and deploying HPC towards meeting GM’s need in various areas ranging from car design using CAE methods to pure R&D of future technologies. Sharan has over two decades worth of experience in HPC. His areas of expertise include diverse industry applications such as medical research, industrial applications, engineering, basic and applied research, computer vision and chemical technologies. He is a member of IEEE Computer Society, USENIX, SAE, SIAM and SAGE. His academic and educational background span both Mechanical Engineering and Computer Science.

Tuesday, Nov. 13

Toward Perfect Product Design
Room: C1 / C2 / C3
Session Chair: Stanley Ahalt (Ohio Supercomputer Center)

3:30 p.m. - 4:15 p.m.

A Grand Challenge: MultiCore and Industrial Modeling and Simulation
Jamil Appa (BAE Systems)

The impact of computational-based modeling and simulation in providing industry with a competitive edge delivering higher performance and cost-effective products is well documented and cannot be underestimated. In the past, this has been enabled by the yearly floating point performance improvements delivered by the processor and system manufacturers and has been the primary end-user experience of Moore’s law. The development of MultiCore technologies in its wide variety of forms challenges this end-user effect of Moore’s law, requiring a radical rethink of algorithms and their implementations just to maintain the same progress seen over the previous decades. Current examples of the industrial use of modeling and simulation in the design and life support of a range of products is presented, followed by recent results illustrating the effects of current MultiCore architectures and concluding with how industry is looking to proactively address this challenge to our business.
**Biography:** Jamil Appa is Group Leader for Technology and Engineering Services at BAE Systems. As an aeronautical engineer, he initially worked on the Flite3D Computational Fluid Dynamics suite based on then state-of-the-art unstructured methods for Euler simulations. This was the main CFD suite used by the Airbus wing design team to perform the integrated design for the A380. This led to heading up the technical work on Solar, a UK Unstructured Navier Stokes CFD suite developed by BAE Systems, ARA, QinetiQ and Airbus. He is currently responsible for IT and HPC-related research in the $36M CFMS research program and leads the Integration Technologies Group in the Mathematical Modeling Department at the ATC. He is project manager and strategist for the BAE Systems corporate e-Engineering initiative aimed at developing and demonstrating Grid Solutions for a range of business activities both within the company and with external collaborators, partners and suppliers.

**Biography:** Thomas P. Gielda, Ph.D. was named the director of Global Mechanical Structures and Systems shortly after joining Whirlpool Corporation in February 2004. Prior to this assignment, he served as the director, Innovation and Technology, Simulation Based Design. Gielda joined Whirlpool from Visteon Automotive Systems, a leading full-service supplier that delivers consumer-driven technology solutions to automotive manufacturers. His position at Visteon was as a Distinguished Technical Fellow. Before joining Visteon, Gielda was a senior scientist at the McDonnell Douglas Research Laboratory. While at MDRL he developed interdisciplinary Computational Fluid Dynamics analysis codes. These codes were used extensively on the National Aerospace Plane and Single-Stage-to-Orbit (SSTO) Delta Clipper Rocket. He holds a Ph.D. in Aeronautical Engineering from North Carolina State University and bachelor and master’s degrees in Mechanical Engineering from Michigan State University.

**4:15 p.m. - 5:00 p.m.**

**HPC in the Kitchen and Laundry Room: Optimizing Everyday Appliances for Customer Satisfaction and Market Share**

*Thomas P. Gielda (Whirlpool Corporation)*

From washers to dryers, from microwaves to freezers, Whirlpool is engaged in intense global competition to deliver the highest quality appliances to consumers. Doing this successfully requires Whirlpool to take a systems approach to design, considering customer desires, regulatory requirements, safety and manufacturing specifications and end-of-life guidelines. Optimizing these multiple, sometimes conflicting demands in order to find the "design sweet spot" is impossible without HPC. I will discuss the critical role HPC and simulation has played across various product lines in "designing out" potential conflicts at the start of the process with virtual prototyping so that our physical prototypes can be used for validation instead of traditional engineering development. In addition, I will discuss how we used HPC to design more secure packaging to reduce shipping damage. In an intensively competitive marketplace, design simulation with HPC is an essential ingredient in Whirlpool’s recipe for maximum market share and profitability.
DOE’s “Innovative and Novel Computational Impact on Theory and Experiment” (INCITE) program awards substantial amounts of time on DOE’s leadership class HPC systems through a peer reviewed proposal process. This session focuses on two such awards.

**10:30 a.m. - 11:15 p.m.**

**Rendering at the Speed of Shrek**

*Evan Smyth (DreamWorks Animation, SKG)*

Generating computer graphics imagery for feature films has long exploited the fact that any two images can be rendered independently. However, this parallelism does not address the time required to generate one image: When an artist tweaks a setting, we must render one image while they wait. To address this, DreamWorks Animation is working on a system that will deliver real-time rendering to the desktop. This requires a rendering speedup of five orders-of-magnitude and we are thus driven to a multi-system approach: A scalable distributed-memory rendering architecture that delivers the uncompromised quality our creative process demands. In this effort, we are taking advantage of our substantial Department of Energy INCITE Award for “Real-Time Ray-Tracing” to explore the highest end of performance and scalability. Here, we provide insight into the design and implementation of our rendering system as well as the challenges we face and our approaches to resolving them.

**Biography:** Evan Smyth is the Principal Software Engineer at DreamWorks Animation, SKG. At DreamWorks, he has been actively involved in developing supercomputing strategies as well as exploring various co-processing technologies. Since joining the company in 2002 Evan has worked on both Shark Tale and Madagascar. Prior to coming to DreamWorks, Evan was the Software Architect at Sony Pictures Imageworks where he worked on many films, including Stuart Little, Harry Potter and the Sorcerer’s Stone, Spider-Man, I Spy and Hollow Man. Before transitioning to the film industry, Evan worked at ElectricImage, Alias|Wavefront, ComputerVision and BMW. Smyth holds a Ph.D. from the Massachusetts Institute of Technology as well as a master’s degree from the Harvard University Graduate School of Design and both a B.Sc. and a B.Arch. from the University of Notre Dame.

**11:15 a.m. - 12:00 p.m.**

**Turbo Charging Gas Turbine Engine Development: How Pervasive Supercomputing Is Helping Pratt & Whitney Develop a “Greener” Jet Engine**

*Peter Bradley (Pratt & Whitney)*

Pratt & Whitney is a world leader in the design, manufacture, and support of turbine engines for commercial and military applications. From aerodynamic design to the simulation of bird impact, high performance
computing has become indispensable to gas turbine engine development. In addition to extensive use of grid computing and clustering, Pratt & Whitney is in its second year of a DOE INCITE award to explore leading edge combustor analysis using leadership-class computing. Our presentation will focus on how Pratt & Whitney is combining HPC-integrated production design systems with INCITE to develop the next generation of fuel-efficient, low emission jet engines.

**Biography:** Peter Bradley is a Fellow in High Performance Technical Computing at United Technologies Pratt & Whitney. In his 20 years with P&W, Pete has driven technology development to apply distributed, parallel, and high performance computing to production design and manufacturing. He is a longtime advocate for applied HPC and grid computing for message passing applications. He also participated in the MPI-2 forum and is now in the second year of a U.S. Department of Energy INCITE (Innovative and Novel Computational Impact on Theory and Experiment) Award.

---

**Wednesday, Nov. 14**

**HPC in Entertainment**

**Room: C1 / C2 / C3**

*Session Chair: John Grosh (Lawrence Livermore National Laboratory)*

**3:30 p.m. - 4:15 p.m.**

**HPC Comes to Hollywood: Birth of a New Industry**

*Robert Eicholz (EFILM, LLC)*

For years, we’ve been dazzled by visual effects in movies. That’s not new. But with the advent of new HPC platforms, a new industry is being created that digitizes, enhances, and stores not just visual effects, but entire movies. This new industry starts with petabytes and goes up from there. This talk will address the business and technical aspects of the new Digital Laboratory industry, focusing on using HPC and other technologies to create competitive advantage by leveraging: Massive Data Storage - Managing petabytes of data; High Performance Computing / Graphics - Providing the processing to render, display, and store large amounts of complex image data; High Bandwidth Telecommunications; and deploying high-speed worldwide bandwidth to move large files quickly to virtualize the movie-making process. The talk will include clips from several recent blockbuster movies and a discussion of the technologies and platforms used to make them.
Biography: As Vice President of Technology / Corporate Development, Robert Eicholz of EFILM, LLC, A Deluxe Company oversees software development, hardware design and procurement, storage technology, data asset management systems, and motion picture imaging science. Mr. Eicholz and his teams create the data, application software, and hardware systems that have been used in major motion pictures as a Digital Laboratory - successor to the photochemical Film Laboratory. Every single frame of films such as Hairspray, Spiderman2, Bourne Supremacy, and Lemony Snicket, was stored and manipulated on EFILM’s high-speed computers, network, and storage arrays, with over 1.5 Petabytes of storage. Mr. Eicholz has also worked in the entertainment industry as a technology and business planning consultant to post-production facilities and manager of computing for 235 schools in Houston, Texas. In addition, he has served as CIO of a mortgage bank, and worked for Deloitte & Touche as a systems integration consultant and large-scale software developer for numerous corporations. He holds a Masters Degree from UCLA (MBA) with Honors, focused on Information Systems and Finance / Accounting.

4:15 p.m. - 5:00 p.m.

A Tiger by the Tail...HPC Designs High Performance Golf Clubs

Eric Morales (PING Golf)

Tools to take full advantage of computer aided engineering products have been on the market for years; the problem was that it took too long for standard desktops to analyze the data and get meaningful data back from the analysis. HPC has changed that for PING. We can now get meaningful analysis from prototype designs by fine tuning with virtual testing. This allows us more time to hone the design while still getting a quality product to market on time. PING engineers design golf clubs with specific performance goals in mind. Finite element analysis allows us to accurately simulate performance and durability of a new design. Clubs cycle through a design/analysis loop before being certified for prototyping. We create prototype club models using “printers” that build up layers of plastic or wax. These models are sent to a foundry where metal club heads are made within a few days.

Biography: Eric J. Morales has worked at PING Golf as a staff test engineer since he graduated from Arizona State University in 2002 with a B.S. in Aeronautical Engineering. He has been working with Computer Aided Engineering (CAE) or Finite Element Analysis (FEA) at PING and is responsible for the virtual prototyping of their golf clubs. He engages in theoretical impact testing, modal analysis, and investment casting simulations using CAE tools and PING’s computing cluster. Simulations of product changes that once took a full day to run can now be processed in 20 minutes or less, allowing the company to get products to market faster.
Thursday, Nov. 15

Large Scale Data Analysis and Visualization
Room: C1 / C2 / C3
Session Chair: Rick Stevens (Argonne National Laboratory)

10:30 a.m. - 11:15 p.m.

High Performance Computing - Pragmatism in the Corporate World
Anthony L. Abbattista (Allstate Insurance Company), Catherine S. Brune (Allstate Insurance Company)

In this session, Cathy Brune and Anthony Abbattista will discuss how Allstate Insurance Company is using high performance computing techniques to drive business results and increase the utilization of its technology investment. Allstate’s computing roots are in typical transaction systems and using technology to directly serve its customers. However, they have been aggressively adopting system architectures that take advantage of advances in farm and grid based computing, while at the same time creating an information advantage by deploying technology that allows employees to turn data into information. Cathy and Anthony will highlight practical ways to embrace high performance computing and provide measurable value to the business.

Biography: Anthony Abbattista is vice president of technology solutions for Allstate Insurance Company. He joined Allstate in 2003 and has responsibility for enterprise technology strategy, information services, technical architecture, technology governance and infrastructure. Prior to Allstate, Abbattista was a senior partner at DiamondCluster International, where he was managing director of its global competency organization. He also spent eight years at Andersen Consulting in its emerging technology group, where he was an associate partner. Abbattista has extensive expertise in the design and implementation of high-performance architectures, the execution of technical strategy projects and management of complex implementation projects. Abbattista received his bachelor’s degree from Northwestern University where he has served on their guest faculty and currently serves on the advisory board of the McCormick School of Engineering and Applied Science. He is also a technology working group member of the Executives’ Club of Chicago and serves on Sun Microsystems Executive Advisory Council.

Biography: Catherine S. Brune is senior vice president and chief information officer for Allstate Insurance Company. She is also a member of the Allstate senior management team. Brune is responsible for enterprise-wide technology strategy, network infrastructure, enterprise applications and technology-related governance, security and compliance activities. Under her leadership, Allstate developed an enterprise-wide business-aligned technology strategy, created a virtual data center environment, developed best-in-class monitoring systems for its web operations, consolidated knowledge management processes, and brought an enterprise focus to information security and business resumption platforms while generating efficiency.
gains. Brune is an Inductee of the YWCA’s Academy of Women Achievers and has been recognized as one of the Premier 100 Information Technology leaders by Computerworld magazine. She is also a recipient of the CIO of the Year Award from the Executives’ Club of Chicago, the Excellence in Corporate IT Leadership Award from Women in Technology International and the Moore School of Business Distinguished Alumni Award from the University of South Carolina. Brune graduated from the University of South Carolina with a B.S. degree in management.

11:30 a.m. - 12:00 p.m.

From the Earth to the Stars: Supercomputing at the Hayden Planetarium
Mordecai-Mark Mac Low (American Museum of Natural History)

Classical planetarium shows have been centered around a terrestrial view of the two-dimensional night sky. Such a perspective neither captures our vastly expanded scientific understanding of the cosmos, nor is it competitive with other location-based leisure activities. At the Hayden Planetarium in the Rose Center for Earth and Space of the American Museum of Natural History in New York City, high-performance computing coupled to high-resolution video projection enables travel in the third dimension, off the Earth and into the observable universe. We here describe the techniques used to allow us to visualize large-scale scientific data sets, and how they have allowed us to bring the latest research to audiences numbering in the millions.

Biography: Mordecai-Mark Mac Low, Ph.D., is Chair, Division of Physical Sciences and Curator-in-Charge, Department of Astrophysics at the American Museum of Natural History. He is also adjunct professor, Department of Astronomy, at Columbia University Mac Low’s work focuses on understanding the causes and results of the formation of stars from interstellar gas. This is a fundamental problem in modern astrophysics, as stars produce all the elements heavier than helium, determine the possibilities for life to occur, and shape the fates of galaxies. Working with fellow astronomers at several universities and observatories, Dr. Mac Low has developed numerical models at several different scales to attack this problem. Mac Low also is collaborating on projects in three other areas: the structure of the shells produced by nova explosions; the behavior of magnetized gas in the very early universe, shortly after the formation of electrons and protons; and the impact of asteroids on Venus. He received his doctorate in physics from the University of Colorado at Boulder.
Thursday, Nov. 15

**Supply Chain Optimization**

Room: C1 / C2 / C3

Session Chair: Thomas Zacharia (Oak Ridge National Laboratory)

1:30 p.m. - 2:15 p.m.

**From the Molecule to the Pump: Global Energy Supply Chain Optimization with HPC for Maximum Energy Security**

Lynn Chou (Chevron Information Technology Company)

The energy sector has entered a new era characterized by supply and demand uncertainties, increased concern about global warming, and changing roles for oil companies. International oil companies find themselves in new and complex relationships with their government-owned counterparts, who control the majority of the remaining natural energy resources. The international oil companies will not only need to continuously developing technologies to remain valuable to the resource owning governments, but also adjust their focus further down the supply chain and move more into downstream activities, building and expanding refineries and retail operations outside the US and optimize supply and manufacturing efficiency in the growth markets. This presentation will cover the computing technology trends regarding optimizing the immense and diverse energy supply chain, and unique information technology challenges facing the oil and gas industry.

2:15 p.m. - 3:00 p.m.

**High Performance Computing in a 24x7 Operational Environment**

Don Fike (FedEx Information Services)

FedEx Corporation, with annual revenue of $35 billion, provides the transportation industry’s broadest range of services. FedEx applications, advanced networks and data centers provide around-the-clock and around-the-globe support for the information intensive transportation, logistics and business related product offerings of FedEx Corporation. For example, the company delivers 6.5 million packages daily to more than 220 countries and territories. It maintains a fleet of about 670 aircraft and more than 75,000 motor vehicles. In his presentation, Don Fike will discuss FedEx’s cutting-edge use of high performance computing.

**Biography:** Lynn Chou is GM, Global Technology and Strategy at Chevron Information Technology Company. She received her PhD and bachelor’s degree in Agricultural Science at the University of Queensland, as well as a graduate certificate in education. She has authored and co-authored over 100 journal papers, 25 book chapters, and 50 conference papers and a number of commissioned Research Reports. Her awards include The Australian Institute of Agricultural Science and Technology (AIAST) Young Professional in Agriculture and the Commonwealth Postgraduate Research Award.
Biography: Don Fike is Vice President of Common Services and Chief Technical Architect for FedEx Information Services. Fike joined FedEx in 1984 as the company was undergoing enormous growth. He has specialized in high volume transaction and real-time event systems, with particular focus on platforms that involve operating systems, the Web, applications, databases and real-time event messaging. Fike was involved in the architecture of the company’s high volume storage and server hardware systems. His current responsibilities include oversight of FedEx enterprise architecture and technical design, database management, applications, messaging, wireless platforms and support. Previous positions with FedEx include Sr. Technical Fellow and Technical Director. Before joining FedEx, Fike worked for Holiday Inns Inc. During his five years there, he worked as a Sr. Systems Programmer, specializing in operating systems, TP monitors and communications software.

Thursday, Nov. 15

CTO Roundtable
Room: C1 / C2 / C3
Moderator: Suzy Tichenor (Council on Competitiveness)

3:30 p.m. - 5:00 p.m.

CTO Roundtable: Strategic Decisions in Industrial HPC
Anna Ewing (NASDAQ Stock Market), Kevin Humphries (FedEx), Reza Sadeghi (MSC Software), David E. Shaw (D.E. Shaw Research, LLC), Nancy Stewart (Wal-Mart Stores, Inc.)

This capstone Masterworks session explores the strategic value of HPC and how this critical asset is best positioned to contribute maximum business return for competitive gain. While the session will be moderated, the format will be free flowing and highly interactive to encourage maximum interplay among the participants and even the audience. Topics to explore could include: What are the tradeoffs that companies must make in order to maximize profitability while ensuring that investment in a strategic asset like HPC is at an appropriate level to ensure competitive advantage? How are rising energy costs coupled with increasing system power requirements driving decisions about HPC purchases and HPC data center locations? Should leading user companies use their market leverage to drive HPC through their supply chain? Are they at risk if they don’t and their suppliers fail to adopt modeling and simulation with HPC?
Biography: Anna Ewing is Executive Vice President of Operations and Technology and Chief Information Officer of The Nasdaq Stock Market, Inc, the largest U.S. electronic stock market. She is responsible for all technology development and network operations. As NASDAQ’s Chief Information Officer, she has led the technology integration of the INET platform within NASDAQ’s existing trading systems and is overseeing NASDAQ’s cost-saving technology roadmap. Ewing was also the Technical Project Manager for the implementation of the NASDAQ Market Center and the integration of the Brut ECN. Prior to joining NASDAQ, Ewing was employed at CIBC World Markets in New York, where she served as Managing Director of Electronic Commerce. Before that, Ewing served as Vice President at Merrill Lynch, holding various leadership positions within the Corporate and Institutional Client Group Technology Division.

Biography: Kevin Humphries is the senior vice president of Technology Systems for FedEx Corporate Services. In this role, Humphries is responsible for setting technology direction as well as providing data center, network and field infrastructure support. Humphries is also responsible for the technology systems used to capture information important to package movement including scanning systems, label strategies, and barcode formats. Other key areas of responsibility include Information Security, Business Alignment and Customer Technology Services. Prior to his current position, Humphries served as senior vice president of the Information Technology Division and chief information officer for FedEx Express. As such, Humphries was responsible for the development and maintenance of the company’s global information systems and shipping technology. Prior to joining FedEx, Humphries worked for Andersen Consulting, holding a variety of management and staff positions during his 11-year tenure. Humphries earned his bachelor’s degree in Business Administration from Texas A&M University.

Biography: Reza Sadeghi is currently the CTO of MSC Software, with responsibility for strategy and development of all MSC core technologies (Nastran, Marc, Adams and Dytran). Prior to the merger of MARC Analysis Research Corp. and MSC Software, he was responsible for the company’s overall operations. Before joining MARC Analysis Research Corporation, he was responsible for the simulation methods group at Goodrich aerospace, where he led the development of a number of math based tools for design and manufacturing of commercial and military jet engine nacelle structures. He has more than 10 years of teaching experience in the field of Computational mechanics and math based modeling. He has been asked to serve on a number of science and technology review boards, among them the U.S. Department of Energy and the U.S. Department of Defense, as well as authored a number of papers in the field of computational mechanics and math based modeling.

Biography: David E. Shaw serves as chief scientist of D. E. Shaw Research, LLC, and as a senior research fellow at the Center for Computational Biology and Bioinformatics
at Columbia University. He received his Ph.D. from Stanford University in 1980 and served on the faculty of the Computer Science Department at Columbia University until 1986 when he turned his attention to the emerging field of computational finance. In 1988, he founded the D. E. Shaw group, an investment and technology development firm that now has approximately 1,200 employees and $30 billion in aggregate investment capital. In 2001, Dr. Shaw returned to full-time, hands-on scientific research and now leads a research group in the field of computational biochemistry. His lab is currently involved in the design of novel algorithms and machine architectures for high-speed molecular dynamics simulations of proteins and other biological macromolecules, and in the application of such simulations to basic scientific research in structural biology and biochemistry and to the process of computational drug design. In 1994, President Clinton appointed Dr. Shaw to the President’s Council of Advisors on Science and Technology, in which capacity he served as chairman of the Panel on Educational Technology. He has since testified before the National Science Board and several Congressional committees on various topics related to science and technology policy. Dr. Shaw is a fellow of the American Academy of Arts and Sciences and was elected to the board of directors of the American Association for the Advancement of Science in 2000.

**Biography:** Nancy Stewart is the Senior Vice President and Chief Technology Officer in the Information Systems Division. She began her career with Wal-Mart Stores, Inc. in March 2004 and she is responsible for all of the company’s infrastructure and technology, as well as operations, facilities and systems implementation. Prior to joining Wal-Mart, Nancy worked for General Motors Corp as the Business Services Information Officer for the Information Systems and Services Division and held multiple officer positions at IBM Corporation. Her key accomplishments at GM include setting global direction for enterprise application development, SAP standards and GM’s direction for its enterprise data warehouse architecture. At IBM, Stewart successfully managed a $600 million Information Technology budget for Corporate Information business functions, Year 2000 internal readiness compliance, and worldwide focus on application reengineering enhancements. She has received several industry awards, including GM’s highest award, the Chairman’s Honors Award, for implementing GM’s global portal. The portal also received several key industry awards. Stewart received her M.S. degree at the Massachusetts Institute of Technology and her B.S. degree in mathematics at Dominican University.
The SC07 Papers program is the premier forum for disseminating innovative and important advances in high performance computing, networking, storage, and analytics from academic, government, and corporate institutions around the world. Fifty-four papers were selected for presentation in Reno from 268 submissions contributed by 918 authors. They span theory, practice, modeling, experimentation, infrastructure, and application of high performance computing, advanced networking, innovative storage solutions, systems engineering and grid technologies. The papers are presented in 18 sessions of three papers each covering the following topics:

- Computational Biology
- Network Switching and Routing
- System Performance
- Grid Scheduling
- Security and Fault Tolerance
- System Architecture
- Microarchitecture
- PDE Applications
- File Systems
- Performance Tools and Methods
- Grid Management
- Network Interfaces
- Benchmarking
- Grid Performance
- Storage, File Systems and GPU Hashing
- Modeling in Action
- Performance Optimization
- Scheduling

Awards are given for best technical and best student papers. This year five papers are candidates for Best Paper and five for Best Student Paper. The candidate papers, which will be published in a special edition of *Journal of Scientific Programming*, are identified below. Winners will be announced at the conference Awards Ceremony on Thursday afternoon at 1:30 p.m.

---

Reno Fact

Reno is situated in a high desert valley of approximately 4,400 feet above sea level. Winter has snowfall but typically it is light. Summer highs are generally in the low to mid 90s but temperatures above 100 degrees occur occasionally. July daytime and nighttime temperatures average 92 degrees and 51 degrees, respectively; while January day and night temperatures average 46 degrees and 22 degrees. Most precipitation occurs in winter and spring.
Tuesday, Nov. 13

**Computational Biology**
Room: A1 / A6

*Session Chair: Ann L. Chervenak (University of Southern California Information Sciences Institute)*

### 10:30 a.m. - 11:00 a.m.

**A Preliminary Investigation of a Neocortex Model Implementation on the Cray XD1**

*Kenneth L. Rice, Christopher N. Vutsinas, Tarek M. Taha (Clemson University)*

In this paper we study the acceleration of a new class of cognitive processing applications based on the structure of the neocortex. Specifically we examine the speedup of a visual cortex model for image recognition. We propose techniques to accelerate the application on general purpose processors and on reconfigurable logic. We present implementations of our approach on a Cray XD1 and compare the performance potential of scaling the design utilizing reconfigurable logic based acceleration to a software only design. Our results indicate that acceleration using reconfigurable logic can provide a significant speedup over a software only implementation.

### 11:00 a.m. - 11:30 a.m.

**Anatomy of a Cortical Simulator**

*Rajagopal Ananthanarayanan, Dharmendra Modha (IBM Research)*

Insights into brain’s high-level computational principles will lead to novel cognitive systems, computing architectures, programming paradigms, and numerous practical applications. An important step towards this end is the study of large networks of cortical spiking neurons. We have built a cortical simulator, C2, incorporating several algorithmic enhancements to optimize the simulation scale and time, through: computationally efficient simulation of neurons in a clock-driven and synapses in an event-driven fashion; memory efficient representation of simulation state; and communication efficient message exchanges. Using phenomenological, single-compartment models of spiking neurons and synapses with spike-timing dependent plasticity, we represented a rat-scale cortical model (55 million neurons, 442 billion synapses) in 8TB memory of a 32,768-processor BlueGene/L. With 1 millisecond resolution for neuronal dynamics and 1-20 milliseconds axonal delays, C2 can simulate 1 second of model time in 9 seconds per Hertz of average neuronal firing rate. In summary, by combining state-of-the-art hardware with innovative algorithms and software design, we simultaneously achieved unprecedented time-to-solution on an unprecedented problem size.
**Tuesday, Nov. 13**

### 11:30 a.m. - 12:00 a.m.

**Large-scale Maximum Likelihood-based Phylogenetic Analysis on the IBM BlueGene/L**

Michael Ott (Technical University of Munich); Jaroslaw Zola, Srinivas Aluru (Iowa State University); Alexandros P. Stamatakis (Swiss Federal Institute of Technology, Lausanne)

**Best Paper Nominee**

Phylogenetic inference is a grand challenge in Bioinformatics due to immense computational requirements. The continuous accumulation of sequence data poses new challenges for high performance computing. We demonstrate how state-of-the-art Maximum Likelihood (ML) programs can be efficiently scaled to the IBM BlueGene/L architecture, by example of RAxML. Performance is assessed using datasets consisting of 212 sequences and 566,470 base pairs, and 2,182 sequences and 51,089 base pairs, respectively. To the best of our knowledge, these are the largest datasets analyzed under ML to date. The analysis of such large datasets allows us to address novel biological questions. We exploit coarse-grain and fine-grain parallelism that is inherent in every ML-based analysis. Fine-grained parallelism scales well up to 1,024 processors. A larger number of processors can be efficiently exploited by a combination of coarse-grained and fine-grained parallelism. Our approach scales equally well on an AMD Opteron cluster with a less favorable network latency to processor speed ratio and partially yields super-linear speedups due to increased cache efficiency.

### 10:30 a.m. - 11:00 a.m.

**Network Switching and Routing**

**Room: A2 / A5**

Session Chair: Keith Underwood (Intel Corporation)

#### Session: Age-Based Packet Arbitration in Large-Radix k-ary n-cubes

Dennis Abts (Cray Inc.), Deborah Weisser (Google)

**Best Paper Nominee**

As applications scale to increasingly large processor counts, the interconnection network is frequently the limiting factor in application performance. In order to achieve application scalability, the interconnect must maintain high bandwidth while minimizing variation in packet latency. As the offered load in the network increases with growing problem sizes and processor counts, so does the expected maximum packet latency in the network, directly impacting performance of applications with any synchronized communication. Age-based packet arbitration reduces the variance in packet latency as well as average latency. This paper describes the Cray XT router packet aging algorithm which allows globally fair arbitration by incorporating "age" in the packet output arbitration. We describe the parameters of the aging algorithm and how to arrive at appropriate settings. We show that an efficient aging algorithm reduces both the average packet latency and the variance in packet latency on communication-intensive benchmarks.
Performance Adaptive Power-aware Reconfigurable Optical Interconnects for HPC Systems
Avinash Kodi (Ohio University), Ahmed Louri (University of Arizona)

As communication distances and bit rates increase, opto-electronic interconnects are being deployed for designing high-bandwidth low-latency interconnection networks for high performance computing (HPC) systems. While bandwidth scaling with efficient multiplexing techniques are available, static assignment of wavelengths can be detrimental to network performance for non-uniform workloads. Dynamic bandwidth re-allocation based on actual traffic pattern can lead to improved network performance by utilizing idle resources. While dynamic bandwidth re-allocation (DBR) techniques can alleviate interconnection bottlenecks, power consumption also increases considerably. In this paper, we propose to improve the performance of optical interconnects using DBR techniques and simultaneously optimize the power consumption using Dynamic Power Management (DPM) techniques. A reconfigurable opto-electronic architecture and a performance adaptive algorithm for implementing DBR and DPM are proposed in this paper. Our proposed reconfiguration algorithm achieves a significant reduction in power consumption and considerable improvement in throughput with a marginal increase in latency for various traffic patterns.

Evaluating Network Information Models on Resource Efficiency and Application Performance in Lambda-Grids
Nut Taeombut (University of California, San Diego), Andrew A. Chien (Intel Corporation / University of California, San Diego)

A critical challenge for wide-area configurable networks is definition and widespread acceptance of Network Information Model (NIM). When a network comprises multiple domains, intelligent information sharing is required for a provider to maintain a competitive advantage and for customers to use a provider’s network and make good resource selection decisions. We characterize the information that can be shared between domains and propose a spectrum of network information models. To evaluate the impact of the proposed models, we use a trace-driven simulation under a range of real providers’ networks and assess how the available information affects application’s and providers’ ability to utilize network resources. We find that domain topology information is crucial for achieving good resource efficiency, low application latency and network configuration cost, while domain link state information contributes to better resource utilization and system throughput. These results suggest that collaboration between service providers can provide better overall network productivity.
Tuesday, Nov. 13

System Performance
Room: A3 / A4
Session Chair: Bronis R. de Supinski
(Lawrence Livermore National Laboratory)

10:30 a.m. - 11:00 a.m.

Using MPI File Caching to Improve Parallel Write Performance for Large-scale Scientific Applications
Wei-keng Liao, Avery Ching, Kenin Coloma, Arifa Nisar, Alok Choudhary (Northwestern University); Jacqueline Chen (Sandia National Laboratories); Ramanan Sankaran, Scott Klasky (Oak Ridge National Laboratory)

Typical large-scale scientific applications periodically write checkpoint files to save the computational state throughout execution. Existing parallel file systems improve such write-only I/O patterns through the use of client-side file caching and write-behind strategies. In distributed environments where files are rarely accessed by more than one client concurrently, file caching has achieved significant success; however, in parallel applications where multiple clients manipulate a shared file, cache coherence control can serialize I/O. We have designed a thread based caching layer for the MPI I/O library, which adds a portable caching system closer to user applications so more information about the application’s I/O patterns is available for better coherence control. We demonstrate the impact of our caching solution on parallel write performance with a comprehensive evaluation that includes a set of widely used I/O benchmarks and production application I/O kernels.

11:00 a.m. - 11:30 a.m.

Virtual Machine Aware Communication Libraries for High Performance Computing
Wei Huang, Matthew Koop, Qi Gao, Dhabaleswar Panda (Ohio State University)

Best Student Paper Nominee, Best Paper Nominee

While Virtual Machine (VM) technology provides various features that target emerging manageability issues on large-scale systems, performance concerns have largely blocked the deployment of VM-based environments for High-Performance Computing (HPC). In this paper, we follow three steps to demonstrate how performance and manageability can co-exist in a VM-based environment. First, we propose Inter-VM Communication (IVC), a VM-aware communication library to allow efficient shared memory communication among VMs on the same physical host. This is critical as multi-core systems are becoming popular for HPC. Second, we design MVAPICH2-ivc, an MPI library based on MVAPICH2 modified to use IVC, allowing MPI applications to benefit from IVC transparently. Finally, we carry out detailed performance evaluations of MVAPICH2-ivc. We demonstrate that MVAPICH2-ivc can improve NAS Parallel Benchmarks performance by up to 11% compared with MVAPICH2 on eight-core systems in a VM-based environment, and MVAPICH2-ivc incurs negligible overhead compared with a native environment running MVAPICH2.
11:30 a.m. - 12:00 p.m.

Investigation of Leading HPC I/O Performance using a Scientific-application-derived Benchmark
Julian Borrill, Leonid Oliker, John Shalf, Hongzhang Shan (Lawrence Berkeley National Laboratory)

With the exponential growth of high-fidelity sensor and simulated data, the scientific community is increasingly reliant on ultra-scale HPC resources to handle their data analysis requirements. To understand balanced I/O performance of data-intensive applications in realistic computational settings, we develop a lightweight, portable benchmark called MADbench2, which is derived directly from a large-scale Cosmic Microwave Background (CMB) data analysis package. Our study examines a broad array of cluster filesystems including Lustre, GPFS, and PVFS2 on 6 distinct system architectures for both default environment and with hand-tuning of the code and system. We compare a number of key I/O performance parameters comparisons including concurrency, POSIX- versus MPI-IO, and unique- versus shared-file accesses and explore the potential of asynchronous I/O. Overall, our study quantifies the vast differences in performance and functionality of parallel filesystems across state-of-the-art platforms, while providing system designers and computational scientists a lightweight tool for conducting further analyses.

Tuesday, Nov. 13

1:30 p.m. - 3:00 p.m.

Grid Scheduling
Room: A1 / A6
Session Chair: Satoshi Matsuoka (Tokyo Institute of Technology)

Richard Y. Huang (University of California, San Diego), Henri Casanova (University of Hawaii at Manoa), Andrew A. Chien (University of California, San Diego)

With an increasing number of available resources in large-scale distributed environments, a key challenge is resource selection. Fortunately, several middleware systems provide resource selection services. However, a user is still faced with a difficult question: “What should I ask for?” Since most users end up using naïve and suboptimal resource specifications, we propose an automated way to answer this question. We present an empirical model that given a workflow application (DAG-structured) generates an appropriate resource specification, including number of resources, the range of clock rates among the resources, and network connectivity. The model employs application structure information as well as an optional utility function that trades off cost and performance. With extensive simulation experiments for different types of applications, resource conditions, and scheduling heuristics, we show that our model leads consistently to close to optimal application performance and often reduces resource usage.
Scheduling large-scale applications on the Grid is a fundamental challenge and is critical to application performance and cost. Large-scale applications typically contain a large number of homogeneous and concurrent activities which are main bottlenecks, but open great potentials for optimization. This paper presents a new formulation of the well-known NP-complete problems and two novel algorithms that addresses the problems. The optimization problems are formulated as sequential cooperative games among workflow managers. Experimental results indicate that we have successfully devised and implemented one group of effective, efficient, and feasible approaches. They can produce solutions of significantly better performance and cost than traditional algorithms. Our algorithms have considerably low time complexity and can assign 1,000,000 activities to 10,000 processors within 0.4 second on one Opteron processor. Moreover, the solutions can be practically performed by workflow managers, and the violation of QoS can be easily detected, which are critical to fault tolerance.
Tuesday, Nov. 13

Security and Fault Tolerance
Room: A2 / A5
Session Chair: Karsten Schwan (Georgia Institute of Technology)

1:30 p.m. - 2:00 p.m.

Automatic Software Interference Detection in Parallel Applications
Vahid Tabatabaee, Jeffrey K. Hollingsworth (University of Maryland)

We present an automated software interference detection methodology for Single Program, Multiple Data (SPMD) parallel applications. Interference comes from the system and unexpected processes. If not detected and corrected such interference may result in performance degradation. Our goal is to provide a reliable metric for software interference that can be used in software failure protection and recovery systems. A unique feature of our algorithm is that we measure the relative timing of application events (i.e. time between MPI calls) rather than system level events such as CPU utilization. This approach lets our system automatically accommodate natural variations in an application’s utilization of resources. We use performance irregularities and degradation as signs of software interference. However, instead of relying on temporal changes in performance, our system detects spatial performance degradation across multiple processors. We also include a case study that demonstrates our technique’s effectiveness, resilience and robustness.

2:00 p.m. - 2:30 p.m.

DMTracker: Finding Bugs in Large-scale Parallel Programs by Detecting Anomalies in Data Movements
Qi Gao, Feng Qin, Dhabaleswar K. Panda (Ohio State University)

Best Student Paper Nominee, Best Paper Nominee

While software reliability in large-scale systems becomes increasingly important, debugging large-scale parallel programs remains as a daunting task. This paper proposes an innovative technique to automatically find hard-to-detect software bugs in parallel programs via detecting abnormal behaviors in data movements. Based on the observation that data movements in parallel programs typically follow certain patterns, our idea is to extract data movement (DM)-based invariants at runtime and check violations of them. These violations can indicate the potential bugs such as data races and memory corruptions that manifest themselves in data movements. We have built a tool, DMTracker, to implement our idea. Our experiments with two real-world bug cases in MVAPICH/MVAPICH2, a popular MPI library, have shown that DMTracker can effectively detect them and report abnormal data movements for further diagnosis of the root causes. Moreover, DMTracker incurs very low runtime overhead, 0.9%-6.0% with High Performance Linpack and NAS Parallel Benchmarks.
Petascale, high-performance file systems often hold sensitive data and thus require security, but authentication and authorization can dramatically reduce performance. Existing security solutions perform poorly in these environments because they cannot scale with the number of nodes, highly distributed data, and demanding workloads. To address these issues, we developed Maat, a security protocol designed to provide strong, scalable security to these systems. Maat introduces three new techniques. Extended capabilities limit the number of capabilities needed by allowing a capability to authorize I/O for any number of client-file pairs. Automatic Revocation uses short capability lifetimes to allow capability expiration to act as global revocation, while supporting non-revoked capability renewal. Secure Delegation allows clients to securely act on behalf of a group to open files and distribute access, facilitating secure joint computations. Experiments on the Maat prototype in the Ceph petascale file system show an overhead as little as 6-7%.

This paper describes the system architecture of the Cray BlackWidow scalable vector multiprocessor. The BlackWidow system is a distributed shared memory (DSM) architecture that is scalable to 32K processors, each with a 4-way dispatch scalar execution unit and an 8-pipe vector unit capable of 20.8 Gflops for 64-bit operations and 41.6 Gflops for 32-bit operations at the prototype operating frequency of 1.3 GHz. Global memory is directly accessible with processor loads and stores and is globally coherent. The system supports thousands of outstanding references to hide remote memory latencies, and provides a rich suite of built-in synchronization primitives. Each BlackWidow node is implemented as a 4-way SMP with up to 128 Gbytes of DDR2 main memory capacity. The system supports common programming models such as MPI and OpenMP, as well as global address space languages such as UPC and CAF. We describe
the system architecture and microarchitecture of the processor, memory controller, and router chips. We give preliminary performance results and discuss design tradeoffs.

2:00 p.m. - 2:30 p.m.

**GRAPE-DR: 2-Pflops Massively-Parallel Computer with 512-Core, 512-Gflops Processor Chips for Scientific Computing**  
*Junichiro Makino (National Astronomical Observatory of Japan); Kei Hiraki, Mary Inaba (University of Tokyo)*

We describe the GRAPE-DR (Greatly Reduced Array of Processor Elements with Data Reduction) system, which will consist of 4096 processor chips each with 512 cores operating at the clock frequency of 500 MHz. The peak speed of a processor chip is 512 Gflops (single precision) or 256 Gflops (double precision). The GRAPE-DR chip works as an attached processor to standard PCs. Currently, a PCI-X board with single GRAPE-DR chip is in operation. We are developing a 4-chip board with PCI-Express interface, which will have the peak performance of 1 Tflops. The final system will be a cluster of 512 PCs each with two GRAPE-DR boards. We plan to complete the final system by early 2009. The application area of GRAPE-DR covers particle-based simulations such as astrophysical many-body simulations and molecular-dynamics simulations, quantum chemistry calculations, various applications which require dense matrix operations, and many other compute-intensive applications.

2:30 p.m. - 3:00 p.m.

**A Case for Low-complexity MP Architectures**  
*Håkan Zeffer, Erik Hagersten (Uppsala University)*

Advances in semiconductor technology have driven shared-memory servers toward processors with multiple cores per die and multiple threads per core. This paper presents simple hardware primitives enabling flexible and low-complexity multi-chip designs supporting an efficient inter-node coherence protocol implemented in software. We argue that our primitives and the example design presented in this paper have lower hardware overhead, have easier (and later) verification requirements and provide the opportunity for flexible coherence protocols and simpler protocol bug corrections than traditional designs. Our evaluation is based on detailed full-system simulations of modern chip-multiprocessors and both commercial and HPC workloads. We compare a low-complexity system based on the proposed primitives with aggressive hardware multi-chip shared-memory systems and show that the performance is competitive across a large design space.
Tuesday, Nov. 13

**Microarchitecture**
**Room: A1 / A6**
*Session Chair: Dennis Abts (Cray Inc.)*

**3:30 p.m. - 4:00 p.m.**

**Variable Latency Caches for Nanoscale Processor**
*Serkan Ozdemir, Arindam Mallik, Ja Chun Ku, Gokhan Memik, Yehea Ismail (Northwestern University)*

**Best Student Paper Nominee, Best Paper Nominee**

Variability is one of the important issues in nanoscale processors. Due to increasing importance of interconnect structures in submicron technologies, the physical location and phenomena such as coupling have an increasing impact on the latency. The traditional view of rigid access latencies to components will result in suboptimal architectures. In this paper, we build a 45nm SPICE model for an L1 cache to find access latencies with different address transitions and environmental conditions. Motivated by the large difference in latency, we change the cache architecture to allow variable latency accesses. We then modify the functional units by adding special queues to store dependent instructions and allow data to be forwarded from the cache to the functional units correctly to allow for variation. Simulations based on SPEC2000 benchmarks show that our variable latency cache can reduce the execution time up to 19.4% and 10.7% on average compared to a conventional cache. (Best Paper Nominee, Best Student Paper Nominee)

**4:00 p.m. - 4:30 p.m.**

**Data Access History Cache and Associated Data Prefetching Mechanisms**
*Yong Chen, Surendra Byna, Xian-He Sun (Illinois Institute of Technology)*

Data prefetching is an effective way to bridge the increasing performance gap between processor and memory. As computing power is increasing much faster than memory performance, we suggest that it is time to have a dedicated cache to store data access histories and to serve prefetching to mask data access latency effectively. We thus propose a new cache structure, named Data Access History Cache (DAHC), and study its associated prefetching mechanisms. DAHC behaves as a cache for recent reference information instead of as a traditional cache for instructions or data. Theoretically, it is capable of supporting many well known history-based prefetching algorithms, especially adaptive and aggressive approaches. We have carried out simulation experiments to validate DAHC design and DAHC-based data prefetching methodologies and to demonstrate performance gains. DAHC provides a practical approach to reaping data prefetching benefits and its associated prefetching mechanisms are proven more effective than traditional approaches.
In this paper we describe parallelization of interior-point method (IPM) aimed at achieving high scalability on large-scale chip-multiprocessors (CMPs). IPM is an important computational technique used to solve optimization problems in many areas of science, engineering and finance. IPM spends most of its computation time in a few sparse linear algebra kernels. While each of these kernels contains a large amount of parallelism, sparse irregular datasets seen in many optimization problems make parallelism difficult to exploit. As a result, most researchers have shown only a relatively low scalability of 4X-12X on medium to large scale parallel machines. This paper proposes and evaluates several algorithmic and hardware features to improve IPM parallel performance on large-scale CMPs. Through detailed simulations, we demonstrate how exploring multiple levels of parallelism with hardware support for low overhead task queues and parallel reduction enables IPM to achieve up to 48X parallel speedup on a 64-core CMP.

We describe a new environment for the exploration of turbulent flows that uses a cluster of databases to store complete histories of Direct Numerical Simulation (DNS) results. This allows for spatial and temporal exploration of high-resolution data that were traditionally too large to store and too computationally expensive to produce on demand. We perform analysis of these data directly on the databases nodes, which minimizes the volume of network traffic. The low network demands enable us to provide public access to this experimental platform and its datasets through Web services. This paper details the system design and implementation. Specifically, we focus on hierarchical spatial indexing, cache-sensitive spatial scheduling of batch workloads, localizing computation through data partitioning, and load balancing techniques that minimize data movement. We provide real examples of how scientists use the system to perform high-resolution turbulence research from standard desktop computing environments.
We present the design of a scalable parallel pathline construction method for visualizing large time-varying 3D vector fields. A 4D (i.e., time and the 3D spatial domain) representation of the vector field is introduced to make a time-accurate depiction of the flow field. This representation also allows us to obtain pathlines through streamline tracing in the 4D space. Furthermore, a hierarchical representation of the 4D vector field, constructed by clustering the 4D field, makes possible interactive visualization of the flow field at different levels of abstraction. Based on this hierarchical representation, a data partitioning scheme is designed to achieve high parallel efficiency. We demonstrate the performance of parallel pathline visualization using data sets obtained from terascale flow simulations. This new capability will enable scientists to study their time-varying vector fields at the resolution and interactivity previously unavailable to them.
Noncontiguous Locking Techniques for Parallel File Systems
Avery Ching (Northwestern University), Rob Ross (Argonne National Laboratory), Wei-keng Liao (Northwestern University), Lee Ward (Sandia National Laboratories), Alok Choudhary (Northwestern University)

Many parallel scientific applications use high-level I/O APIs that offer atomic I/O capabilities. Atomic I/O in current parallel file systems is often slow when multiple processes simultaneously access interleaved, shared files. Current atomic I/O solutions are not optimized for handling noncontiguous access patterns because current locking systems have a fixed file system block-based granularity and do not leverage high-level access pattern information. In this paper we present a hybrid lock protocol that takes advantage of new list and datatype byte-range lock description techniques to enable high performance atomic I/O operations for these challenging access patterns. We implement our scalable distributed lock manager (DLM) in the PVFS parallel file system and show that these techniques improve locking throughput over a naive noncontiguous locking approach by several orders of magnitude in an array of lock-only tests.

Additionally, in two scientific I/O benchmarks, we show the benefits of avoiding false sharing with our byte-range granular DLM when compared against a block-based lock system implementation.

Integrating Parallel File Systems with Object-based Storage Devices
Ananth Devulapalli, Dennis Dalessandro, Pete Wyckoff (Ohio Supercomputer Center); Nawab Ali, P. Sadayappan (Ohio State University)

As storage systems evolve, the block-based design of today’s disks is becoming inadequate. As an alternative, object-based storage devices (OSDs) offer a view where the disk manages data layout and keeps track of various attributes about data objects. By moving functionality that is traditionally the responsibility of the host OS to the disk, it is possible to improve overall performance and simplify management of a storage system. The capabilities of OSDs will also permit performance improvements in parallel file systems, such as further decoupling metadata operations and thus reducing metadata server bottlenecks. In this work we present an implementation of the Parallel Virtual File System (PVFS) integrated with a software emulator of an OSD and describe an infrastructure for client access. Even with the overhead of emulation, performance is comparable to a traditional server-fronted implementation, demonstrating that serverless parallel file systems using OSDs are an achievable goal.
Evaluation of Active Storage Strategies for the Lustre Parallel File System
Juan Piernas, Evan J. Felix (Pacific Northwest National Laboratory)

Active Storage provides an opportunity for reducing the amount of data movement between storage and compute nodes of a parallel filesystem such as Lustre, and PVFS. It allows certain types of data processing operations to be performed directly on the storage nodes of modern parallel filesystems, near the data they manage. This is possible by exploiting the underutilized processor and memory resources of storage nodes that are implemented using general purpose servers and operating systems. In this paper, we present a novel user-space implementation of Active Storage for Lustre, and compare it to the traditional kernel-based implementation. Based on microbenchmark and application level evaluation, we show that both approaches can reduce the network traffic, and take advantage of the extra computing capacity offered by the storage nodes at the same time. However, our user-space approach has proved to be faster, more flexible, portable, and readily deployable than the kernel-space version.

Performance Tools and Methods

The Ghost in the Machine: Observing the Effects of Kernel Operation on Parallel Application Performance
Aroon Nataraj, Alan Morris, Allen D. Malony (University of Oregon); Matthew Sottile (Los Alamos National Laboratory); Pete Beckman (Argonne National Laboratory)

The performance of a parallel application on a scalable HPC system is determined by user-level execution of the application code and system-level (OS kernel) operations. To understand the influences of system-level factors on application performance, the measurement of OS kernel activities is key. We describe a technology to observe kernel actions and make this information available to application-level performance measurement tools. The benefits of merged application and OS performance information and its use in parallel performance analysis are demonstrated, both for profiling and tracing methodologies. In particular, we focus on the problem of kernel noise assessment as a stress test of the approach. We show new results for characterizing noise and introduce new techniques for evaluating noise interference and its effects on application execution. Our kernel measurement and noise analysis
technologies are being developed as part of Linux OS environments for scalable parallel systems.

11:00 a.m. - 11:30 a.m.

P^nMPI Tools: A Whole Lot Greater than the Sum of Their Parts

Martin Schulz, Bronis R. de Supinski
(Lawrence Livermore National Laboratory)

P^nMPI extends the PMPI profiling interface to support multiple concurrent PMPI-based tools by enabling users to assemble tool stacks. We extend this basic concept to include new services for tool interoperability and to switch between tool stacks dynamically. This allows P^nMPI to support modules that virtualize MPI execution environments within an MPI job or that restrict the application of existing, unmodified tools to a dynamic subset of MPI calls or even call sites. Further, we extend P^nMPI to platforms without dynamic linking, such as BlueGene/L, and we introduce an extended performance model along with experimental data from microbenchmarks to show that the performance overhead on any platform is negligible. More importantly, we provide significant new MPI tool components that are sufficient to compose interesting MPI tools. We present three detailed P^nMPI usage scenarios that demonstrate that it significantly simplifies the creation of application-specific tools.

11:30 a.m. - 12:00 p.m.

Multi-threading and One-sided Communication in Parallel LU Factorization

Parry Husbands (Lawrence Berkeley National Laboratory), Katherine Yelick (University of California, Berkeley / Lawrence Berkeley National Laboratory)

Dense LU factorization has a high ratio of computation to communication and, as evidenced by the High Performance Linpack (HPL) benchmark, this property makes it scale well on most parallel machines. Nevertheless, the standard algorithm for this problem has non-trivial dependence patterns which limit parallelism, and local computations require large matrices in order to achieve good single processor performance. We present an alternative programming model for this type of problem, which combines UPC’s global address space with lightweight multithreading. We introduce the concept of memory-constrained lookahead where the amount of concurrency managed by each processor is controlled by the amount of memory available. We implement novel techniques for steering the computation to optimize for high performance and demonstrate the scalability and portability of UPC with Teraflop level performance on some machines, comparing favourably to other state-of-the-art MPI codes.
Workstation Capacity Tuning Using Reinforcement Learning
Aharon Bar-Hillel, Amir Di-Nur, Liat Ein Dor, Ran Gilad-Bachrach, Yossi Ittach (Intel Research)

Computer grids are complex, heterogeneous, and dynamic systems, whose behavior is governed by hundreds of manually-tuned parameters. As the complexity of these systems grows, automating the procedure of parameter tuning becomes indispensable. In this paper, we consider the problem of auto-tuning server capacity, i.e. the number of jobs a server runs in parallel. We present three different reinforcement learning algorithms, which generate a dynamic policy by changing the number of concurrent running jobs according to the job types and machine state. The algorithms outperform manually-tuned policies for the entire range of checked workloads, with average throughput improvement greater than 20%. On multi-core servers, the average throughput improvement is approximately 40%, which hints at the enormous improvement potential of such a tuning mechanism with the gradual transition to multi-core machines.

Anomaly Detection and Diagnosis in Grid Environments
Lingyun Yang (University of Chicago), Chuang Liu (Microsoft Corporation), Jennifer M. Schopf (Argonne National Laboratory), Ian Foster (University of Chicago)

Identifying and diagnosing anomalies in application behavior is critical to delivering reliable application-level performance. In this paper we introduce a strategy to detect anomalies and diagnose the possible reasons behind them. Our approach extends the traditional window-based strategy by using signal-processing techniques to filter out recurring, background fluctuations in resource behavior. In addition, we have developed a diagnosis technique that uses standard monitoring data to determine where related changes in behavior occur at the times of the anomalies. We evaluate our anomaly detection and diagnosis technique by applying it in three contexts and inserting anomalies into the system at random intervals. The experimental results show that our strategy detects up to 96% of anomalies while reducing the fault positive rate by up to 90% compared to the traditional window average strategy. In addition, our strategy can diagnose the reason for the anomaly approximately 75% of the time.
Writing grid applications is hard. First, interfaces to existing grid middleware often are too low-level for application programmers who are domain experts rather than computer scientists. Second, grid APIs tend to evolve too quickly for applications to follow. Third, failures and configuration incompatibilities require applications to use different solutions to the same problem, depending on the actual sites in use. This paper describes the Java Grid Application Toolkit (JavaGAT) that provides a high-level, middleware-independent and site-independent interface to the grid. The JavaGAT uses nested exceptions and intelligent dispatching of method invocations to handle errors and to automatically select suitable grid middleware implementations for requested operations. The JavaGAT’s adaptor writing framework simplifies the implementation of interfaces to new middleware releases by combining nested exceptions and intelligent dispatching with rich default functionality. The many applications and middleware adaptors that have been provided by third-party developers indicate the viability of our approach.

Due to the growing need to tolerate network faults and congestion in high-end computing systems, supporting multiple network communication paths is becoming increasingly important. However, multi-path communication comes with the disadvantage of out-of-order arrival of packets (because packets may traverse different paths). While modern networking stacks such as the Internet Wide-Area RDMA Protocol (iWARP) over 10-Gigabit Ethernet (10GE) support multi-path communication, their current implementations do not handle out-of-order packets primarily owing to the overhead on in-order communication that it adds. Thus, in this paper, we analyze the trade-offs in designing a feature-complete iWARP stack, i.e., one that provides support for out-of-order arriving packets, and thus, multi-path systems, while focusing on the performance of in-order communication.
We propose three feature-complete designs of iWARP and analyze the pros and cons of each of these designs using performance experiments based on several micro-benchmarks as well as an iso-surface visual rendering application.

4:00 p.m. - 4:30 p.m.

Evaluating NIC Hardware Requirements to Achieve High Message Rate PGAS Support on Multi-Core Processors
Keith Underwood (Intel Corporation); Michael Levenhagen, Ron Brightwell (Sandia National Laboratories)

In recent years, partitioned global address space (PGAS) programming models have begun to receive more attention. These models tend to generate many small messages, which requires specific support from the network interface hardware to enable efficient execution. In the past, Cray included E-registers on the Cray T3E to support the SHMEM API; however, with the advent of multi-core processors, the balance of computation to communication capabilities has shifted toward computation. This paper explores the message rates that are achievable with multi-core processors and simplified PGAS support on a more conventional network interface. For message rate tests, we find that simple network interface hardware is more than sufficient. We also find that even simple data distributions, such as cyclic or block-cyclic, do not need specialized hardware support. Finally, we assess the impact of such support on the well known RandomAccess benchmark.

4:30 p.m. - 5:00 p.m.

High-performance Ethernet-based Communications for Future Multi-core Processors
Michael Schlansker (Hewlett-Packard), Nagabhushan Chitlur (Intel Corporation), Erwin Oertli (VMware, Inc.); Paul M. Stillwell, Jr., Linda Rankin, Dennis Bradford (Intel Corporation); Richard J. Carter (retired), Jayaram Mudigonda, Nathan Binkert, Norman P. Jouppi (Hewlett-Packard)

Data centers and HPC clusters often incorporate specialized networking fabrics to satisfy system requirements. However, Ethernet’s low cost and high performance are causing a shift from specialized fabrics toward standard Ethernet. Although Ethernet’s low-level performance approaches that of specialized fabrics, the features that these fabrics provide such as reliable in-order delivery and flow control are implemented in the case of Ethernet, by endpoint hardware and software. Unfortunately, current Ethernet endpoints are either slow (commodity NICs with generic TCP/IP stacks) or costly (offload engines). To address these issues, the JNIC project developed a novel Ethernet endpoint. JNIC’s hardware and software were specifically designed for the requirements of high-performance communications within future data-centers and compute clusters. The architecture combines capabilities already seen in advanced network architectures with new innovations to create a comprehensive solution for scalable and high-performance Ethernet. We envision a JNIC architecture that is suitable for most in-data-center communication needs.
Thursday, Nov. 15

**Benchmarking**

*Room: A1 / A6*

_Session Chair: Allan Snavely (San Diego Supercomputer Center)*

---

**10:30 a.m. - 11:00 a.m.**

**Optimization of Sparse Matrix-vector Multiplication on Emerging Multicore Platforms**

Samuel W. Williams (University of California, Berkeley / Lawrence Berkeley National Laboratory), Leonid Oliker (Lawrence Berkeley National Laboratory), Richard Vuduc (Lawrence Livermore National Laboratory), Katherine Yelick (University of California, Berkeley / Lawrence Berkeley National Laboratory), James Demmel (University of California Berkeley), John Shalf (Lawrence Berkeley National Laboratory)

We are witnessing a dramatic change in computer architecture due to the multicore paradigm shift, as every electronic device from cell phones to supercomputers confronts parallelism of unprecedented scale. To fully unleash the potential of these systems, the HPC community must develop multicore-specific optimization methodologies for important scientific computations. In this work, we examine sparse matrix-vector multiply (SpMV), one of the most heavily used kernels in scientific computing, across a broad spectrum of multicore designs. Our experimental platform includes the AMD dual-core, the Intel quad-core, the heterogeneous STI Cell, as well as the first scientific study of the highly multithreaded Sun Niagara2. We present several optimization strategies especially effective for the multicore environment, and demonstrate significant performance improvements compared to existing state-of-the-art serial and parallel SpMV implementations. Additionally, we present key insights into the architectural tradeoffs of leading multicore design strategies, in the context of demanding memory-bound numerical algorithms.

---

**11:00 a.m. - 11:30 a.m.**

**Cray XT4: An Early Evaluation for Petascale Scientific Simulation**

Sadaf R. Alam, Richard F. Barrett, Mark R. Fahey, Jeffery A. Kuehn, Ramanan Sankaran, Patrick H. Worley (Oak Ridge National Laboratory); Jeff M. Larkin (Cray Inc.)

The scientific simulation capabilities of next generation high-end computing technology will depend on striking a balance among memory, processor, I/O, and local and global network performance across the breadth of the scientific simulation space. The Cray XT4 combines commodity AMD dual core Opteron processor technology with the second generation of Cray’s custom communication accelerator in a system design whose balance is claimed to be driven by the demands of scientific simulation. This paper presents an evaluation of the Cray XT4 using micro-benchmarks to develop a controlled understanding of individual system components, providing the context for analyzing and comprehending the performance of several petascale-ready applications. Results gathered from several strategic application domains are compared with observations on the previous generation Cray XT3 and other high-end computing systems, demonstrating performance improvements across a wide variety of application benchmark problems.
We present an Adaptive Mesh Refinement benchmark for evaluating programmability and performance of modern parallel programming languages. Benchmarks employed today by language developing teams, originally designed for performance evaluation of computer architectures, do not fully capture the complexity of state-of-the-art computational software systems running on today’s parallel machines or to be run on the emerging ones from the multi-cores to the peta-scale High Productivity Computer Systems. This benchmark, extracted from a real application framework, presents challenges for a programming language in both expressiveness and performance. It consists of an infrastructure for finite difference calculations on block-structured adaptive meshes and a solver for elliptic Partial Differential Equations built on this infrastructure. Adaptive Mesh Refinement algorithms are challenging to implement due to the irregularity introduced by local mesh refinement. We describe those challenges posed by this benchmark through two reference implementations (C++/Fortran/MPI and Titanium) and in the context of three programming models.
Advanced Data Flow Support for Scientific Grid Workflow Applications

Jun Qin, Thomas Fahringer (University of Innsbruck)

Existing work does not provide a flexible dataset-oriented data flow mechanism to meet the complex requirements of scientific Grid workflow applications. In this paper we present a sophisticated approach to this problem by introducing a data collection concept and the corresponding collection distribution constructs, which are inspired by HPF, however applied to Grid workflow applications. Based on these constructs, more fine-grained data flows can be specified at an abstract workflow language level, such as mapping a portion of a dataset to an activity, independently distributing multiple datasets, not necessarily with the same number of data elements, onto loop iterations. Our approach reduces data duplication, optimizes data transfers as well as simplifies the effort to port workflow applications onto the Grid. We have extended AGWL with these concepts and implemented the corresponding runtime support in ASKALON. We apply our approach to some real world scientific workflow applications and report performance results.

Falkon: Fast and Light-weight task execution framework

Ioan Raicu, Yong Zhao, Catalin Dumitrescu, Ian Foster, Mike Wilde (University of Chicago)

To enable the rapid execution of many tasks on compute clusters, we have developed Falkon, a Fast and Light-weight task execution framework. Falkon integrates (1) multi-level scheduling to separate resource acquisition (via, e.g., requests to batch schedulers) from task dispatch, and (2) a streamlined dispatcher. Falkon’s integration of multi-level scheduling and streamlined dispatchers delivers performance not provided by any other system. We describe Falkon architecture and implementation, and present performance results for both microbenchmarks and applications. Microbenchmarks show that Falkon throughput (487 tasks/sec) and scalability (to 54,000 executors and 2,000,000 tasks processed in just 112 minutes) are one to two orders of magnitude better than other systems used in production Grids. Large-scale astronomy and medical applications executed under Falkon by the Swift parallel programming system achieve up to 90% reduction in end-to-end run time, relative to versions that execute tasks via separate scheduler submissions.
Thursday, Nov. 15

**Storage, File Systems, and GPU Hashing**

**Room: A1 / A6**

*Session Chair: Brett M. Bode (Ames Laboratory)*

1:30 p.m. - 2:00 p.m.

**RobuSTore: A Distributed Storage Architecture with Robust and High Performance**

*Huaxia Xia (Google), Andrew Chien (Intel Corporation / University of California, San Diego)*

Emerging large-scale scientific applications require to access large data objects in high and robust performance. We propose RobuSTore, a storage architecture combining erasure codes and speculative access mechanisms for parallel write and read in distributed environments. The mechanisms can effectively aggregate the bandwidth from a large number of distributed disks and statistically tolerate per-disk performance variation. Our simulation results affirm the high and robust performance of RobuSTore in both write and read operations compared to traditional parallel storage systems. For example, for a 1GB data access using 64 disks, RobuSTore achieves average bandwidth of 186MBps for write and 400MBps for read, nearly 6x and 15x that achieved by a RAID-0 system. The standard deviation of access latency is only 0.5 second, a 5-fold improvement from RAID-0. The improvements are achieved at moderate cost: about 40% increase in I/O operations and 2x-3x increase in storage capacity utilization.

2:00 p.m. - 2:30 p.m.

**A User-level Secure Grid File System**

*Ming Zhao, Renato Figueiredo (University of Florida)*

A grid-wide distributed file system provides convenient data access interfaces that facilitate fine-grained cross-domain data sharing and collaboration. However, existing widely-adopted distributed file systems do not meet the security requirements for grid systems. This paper presents a Secure Grid File System (SGFS) which supports GSI-based authentication and access control, end-to-end message privacy, and integrity. It employs user-level virtualization of NFS to provide transparent grid data access leveraging existing, unmodified clients and servers. It supports user and application-tailored security customization per SGFS session, and leverages secure management services to control and configure the sessions. The system conforms to the GSI grid security infrastructure and allows for seamless integration with other grid middleware. A SGFS prototype is evaluated with both file system benchmarks and typical applications, which demonstrates that it can achieve strong security with an acceptable overhead, and substantially outperform native NFS in wide-area environments by using disk caching.
Efficient Gather and Scatter Operations on Graphics Processors

Bingsheng He (Hong Kong University of Science and Technology), Naga Govindaraju (Microsoft Corporation), Qiong Luo (Hong Kong University of Science and Technology), Burton Smith (Microsoft Corporation)

Gather and scatter are two fundamental data-parallel operations, where a large number of data items are read (gathered) from or are written (scattered) to given locations. In this paper, we study these two operations on graphics processing units (GPUs). With superior computing power and high memory bandwidth, GPUs have become a commodity multiprocessor platform for general-purpose high-performance computing. However, due to the random access nature of gather and scatter, a naive implementation of the two operations suffers from low utilization of the memory bandwidth. Therefore, we design multi-pass gather and scatter operations to improve their data access locality, and develop a performance model to help understand and optimize these two operations. We have evaluated our algorithms in sorting, hashing, and the sparse matrix-vector multiplication. Our results show that these optimizations yield 2-4X improvement on the GPU bandwidth utilization. Overall, our optimized GPU implementations are 2-7X faster than their optimized CPU counterparts.

Modeling in Action

Room: A2 / A5

Session Chair: Vladimir Getov (University of Westminster)

A Genetic Algorithms Approach to Modeling the Performance of Memory-bound Computations

Mustafa Tikir, Laura Carrington (San Diego Supercomputer Center); Erich Strohmaier (Lawrence Berkeley National Laboratory), Allan Snively (San Diego Supercomputer Center)

Benchmarks that measure memory bandwidth, such as STREAM, Apex-MAPS, and Multi-MAPS, are increasingly popular due to the "Von Neuman" bottleneck of modern processors which causes many calculations to be memory-bound. We exhibit a scheme for predicting the performance of HPC applications based on the results of such benchmarks. A Genetic Algorithm approach is used to "learn" bandwidth as a function of cache hit rates per machine with Multi-MAPS as the fitness test. The specific results are 56 individual performance predictions including 3 full-scale parallel applications run on 5 different modern HPC architectures, with various cpu counts and inputs, predicted within 10% average difference with respect to independently verified runtimes.
Performance under Failure of High-end Computing
Ming Wu, Xian-He Sun, Hui Jin (Illinois Institute of Technology)

Modern high-end computers are unprecedentedly complex. Occurrence of faults is an inevitable fact in solving large-scale applications on these systems and future Petaflop machines. Many methods are proposed to mask faults. These methods, however, impose various performance and production costs. A better understanding of faults’ influence on application performance is necessary to use existing fault masking methods wisely. In this study, we first introduce some practical and effective performance models to predict the application completion time under system failures. These models separate the influence of failure arrival, failure repair, checkpointing period, checkpointing cost, and parallel task allocation on parallel and sequential execution times. They provide a guideline to choose the best performance/cost fault tolerance methods for a given application. We then develop effective failure-aware task scheduling algorithms to optimize application performance under system failures. Finally, extensive experiments are conducted to evaluate our prediction models and scheduling strategies with actual failure data.

Bounding Energy Consumption in Large-scale MPI Programs
Barry Rountree, David K. Lowenthal, Shelby Funk (University of Georgia); Vincent W. Freeh (North Carolina State University); Bronis R. de Supinski, Martin Schulz (Lawrence Livermore National Laboratory)

Power is now a first-order design constraint in large-scale parallel computing. Used carefully, dynamic voltage scaling can execute parts of a program at a slower CPU speed to achieve energy savings with a relatively small (possibly zero) time delay. However, the problem of when to change frequencies in order to optimize energy savings is NP-complete, which has led to many heuristic energy-saving algorithms. To determine how closely these algorithms approach optimal savings, we developed a system that determines a bound on the energy savings for an application. Our system uses a linear programming solver that takes as inputs the application communication trace and the cluster power characteristics and then outputs a schedule that realizes this bound. We apply our system to three scientific programs, two of which exhibit load imbalance—particle simulation and UMT2K. Results from our bounding technique show particle simulation is more amenable to energy savings than UMT2K.
Thursday, Nov. 15

**Performance Optimization**

**Room: A1 / A6**

*Session Chair: Derek Chiou (University of Texas-Austin)*

3:30 p.m. - 4:00 p.m.

**Application Development on Hybrid Systems**

Roger D. Chamberlain, Eric J. Tyson, Saurabh Gayen, Mark A. Franklin, Jeremy Buhler, Patrick Crowley, James Buckley (Washington University)

Hybrid systems consisting of a multitude of different computing device types are interesting targets for high-performance applications. Chip multiprocessors, FPGAs, DSPs, and GPUs can be readily put together into a hybrid system; however, it is not at all clear that one can effectively deploy applications on such a system. Coordinating multiple languages, especially very different languages like hardware and software languages, is awkward and error prone. Additionally, implementing communication mechanisms between different device types unnecessarily increases development time. This is compounded by the fact that the application developer, to be effective, needs performance data about the application early in the design cycle. We describe an application development environment specifically targeted at hybrid systems, supporting data-flow semantics between application kernels deployed on a variety of device types. A specific feature of the development environment is the availability of performance estimates (via simulation) prior to actual deployment on a physical system.

4:00 p.m. - 4:30 p.m.

**Multi-level Tiling: M for the Price of One**

DaeGon Kim, Lakshminarayanan Renganarayana, Dave Rostron, Sanjay Rajopadhy, Michelle Strout (Colorado State University)

Tiling is a widely used loop transformation for exposing/exploiting parallelism and data locality. High-performance implementations use multiple levels of tiling to exploit the hierarchy of parallelism and cache/register locality. Efficient generation of multi-level tiled code is essential for effective use of multi-level tiling. Parameterized tiled code, where tile sizes are not fixed but left as symbolic parameters, can enable several dynamic and run-time optimizations. Previous solutions to multi-level tiled loop generation are limited to the case where tile sizes are fixed at compile time. We present an algorithm that can generate multi-level parameterized tiled loops at the same cost as generating single-level tiled loops. The efficiency of our method is demonstrated on several benchmarks. We also present a method—useful in register tiling—for separating partial and full tiles at any arbitrary level of tiling. The code generator we have implemented is available as an open source tool.
Collective operations and non-blocking point-to-point operations have always been part of MPI. Although non-blocking collective operations are an obvious extension to MPI, there have been no comprehensive studies of this functionality. We present LibNBC, a portable high-performance library for implementing non-blocking collective MPI communication operations. LibNBC provides non-blocking versions of all MPI collective operations, is layered on top of MPI-1, and is portable to nearly all parallel architectures. To measure the performance characteristics of our implementation, we also present a microbenchmark for measuring both latency and overlap of computation and communication. Experimental results demonstrate that the blocking performance of the collective operations in our library is comparable to that of collective operations in other high-performance MPI implementations. Our library introduces a very low overhead between the application and the underlying MPI and thus, in conjunction with the potential to overlap communication with computation, offers the potential for optimizing real-world applications.
Our results also show that AMPS improves fairness and repeatability of application performance measurements.

4:00 p.m. - 4:30 p.m.

A Job Scheduling Framework for Large Computing Farms
Gabriele Capannini, Ranieri Baraglia, Diego Puppin, Marco Pasquali (Italian National Research Council); Laura Ricci (University of Pisa)

In this paper, we propose a new method, called Convergent Scheduling, for scheduling a continuous stream of batch jobs on the machines of large-scale computing farms. This method exploits a set of heuristics that guide the scheduler in making decisions. Each heuristic manages a specific problem constraint, and contributes to carry out a value that measures the degree of matching between a job and a machine. Scheduling choices are taken to meet the QoS requested by the submitted jobs, and optimizing the usage of hardware and software resources. We compared it with some of the most common job scheduling algorithms, i.e. Backfilling, and Earliest Deadline First. Convergent Scheduling is able to compute good assignments, while being a simple and modular algorithm.

4:30 p.m. - 5:00 p.m.

Optimizing Center Performance through Coordinated Data Staging, Scheduling and Recovery
Zhe Zhang, Chao Wang (North Carolina State University); Sudharshan S. Vazhkudai (Oak Ridge National Laboratory); Xiaosong Ma (North Carolina State University); Gregory G. Pike, John W. Cobb (Oak Ridge National Laboratory); Frank Mueller (North Carolina State University)

Procurement and the optimized utilization of Petascale supercomputers and centers is a renewed national priority. Sustained performance and availability of such large centers is a key technical challenge significantly impacting their usability. Storage systems are known to be the primary fault source leading to data unavailability and job resubmissions. This results in reduced center performance, partially due to the lack of coordination between I/O activities and job scheduling. In this work, we propose the coordination of job scheduling with data staging/offloading and on-demand staged data reconstruction to address the availability of job input data and to improve center-wide performance. Fundamental to both mechanisms is the efficient management of transient data: in the way it is scheduled and recovered. Collectively, from a center’s standpoint, these techniques optimize resource usage and increase its data/service availability. From a user’s standpoint, they reduce the job turnaround time and optimize the allocated time usage.
The Posters track is a highlight of the SC conference every year. Chosen from nearly 150 submissions, 39 regular posters and six student posters show up-to-the-minute results in all areas of high performance computing, networking, and applications. Come see the future of multi-core processing, computing on GPUs, grid computing, optical networking, performance analysis tools, computational nanoscience, fluid dynamics, and physics, plus many other topics.

The posters reception (Tuesday, 5:15 p.m. - 7:00 p.m.) allows conference attendees to discuss the research displays with presenters in a casual setting. Presenters also get the chance for greater in-depth, one-on-one dialogue about their work.

Student posters are part of ACM’s Student Research Competition. Finalists present their work in a special session during the Technical Program (Wednesday, 10:30 a.m. - 12:00 p.m.).

As of the 2000 census, the city population of Reno was 180,480, making it the second largest city in Nevada. Current census estimates, however, show the city’s population has grown to approximately 214,000, but the city is now the third largest in the state, following Las Vegas and Henderson. There were 73,904 households out of which 27.6% included children under the age of 18 living with them, 40.5% were married couples living together, 10.6% had a female householder with no husband present, and 43.6% were single individuals. The racial makeup of the city was 77.46% White, 2.58% African American, 1.26% Native American, 1.29% Asian, 0.56% Pacific Islander, 9.26% from other races, and 3.60% from two or more races.
Tuesday, Nov. 13

Posters Reception
Room: Ballroom Lobby
Session Chair: Tamara K. Grimmett (Idaho National Laboratory)
5:15 p.m. - 7:00 p.m.

Algorithms and Applications

Billion Vortex Particle Direct Numerical Simulations of Wake Vortices
Philippe Chatelain (ETH Zurich), Alessandro Curioni (IBM Research), Michael Bergdorf (ETH Zurich), Wanda Andreoni (IBM Research), Petros Koumoutsakos (ETH Zurich)

We present the Direct Numerical Simulations of high Reynolds numbers aircraft wakes employing adaptive vortex particle methods. The simulations involve a highly efficient implementation of vortex methods on massively parallel computers, enabling unprecedented simulations using billions of particles. The method relies on the Lagrangian discretization of the Navier-Stokes equations in vorticity-velocity form, along with a remeshing of the particles in order to ensure the convergence of the method. The remeshed particle locations are utilized for the computation of the field quantities, the discretization of the differential operators for diffusion and vortex stretching, and the solution of the Poisson equation for the stream-function. The methods exhibit very good weak scaling up to 16K BG/L processors. The results include unprecedented direct numerical simulations of the onset and the evolution of long wavelength instabilities induced by ambient noise in aircraft vortex wakes at Re=6000.

Feasibility Study of CFD Code Acceleration using FPGA
Naoyuki Fujita, Takashi Nakamura, Yuichi Matsuo (Japan Aerospace Exploration Agency); Katsumi Yazawa (Fujitsu Ltd); Yasuyuki Shiromizu, Hiroshi Okubo (FFC Ltd)

We try to define technical issues that arise when we use FPGA on CFD (Computational Fluid Dynamics) to explore feasibility issues. Our approach is: (1) Translate an existing R&D Navier-Stokes CFD solver into an FPGA block diagram; (2) Incorporate part of the block diagram into an FPGA circuit by VHDL, actually by hand; (3) Assume the circuit’s calculation performance with “block diagram analysis” and an RTL simulation tool; (4) By performing these steps, dig out technical issues that exist when we use FPGA. Some technical issues are listed, and we are suggesting a few directions and ideas to deal with these issues. We’ll introduce our approach and discuss listed technical issues, our approaches and ideas. Also, by way of applying a few of our approaches or ideas on FPGA circuit design, a proposed FPGA-based custom computer for CFD will be presented.
Performance Evaluation of a Coupled System with Multiple Spatial Domains and Multiple Temporal Scales
Jing-Ru C. Cheng, Hwai-Ping Cheng, Robert M. Hunter, David R. Richards (U.S. Army ERDC)

The Corps Engineer Research and Development Center has developed a parallel comprehensive, physics-based watershed model accounting for one-dimensional (1-D) cross-section-averaged channel flow, 2-D depth-averaged overland flow, 3-D subsurface flow, surface-subsurface interaction, and other major hydrological processes. It adopted the multiframe approach to resolve different hydrological processes effectively and efficiently. Different partitioning strategy is employed in each component to achieve load balance and avoid high communication overhead. A coupler has been developed to coordinate the interaction/communication between 2- and 3-D domains, while a different one is required when 1-D interaction is involved. Three meshes are constructed to evaluate the performance of such a complex software system. Various numbers of processors are considered to examine the parallel efficiency on the Cray XT3 with single- and dual-core processors. The poster details the test example covering an area of 570 square miles in South Florida and shows the performance results obtained from 44 runs.

Parallel Scalable Algorithm for 3D Nonlinear Simulations of Plasma Instabilities in Thermonuclear Fusion Devices
Nina Popova (Moscow State University)

Parallel scalable algorithm for solution of 3D nonlinear magnetohydrodynamic equations is suggested. Parallel program is developed on the base of 3D nonlinear code NFTC used in simulation of plasma instabilities in thermonuclear fusion. Numerical method of NFTC code is based on semi-spectral representation of solution. Fourier expansion in toroidal and poloidal angles of toroidal coordinate system and quasi-radial direction finite difference scheme are used. Fully implicit scheme in time needs to invert block tridiagonal matrix in each step. The most time-consuming part of the code is calculation of convolutions sums of Fourier coefficients. A parallel algorithm using different truncation of convolution sums is proposed. Special adaptive algorithms permitting optimal distribution of memory space during calculations is suggested. Effectiveness and scalability of the new parNFTC code for the IBM pSeries 690 Regatta system are demonstrated. New advantages of parallel code for solving optimization plasma parameters problem for reactor devices are shown.
Decentralized Replica Exchange Parallel Tempering: An Efficient Implementation of Parallel Tempering using MPI and SPRNG
Yaohang Li (North Carolina A&T State University), Michael Mascagni (Florida State University), Andrey Gorin (Oak Ridge National Laboratory)

Parallel Tempering (PT), also known as Replica Exchange, is a powerful Markov Chain Monte Carlo sampling approach which aims at reducing the relaxation time in simulations of physical systems. In this paper, we present a novel implementation of PT, so-called decentralized replica exchange PT, using MPI and the Scalable Parallel Random Number Generator (SPRNG) libraries. By adjusting the replica exchange operations in the original PT algorithm, and taking advantage of the characteristics of pseudorandom number generators, this implementation minimizes the overhead caused by interprocessor communication in replica exchange in PT. This enables one to efficiently apply PT to large-scale massively parallel systems. The efficiency of this implementation has been demonstrated in the context of various real value benchmark functions.

Large Scale Micro-Finite Element Analysis of Human Bone Structure on the IBM BlueGene/L Supercomputer
Peter Arbenz (ETH Zurich); Costas Bekas, Alessandro Curioni (IBM Research); G. Harry van Lenthe (Katholieke Universiteit Leuven); Ralph Mueller, Andreas Wirth (ETH Zurich)

Coupling recent imaging capabilities with microstructural finite element (microFE) analysis offers a powerful means to determine bone stiffness and strength. It shows high potential to improve individual fracture risk prediction, a tool much needed in the diagnosis and treatment of osteoporosis that is, according to the WHO, second only to cardiovascular disease as a leading health care problem. We adapted a multilevel preconditioned conjugate gradient method to solve the large voxel models that arise in microFE bone structure analysis. We targeted the IBM BG/L supercomputer and improved both algorithms and implementation techniques in order to exploit its excellent scale-out potential. We conducted a study of real (as well as of artificial) human bone models resulting in very large sparse systems of up to about 3 billion unknowns. These runs required less than half an hour, using up to 8 racks of an IBM BG/L system (8192 nodes).
Large-scale FE Software IPSAP for High Performance Computing
Min Ki Kim, Seung Jo Kim (Seoul National University)

This paper introduces large scale finite element structural analysis software, IPSAP, for high performance computing. IPSAP (Internet Parallel Structural Analysis Program) was developed by ASTL. IPSAP is able to solve linear static analysis, thermal conduction analysis, and vibration analysis of structures. IPSAP has three type solution modules, two linear solution methods and one eigensolution method. Linear solver in IPSAP contains both serial/parallel version of Multifrontal method and hybrid domain decomposition method, and eigensolver is block-Lanczos method. Multifrontal solver is the best direct solution method in terms of requirements of computations, memory, and parallel efficiency. It is extremely efficient in the both serial and parallel machines. Hybrid domain decomposition method is based on FETI-DP method which uses multifrontal method to computational efficiency. Last, block-Lanczos eigensolver was developed for multiple eigenvalues/eigenvectors of structures. By using developed solver in IPSAP, we can achieve good parallel performance and efficiency with direct solution method.

A New O(N) Method for Petascale Nanoscience Simulations
Zhengji Zhao, Juan Meza, Lin-Wang Wang (Lawrence Berkeley National Laboratory)

We present a new linear scaling 3-dimensional fragment (LS3DF) method for ab initio electronic structure calculations. The method is based on a divide-and-conquer approach with a novel scheme to patch up the divided fragments. LS3DF has excellent numerical agreement with direct local density approximation (LDA) calculations and can be used to simulate systems with tens of thousands of atoms. In addition, the method scales to thousands of processors and is thousands of times faster than the comparable direct LDA calculations. We will present performance results of LS3DF on different computer platforms, including IBM SP3 and Cray-XT3/XT4 machines. We will focus on parallel scaling, algorithmic scaling, IO rate, global convergence rate, and the error analysis with different fragment sizes. We will also present results for the computation of dipole moments of CdSe quantum dots, which have shed light on a decades old problem in physics.

A Massively Parallel Simulator for Nano-Electronics
Hansang Bae, Steve Clark, Gerhard Klimeck, Sunhee Lee, Maxim Naumov, Faisal Saied (Purdue University)

Simulating realistic nano-electronic devices using an atomistic, quantum mechanical model leads to very large scale problems for the electronic structure. NEMO 3-D is scal-
able code that can model devices such as quantum dots. We present new performance analyses that demonstrate that our electronic structure code scales effectively on the Cray XT3, the BlueGene, and an Intel Woodcrest cluster. We are especially interested in exploiting petascale compute power for nano-electronic applications. Through an in-depth performance analysis of NEMO3d on the most advanced architectures, we show that our code achieves a very favorable balance between computation and communication for a relatively small number of atoms per core, in the weak scaling context. Our results, including some up to 8192 processors, show that NEMO 3-D is very well-suited for scaling up significantly and computations with 50 to 100 millions of atoms could be routinely performed on a petascale machine.

The LQCD Workflow Experience: What We Have Learned
Luciano Piccoli, Xian-He Sun (Illinois Institute of Technology); James N. Simone (Fermi National Laboratory); Alaknantha Eswaradass (Illinois Institute of Technology); Donald J. Holmgren (Fermi National Laboratory); Hui Jin (Illinois Institute of Technology); James B. Kowalkowski, Nirmal Seenu, Amitaj G. Singh (Fermi National Laboratory)

Lattice Quantum Chromo Dynamics (LQCD) is an identified grand challenge scientific application employing large-scale numerical calculations to extract predictions of the standard model of high-energy physics. With the collaboration of several scientific workflow teams, we have conducted an in-depth study of workflow issues of LQCD. We studied workflow requirements of LQCD, investigated functionalities of existing systems, and worked closely with two teams to test their systems. We find there is no existing workflow system ready to use. This leads us to investigate the differences between scientific workflows and general service-oriented workflows, and the differences between scientific workflows and conventional parallel process scheduling. Grid service is the integration of service oriented computing and HPC. We find current workflow systems lack the ability to integrate these two computing structures seamlessly. Based on our experience we propose a two-level workflow solution to separate service dependence and execution scheduling.

Multi-Core Applications

Co-Processor Acceleration of an Unmodified Parallel Structural Mechanics Code with FEAST-GPU
Dominik Goeddeke, Hilmar Wobker (University of Dortmund); Robert Strzodka (Stanford University); Jamaludin Mohd-Yusof, Patrick McCormick (Los Alamos National Laboratory); Stefan Turek (University of Dortmund)

FEAST is a hardware-oriented MPI-based finite element solver package capable of leveraging graphics cards as scientific coprocessors. Previously, the authors have demonstrated significant speedups in the solution of the scalar Poisson problem through the addition of GPUs to a commodity based cluster [Goeddeke:2007:UGT]. In this paper we
work with a real-world application code built on top of FEAST that uses a more complex data-flow, puts higher requirements on the solver, and has a more diverse CPU/co-processor interaction. In particular, we verify the claims that the restricted precision of the GPU does not compromise the accuracy of the final result in any way, the GPU accelerated nodes clearly improve important metrics (price, power, space/performance) of the cluster, and that all this is achieved with the unmodified structural mechanics code that previously only executed on the CPUs. We demonstrate a scalability series and results for different solid objects under load.

Optimization, Parallelization and Characterization of an Probabilistic Latent Semantic Analysis Implementation
Chuntao Hong (Tsinghua University), Jiulong Shan (Intel China Research Center), Wenguang Chen (Tsinghua University), Yurong Chen (Intel China Research Center), Weimin Zheng (Tsinghua University), Yimin Zhang (Intel China Research Center)

Probabilistic Latent Semantic Analysis (PLSA) is one of the most popular statistical techniques for the analysis of two-model and co-occurrence data. However, the fact remains that PLSA is rarely applied to large datasets due to its high computational complexity. This paper presents an optimized Tempered Expectation-Maximization (TEM) implementation of PLSA. The optimized implementation uses the same memory size but reduces the computational complexity significantly compared to a well-known simple PLSA implementation used in Lemur. According to our experiments, it is 1000 times faster than the Lemur implementation on datasets containing more than 5,000 documents. We also parallelized the implementation with OpenMP. To achieve better load balance, we propose a block dividing algorithm and a job scheduling scheme and get fair speedup on multicore systems. The performance analysis of the parallel implementation indicates that this program is memory intensive and the limited memory bandwidth is the bottleneck for better speedup.

Calculation of the Flow over a Hypersonic Vehicle using a GPU
Eric Darve, Patrick LeGresley, Erich Elsen (Stanford University)

Graphics processing units are capable of impressive computing performance up to 345 Gflops peak performance. Various groups have been using these processors for general purpose computing; most efforts have focused on demonstrating relatively basic calculations, e.g., numerical linear algebra or physical simulations for visualization purposes with low accuracy. This poster will describe the simulation of a hypersonic vehicle configuration using the compressible Euler equations. To the authors’ knowledge, this is the most sophisticated calculation of this kind in terms of complexity of the geometry, the physical model, the numerical methods employed, and the accuracy of the solution. Based on a comparison of the Intel Core 2 Duo and NVIDIA 8800GTX, speed-ups of over 40x were demonstrated for simple test geometries and 20x for complex
Implementation of an NAMD Molecular Dynamics Non-bonded Force-field on the Cell Broadband Engine Processor
Guochun Shi, Volodymyr Kindratenko (National Center for Supercomputing Applications)

The Cell/B.E. processor has been evaluated for implementing and running the NAMD molecular dynamics code with both the single-precision and double-precision kernels. The non-bonded force-field kernel, as implemented in the NAMD SPEC 2006 CPU benchmark, has been implemented on a pre-production 2.4 GHz Cell/B.E. blade system, showing linear speedups when using multiple synergistic processing elements. We observe that performance of the double-precision floating-point kernel differs only by a factor of ~2 from the performance of the single-precision floating-point kernel. Since the peak performance of the Cell’s single-precision floating point SIMD engine is 14 times the peak performance of the double-precision floating-point SIMD engine, our findings underscore the disconnect between the theoretical peak performance and the actual application performance achievable on this architecture. Our results point out the potential of the Cell/B.E. processor in accelerating applications that use the double-precision floating-point math beyond what is achievable on the mainstream microprocessors.

The Sony PlayStation 3 and the NVIDIA 8800 GPU: Performance and Programmability Evaluation for Machine Learning
Ahmed El Zein, Eric McCreath, Alistair Rendell (Australian National University); Alex Smola (NICTA)

This poster outlines our efforts to exploit the Sony PlayStation3 (PS3) and the NVIDIA GeForce 8800 for machine learning (ML) applications. It will compare both the performance obtained and the programmability of these two different systems. The ML algorithm is iterative requiring two large matrix vector operations per iteration. Our implementation strategy is different on each machine. On the 8800 we have explored the use of NVIDIA’s Compute Unified Device Architecture (CUDA) and its associated BLAS library (cuBlas). On the PS3 we have programmed our own routines. Comparison will be made between the use of these non-conventional compute platforms with alternatives (the PowerPC processor and a dual core Athlon64). This work is on going, and while we have preliminary results for both systems we expected to have additional data by the time of SC07. Results will be presented in a standard poster format using a variety of tables and figures.

GPU-Enhanced Conjugate Gradient Solver
Serban Georgescu, Hiroshi Okuda (University of Tokyo)

The Conjugate Gradient (CG) method is one of the most commonly used iterative methods for solving very large systems of equations. Memory-bound and with irregu-
lar memory access patterns, CG performs very poorly on standard processors.
Motivated by the widening performance gap and their recent entrance in the HPC market, we investigate the performance of a last generation GPU as an accelerator to the CG method. In this poster we present performance results on a wide range of matrices taken from real world problems. We work directly in compressed format and obtain what we believe to be the highest sparse matrix-vector multiplication speed reported so far. We discuss the biggest issues we encountered while working in single precision, most important the need for good preconditioners. Without losing any accuracy, we obtain a large speed boost for the larger matrices, with one order of magnitude speedup within reach.

Architecture

Collective Algorithms for Kautz Bus Networks
Robert B. Thayer (Sun Research)

As the processor count in cluster computers continues to grow, the scalability of both the interconnect network(s) and, increasingly, the associated collective algorithms become critical to system performance. Families of highly scalable networks based on Kautz graph and hypergraph topologies are briefly reviewed and compared with incumbent networks. Novel algorithms for essential global collectives—barrier, broadcast, reduce-all, all-to-all, etc.—for a Kautz hypergraph with node degree d and bus size s are developed and found to complete in \(\log dsP\) versus the \(\log 2P\) required for hypercube networks. Initial adaptations of algorithms for vector reduction, matrix transposition, and mergesort are also developed with emphasis on exploiting the Kautz topologies while minimizing implementation costs. Cost tradeoffs appear to be in reducing the number of communication phases versus modest increases in data duplication due to the inherent ds-ary tree structure of the Kautz topologies.

Cluster Design Space Exploration with the CDR: Evaluation and Observations using the Top500 Supercomputers
William R. Dieter, Henry G. Dietz (University of Kentucky)

Design of cluster supercomputers is complicated by the combinatorial explosion in the number of ways that commodity components can be combined. Component and architectural effects on application performance, power, cooling, and operating cost constraints, as well as frequent component price changes further complicate design decisions. The Cluster Design Rules (CDR) can solve this constrained optimization problem using a branch-and-bound search algorithm that evaluates designs using application performance models. CDR designs generated by modeling High Performance Linpack execution time and optimizing for acquisition cost, power consumption, operating cost, and total cost of ownership are compared to designs in the Top 500 list. Depending on the optimization criteria, designs vary widely in cost, number of nodes, amount of memory, type of network, and type of processors.
A Methodology for Coping with Heterogeneity of Modern Accelerators on a Massive Supercomputing Scale
Toshibo Endo, Satoshi Matsuoka (Tokyo Institute of Technology)

Heterogeneous supercomputers with combined general-purpose and accelerated CPUs promise to be the future major architecture due to their wide ranging generality and superior performance-power ratio. However, despite the “hype,” developing applications that achieve effective scalability is still very difficult. We show that an effective strategy for such heterogeneous machines entails careful analysis of the application algorithm, and virtualizing the underlying compute resources, so that the porting from applications written with homogeneous assumptions could be achieved. We demonstrate our methodology with our modified heterogeneous HPL on the TSUBAME heterogeneous supercomputer. We efficiently load balanced between over 10,000 general purpose CPU cores and 360 SIMD accelerators so that both resources that are effectively utilized in a combined fashion—he resulting 48.88 TFlops utilized the entire machine, and not only ranked 14th in the world on the Top500, but also became the first heterogeneous machine to be ranked on the list.

Early Evaluation of On-Chip Vector Caching for the NEC SX Vector Architecture
Akihiro Musa (Tohoku University/NEC Corporation); Yoshiei Sato, Ryusuke Egawa, Hiroyuki Takizawa, Koki Okabe, Hiroaki Kobayashi (Tohoku University)

The sustained performance of supercomputers strongly depends on their memory performance. Modern vector supercomputers deliver a high computational efficiency to real applications. However, non-stop demands for accelerating processor performance give a big pressure on the off-chip memory bandwidth requirement. To relax the off-chip memory requirement for vector supercomputers, we are designing an on-chip vector cache with a selective caching and bypassing mechanism for the NEC SX supercomputers. We present an early evaluation of on-chip vector caching for the NEC SX-7 vector processor when its off-chip memory bandwidth is reduced to half and quarter, by using a kernel and five real application codes. The experimental results indicate that even with a 2MB vector cache, performance loss due to a half memory bandwidth can be recovered by 20% to 98% in five applications. In addition, selective vector caching is effective for efficient use of the limited on-chip cache capacity.

Thermal-aware High Performance Computing using TEMPEST
Hari K. Pyla, Dong Li, Kirk W. Cameron (Virginia Tech)

Despite the critical impact of heat on the operation and availability of high-end clusters, there are no tools in the extant literature that provide significant insight into the thermal characteristics of parallel scientific applications. Thermal simulators are difficult to validate and incapable of efficiently profiling thermals at scale. Thermal sensors provide raw data not particularly useful for analysis.
unless correlated to workload. We created an infrastructure to gather data from temperature sensors, correlate the data to source code, and control the thermal characteristics of an application at runtime. Our direct measurements show distributed scientific application thermal profiles are often affected by adjacency of code hot spots. We tested control theoretic techniques to maintain constant temperature and analyzed impact on parallel application performance. Our results indicate that thermal throttling can be accomplished using a systemic controller with less than 10% performance impact for the NAS parallel benchmark codes we measured.

Evaluating the Role of Scratchpad Memories in Multicore for Sparse Matrix Computations

Aditya Yanamandra, Bryan Cover, Konrad Malkowski, Padma Raghavan, Mahmut Kandemir, Mary J. Irwin (Pennsylvania State University)

We consider hardware acceleration for sparse matrix vector multiplication (SpMV), a kernel that is used widely in iterative linear solvers, modeling and simulation applications. In particular, we consider how scratchpad memory can be used for increasing the performance and the energy efficiency of SpMV in a multi-core system. Scratchpad memories (SPM) are more energy efficient than traditional caches. This, coupled with the predictability of data presence, makes SPM an attractive alternative to a cache. We ensure the efficient utilization of the SPM by using it to store data which doesn’t perform well in the traditional cache. We evaluate the impact of using an SPM at all levels of the on-chip memory hierarchy. Depending on the level of the hierarchy in which the SPM is utilized, we observe on an eight core system an average increase in performance of 13.5%-15% at an average decrease in energy consumption of 23%-28%.

CellFS: Taking The "DMA" Out Of Cell Programming

Latchesar Ionkov, Aki Nyrhinen (University of Helsinki), Andrey Mirchovski (Los Alamos National Laboratory)

In this paper we present a new programming model for the Cell BE architecture. We call this programming model CellFS. CellFS aims at simplifying the task of managing I/O between the local store of the processing units and main memory. The CellFS support library provides the means for transferring data via simple file I/O operations between the PPE and the SPE.

Checkpointing Parallel Applications using Aspect Oriented Programming

Ritu Arora, Purushotham Bangalore (University of Alabama at Birmingham)

One of the key ingredients in writing self-healing parallel applications that can be migrated from one resource to another in a heterogeneous computing environment is checkpointing. Checkpointed applications
are reliable, fault tolerant and flexible and hence are useful in ensuring desired Quality of Service to the user. Since complete reengineering of an already existing parallel application to embed checkpointing logic is a daunting task, mainly due to the time and cost overheads involved in the debugging and reengineering process, we opted for Aspect Oriented Programming (AOP) which promotes code reusability and adaptability. Using AOP and code refactoring, we seamlessly injected the checkpointing logic into an existing parallel application. This poster demonstrates the use of AOP techniques to reengineer a parallel genetic algorithm to introduce checkpointing without requiring modifications to the original program.

**Parallel Performance Wizard: A Generalized Performance Analysis Tool**  
Hung-Hsun Su, Max Billingsley III, Seth Koehler, John Curreri, Alan D. George  
(University of Florida)

Scientific programmers using any of the existing parallel programming models often must rely on performance analysis tools to help them optimize the performance of their programs. There are many existing tools designed to fill this need, but they generally support a limited selection of programming models and cannot be easily extended to support additional models. In tackling this problem, we present the extension of Parallel Performance Wizard (PPW), a tool originally designed for partitioned global-address-space (PGAS) models, to employ a generic operation-type-based framework that makes it possible to add support for additional models with minimal effort. As a proof of concept, we show how support for the Message Passing Interface (MPI) was quickly added to PPW and discuss work currently underway to support reconfigurable computing applications employing both traditional CPUs and FPGAs.

**Improving All-to-All Communication for Parallel MATLAB**  
David E. Hudak, Neil Ludban, Vijay Gadepally, Ashok Krishnamurthy (Ohio Supercomputer Center)

ParaM is a software distribution for HPC systems that provides parallel execution of MATLAB scripts. ParaM supports the “message passing” programming model via its MPI binding layer (bcMPI) which supports modern interconnects such as InfiniBand and Myrinet. ParaM also supports the partitioned global address space (PGAS) programming model via pMatlab, a PGAS implementation for MATLAB programs developed at MIT Lincoln Labs. In this poster, we will present an optimization strategy for pMatlab in which the implementation of an all-to-all communication pattern is moved from a MATLAB script to a compiled C extension. This strategy eliminates the overhead of creating, copying, and destroying many temporary matrices and allows use of MPI_Alltoall in place of a sequence of MPI_Send and MPI_Recv calls. Our simulations show improvements of 3x-8x for large matrices and up to 5x for small matrices.
I/O is a known bottleneck in high-end computing (HEC). Although advanced parallel file systems have been developed in recent years, they provide high bandwidth only for large, well-formed data streams, and perform poorly for accessing small, noncontiguous data. However, many HEC applications make a large number of requests for small and noncontiguous pieces of data. In this poster, we present a novel Server-Push Architecture to tackle the parallel I/O problem. Unlike traditional I/O designs where data is stored and retrieved by request, in our architecture, a File Access Server (FAS) proactively pushes data from a file server to its client node’s memory. FAS is responsible for identifying I/O access patterns, predicting future references, and pushing data before an I/O request is generated by a client. Our initial results show that the Server-Push Architecture has significant potential in improving I/O performance of parallel applications.

Availability of graphics-processors based compute devices and multi-core host architectures, with larger memories on both, means that it is possible to run large scientific computing problems on "personal" machines. For wide adoption by scientists and achieve an increase in their productivity these architectures must be relatively easy to use in the languages scientists use, such as FORTRAN-9x, and without having to retranslate their thinking and algorithms into graphics metaphors. At the same time, to actually achieve good performance developers must be aware of issues, such as the large cost for device-host communications, the need to achieve optimal occupancy, the need to use local memory, etc. Using NVIDIA’s CUDA architecture on 8800GTX GPUs hosted in a multicore Intel host, we develop middleware that allows use of the architecture in conventional Fortran-9x. Applications in plasma turbulence and radial-basis-function interpolation are ported and achieve speedups of 25 and 662 on the GPU.

Petascale application developers will need monitoring and analysis tools capable of processing enormous data volumes. For 100,000 or more processors, an instrumented application can generate terabytes of event trace data during a single execution. At
such scales, the transfer and storage of complete traces is infeasible, necessitating tools that reduce, process and analyze data on-line before human exploration and optimization. We are developing a scalable, reconfigurable infrastructure for performance analysis on large-scale machines. One of its central capabilities is the collection of near-constant size communication traces. In this poster, we build on this capability with new mechanisms that annotate runtime traces with timing statistics and computational load measures. We describe our experiences using these techniques on current scientific applications. We also give preliminary results showing these techniques can be used to visualize time-evolution of load imbalance and to replay codes for accurate postmortem analysis.

Using MPI Communication Patterns to Guide Source Code Transformations
Robert Preissl (Johannes Kepler University); Martin Schulz (Lawrence Livermore National Laboratory); Dieter Kranzlmueller (Johannes Kepler University); Bronis R. de Supinski, Daniel J. Quinlan (Lawrence Livermore National Laboratory)

Optimizing the performance of HPC software requires a high-level understanding of communication patterns as well as their relation to source code structures. Such patterns can be either user-defined or automatically extracted from program traces. In both cases they must be associated with static program information to enable efficient and automatic source code transformations. In this poster we describe an algorithm to detect communication patterns from parallel traces as well as early experiences in using this information to enhance static code analysis and to subsequently optimize MPI codes. First we detect patterns that identify potential bottlenecks in MPI communication traces. Then we associate them with the corresponding nodes in an Abstract Syntax Tree (AST) using the ROSE compiler framework. This combined information is afterwards used together with static analysis, like System-Dependence-Graphs (SDG) and Control-Flow-Graphs (CFG) to guide static optimizations like code motion or the automatic introduction of MPI collectives.

Characterization of Intra-node Topology and Locality
Kevin T. Pedretti (Sandia National Laboratories)

This poster presents the results of a study examining the effects of intra-node locality and topology on an Intel dual quad-core system. MPI micro-benchmarks and a CG solver mini-application are used to characterize the impact of various rank to physical-CPU mappings. To accomplish this, the MPICH2 mpiexec launcher was modified to allow load-time specification of the mapping. Significant differences are observed between on-die, on-package, and inter-socket communication. A surprising discovery from this study is that the physical CPUs on the platform examined were numbered in a non-intuitive way, actually harming performance if straight-forward ascending order CPU affinity is enabled. Since there is no one-size-fits-all solution, an application-level interface is proposed for controlling the
rank-to-CPU mapping at runtime. This interface is currently being implemented in LIBSM, a library solution for augmenting MPI applications with intra-node shared memory operations. Early performance data for the new LIBSM Linux port is also presented.

Networking and Grids

Performability Modeling for Scheduling and Fault Tolerance Strategies for Grid Workflows
Lavanya Ramakrishnan (Indiana University), Daniel A. Reed (Renaissance Computing Institute)

Grid applications have diverse characteristics and resource requirements. When combined with the complexity of underlying Grid resources, these applications can experience significant performance and reliability variations. Although the performance and reliability of Grid systems have been studied separately, there has been little analysis of the lost Quality of Service (QoS) with varying failure levels. Next generation Grid tools need extensible application interfaces that allow users to qualitatively express combined performance and reliability requirements for the underlying systems. In this poster, we use the concept of performability to capture and analyze the degraded performance that might result from varying resource reliability levels. We also present the use of performability as a basis for workflow scheduling and fault tolerance strategies.

Co-Processor Acceleration of an Unmodified Parallel Structural Mechanics Code with FEAST-GPU
Dominik Goeddeke, Hilmar Wobker (University of Dortmund); Robert Strzodka (Stanford University); Jamaludin Mohd-Yusof, Patrick McCormick (Los Alamos National Laboratory); Stefan Turek (University of Dortmund)

FEAST is a hardware-oriented MPI based finite element solver package, capable of leveraging graphics cards as scientific co-processors. Previously, the authors have demonstrated significant speedups in the solution of the scalar Poisson problem through the addition of GPUs to a commodity based cluster [Goeddeke: 2007:UGT]. In this paper we work with a real-world application code built on top of FEAST that uses a more complex data-flow, puts higher requirements on the solver, and has a more diverse CPU/co-processor interaction. In particular, we verify the claims that the restricted precision of the GPU does not compromise the accuracy of the final result in any way, the GPU accelerated nodes clearly improve important metrics (price, power, space/performance) of the cluster, and that all this is achieved with the unmodified structural mechanics code that previously only executed on the CPUs. We demonstrate a scalability series and results for different solid objects under load.
Parallel Streaming: Tenfold Accelerations of Office/Database/Web/Media Applications over the Internet/Grids
Frank Wang (Cambridge-Cranfield HPCF); Na Helian (London Metropolitan University); Sining Wu, Yuhui Deng, Vineet R. Khare, Chenhan Liao (Cambridge-Cranfield HPCF); Amir Nathoo, Rodric Yates, Paul Fairbairn (IBM); Jon Crowcroft, Jean Bacon, Michael Andrew Parker (Cambridge University); Zhiwei Xu (Institute of Computer Technology); Yike Guo (Imperial College)

g-Jet is the first demonstration that a Grid-enabled data communication protocol can accelerate tenfold distributed applications, including OpenOffice, MySQL/IBM DB2, Firefox, MPlayer, and Google Earth, in real-world tests. Inspired by the success of GridFTP, g-Jet integrates a parallel stream engine and Grid Security Infrastructure (GSI). Conforming to the universal VFS (Virtual Filesystem Switch) semantics, g-Jet can be pervasively used as an underlying platform to accelerate other applications. Based on a parallel stream model in a lossy network, a dynamic numbering optimizer is designed to automatically maintain high throughput over the life time of the connection. A cache layer is introduced in the “Data Window” to accumulate the file system calls with default size and then flush them to the underlying g-Jet, resulting in more data sent in fewer trips. The source code is 40,000 lines in length and we have spent six man years in developing and revamping it.

Performance Analysis of Volunteer Computing Traces
Trilce Estrada, Michela Taufer (University of Delaware); Kevin Reed (IBM)

Volunteer Computing (VC) projects deploy computers on the Internet owned by volunteers. Their performance is measured in terms of throughput (valid results delivered to the scientists in a given time interval). The path that leads from the generation of work-units to the result validation is error-prone and is characterized by delays hidden in the different phases of a work-unit’s lifetime. The proposed work aims: (1) to identify the critical phases of a work-unit’s lifetime across projects, from its generation to its final result validation, which take more time than anticipated, and (2) tailor scheduling polices that target these phases so that ultimately the overall throughput increases. This includes understanding whether volunteers’ hosts are more active at particular times of the day or week and what host features are vital for the success of a work-unit as well as integrate this knowledge in the scheduling policy applied by the VC project.

XML Data Unification for Visualization
Svetlana Shasharina, Paul Hamill (Tech-X Corporation)

Under the auspices of the US DOE SciDAC program, projects such as CEMM, SWIM, FACETS and COMPASS are using high-end supercomputing to expand our knowledge in fusion and accelerator physics. Such simulations generate data using very different formatting conventions, even if they use
Posters

Towards Terabit/s Systems: Performance Evaluation of Multi-Rail Systems
Venkatram Vishwanath (University of Illinois at Chicago); Takashi Shimizu, Makoto Takizawa, Kazuaki Obana (NNT Network Innovations Laboratory); Jason Leigh (University of Illinois at Chicago)

We present a novel multi-rail approach that is necessary in order for future E-Science applications to effectively exploit Terabit/s networks. The multi-rail approach consists of creating parallel "rails" through every aspect of an end-system: from processing on the multiple and many cores, generation of multiple application data flows, and streaming over multiple-lanes, multi-wavelength NICs connected via a parallel interconnect. In the poster, we present the evaluation of end-systems parameters that impact the efficiency of multi-rail systems such as interrupt, memory, thread, and core affinities. These evaluations were tested on the ability of individual cluster nodes to achieve TCP and UDP throughput at 10Gbps and 20Gbps rates. We analyze the additive effects of the parameters - a key property for achieving scalable performance towards Terabits/s. Thread and Interrupt affinity together was found to have an additive effect and plays a critical role in achieving maximum throughput.

GSIMF: A Service Based Software and Database Management System for the Next Generation Grids
Nanbor Wang, Balamurali Ananthan (Tech-X Corporation); Alexandre Vaniachine, Gerald Gieraltowski (Argonne National Laboratory)

To process the vast amount of data from scientific research experiments, scientists rely on Computational and Data Grids; yet, the distribution, installation, and updating of a myriad of different versions of different programs over the Grid environment is complicated, time-consuming, and error-prone. Our Grid Software Installation Management Framework (GSIMF) is a set of Grid Services that has been developed for managing versioned and interdependent software applications and file-based databases over the Grid infrastructure. This set of Grid services provide mechanism to install software packages on distributed Grid computing elements, thus automating the software and database installation management process on behalf of the users. The GSIMF will contribute to a robust environment for various
data intensive and collaborative applications, such as nuclear physics experiments, space science observations, and climate modeling. Our poster covers the architectural details of our GSIMF system and several case studies where it is been successfully deployed.

A High-Performance GridFTP Server at Desktop Cost
Samer Al Kiswany, Armin Bahramshahry, Hesam Ghasemi, Matei Ripeanu (University of British Columbia); Sudharshan S. Vazhkudai (Oak Ridge National Laboratory)

We prototype a storage system that provides the access performance of a well endowed GridFTP deployment (e.g., using a cluster and a parallel file-system) at the modest cost of single desktop. To this end, we integrate GridFTP and a combination of dedicated but low-bandwidth (thus cheap) storage nodes and scavenged storage from LAN-connected desktops that participate intermittently to the storage pool. The main advantage of this setup is that it alleviates the server I/O access bottleneck. Additionally, the specific data access pattern of GridFTP, that is, the fact that data accesses are mostly sequential, allows for optimizations that result in a high-performance storage system. To provide data durability when facing intermittent participation of the storage resources, we use an intelligent replication scheme that minimizes the volume of internal transfers that impact the low-bandwidth storage nodes.

Data Stream Management in Global-Scale Ecological Observatory Networks
Ebbe Strandell, Hsiu-Mei Chou, Yao-Tsung Wang, Fang-Pang Lin (Taiwan National Center for HPC); Sameer Tilak, Peter Arzberger (University of California, Davis)

Ecological Observatory Networks are playing an important role to drive new scientific discovery. One of the current key issues is that the size of collected data is growing exponentially as networks expand, and over time diverse data sources can easily accumulate huge amounts of data. How to access, manage, and archive this data effectively has emerged to a challenging field in IT development. In this work, a robust data stream management system is demonstrated. The system allows users to subscribe to two-way data channels on the fly, either for sources of dynamic data or for legacy data. The system adapts RBNB DataTurbine and SRB technologies to bridge users, sensors and persistent storage pools in a scalable manner. Parallel data streaming from two ecological observatories in Taiwan are used as proof-of-concept, allowing us to demonstrate how the system can be effectively used for sharing data amongst global science and IT communities.
ACM Student Research Competition

Following a successful introduction last year, students again were invited to submit posters as part of the internationally recognized ACM Student Research Competition (SRC) sponsored by Microsoft Research. The SRC venue allows students the opportunity to experience the research world; share research results with other students, judges, and conference attendees; and rub shoulders with academic and industry leaders. Six student posters were selected and are displayed alongside the research posters. In a special technical program session, students will have the opportunity to present their work before a panel of judges. Winners will receive cash prizes and are eligible to enter the annual Grand Finals, a culmination of the ACM Student Research Competitions during the academic year. To learn more about the SRC, visit www.acm.org/src.

Wednesday, Nov. 14

Student Posters
Room: A10 / A11
10:30 a.m. - 12:00 p.m.

A Dynamic Programming Approach to Kd-Tree Based Data Distribution
Susan Frank (Stony Brook University)

We propose a dynamic programming approach for finding a kd-tree partition which gives a balanced data distribution among rendering nodes for distributed ray tracing. We define a stage as the portion of the scene between a start and end set of cut partition planes. Stages are ordered in increasing z, y, x value. We determine the minimum cost for stages which may be part of the optimal solution starting with the last qualifying stage and moving toward the first. The cost of a stage includes the rendering cost of the assigned subvolume and optimal costs of subsequent stages associated with a specific partition. The partition is defined by a set of non-overlapping subsequent stages, one for each positive x, y and z direction.

Our algorithm has been successfully used to create partitions and data distributions for several data sets.

Storing and Searching Massive Scale-free Graphs
Timothy Hartley (Ohio State University)

There is a strong demand for scalable massive graph analysis technology from a wide range of areas such as national security to business decision making. Currently, there are few practical solutions. To answer this challenge, we have developed MSSG, a parallel system capable of storing and accessing extreme-scale graphs with tens to hundreds of billions of edges, and which enables on-demand data analysis with short query execution times. Through our use of a novel external-memory graph data structure, MSSG can store massive scale-free graphs which present severe challenges for systems based on traditional database management systems. Moreover, MSSG is specifically designed for cluster-based systems with local disks, obviating the need for application-specific and expensive custom database systems.
In this work we show our results from searching a massive scale-free graph with 120 billion edges, which we believe is the largest graph searched for source-target path lengths on a commodity cluster.

**Obtaining High Performance via Lower-Precision FPGA Floating Point Units**

*Junqing Sun (University of Tennessee, Knoxville)*

Experimental results and vendor specifications reveal that lower-precision floating point components on FPGAs cost fewer resources, require lower memory bandwidth, and can achieve higher frequency compared to higher-precision components. The idea of this paper is to seek high performance for linear equation solvers by using lower-precision floating point arithmetic whenever possible. The high accuracy of final solutions is achieved by higher-precision iterative refinements on lower-precision intermediate results. Our mixed-precision algorithm and design achieve a speedup of 37x for the LU decomposition and 10x for direct solvers over a 2.2GHz Opteron processor on a Cray-XD1 supercomputer. The mixed-precision algorithm, LU and solver architecture, and profiled performance data will be presented in the poster session. Real tests on a Cray-XD1 supercomputer can also be demonstrated.

**Performance Analysis and Optimization of Large-scale Scientific Applications on Clusters with CMPs**

*Charles Lively (Texas A&M University)*

The current trend in computer systems is shifting towards the use of chip multiprocessors (CMPs). A major challenge is efficient use of the CMPs that are used to comprise a cluster system. In this poster, we conduct detailed experiments to identify methods for efficient execution on such systems. In particular, we focus on three scientific applications: a simulation of four dimensional SU(3) lattice gauge theory (MILC), an advanced Eulerian gyrokinetic-Maxwell equation solver for simulating microturbulent transport in plasma (GYRO), and a Lattice Boltzmann Method for simulating fluid dynamics (LBM). Experimental results indicate that performance optimization on these applications can be obtained from methods, such as loop unrolling and loop blocking, and utilizing alternative global communication schemes. For example, with the LBM, a 15.77% improvement occurred when the code was modified to use loop blocking on 32 processors of the SDSC DataStar p655, when using all 8 processors per node.

**Evolving the GPU-based Cluster**

*Jay E. Steele (Clemson University)*

Today’s current graphics processing units (GPUs) are programmable, highly-parallel, floating point processors that are ideally suited for use in high performance clusters. GPU-based clusters have been successfully used by researchers to solve large, computationally intensive numerical problems. No current software system exists that provides researchers with high-level tools to fully exploit the performance GPU-based clusters. We have begun development of such tools. To this end, we have developed a C library, Purple, that facilitates solving numerical problems with GPU-based clusters. Though far from complete, Purple is currently
ACM Student Research Competition

Doctoral Research Showcase

A new feature this year is the Doctoral Research Showcase. These sessions provide venues for PhD students who will be graduating in the next 12 months to present a short summary of their research. We selected 12 students to make presentations in two different sessions from the 48 proposals received. To accommodate as many people as possible, presentations have been limited to 15 minutes each. Everyone attending the technical program is encouraged to attend these sessions. Come see what the next generation of HPC researchers are up to.

Wednesday, Nov. 14

Doctoral Research I
Room: A10 / A11
Session Chair: Jeffrey K. Hollingsworth
(University of Maryland)

3:30 p.m. - 5:00 p.m.

Improving Power-Performance Efficiency in High-End Computing
Rong Ge (Virginia Tech)

Today, power consumption costs supercomputer centers millions of dollars annually and the heat produced can reduce system reliability and availability. Achieving high performance while reducing power consumption is challenging since power and performance are inextricably interwoven;

capable of automatically decomposing certain stream processing problems along the major dimension, distributing the workload to individual nodes, handling necessary inter-node communication, and reassembling the results. From the researcher’s point of view, the cluster appears as a single large stream processor. We will present Purple, including example code and performance results for certain numerical problems.

GrenchMark: A Framework for Testing Large-Scale Distributed Computing Systems
Alexandru Iosup (Delft University of Technology)

The dynamicity, the heterogeneity, or simply the scale of today’s grids expose problems in the grid performance, reliability, and functionality. Thus, an important research question arises: How to gain insights into the performance and the reliability of large-scale distributed computing systems such as grids? In this work we attempt to address this question with the GrenchMark framework for analyzing, testing, and comparing grids. The framework focuses on realistic and repeatable testing, and on obtaining comparable testing results. We have used the reference implementation in a wide variety of testing scenarios in grids (e.g., Globus-based), peer-to-peer (e.g., BitTorrent-based), and heterogeneous computing environments (e.g., Condor-based), some of which have been briefly presented here. Finally, we briefly describe three future research directions towards a complete framework for the performance evaluation of large-scale distributed systems under realistic workloads, in real and simulated environments.
reducing power often results in degradation in performance. This thesis aims to address this challenge by providing theories, techniques, and tools to 1) accurately predict performance and improve it in systems with advanced hierarchical memories, 2) understand and evaluate power and its impacts on performance, 3) control power and performance for maximum efficiency. Our theories, techniques, and tools have been applied to high-end computing systems. Results show that our theoretical models can improve algorithm performance by 59%; our hardware/software toolset for power profiling provides previously unavailable insight to parallel scientific application power consumption; our power-aware techniques can save 36% energy with little performance degradation.

Qualitative Performance Analysis for Large-Scale Scientific Workflows
Emma S. Buneci (Duke University / Renaissance Computing Institute)

Computational grids promote new modes of scientific collaboration and discovery by connecting distributed instruments, data, and computing facilities. Because resources are shared, application performance can vary widely and unexpectedly. Teresa is a performance analysis framework for large-scale scientific grid applications that reasons temporally and qualitatively about performance data from multiple monitoring levels. Teresa generates qualitative temporal signatures from these real-time, multi-level performance data and compares these signatures to expected behavior. In any mismatch, the framework hints at causes of altered performance and suggests potential solutions to an application user. Experiments with grid workflows from meteorology and astronomy reveal common qualitative temporal signatures characterizing successful and well-performing execution. This allows the framework to learn signatures of typical temporal behaviors. Furthermore, experiments with stimuli affecting performance result in signatures with distinct characteristics. The ability to distinguish qualitative signatures of expected states supports the performance validation and diagnosis of long-running scientific grid applications.

Parallel Performance Wizard: An Infrastructure and Tool for Analysis of Parallel Application Performance
Hung-Hsun Su (University of Florida)

Scientific programmers using any of the existing parallel programming models often must rely on performance analysis tools to help them optimize the performance of their programs. There are many existing tools designed to fill this need, but they generally support a limited selection of programming models and cannot be easily extended to support additional models. To tackle this problem, a new performance analysis tool infrastructure, Parallel Performance Wizard (PPW), was developed so multiple programming models can be supported with minimal effort. In the presentation, this infrastructure, which was designed around generic operation types typically found in parallel programming models, is first introduced. Next, a new bottleneck detection and resolution model, based also on the generic operation types, that supplements the infrastruc-
Compiler Techniques for Efficient Communication in Multiprocessor Systems
Shuyi Shao (University of Pittsburgh)

Multiprocessor interconnection can assume a disproportionately high portion of the system’s cost and circuit switching has been identified as a promising alternative to packet and wormhole switching for achieving higher efficiency with lower cost. However, the overhead of circuit establishment can be large. The benefits of circuit switching can only outweigh its drawbacks when communication exhibits locality and this locality is appropriately explored, ideally, before the application starts execution. Hence, I explore compiler techniques for efficient communication on circuit switching interconnects. A compilation framework is proposed for identifying communication patterns and composing network configuration directives to reduce circuit switching overhead. To overcome the network capacity limits of circuit switching, communication within the application is partitioned into phases. By applying my compiler techniques to applications, it is identified that static and persistent communication dominates in many cases. Simulation-based performance analysis demonstrates the benefits of using my techniques for achieving efficient communication in multiprocessor systems.

On Economics and the User-Scheduler Relationship in HPC and Grid Systems
Cynthia Bailey Lee (University of California, San Diego)

Effective management of Grid and HPC resources is essential to maximizing return on the substantial infrastructure investment these resources entail. An important prerequisite to effective resource management is productive interaction between the user and scheduler. My work analyzes several aspects of the user-scheduler relationship and develops solutions to three of the most vexing barriers between the two. First, users’ monetary valuation of compute time and schedule turnaround time is examined in terms of a utility function. Second, responsiveness of the scheduler to users’ varied valuations is optimized via a genetic algorithm heuristic, creating a controlled market for computation. Finally, the chronic problem of inaccurate user runtime requests, and its implications for scheduler performance, is examined, along with mitigation techniques.

Reliability for Scalable Tree-based Overlay Networks
Dorian C. Arnold (University of Wisconsin)

As HPC’s petascale era approaches, there exists a need for efficient failure recovery models to mitigate the high failure rates of large, complex systems. Tree-based overlay networks (TBONs, hierarchically organized process networks) provide a powerful model for scalable data multicast, gather, and aggregation. We study state compensation, a recovery model for extreme-scale TBONs.
with high throughput, low latency requirements. State compensation leverages information redundancies inherent to TBON computations and weak consistency models to yield a general recovery model that consumes no resources in the absence of failures. We use a formal definition of TBON computations to prove that state compensation preserves computational semantics across failures. This formalization also directs our prototype implementation, which will benchmark recovery performance and demonstrate graceful application performance degradation as failures occur. We also study TBON performance models and compare TBON-based information dissemination protocols with other protocols like epidemic style dissemination.

Thursday, Nov. 15

Doctoral Research II
Room: A10 / A11
Session Chair: Martin Schulz (Lawrence Livermore National Laboratory)

10:30 a.m. - 12:00 p.m.

Efficiently Solving Large-scale Graph Problems on High Performance Computing Systems
Kamesh Madduri (Georgia Institute of Technology)

Graph-theoretic problems have emerged as a prominent computational workload in the petascale computing era, and are representative of fundamental kernels in biology, scientific computing, and applications in national security. However, they pose serious challenges on current parallel machines due to non-contiguous, concurrent accesses to global data structures with low degrees of locality. The focus of my doctoral dissertation is on efficiently solving large-scale graph-theoretic applications arising in complex network analysis, on current high performance computing systems. Exploiting typical topological characteristics of real-world networks, we present novel parallel algorithms and efficient implementations for fundamental kernels in graph traversal and connectivity. We also present fast parallel algorithms for key network analysis applications such as centrality and community detection, and demonstrate the ability to process large-scale graphs with billions of vertices and edges.
High Performance Non-rigid Registration for Image-Guided Neurosurgery
Andriy Fedorov (College of William and Mary)

Non-rigid registration of brain imagery is an essential procedure in medical image processing. It allows us to account for intra-operative brain shift and use pre-operative MRI during the neurosurgical resection of tumors. In my research, I study how HPC can be used to: (1) perform non-rigid registration within the time constraints of the surgery, and (2) improve the accuracy of registration. Using HPC, we were able to reduce the time required for registration to minutes. In our current work we study how large-scale computing resources like TeraGrid can facilitate intra-operative search for optimal parameter setting of a registration algorithm to reduce the alignment error. I will also briefly discuss the educational aspects of my PhD thesis work. For the last two years, non-rigid registration and related research have been providing materials for building a platform for teaching medical imaging and image processing at the College of William and Mary.

Statistical and Pattern Recognition Techniques Applied to Algorithm Selection for Solving Linear Systems
Erika Fuentes (University of Tennessee, Knoxville)

There are many applications in science and engineering where the issue of choosing an appropriate algorithm is common but most of the times too hard for humans to make (most require large-scale numerical computations). In this case, statistical approaches are very useful because they can experimentally help us to uncover relations, identify patterns and extract knowledge from these applications. We propose a statistical approach to develop a methodology to automate such decision-making process in numerical problems. This research centers on using the non-numerical techniques, such as pattern classification and data mining, to experimentally analyze and find the most suitable solution method to a given numerical problem based on its numerical properties, and ultimately build a recommendation system based on the experimental findings.

Migratable and Reparallelizable OpenMP Programs
Michael Klemm (University of Erlangen-Nuremberg)

Most scientists can access various clusters in a computational Grid. Besides having to deal with different architectures, network interconnects, etc., users must estimate the application’s resource needs for the job schedulers to reserve CPUs. Underestimations cause loss of computation; overestimations are penalized by long queue times. We solve this problem by transparently reparallelizing and migrating OpenMP applications. Before an application exceeds its reservation, a distributed heterogeneous checkpointing algorithm automatically migrates the application to a new reservation either on the local or on a remote system. Reparallelization automatically adapts the application’s degree of parallelism to the number of CPUs available on the target cluster. Our prototype is part of
the Jackal compiler and object-based DSM that handle an OpenMP binding for Java. On a compute-intensive Lattice-Boltzmann Method for fluid simulation the overhead of a smooth migration and reparallelization of OpenMP applications across heterogeneous clusters is just about 1%.

Runtime Coupling Support for Component Scientific Simulations
Joe Shang-Chieh Wu (University of Maryland)

Allowing loose coupling between the components of complex applications has many advantages, such as flexibility in the components that can participate and making it easier to model multiscale physical phenomena. I have designed and implemented a loosely coupled framework to support coupling of parallel and sequential application components that has the following characteristics: (1) connections between participating components are separately identified from the individual components; (2) all data transfers between data exporting and importing components are determined by a runtime-based low overhead method (approximate match); (3) two runtime-based optimization approaches, collective buffering and inverse-match cache, are applied to speed up the applications in many common coupling modes; and (4) underlining multi-threaded multi-process control protocol can be systematically constructed by validating the composition of sub-tasks’ protocols. Experimental results show the overhead of approximate match, and the effectiveness of the optimizations will also be presented.

Adaptive Fault Management for High Performance Computing
Yawei Li (Illinois Institute of Technology)

Reliability becomes a critical concern in the ever-growing HPC environment as failures cause serious application and system performance loss. In this doctoral thesis, an adaptive fault management framework called APRO is presented. Exploiting failure prediction, APRO makes runtime decisions by coordinating multiple fault tolerance actions. The principle of APRO is to enable applications to avoid anticipated faults through proactive migration, and in the case of unforeseeable faults, to minimize their impact via selective checkpointing. The thesis makes the following contributions: (1) Present meta-learning techniques to improve failure prediction in large-scale clusters by combining the strengths of multiple data mining techniques; (2) Design an adaptation manager for integration of proactive migration and checkpointing in response to failure prediction; and (3) Design a failure-aware job runtime management method to improve productivities of HPC systems. The ultimate objective is to build an end-to-end fault management framework to improve fault resilience of HPC applications and systems.
Birds-of-a-Feather

Birds-of-a-Feather (BOF) sessions provide forums for conference attendees to discuss topics of mutual interest. BOFs are open to all conference attendees, including exhibitors and exhibits-only badge holders.

A record number of BOF submissions, nearly 90, were received for SC07. Highlights include the Top500 and HPC Challenge awards, several BOFs on major federal programs, career paths in HPC, and a variety of sessions on technology and software.

Tuesday, Nov. 13

12:15 p.m. - 1:15 p.m.

Federal Plan for Advanced Networking Research and Development

Room: A1 / A6

Dan Hitchcock (DOE Office of Science), Suzi Iacono (National Science Foundation)

The Federal agencies need your vision, knowledge, and guidance to identify and prioritize future Federal networking R&D needs and plans. On January 30, 2007 the Director of the Federal Office of Science and Technology Policy established the Interagency Task Force for Advanced Networking (ITFAN) and charged it with developing a Federal Plan for Advanced Networking Research and Development.

A draft Interim Plan, delivered May 15, includes:

- A strategic vision for networking capabilities by mid next-decade
- Scope and objectives for Federal advanced networking R&D
- Identification of Federal networking R&D
- Identification and prioritization of networking R&D needs

To finalize this plan and assure it meets networking community needs, ITFAN needs your review, comments, suggestions, estimates of timing, and your prioritization of networking R&D areas. Please access the draft Interim Plan and provide your comments and inputs at: www.nitrd.gov/advanced_networking_plan.

Parallel File Systems

Room: A10 / A11

Brent Gorda (Lawrence Livermore National Laboratory); Jacques-Charles Lafoucriere, Philippe Gregoire (Commissariat a l’Energie Atomique)

As systems grow faster and get more memory, there is a looming critical issue associated with moving data into and out of the application. The necessary solution involves a high performance parallel file system. Several vendors exist in this space and their products are in use at installations around the world. This is an area that is growing rapidly and maturing as large sites push toward targets of 100-500 Gigabytes per second to the application. This BOF features speakers from private and public institutions who are using parallel file systems in production. Speakers were chosen to represent the available file system solutions and will provide their particular use-model, setup and achieved results.
Adapting Legacy Software to Hybrid Multithreaded Systems hosted by John Gustafson  
Room: A7

Jamil Appa (BAE Systems), Arch Robison (Intel Corporation), Vincent Heuveline (University of Karlsruhe Department of Mathematics)

Hardware architectures are evolving to a hybrid mix of CPU types with multiple cores of each type, and this presents the biggest challenge yet for preserving our massive investment in legacy HPC software. Can the architectural complexity be hidden away in library routines? If so, how do we define and supply those libraries? Or are there ways to extend existing languages to gracefully handle a heterogeneous mix of processor resources that varies from run to run, and modify legacy software to use those extensions? Are there other, perhaps more radical approaches we should consider? In this birds of feather panel session you’ll hear from industry leaders in multiple disciplines (Intel Principle Engineer, Dr. Arch Robison; BAE Systems Group Leader for Technology and Engineering Services, Jamil Appa; University of Karlsruhe Department of Mathematics, Prof. Dr. Vincent Heuveline) looking to tackle these issues.

Converged Fabrics: Opportunities and Challenges  
Room: A20

Zarka Cvetanovic (Hewlett-Packard)

Converged Fabrics is the technology that consolidates all communication (server, storage, and network) into a single high-speed fabric (currently InfiniBand or 10 GigE).

This technology can be applied within HPC to offer high-performance scalable access for storage and sever networking. Not only does this technology reduce the power and cost of compute clusters (less adapters, switches, cables), it also simplifies the cluster configuration and management infrastructure and enhances virtualization and provisioning (SOI) capabilities. The main discussion topics in this session will include: (1) Technology review and system architecture presented by the leading interconnect vendors. (2) Software capabilities from the vendors and OFED community. (3) Reliability and management requirements for seamless integration with different interconnect networks. (4) Potential challenges for converged fabrics including performance, QoS management and interoperability. (5) Usage models: inputs from the users will be encouraged on how this technology can be best applied to enhancing their cluster environments.

TORQUE Resource Manager and Moab: New Capabilities and Roadmap Forum  
Room: A3 / A4

David Jackson (Cluster Resources)

TORQUE, a free, market-leading resource manager, offers control over batch jobs and distributed compute nodes. This BOF will include discussions on recent enhancements and planned developments, including initial high-availability, added job array support, Cray systems support, Altix CPU-sets and Darwin-based systems. We will initiate a dialogue on observed needs, and then solicit your requirements for future development. We will then discuss Moab Cluster, Grid and Utility Suite roadmaps, including
improved virtualization support, scheduling enhancements, partner-centric auto-acceptance test tools and an out-of-the-box, SLES-based, full cluster stack solution (cluster deployment kit) that installs as easily as SLES, plus the answers to three questions. Cluster Resources continues to enhance the Moab suites, helping organizations better control and optimize their computer resources. We will discuss recent and project- ed developments including hierarchical fair-share, advanced workflow support for data center and HPC environments, statistical enhancements and improved support for Cray, Blue Gene and Grid Engine-based systems.

NSF Cyber Enabled Discovery (CDI): Challenges for the Scientific Community
Room: A2 / A5
Lenore M. Mullin (National Science Foundation), Robert R. Borchers (Maui Supercomputing Center)

NSF CDI Objectives include the broadening of our nation’s capability for innovation by developing a new generation of computationally-based discovery concepts and tools to deal with complex, data-rich, and interacting systems. With CDI, we expect to have enhanced ability to deal with research requiring petascale cyberinfrastructure, a strengthened technical basis for a new generation of computational discovery in all science and engineering, and significant progress in educating computational discoveries. This BOF hopes to address challenges facing languages, compilers, and operating systems in support of this venture. Examples include addressing Proebsting’s Law and software keeping up with Moore’s Law, verification of scientific software both semantically and operationally, co-design of machines in support of scientific software, symbolic, numeric, and algebraic software systems, team building, education and research training, as well as issues of how to address changes required in academia is support of these goals.

Multi-core Support in Resource Managers and Job Schedulers
Room: A8
Susanne Balle (Hewlett-Packard), Moe Jette (Lawrence Livermore National Laboratory)

Multi-core processing is a growing industry trend as single-core processors rapidly reach the physical limits of possible complexity and speed. Both AMD and Intel are currently marketing dual-core processors. Quad cores are currently available and CPUs with larger core count will be available in 2008. In a clustered environment, an important factor contributing to an application’s performance is the layout of its processes onto the nodes. Developers currently worry about finding the optimal layout for their application on CPUs and nodes. With the introduction of multi-core and hyper-threaded processors, the level of complexity increases considerably. In this BOF we discuss the problems encountered by users when using resource managers and job schedulers that do not support multi-core explicitly as well as how these problems can be resolved. The multi-core support in the open source resource manager SLURM will be presented as well as strategies from major resource manager vendors.
Grid Operating Systems Community Meeting
Room: A9
Christine Morin (IRISA-INRIA)

Grid middleware is typically built on top of existing operating systems. Little has been done to extend the underlying operating systems to enable and facilitate Grid computing. In June 2006 the EC-funded open source project XtreemOS (www.xtreemos.eu) started to develop a Linux-based grid operating system to directly support virtual organizations for next generation grids. This BoF will be a focal point for the grid operating system community at SC07 where both developers and users may gather to discuss the "current state" as well as future directions for grid operating systems, XtreemOS, and middleware.

The 2007 HPC Challenge Awards
Room: C1 / C2 / C3
Jack Dongarra (University of Tennessee, Knoxville)

The 2007 HPC Challenge Awards are given in two classes: Class 1—Best Performance awards best run submitted to the HPC Challenge website as of November 1. Since there are multiple tests, "best" is a subjective term. It has been decided by the committee that winners in four categories will be announced: HPL, Global-RandomAccess, STREAM-Triad and Global-FFT. Class 2—Most Elegant awards implementation of three or more of the HPC Challenge benchmarks with special emphasis being placed on: HPL, Global-RandomAccess, STREAM-Triad and Global-FFT. This award would be weighted 50% on performance and 50% on code elegance/clarity/size.

Career Paths: High Performance Computing Education and Training
Room: D4
Laura McGinnis (Pittsburgh Supercomputing Center)

Within the next five years, the amount of computing resource cycles available for scientific research will increase significantly. Exploiting this explosion in computing power poses opportunities and challenges. The early adopters of these systems are in our classrooms today. This year's freshman will graduate mere months before NSF's target date for production petascale computing. Graduate students are already in place, learning from pioneers in computational science. These are the users who will advance their fields of science into the next generation of computation. The NSF TeraGrid community is working to prepare the next generation of users for the next generation of resources. This session invites computational scientists from all disciplines and agencies to join in a discussion of HPC career development pathways. Topics will include pedagogies, curricula, and training and education resources that are available or need to be developed to meet the needs of future computational scientists.
Cyberinfrastructure and Society: Creating Outreach Programs for the Public
Room: D5
Rebeka Villarreal Martinez (Texas Advanced Computing Center), Julia White (Oak Ridge National Laboratory), Edee Wiziecki (National Center for Supercomputing Applications)

The purpose of this BOF is to provide a forum for cyberinfrastructure professionals interested in community outreach to share current efforts and future ideas for initiating, executing, and maintaining a comprehensive program of activities that contribute to informing the public about the importance and value of advanced computing technologies in society. These activities will promote the discovery and application of new knowledge. The goal of community outreach activities and programs is to develop relationships and partnerships with new and diverse communities, enabling them to benefit from advanced computing resources and technologies, without having to develop significant computational expertise. Participants should be prepared to discuss, in non-technical terms, projects designed to increase public awareness of, interest in, and understanding of STEM and advanced computing technologies. Methods for designing, deploying, and evaluating activities that attract, inform, and engage broad and diverse communities are critical.

Open Standards for Reconfigurable Computing in a Hybrid Computing Environment
Room: A1 / A6
Wallid Najjar (University of California, Riverside), Martin Herboldt (Boston University), Thomas Steinke (Zuse Institute Berlin)

High-performance computing is undergoing a revolution in system composition. Increasingly, accelerators of many types are being relied upon to provide a desired performance boost to complement the strengths of general purpose multi-core processors. This fact is exemplified in the AMD Torrenza initiative, the Quick Assist effort from Intel and systems from traditional supercomputing vendors including Cray and SGI. With efforts of established and emerging companies embracing Field Programmable Gate Arrays in future systems, the role of open standards for reconfigurable computing is crucial for continued performance improvements on hybrid systems. The BoF will provide a forum for industry leaders and stakeholders to share directions and priorities, progress and needs, in their respective initiatives and provide a context for harmonizing and establishing open standards among the various efforts in such areas as application libraries, functional core definition, debugging, development environments, algorithm specification in high-level languages, intercomponent communication, validation and verification.
Petascale System Interconnect Project

Room: A10 / A11
Kazuaki J. Murakami (Kyushu University), Yasunori Kimura (Fujitsu, Ltd.)

Petascale System Interconnect (PSI) project is one of the Japanese national projects on "Fundamental Technologies for the Next Generation Supercomputing." The project started in 2005. The goal of PSI project is to develop technologies enabling petascale supercomputing systems with hundreds of thousands of computing nodes. To achieve a ten times higher cost/performance ratio, the project tackles with the following three fundamental technologies: (1) small and efficient optical packet switches, (2) low-cost, high-performance MPI communications, and (3) methodologies for evaluating and estimating the performance of petascale systems. The purpose of this BOF is to present what PSI project obtained for these fundamental technologies in the last three years and discuss the application of these results to petascale computing.

Developing Applications for Petascale Computers

Room: A2 / A5
Thom Dunning (National Center for Supercomputing Applications), Dan Reed (Renaissance Computing Institute), Ed Seidel (Louisiana State University)

As applications are scaled to machines with hundreds of thousands of cores, multiple issues arise to hinder achieving the desired performance level. Identifying and responding to the major issues from the point of view of the applications scientist is essential to the success of petascale systems. Areas in which applications capabilities will need to be developed or enhanced include: Application algorithms: libraries, programming techniques. Performance analysis and optimization: identifying performance bottlenecks and projecting future performance, and optimizing and measuring performance. Modeling and projection: application analysis and developing models of performance to explain current or future performance. Application simulation: execution-driven, trace-driven, whole system, and network simulations. Each of these issues has multiple facets and numerous subtleties. In this BOF, the goal is to draw out a high level view of the expectations of the applications communities within the context of near term and long term plans for petascale computing systems.
Open Standards for Accelerated Computing  
Room: A20  
Douglas O’Flaherty, Patti Harrell (AMD)

Accelerated Computing—running applications on specialized processors, such as GPUs, FPGAs and ASICs—is gaining momentum in the HPC market. However, open standards to encourage next generation tool development, languages and applications are not keeping pace. This BOF will include short presentations on emerging topics in the area of open standards for Accelerated Computing and strongly encourages community feedback and discussion.

Reliability of High-Speed Networks in Large-Scale Clusters  
Room: A3 / A4  
Gilad Shainer (Mellanox Technologies), Darren Kerbyson (Los Alamos National Laboratory)

The increasing demand for computing power in scientific and engineering applications has spurred deployment of high-performance computing clusters. According to the TOP500 list, the HPC market will be entering the Petaflop era in 2015. Future HPC systems will span thousands to tens-of-thousands of nodes; all connected together via high-speed connectivity solutions to form a Tera to multi-Petaflop clusters. The overall number of communication links grows with the size of the cluster. Link data errors, resulting from bit-flips, have become a growing concern for large-scale platforms, as they tend to have an adverse affect on the performance. Several approaches exist for dealing with bit-flips, such as end-to-end retry, forward error correction, switch level retransmission and others. The session will discuss the benefits and the performance impacts of each of various options, including case studies, and what should be the preferred solution for future large-scale clusters.

Parallel Program Development Tools Users BOF  
Room: A7  
Bernd Mohr (Forschungszentrum Jülich), Daniel Terpstra (University of Tennessee), Sameer Shende (University of Oregon)

Two major new projects, an NSF-funded SDCI project focusing on integration and improvement of performance analysis tools, and the international Virtual Institute for High Productivity Supercomputing (VI-HPS), focusing on scalable correctness and performance tools, aim to provide applications developers with a robust, scalable, and integrated development tools infrastructure. Tools involved include PAPI, PerfSuite, TAU, KOJAK/Scalasca, VampirServer, and Marmot. PAPI is a portable interface to hardware performance counters, PerfSuite and TAU are profiling tools, KOJAK/Scalasca and VampirServer are scalable trace analysis tools, and Marmot is an MPI correctness tool. Following a brief overview of the collaboration projects and presentation on the status of these tools, this BOF will continue the tradition of past well-attended PAPI Users BOFs by soliciting input from users on desirable tool capabilities and features. Such input has been invaluable in the past for prioritizing work on the PAPI project and ensuring that the tool meets user needs.
Globus and Community: Today’s Cyberinfrastructure
Room: A8
Jennifer M. Schopf (Argonne National Laboratory)

The Globus Project is developing fundamental technologies needed to build Grids -- persistent environments that enable software applications to integrate instruments, displays, computational and information. This paradigm has become common in many application communities that use distributed resources in a coordinated manner. In this BOF we will give a brief overview of the components which make up the Globus Toolkit, an open source software toolkit that includes services and libraries for security, resource management, monitoring and discovery, file transfer and data management. We summarize future plans, and give details on how others can contribute to this effort. We will then open up the session to community concerns and needs through discussion with the audience members.

TOP500 Supercomputers
Room: C1 / C2 / C3
Erich Strohmaier (Lawrence Berkeley National Laboratory)

Now in its 15th year, the TOP500 list of supercomputers serves as a “Who’s Who” in the field of High Performance Computing (HPC). The TOP500 list was started in 1993 as a project to compile a list of the most powerful supercomputers in the world. It has evolved from a simple ranking system to a major source of information to analyze trends in HPC. The 30th TOP500 list will be published in November 2007 just in time for SC07. This BoF will present detailed analyses of the TOP500 and discuss the changes in the HPC marketplace during the past years. This includes the presentation of a new metric to track power consumption and updates on various benchmark initiatives. The BoF is meant as an open forum for discussion and feedback between the TOP500 authors and the user community.

Scaling I/O Capability in Parallel File Systems
Room: D4
Nikhil Kelshikar (Cisco Systems, Inc.), Brent Welch (Panasas, Inc.), Steve Jones (Stanford University)

HPC clusters have been growing in size and capacity. To scale application performance to meet these advances, data storage and server I/O capability needs to scale proportionately. Innovations in data storage technology are enabling applications to achieve performance levels demanded by HPC users. Additionally, new multi-fabric switches are enabling the use of high performance interconnects such as InfiniBand and 10GigE to provide new levels of bandwidth between compute clusters and storage systems. In this session we will discuss the following: (1) Innovative cluster architectures for scaling I/O over high-speed interconnects such as InfiniBand or 10 Gigabit Ethernet to Parallel File Systems; (2) Use of InfiniBand to Ethernet & Storage Gateway Devices and benchmarks in real world environment with applications; (3) How Parallel File System innovations achieve maximum I/O application performance at scale.
The purpose of this BOF is to allow HPC Centers’ staff (NSF, DOE, DOD, and others from the USA and around the world), including front-line support, system administrators, and managers an opportunity to discuss and share common problems and concerns. Attendees should expect to share, discuss and compare Best Practices in order to further the goals of HPC centers. Attendees will have an opportunity to meet their peers from different user facilities and discuss various topics such as issue tracking, account management, security, and documentation techniques. This BOF will be a benefit to anyone who is responsible for providing services to users of high performance computing resources. This BOF will also provide an opportunity for participants to discuss the possibility of an HPC Centers’ Panel at SC08. Representatives from user services organizations at TACC, NCAR, NERSC and PSC have agreed to participate.

Rocks Clusters
Room: A1 / A6
Steve Jones (Stanford University), Greg Bruno (University of California, San Diego), Mason Katz (University of California, San Diego)

The Rocks Cluster Toolkit is quickly growing into the leading Linux-based cluster distribution. This second annual BOF serves as a forum to discuss new technologies and trends in designing, deploying and managing large-scale Rocks Clusters. Meet and discuss developer and end-user issues with members of the core Rocks development team from the San Diego Supercomputer Center, leading Rocks system administrators, and key leaders developing Rolls and providing support in the commercial space. Discussions will include new elements in recently released Rocks versions, Rocks Rolls, and feedback on the roadmap for the future of Rocks—focusing on an interactive session between users and developers to help bring community needs into focus.

Bayesian Network Awareness
Room: A10 / A11
Bill Rutherford (RRX Network), Loki Jorgenson (Apparent Networks)

A short presentation and open discussion of distributed learning approaches to network awareness for supercomputing. Time permitting we will consider: the potential difficulties of applying emerging artificial
intelligence techniques to learn normal behavior and identify problematic activities; possible strategies for feeding sensor nodes with summarized parameters; issues around timely application of dynamic views. The use of learning to, in an ongoing unsupervised manner, characterizes the detailed normal behavior of packets, conversations and data streams which may be represented as vectors of scalars, opens up the possibility of rapid identification of subtle abnormal activity. Clearly derived views depend heavily on monitor stack ability to supply sensor nodes with appropriate and timely data. Practical scalability and performance requirements imply a high degree of autonomous adaptation in a complex distributed environment.

**OSCAR Community Meeting**  
**Room: A2 / A5**  
*Stephen Scott, Thomas Naughton (Oak Ridge National Laboratory)*

Since the first public release in 2001, there have been well over 190,000 downloads of the Open Source Cluster Application Resources (OSCAR) software stack, and it is presently ranked as #4 in downloads of cluster-related software at SourceForge.net. OSCAR is a self-extracting cluster configuration, installation, maintenance, and operation suite consisting of “best known practices” for cluster computing. OSCAR has been used on highly ranking clusters in the TOP500 list and is available in both a freely downloadable version as well as commercially supported instantiations. The OSCAR team is comprised of an international group of developers from research laboratories, universities, and industry cooperating in the open source effort. This BOF will be a focal point for the OSCAR community at SC07 where both developers and users may gather to discuss the “current state” as well as future directions for the OSCAR software stack. New and potential users and developers welcome.

**Fortran**  
**Room: A20**  
*Craig Rasmussen (Los Alamos National Laboratory)*

Fortran is an important language for scientific application development, and there is a lot going on in the Fortran world. A new standard is arriving in 2008 (hopefully) and will include co-array extensions for parallel application development. An integrated development environment (Photran) has been created in Eclipse, and many tools and program refactorings are being considered for users. In addition, an open-source Fortran 2003 parser now exists (based on ANTLR) and is available for Fortran tool development. Plus, I’m sure there is much more taking place, please come and bring your favorite topics for discussion. Specific topics to be discussed include: (1) What is in the Fortran 2008 standard and should we be thrilled or worried?; (2) What direction should Fortran take in the parallel dimension beyond co-arrays?; (3) What tools should the open source community create in support of Fortran developers?; and (4) Other topics as suggested by attendees.
Coordinated Fault Tolerance in High-end Computing Environments
Room: A3-A4
Pete Beckman, Rinku Gupta (Argonne National Laboratory); Al Geist (Oak Ridge National Laboratory)

The ability to detect and recover from faults on large HPC systems would be greatly aided by a standardized interface to exchange fault information. A standard framework where any component of the software stack can report or be notified of faults through a common interface enables coordinated fault tolerance and recovery. This BOF will present the draft design of such an interface for comment by the HPC community, both users and vendors. The objectives of this BOF session are: (1) To have an open discussion about the usefulness, impact, and adoption of a comprehensive fault-tolerance framework in enterprise and research environments; (2) To better understand fault management and fault-tolerance challenges being faced in today’s environment; and (3) To bring together individuals dealing with high-end, petascale computing infrastructures, who have an interest in developing and tolerance in high-end computing environments.

TeraGrid Operations and Plans in Oak Ridge: User Community Interaction
Room: A7
John W. Cobb (Oak Ridge National Laboratory)

This BOF will be a chance for the user community to get an update of TeraGrid activities and plans at Oak Ridge and across all TeraGrid resource providers*. This will include current experimental facility interaction as well as HPC resources. The main agenda items are: (1) Update on current and plan user services, including how to request and use resources and (2) Feedback from current and prospective users on user experience and possible improvements. *Note: The TeraGrid currently has several large hardware and software programs where proposals have been solicited and received but awards have not yet been announced but are anticipated before SC07. Consequently, the content of this BOF may be affected by announcements between this submission deadline and the conference.

Open MPI State of the Union
Room: A8
Jeffrey Squyres (Cisco Systems), George Bosilca (University of Tennessee, Knoxville)

The Open MPI Project is a world-wide collaboration between academic institutions, research organizations, and industry partners that develops, maintains, and supports an open source implementation of the MPI-2 standard. The Open MPI software is a high performance, scalable MPI implementation, providing a highly modular architecture that not only adapts to a wide variety of environments, but also uniquely lends itself to HPC research. The meeting will consist of three parts: (1) Members of the Open MPI core development team will present the current status of Open MPI. (2) Presentations of real-world "success stories" using Open MPI. (3) Discuss the Open MPI roadmap, to include possible future directions for Open MPI, and actively soliciting feedback
from real-world MPI users and ISVs with MPI-based products (please bring your suggestions!).

**Deploying HPC for Interactive Simulation**

**Room: A9**

*Roger Smith (U.S. Army)*

The community of academia, industry, and government offices that are leading the development of new interactive simulations for training and analysis are reaching a point at which the application of traditional networks of computing assets are no longer able to support simulation scenarios of sufficient scope, breadth, and fidelity. Several organizations are turning to high performance computing in the form of clusters and shared memory machines to create a flexible computing platform that is powerful enough to run realistic models of military activities and very large scenarios as a matter of course. This BOF will discuss the problem-space and experiments that have been conducted in applying HPC to this domain. BOF sponsors: Roger Smith, U.S. Army Simulation and Training; Brian Goldie, UCF Institute for Simulation and Training; Dave Pratt, SAIC Simulation; Eng Lim Goh, SGI; Robert Lucas, USC Information Sciences Institute.

**Federal Activities Impacting Long Term HEC Strategies**

**Room: C4 / D1 / D2 / D3**

*Daniel A. Reed (Renaissance Computing Institute), Phillip Colella (Lawrence Berkeley National Laboratory), George O. Strawn (National Science Foundation)*

2007 activities associated with the $3.1 billion 14-agency Federal Networking and Information Technology Research and Development (NITRD) Program will be discussed: The release of the President’s Council of Advisors on Science and Technology (PCAST) report “Leadership Under Challenge: Information Technology R&D in a Competitive World,” which looks at global competitiveness and makes recommendations about U.S. IT education and the NITRD Program’s funding profile and R&D priorities. The National Academies study “Toward Better Understanding the Potential Impact of High-End Capability Computing on Science and Technology.” Using astrophysics, biology, chemistry, and weather for illustration, the study is: (1) identifying and analyzing aspects of important problems that are difficult or impossible to address without high-end capability computing, (2) identifying their numerical and algorithmic computational characteristics, and (3) categorizing those characteristics, noting categories that cut across disciplines.
Cyberinfrastructure tools will play a critical role in higher education. This BOF will discuss currently available tools and future tool requirements. A panel from academia will address timely topics such as: Using Cyberinfrastructures in Undergraduate/Graduate Education, and Cyberinfrastructures and On-line Education (The Ralph Regula School of Computational Science). The Computational Chemistry Grid (CCG) is one tool being used in Chemistry Education. The CCG is an NSF-supported virtual organization, that provides access to high performance computing resources for computational chemistry and other disciplines, with intuitive interfaces and a portable java-based client. The GridChem middleware gives users a transparent graphical user interface to complex computational chemistry software. This interface allows for easier adoption of the application into training modules and undergraduate/graduate curricula. An overview of GridChem features that help its adoption as a tool for chemistry education will be presented. This discussion will benefit educators, software developers, and those involved with Cyberinfrastructure.

Eclipse Parallel Tools Platform (PTP)
Room: A9
Beth R. Tibbitts, Greg Watson (IBM Research); Craig Rasmussen (Los Alamos National Laboratory)

The Eclipse Parallel Tools Platform (http://eclipse.org/ptp) provides a robust, extensible, open-source platform for parallel application development. Release 1.1 was available in early 2007 and features a parallel runtime, parallel debugger, and analysis tools to aid the parallel application developer. Release 2.0 is in the works and will feature: support for job schedulers/resource managers (LSF, MOAB, & Loadleveler initially supported); remote services including remote build, launch, and debug; additional MPI implementations; performance tools and an upcoming framework to ease tool integration; analysis including detection of possible deadlocks due to MPI Barriers; and Fortran refactoring and other tools. We will summarize the current state of PTP, show upcoming features, showcase users and developers who have developed their own extensions to the PTP platform, and welcome user input for future PTP features. We will also discuss potential opportunities for using PTP to support multi-core application development.
Power, Cooling and Energy Consumption for Petascale and Beyond

Room: A1 / A6

John Shalf (Lawrence Berkeley National Laboratory), Stephen Elbert (Pacific Northwest National Laboratory), Rob Pennington (National Center for Supercomputing Applications)

For decades, the notion of "performance of computer" has been synonymous with "raw speed" in FLOP/S. This has increased power consumption dramatically creating problems economically providing enough power, making the most efficient use of the power, and managing the byproduct of waste heat. Total cost of ownership of supercomputers has increased extraordinarily. Leading research and industry groups will present an overview of power consumption trends and discuss solutions for petascale computing power and cooling. Topics include: Performance, power-related and efficiency metrics for system ranking; Minimization of power consumption; Air or liquid cooling for rooms, systems, racks, nodes, or chips; Dense server packaging and cooling alternatives; AC or DC power feeds with or without UPS; Chiller free cooling and cooling with outside air; ASHRAE allowable conditions; and Efficiency rebate programs. Datacenter case studies will cover acquisition and risk management as well as competitive analysis.

Unleashing the Power of the Cell Broadband Engine Processor for HPC

Room: A10 / A11

David A. Bader (Georgia Institute of Technology), Michael Perrone (IBM Research), Ashok Srinivasan (Florida State University)

The Cell/B.E. is a heterogeneous multi-core processor, capable of over 200 gigaflops. This BOF includes a discussion of techniques and tools that can enable HPC-applications to exploit this power: (1) How can application developers get maximum performance through "heroic" programming efforts? We shall discuss programming tricks and algorithmic techniques for some important computational kernels, and open problems. (2) What productivity tools are available? We shall discuss features of the Cell-SDK-3.0, performance tools, programming models, and developing applications using the PS3. (3) Will a lightweight-OS be suitable for Cell? An embedded real-time Linux may help reduce the load on the PPU, and can be effective in the SPU-centric programming model which is typically used. (4) How can the Cell/B.E. be used as an accelerator for traditional processors? We shall discuss the LANL Roadrunner, which uses Cell/B.E.-processors to augment Opterons, with the former handling the bulk of the computational workload.

PVFS: A Parallel File System for Petascale Computing

Room: A2-A5

Sam Lang, Rob Ross, Rob Latham (Argonne National Laboratory)

PVFS is an open source parallel file system that provides high performance IO for high-
end systems, used in both research and production environments. 500 teraflop systems are coming online within the year and petascale systems soon to follow. These systems will have massive storage capabilities and bandwidth requirements in the range of 50-100 GB/s, in an environment where failures must be tolerated. This presents unique challenges to high performance IO. Join us to learn how PVFS achieves high performance and stability within petascale environments, as we discuss several new developments and features targeted at petascale systems. As always, PVFS developers will be present for an open forum discussion. Users, researchers, or the merely curious are all encouraged to attend.

**FAST-OS (Forum to Address Scalable Technology for runtime and Operating Systems)**

**Room: A20**

Arthur B. Maccabe (University of New Mexico)

FAST-OS (Forum to Address Scalable Technology for runtime and Operating Systems) was established to provide a forum to discuss issues related to scalability in the development of runtime and operating systems for next generation, high-end computing systems. As high end computing systems continue to increase in size, runtime and operating systems are becoming a critical bottleneck in application scalability. In the past, system developers have considered two general approaches to providing the needed runtime/operating system: lightweight approaches and full-featured approaches. Lightweight approaches do not adversely affect application scalability; however, they frequently impose a significant burden on application development as many of the commonly used services are not provided. In contrast, full-featured approaches tend to provide all of the needed services, but have been shown to adversely impact application scalability. For more information, see the FAST-OS web page (www.cs.unm.edu/~fastos/).

**Evaluating Petascale Infrastructure: Benchmarks, Models, and Applications**

**Room: A3 / A4**

Robert J. Fowler (University of North Carolina Chapel Hill), Daniel A. Reed (Renaissance Computing Institute), Allan E. Snavely (San Diego Supercomputer Center)

This BOF is a venue for presentations and discussions of recent progress in evaluating the performance, reliability, energy efficiency and usability of petascale computing systems and in developing applications that scale effectively on such systems. We invite participation by large system vendors, funding agencies, infrastructure operators, application development groups, and others with interests in the design, operation, and successful use of these systems. Technical topics to be addressed include the scaling properties of benchmarks; new and proposed benchmark suites; system balance benchmarks and models; application modeling to predict scaling for future machines; preparing applications for petascale environments; and the role of all of these topics in petascale acquisitions. At this session we will initiate planning for a series of workshops on the subject.
Panel Discussion of Large Data Handling
Room: A7
Yoshinobu Yamade (University of Tokyo), Shigeru Obayashi (Tohoku University), Haruo Terasaka (University of Tokyo)

The aim of the proposed BOF is to hold a panel discussion for the “Large Data-Handling” problem. This problem originates in the very large-scale numerical simulation. Owing to remarkable progress of supercomputers, the very-large-scale numerical simulation is now possible to solve complicated phenomena and/or to obtain very precise results. To do this, we need to handle very large data. This “Large Data-Handling” problem has various aspects: data management such as storage and transmission, pre-processing such as mesh generations and setting of initial and boundary conditions, post-processing such as analysis of results including visualization and data mining, and so on. We plan to discuss these issues as a panel discussion with young researchers who face them daily, such as students, post-docs, engineers and software vendors.

Adaptive Routing in InfiniBand
Room: A8
Yaron Haviv, Ida Parisi (Voltaire)

Adaptive Routing in InfiniBand With the growing size of clusters, the introduction PetaFlop machines, and growing capacity of compute nodes consisting of many CPU cores, more traffic is being generated across the interconnect. InfiniBand is now becoming the leading cluster and storage interconnect in HPC, and must address the challenge of growing traffic load, growing traffic diversity, and congestion behavior. This session will start by presenting the current challenges in large scale clusters, the current InfiniBand architecture and mechanisms to address congestions and hot-spots, and potential future enhancements. An open discussion involving end-users, vendors, and professionals in the field will follow.

Introduction to Ranger: The First NSF "Track 2" Petascale System
Room: C1 / C2 / C3
Dan Stanzione (Arizona State University), Karl Schulz (University of Texas-Austin), Dave Lifka (Cornell University)

This BOF will introduce the community to the architecture, programming environment, applications, and support services of Ranger, the largest open, general-purpose computer in the world, and the first machine deployed under "Track 2" of the National Science Foundation’s “Towards a Petascale Computing Environment” program. Ranger, located at the Texas Advanced Computing Center (TACC) in Austin, Texas, will have over 60,000 processing cores with 1/2 petaflop peak performance, 125 TB memory, and 1.7 PB disk. Ranger will become available to the community through the NSF Teragrid allocation process on January 1, 2008. This BOF will present an overview of the Ranger system, with speakers from lead integrator SUN Microsystems, as well as from TACC and the partner institutions involved in the Ranger project, Arizona State University and Cornell University. In addition to an introduction to Ranger itself, the discussion will include information on
upcoming training opportunities and the allocation process.

**Parallel Network File System (pNFS)**

**Room: C4 / D1 / D2 / D3**

*Peter Honeyman (University of Michigan), Garth A. Gibson (Carnegie Mellon University / Panasas Inc.)*

Parallel NFS is a key component of the emerging NFS version 4 minor version 1 (NFSv4.1). The protocol definition for this draft standard will be complete or near complete at the time of this BOF and a primary use case is large-scale cluster computing and supercomputing. This BOF will discuss the emerging standard protocol, the status of its implementations, its potential implications for the SC community and how the community can get more involved. Parallel NFS (pNFS) extends NFSv4 to allow clients direct access to file, object, or block storage while preserving the operating system and hardware platform independence features of NFSv4. pNFS can distribute I/O across the bisectional bandwidth of the storage network connecting clients and storage devices, removing the single server bottleneck that plagues client/server-based systems. For more information on this BOF, see www.pdsi-scidac.org, and on the protocol draft, see www.ietf.org/html.charters/nfsv4-charter.html.

**Advancements in Distributed Rendering Open New Visualization Capabilities**

**Room: A1 / A6**

*Mike Long (Linux Networx, Inc.)*

Advancements in High Performance Visualization follow those pioneered in clustered computing. The latest in distributed rendering technology allow researchers to visualize extremely large data sets in their entirety. Additionally, costs of hardware and software continue to fall, making high performance visualization more affordable. These trends are providing organizations, from national laboratories to oil and gas companies, product designers and others, with exciting new capabilities. Visualization users now have the ability to render large simulations in real time and achieve insightful results quicker and more efficiently than ever before. Additionally, visualization clusters enable users to dedicate nodes to specific computational and visualization needs, and simultaneously process data and interpret results. As a result, users can analyze greater amounts of data and higher-fidelity models more rapidly and thoroughly, accelerating their research, product design cycles, and time-to-results. This session will have a short presentation describing trends in visualization, followed by an open discussion.

**OpenMP 3.0: Tasks Rule!**

**Room: A10 / A11**

*Larry Meadows (OpenMP ARB)*

The long awaited OpenMP 3.0 standard has been available for public review since September. The ability to generate unstructured work using the tasking extensions are
the most important change to OpenMP since its inception in 1997. The OpenMP standards folks will provide a very brief organizational update and spend the rest of the session presenting the 3.0 changes and encouraging questions and audience participation.

Thursday, Nov. 15

12:15pm - 1:15pm

Supercomputers or Grids: That is the Question!
Room: A2 / A5
Wolfgang Gentzsch (D-Grid), Dieter Kranzlmueller (GUP Linz)

The inception of Grid Computing in the mid-90s has changed the world. PCs, workstations, and supercomputers have been connected through grid middleware into distributed computing and storage environments, which provide computing to the scientific community. Ever since the appearance of grids, the question has been: Now that we have grids, do we still need supercomputers? This BoF will aim at two important answers to this question. First, with a few real-life use cases, we will demonstrate the usefulness and necessity of each, the supercomputer and the grid, independently, and when to use one versus the other. Second, another set of examples will prove that the ultimate solution might lie somewhere in the middle, namely supercomputers being nodes in a compute and data grid. As example use cases, this BoF will gather use cases from grid and supercomputing initiatives such as DEISA, D-Grid, EGEE, PACE, OSG and TeraGrid.

MPICH2: A High-Performance Open-Source MPI Implementation
Room: A3 / A4
Darius Buntinas, Rajeev Thakur, Bill Gropp (Argonne National Laboratory)

MPICH2 is a popular, open-source implementation of the MPI message passing standard. It has been ported to many platforms and used by several vendors and research groups as the basis for their own MPI implementations. This BoF session will provide a forum for users of MPICH2 as well as developers of MPI implementations derived from MPICH2 to discuss experiences and issues in using and porting MPICH2. New features and future plans for MPICH2 will be discussed. Representatives from MPICH2-derived implementations, such as MVMPICH2, Intel MPI, Microsoft MPI and IBM MPI (for BG/P), will provide brief updates on the status of their efforts. MPICH2 developers will also be present for an open forum discussion. All those interested in MPICH2 usage, development, and future directions are encouraged to attend.

Parallel Debugging and Correctness Checking
Room: A7
Bettina Krammer (High Performance Computing Center Stuttgart), Matthias Mueller (Technische Universität Dresden)

Writing parallel code that runs correctly (and efficiently) on large numbers of processors
and cores is a challenging task that hides many pitfalls. While most software developers still try to tackle these issues with the classical debugging method - printf - there are powerful tools out there that support different parallel programming paradigms (MPI, OpenMP, etc.) and different levels of debugging, such as memory debugging, correctness checking, etc. This BOF will provide a forum for tools developers and application developers to discuss issues related to scalability and usability of such parallel tools (open source and vendors), be it parallel debuggers (DDT, Totalview) or correctness checking tools (e.g. Marmot, Threadchecker).

Meeting the Feature Needs of the LSF User Community
Room: A8
Amy Apon (University of Arkansas)

The University of Arkansas invites you to a BOF session that will provide the venue to discuss feature needs of the LSF user community. Participants from the University of Arkansas will describe how LSF has been incorporated into our production environment and the new features that have been added on our behalf. Platform Computing Product Management personnel will be at this session to discuss the new features in recent releases as well as provide a glimpse into LSF’s development roadmap. Users will have an opportunity to ask questions about the new features and also provide direction to the future roadmap. The University of Arkansas currently hosts the LSF user forum at http://platform.uark.edu/. Platform Computing, founded in 1992, is the largest independent vendor of grid software solutions, with more than 2,000 customers worldwide in the financial services, manufacturing, government, Oil/Gas exploration, Academic research and life sciences markets.

TotalView Tips and Tricks
Room: A9
Chris Gottbrath (TotalView Technologies)

The TotalView Debugger is a scalable, flexible, scriptable debugger that has proven to be an essential tool for programmers working in multi-core environments on parallel and clustered systems, with wide acceptance in the high-performance computing community. This BOF will be an opportunity for TotalView users to share clever and interesting ways of using TotalView to shorten the development cycle, adapting TotalView to their unique environment, using TotalView to do something unusual, or simply making the day to day process of debugging multi-core applications easier. This is a 100% community-content BOF so please contact Chris.Gottbrath@TotalViewTech.com if you want us to reserve time for you to pass on your tip or simply show up at the BOF and step forward.

Partitioned Global Address Space (PGAS) Programming Languages
Room: A20
Tarek El-Ghazawi (George Washington University)

The Partitioned Global Address Space (PGAS) model has been gaining rising attention due to its prospects as the basis for pro-
ductive parallel programming. The PGAS model provides for ease-of-use through its global shared address space (GAS) view, while providing for tuned performance through its locality awareness. A number of PGAS languages are now ubiquitous, such as UPC, CAF and Titanium for which compilers became available on most modern high-performance computers. The DARPA HPCS program has also resulted in introducing new promising PGAS languages, such as X10 and Chapel. This BoF will bring together a group of research scientists, developers, and users of PGAS languages from academia, industry and government to address the progress in PGAS languages. The discussion will include PGAS basic principles, language and compiler availability, resources for new users, performance, compiler optimizations, applications and the major challenges that remain.
Wetlands are an important part of the Reno/Tahoe area. They act as a natural filter for the solids that come out of the water treatment plant. Plant roots absorb nutrients from the water and naturally filter it. Wetlands are also a home for over 75% of the species in the Great Basin. However, the area’s wetlands are at risk of being destroyed due to development around the city. To help protect these important ecosystems, Washoe County has devised a plan where, when developers try to build over a wetland, they will be responsible for creating another wetland near Washoe Lake.

The SC conference continues to serve as the venue for announcing professional awards that recognize key contributions to high performance computing, networking and storage. The distinguished Sidney Fernbach Memorial Award, Seymour Cray Computer Science and Engineering Award, and Gordon Bell Prize are highlighted at SC07 in a plenary technical program session on Wednesday at 1:30 p.m.

The Best Paper, Best Student Paper, and Best Poster Awards recognize the finest of the many outstanding papers in a highly competitive technical program. Other recognitions of achievements include awards for Challenges, ACM Student Research Competition, and the HPC Ph.D. Fellowship. These prestigious honors will be presented in a special session during the conference (Thursday, 1:00 p.m. - 3:00 p.m.).
Gordon Bell Prize

Gordon Bell Prizes are awarded each year to recognize outstanding achievement in high-performance computing. Now administered by the ACM, financial support of the $10,000 award is provided by Gordon Bell, a pioneer in high-performance and parallel computing. The purpose of the award is to track the progress over time of parallel computing, with particular emphasis on rewarding innovation in applying high-performance computing to applications in science. The prize has been awarded every year since 1987. Prizes may be awarded for peak performance as well as special achievements in scalability, time-to-solution on important science and engineering problems, and low price/performance.

Four finalists have been identified, from which one or more Gordon Bell Awards will be awarded at SC07.

Wednesday, Nov. 14

Gordon Bell Prize Finalists
Room: A3 / A4

Session Chair: David H. Bailey (Lawrence Berkeley National Laboratory)

10:30 a.m. - 12:00 p.m.

A 281 Tflops Calculation for X-ray Protein Structure Analysis with the Special-Purpose Computer MDGRAPE-3
Yousuke Ohno (RIKEN); Eiji Nishibori (Nagoya University); Tetsu Narumi (Keio University); Takahiro Koishi (University of Fukui); Tabir H. Tabirov, Hideo Ago, Masashi Miyano, Ryutaro Himeno, Toshikazu Ebisuzaki (RIKEN); Makoto Sakata (Nagoya University); Makoto Taiji (RIKEN)

We achieved 281 Tflops sustained calculation speed for the optimization of 3-D structures of proteins from X-ray experimental data by Genetic Algorithm - Direct Space (GA-DS) method. We used MDGRAPE-3, special-purpose computers for molecular simulations, with the peak performance of 752 Tflops. In GA-DS method, a set of selected parameters which define crystal structures of proteins are optimized by Genetic Algorithm. To appreciate the model parameters the diffraction patterns must be calculated by the nonequispaced Discrete Fourier Transformation (DFT) and they are compared with the X-Ray experiment. To accelerate DFT calculation, which dominates the computational time, we used MDGRAPE-3. We adopted a protein molecule and obtained more accurate structures than typical results with Molecular Replacement method. Our results demonstrate that GA-DS method with special-purpose computers is effective for the structure determination of biological molecules.

First-Principles Calculations of Large-Scale Semiconductor Systems on the Earth Simulator
Takahisa Ohno (National Institute for Materials Science); Takenori Yamamoto (Toho University); Takahiro Yamasaki (University of Tokyo), Tatsunobu Kokubo (NEC Corporation); Yuta Sakaguchi (Advanced Soft Engineering, Inc.), Daisuke Fukata (NEC)
First-principles simulations of large-scale semiconductor systems using the PHASE code on the Earth Simulator (ES) demonstrate high performance with respect to the theoretical peak performance. PHASE, designed for vector-parallel systems like the ES, demonstrates excellent parallel efficiency. We simulated an arsenic donor in silicon using up to 8,000 atom unit cell. A sustained peak performance of 14.6 TFlop/s was measured on 3,072 processing elements, which corresponds to 59% of the theoretical peak performance. Preliminary results using 10,648 atom unit cells are also presented.

**Extending Stability Beyond CPU-Millennium: Micron-Scale Atomistic Simulation of Kelvin-Helmholtz Instability**

James N. Glosli, Kyle J. Caspersen (Lawrence Livermore National Laboratory); John A. Gunnels (IBM Corporation); David F. Richards, Robert E. Rudd, Frederick H. Streitz (Lawrence Livermore National Laboratory)

The Kelvin-Helmholtz (KH) instability, occurring when fluid layers undergo shear flow, is responsible for the wave patterns seen, e.g., on a windblown ocean or as billows on cloud tops. Although the transition from smooth to turbulent flow has been studied extensively, the trend towards smaller length scales in both experiments and continuum modeling raises questions concerning applicability of the hydrodynamic approximation as atomic lengths are approached. Molecular dynamics simulations naturally handle the atomic scale, but have been limited to lengths of less than a micron. With BlueGene/L computer, we can model micron sized samples with atomic resolution. We report the first micron-scale simulation of a KH instability modeled using molecular dynamics. A simulation using the ddcMD code to model over 2 billion atoms ran for a week on 131,072 processors of BlueGene/L, requiring over 2.8 CPU-millennia to complete. We measure the performance of our current implementation to be 54.4 Tflop/s.

**WRF Nature Run**

John Michalakes, Josh Hacker, Rich Loft (University Consortium for Atmospheric Research); Michael McCracken, Allan Snively, Nicholas Wright (San Diego Supercomputer Center); Tom Spezze, Brent Gorda (Lawrence Livermore National Laboratory); Robert Walkup (IBM)

The Weather Research and Forecast (WRF) model is a limited-area model of the atmosphere for mesoscale research and operational numerical weather prediction (NWP). A petascale problem is a WRF nature run that provides very high-resolution "truth" against which more coarse simulations or perturbation runs may be compared for purposes of studying predictability, stochastic parameterization, and fundamental dynamics. We carried out a nature run involving an idealized high resolution rotating fluid on the hemisphere to investigate scales that span the k-3 to k-5/3 kinetic energy spectral transition of the observed atmosphere using 65,536 processors of the BG/L machine at LLNL.
We worked through issues of parallel I/O and scalability. The primary result is not just the scalability and high Tflops number, but an important step towards understanding weather predictability at high resolution.

Seymour Cray Computer Science and Engineering Award

The Seymour Cray Computer Science and Engineering Award recognizes innovative contributions to high performance computing systems that best exemplify the creative spirit of Seymour Cray. The award consists of a crystal model, certificate, and $10,000 honorarium.

Sidney Fernbach Memorial Award

The Sidney Fernbach Memorial Award honors innovative uses of high performance computing in problem solving. This award was established in 1992 in memory of Sidney Fernbach, one of the pioneers in the development and application of high-performance computers for the solution of large computational problems. A certificate and $2,000 are awarded for innovative approaches and outstanding contributions in the application of high performance computers.

Wednesday, Nov. 14

1:30 p.m. - 2:15 p.m.

Seymour Cray Computer Science and Engineering Award Winner: Kenneth E. Batcher, Kent State University

Talk title: Fallacies and Pitfalls in Building Supercomputers

This year’s Seymour Cray Computer Science and Engineering Award winner, Kenneth Batcher, is being recognized for his fundamental theoretical and practical contributions to massively parallel computation, including parallel sorting algorithms, interconnection networks, and pioneering designs of the STARAN and MPP computers.

Professor Batcher has made foundational contributions to both the theoretical analysis and the real-world engineering of highly parallel processing. In the former area, he is probably best known for his early work on sorting networks. He developed the odd-even merge sort and bitonic sort, and showed how each could be implemented in hardware. These are hardware contributions as well as fundamental theoretical contributions. His “bitonic sort,” often called the “Batcher sort,” is one of the classic algorithms in the field. His original 1968 paper has approximately 800 citations on Google Scholar.

Dr. Batcher also designed the architectures of two of the earliest single-instruction,
Sidney Fernbach Award

multiple-data (SIMD) parallel computers: the STARAN (1972) and the MPP (1983). He was a consultant for the development of the ASPRO—an airborne STARAN sometimes called “a STARAN in a shoebox.” His SIMD hardware designs were among the first commercially successful massively parallel computers. He also developed a method that utilized a flip network to organize data in a memory array that allowed multidimensional accesses. His flip network approach was a major contribution that solved a common memory-to-processor bandwidth bottleneck in SIMD parallel computers. Batcher’s STARAN computer was the first SIMD parallel computer to support associative computing (i.e., items are located by their content rather than by their location). He subsequently contributed to the development of the associative computing field, including languages, computational models, and algorithms.

Professor Keyes is a world-renowned leader in the development of scalable numerical algorithms, especially “Newton-Krylov-Schwarz” methods for nonlinear PDEs. These methods, which combine efficient and scalable Schwarz domain decomposition algorithms with globalized Newton-Krylov iterative methods, are at the heart of many applications, including aerodynamics, radiation transport, acoustics, and magnetohydrodynamics. They have been incorporated into open mathematical software libraries that have enabled hundreds of users to make efficient use of parallel computers, from small clusters to the world’s largest computers, including for DOE INCITE projects and Gordon Bell Prizes. An extension to “Lagrange-Newton-Krylov-Schwarz” methods accommodates PDE-constrained optimization problems in design, control, and parameter identification in a unified way that exploits underlying scalable PDE solver infrastructure.

2:15 p.m. - 3:00 p.m.

Sidney Fernbach Memorial Award Winner: David E. Keyes, Columbia University

Talk title: A Nonlinearly Implicit Manifesto

This year’s Sidney Fernbach Memorial Award winner, David Keyes, is being recognized for his outstanding contributions to the development of scalable numerical algorithms for the solution of nonlinear partial differential equations (PDEs) and for his exceptional leadership in high-performance computation.

Dr. Keyes has played a major role in the high-performance computing community. He is currently vice president-at-large for the Society for Industrial and Applied Mathematics (SIAM). He serves on many high-profile advisory committees, including for the National Science Foundation and for the Presidential Council of Advisors on Science and Technology. He is well known within the HPC community for both his technical contributions and editing of community reports on simulation in fusion, fission, aerodynamics, nanotechnology, and other areas of science and engineering. He heads the TOPS SciDAC center, a multi-institution project funded by the DOE Office of Science that is focused on the development of scalable parallel solvers.
Keyes has published over one hundred articles, book chapters, and proceedings papers. He also has edited 13 books and has given several hundred seminars and invited presentations.

Thursday, Nov. 16

Conference Awards

Room: A10 / A11
Session Chair: Steven F. Ashby (Lawrence Livermore National Laboratory)

1:30 p.m. - 3:00 p.m.

In this session the awards for Best Paper, Best Student Paper, Best Poster, ACM Student Research Competition, Analytics Challenge, Bandwidth Challenge, Storage Challenge, HPC Ph.D. Fellowship, and Gordon Bell Prize will be presented.

Cluster Challenge

Did you know that a small cluster today (less than 1/2 rack) would top the Top500.org website from just ten years ago? The computational power that is easily within reach today significantly surpasses that available only to the national labs from that time. The Cluster Challenge showcases the significance of this and highlights how accessible clusters are to anyone today. In this challenge, teams of undergraduate students will assemble a small cluster on the Exhibit floor and run benchmarks and applications selected by industry and HPC veterans. The challenge will be judged on the speed of benchmarks and the throughput of application runs over the first three days of the conference.

Team #1: Purdue University
Supervisor: Preston Smith

The team from Purdue University hails from the Electrical and Computer Engineering Technology, Computer Science, and Electrical and Computer Engineering departments at Purdue’s West Lafayette, Indiana, campus. Led by the Rosen Center for Advanced Computing and in partnership with Hewlett-Packard and AMD, these high-performance Boilermakers come to SC07 with a cluster of HP DL145 Opteron Linux systems, interconnected with Gigabit Ethernet and InfiniBand. Boiler up!

Team #2: Stony Brook University
Supervisor: Yan Yu

Our team consists of undergraduates from Stony Brook University on Long Island, New York, two of whom work at two of the university’s cluster computing facilities: Seawulf and Galaxy. Our vendor partner is Dell, who is providing us with a system of our design which consists of twelve Energy Smart PowerEdge 2950 compute nodes with 2-way SMP Quad Core Xeon processors with InfiniBand interconnects plus one head node. Our team and university mascot, and unfortunate victim of the clever pun used in the name ‘seawulf’ (seawolf/beowulf), is the Stony Brook Seawolf.
**Team #3: Indiana University**  
*Supervisor: Andrew Lumsdaine*

Even though Indiana is home to the greatest spectacle in racing, open wheel cars aren’t the only fast hardware to be found here. Indiana University is partnering with Apple, Intel, and Myricom to make a run for the checkered flag at the SC07 Cluster Challenge. Our “Red Delicious” cluster is built around Apple Xserve systems connected with Myrinet 10G and using Intel compilers and software libraries.

---

**Team #4: University of Alberta**  
*Supervisor: Paul Lu*

Team University of Alberta, from Edmonton, Alberta, Canada, is led by students and faculty from the Department of Computing Science. Our vendor partner is SGI, and our system is based on the Altix XE platform running Linux and SGI’s cluster software. The team’s secret sauce is a melange of excellent hardware, talented students, and coaches with strong opinions but with a hint of the ability to agree to disagree.

---

**Team #5: University of Colorado**

From the beautiful Rocky Mountains comes Team CU. A team of six University of Colorado undergraduates in conjunction with industry leader Aspen Systems has designed a high performance computing cluster which combines the best of commercial and open source technologies modified to meet the specifications of the Cluster Challenge.

---

**Team #6: National Tsing Hua University, Taiwan**  
*Supervisor: Yeh-Ching Chung*

Team NTHU, from the Department of Computer Science at National Tsing Hua University, Taiwan, near Hsin-Chu Science Based Industrial Park, consists of one senior and five juniors. Our vendor is ASUSTek Computer Inc., the largest motherboard OEM company in the world. We have designed a system based on CentOS 4.5 (Linux/x86_64) with appropriate optimizations for the power, the I/O file system, and the Linux kernel. We believe that those optimizations are the keys for us to win the contest.
More than ever before, organizations in commercial, government, university, and research sectors are increasingly tasked with making sense of huge amounts of underutilized data. These dynamics have led to the growing area of SC07 Analytics. SC07 Analytics will highlight rigorous and sophisticated methods of data analysis and visualization used in high performance computing by showcasing powerful analytics applications solving complex, real-world problems. Teams who enter the challenge will be judged by a panel of experts when they present their results. Judging by a panel of experts will take place when entry teams present the results of their project at a special session.

## Tuesday, Nov. 13

### Analytics Challenge Finalists

**Room: A10 / A11**  
**Chair: Paul Fussell (Boeing)**

**10:30 a.m. - 12:00 p.m.**

#### Angle: Detecting Anomalies and Emergent Behavior from Distributed Data in Near Real Time

Robert Grossman, Michal Sabala, Shirley Connelly, Yunhong Gu, Matt Handley, Rajmonda Sulo, David Turkington, Anushka Anand (University of Illinois at Chicago); Leland Wilkinson (Northwestern University); Ian Foster (Argonne National Laboratory); Ti Leggett, Mike Papka, Mike Wilde (University of Chicago); Joe Mambretti (Northwestern University); Bob Lucas, John Tran (University of Southern California)

We describe the design of a system called Angle that detects emergent and anomalous behavior in distributed IP packet data. Currently, Angle sensors are collecting IP packet data at four locations, removing identifying information, and building IP-based profiles in temporal windows. These profiles are then clustered to provide high-level summary information across time and across different locations. We associate meaningful changes in these cluster models with emergent or anomalous behavior. Emergent clusters identified in this way are then used to score the collected data in near real time. The system has a visual analytics interface, which allows different emergent clusters to be visualized, selected, and used for scoring of current or historical data. Each Angle sensor is paired with a node on a distributed computing platform running the Sector middleware. Using Sector, data can be easily transported for analysis or reanalysis. Reanalysis is done using the Swift workflow system.

#### Cognitive Methodology-based Data Analysis System for Large-scale Data

Yoshio Suzuki, Chiaki Kino, Noriyuki Kushida, Norihiro Nakajima (Japan Atomic Energy Agency)

We have conducted research and development of the Cognitive methodology based Data Analysis System (CDAS) which sup-
ports researchers to analyze large scale data efficiently and comprehensively. In data analysis, it is important to evaluate the validity of data and judge whether data is meaningful from a scientific viewpoint. Traditionally, the evaluation and judgment have been carried out by humans. However, when the scale of data is extremely large, the evaluation and judgment are beyond the recognition capability of humans. The basic idea of CDAS is that computers execute evaluation and judgment instead of humans. In the present study, we have applied the system to the virtual plant vibration simulator and succeeded in analyzing large scale data reaching to 1TB thoroughly for the first time.

Storage Challenge

HPC systems are comprised of three major subsystems: processing, networking and storage. In different applications, any one of these subsystems can limit the overall system performance. The HPC Storage Challenge is a competition showcasing effective approaches using the storage subsystem, which is often the limiting system, with actual applications. Participants must describe their implementations and present measurements of performance, scalability, and storage subsystem utilization. Judging will be based on these measurements as well as innovation and effectiveness; maximum size and peak performance are not the sole criteria. Finalists will be chosen on the basis of submissions which are in the form of a proposal; submissions are encouraged to include reports of work in progress. Finalists will present their completed results in a technical session during the conference from which the winners will be selected. Participants with access to either large or small HPC systems are encouraged to enter this challenge.

Tuesday, Nov. 13

Storage Challenge Finalists
Room: A10 / A11
Chair: Raymond L. Paden (IBM)
1:30 p.m. - 3:00 p.m.

ParaMEDIC: A Parallel Meta-data Environment for Distributed I/O and Computing
Pavan Balaji (Argonne National Laboratory); Wu-chun Feng, Jeremy Archuleta (Virginia Tech)

mpiBLAST is an open-source parallelization of the BLAST genome sequence search library. It uses database segmentation to allow different worker processors to search unique segments of the database and write the output to a shared filesystem. For distributed systems sharing a file-system through a low-bandwidth and/or high-latency network, writing this output can be a challenging task, eventually forming a performance bottleneck. In this competition, we plan to demonstrate ParaMEDIC, an environment that decouples computation and I/O in applications and drastically reduces I/O overhead through metadata processing. Specifically, for mpiBLAST, ParaMEDIC partitions worker processes into compute and I/O workers. Compute
workers convert their output to metadata, and send it to I/O workers. I/O workers process this metadata to re-create the actual output and write it to the filesystem. This allows ParaMEDIC to cut down on the I/O time, thus accelerating mpiBLAST several fold (demonstrated a 5-fold improvement on the teragrid).

**Zest: The Maximum Reliable TBytes/sec/$ for Petascale Systems**

*Nathan T. B. Stone, Doug Balog, Paul Nowoczyński, Jason Sommerfield, Jared Yanovich (Pittsburgh Supercomputing Center)*

PSC has developed a prototype distributed file system infrastructure that vastly accelerates aggregated write bandwidth on large compute platforms. Write bandwidth, more than read bandwidth, is the dominant bottleneck in HPC I/O scenarios due to writing checkpoint data, visualization data and post-processing (multi-stage) data. We have prototyped a scalable solution on the Cray XT3 compute platform that will be directly applicable to future petascale compute platforms having an of order 10^8 cores. Our design emphasizes high-efficiency scalability, low-cost commodity components, lightweight software layers, end-to-end parallelism, client-side caching and software parity, and a unique model of load-balancing outgoing I/O onto high-speed intermediate storage followed by asynchronous reconstruction to a 3rd-party parallel file system. The absence of a central metadata service further reduces latency, allowing for the maximum reliable performance per unit cost for petascale systems.

**Astronomical Data Analysis with Commodity Components**

*Micahel S. Warren, John Wofford (Los Alamos National Laboratory)*

During the next decade, large astronomical observing projects will generate more than 100 times as much observational data as has been gathered in all of our prior history. Storage, analysis and management of this information will require significant advances in technology. As an initial step in this process, we propose here to develop and deploy an astronomical data system based on open-source software and commodity hardware capable of storing 100 Terabytes of digital information in immediately accessible disk arrays for a total cost of less than $100,000, scalable to a Petabyte of storage for less than $1M. In the same way in which special-purpose telescopes are now required to obtain the best catalogs of objects in the sky, a focused effort involving state-of-the-art parallel computer hardware and software is required in order to analyze this data and model the Universe which led to the observed distribution of stars and galaxies.
Grid-oriented Storage: Parallel Streaming Data Access to Accelerate Distributed Bioinformatics Data Mining

Frank Wang (Cambridge-Cranfield HPCF), Na Helian (London Metropolitan University); Sining Wu, Vineet Khare, Chenhan Liao (Cambridge-Cranfield HPCF); Amir Nathoo, Rodric Yates, Paul Fairbairn (IBM); Jon Crowcroft, Michael Andrew Parker, Jean Bacon (Cambridge University); Zhiwei Xu (Institute of Computer Technology); Yike Guo (Imperial College); Yuhui Deng (Cambridge-Cranfield HPCF)

Computational grids are often characterized with network latencies greater than 2 ms. Driven by this problem, Grid-oriented Storage (GOS) is designed to deal with cross-domain and single-image file operations. GOS behaves like a file server via a file-based GOS-FS protocol to any entities. We experimentally demonstrate that the parallel-streamed GOS-FS can attain a speedup of 4.9 against the classic NFSv4. GOS is expected to be a variant or successor of NAS. It was demonstrated that GOS can accelerate distributed applications by up to tenfold in real-world tests. Using a distributed data mining application in the International Nucleotide Sequence Database Collaboration (INSDC), GOS-FS delivers a performance speedup of 2.4. The performance scales linearly, with 8 GOS servers servicing 22 simultaneous HPC users.

Bandwidth Challenge

The High Performance Bandwidth Challenge is an annual competition for leading-edge network applications developed by teams of researchers from around the world, providing a showcase for the technologies and people who provide the networking capabilities crucial to supercomputing. The Bandwidth Challenge, running across SCinet, is designed to test the limits of network capabilities, and past events have showcased multi-gigabit-per-second demonstrations never before thought possible.

Tuesday, Nov. 13

Bandwidth Challenge Finalists

Room: A10 / A11
Chair: Debbie Montano (Force10 Networks)

3:30 p.m. - 5:00 p.m.

iWarp-based Remote Interactive Scientific Visualization

Scott A. Friedman (University of California, Los Angeles)

We have undertaken an effort to extend the reach of powerful interactive cluster-based visualization resources to researchers by making them available remotely. Because they are used interactively, visualization resources typically have a fixed location researchers must travel to in order to use. Often, this can be inconvenient even if the resource is available across a campus. By leveraging the latest
generation high speed WAN technologies we aim to bring high performance interactive visualization to the researcher. Accomplishing this required us to extend our existing twenty-four node InfiniBand-based visualization cluster with a 10G iWarp capable remote visualization bridge node. Remote users connect to this node establishing an interactive visualization session. The system is designed to support multiple simultaneous HD quality interactive visualizations by sharing the capabilities of the back-end cluster. This work demonstrates a novel use of high performance network resources, as well as a mixed interconnect topology.

**Streaming Uncompressed 4k Video**

Laura Arns, Scott M. Ballew, David Braun, Patrick Finnegan, Ryan Pedela (Purdue University)

High-resolution full-motion color video presents scientists with an unprecedented opportunity to fully visualize the results of their analyses or simulations to better understand the processes involved. Today’s high-powered rendering and recording systems are capable of producing video with a resolution of 4096x3072, 32-bit color, and at 26 fps. Existing tiled displays are being used to show such media. Such a video stream produces approximately 10Gbps of raw data - faster than can be retrieved from conventional storage systems. When combined with the need to view the video at a distant location distant, the demands placed on the network are enormous. Current systems address these issues by using pre-storage lossy compression algorithms to reduce the demands on both these systems at the expense of detail that may prove critical to the scientist. We propose the construction of a system to stream stored media at full non-compressed resolution to a portable display system.

**Distributed Data Processing over Wide Area Networks**

Robert Grossman, Yunhong Gu, Michal Sabala, David Hanley, Shirley Connelly, David Turkington (University of Illinois at Chicago)

We will start a distributed data processing job within the wide area Teraflow Network and collect the results at SC07. The original data set is from the Angle network monitoring project and managed by the Sector distributed storage system. The client side at SC07 is a simple application that computes new features of the Angle data sets. The application uses the new Sphere API that automatically locates data files and available processors in the Teraflow Network to perform the data processing in parallel. Sphere also takes care of load balancing and fault tolerance. All data transfer in the system will use UDTv4, expected to be released in Fall 2007. This demonstration is set up completely with open source software developed at NCDM. Sector and Sphere: sector.sf.net. UDTv4: udt.sf.net. The computing platform is set up in the Teraflow network. We will measure the throughput for the data transfer of the returned result.
**Phoebus**  
*Aaron Brown (University of Delaware)*

Phoebus is a project whose goal is to allow legacy applications to utilize high bandwidth connections over Internet2’s Hybrid Packet Optical Infrastructure (HOPI) backbone. Phoebus works by breaking a single end-to-end TCP connection into a series of transport layer connections. The end connections (from Internet2 edge nodes to the end hosts) still use TCP. However, the connection over the Internet2 backbone is handled by server software running on the edge nodes. These servers can then choose the best transport protocol and protocol settings for sending the data across Internet2. Using this service, end users are no longer required to modify a variety of TCP settings in order to achieve good performance over long distances. The end users only need to be concerned about ensuring adequate speed across the path between themselves and the Internet2 edge nodes. In many cases, the default TCP settings will prove adequate for achieving acceptable speeds.

**A Virtual Earth TV Set via Real-time Data Transfer from a Supercomputer**  
*Ken T. Murata (Ehime University); Yasuichi Kitamura (National Institute of Information and Communications Technology); Eizen Kimura (Ehime University); Keiichiro Fukazawa, Hironori Shimazu (National Institute of Information and Communications Technology)*

One of our dreams is to have a virtual TV set with wide 3D screen at home. Everyone in the morning turns on the virtual TV, and takes a look at a virtual view of today’s Earth (like virtual 3D Google Earth). The data drawn on the TV set are from the Earth monitoring center, which are combined between Earth observation data via worldwide sensor networks and supercomputer simulations. If a big typhoon shows up or a big earthquake has taken place, you can easily get views and information on it. For this virtual Earth system, 3D data transfer plays an important role. Since the data-size is huge (for example, 1GB for each second), real-time data transfer techniques are required. We challenge high speed data transfer from a supercomputer located in Japan to the SC07 conference hall to demonstrate a virtual Earth TV.

**Using the Data Capacitor for Remote Data Collection, Analysis, and Visualization**  
*Stephen C. Simms, Matthew Davy, Bret Hammond, Matt Link, Craig Stewart, S. Teige, Mu-Hyun Baik, Yogita Mantri, Richard Lord, Rick McMullen, John C. Huffman, Kia Huffman (Indiana University); Guido Juckeland, Michael Kluge, Robert Henschel, Holger Brunst, Andreas Kneepf, Matthias Mueller (Technical University Dresden); P.R. Mukund, Andrew Elble, Ajay Pasupuleti, Richard Bohn, Sripriya Das, James Stefano (Rochester Institute of Technology); Gregory G. Pike (Oak Ridge National Laboratory); Douglas A. Balog (Pittsburgh Supercomputing Center)*

Indiana University provides powerful compute, storage, and network resources to a diverse local and national research community. In the past year, through the use of Lustre across the wide area network, IU has
been able to extend the reach of its advanced cyberinfrastructure across the nation and across the ocean to Technische Universitaet Dresden. For this year’s bandwidth challenge, a handful of researchers from IU, Rochester Institute of Technology, and the Technische Universitaet Dresden will run a set of data-intensive applications crossing a range of disciplines from the digital humanities to computational chemistry. Using IU’s 535 TB Data Capacitor and an additional component installed on the exhibit floor, we will mount Lustre across the wide area network to demonstrate data collection, analysis, and visualization across distance.

TeraGrid Data Movement with GPFS-WAN and Parallel NFS

Patricia Kovatch (University of California); Phil Andrews (San Diego Supercomputer Center); Roger Haskin, Marc Eshel (IBM); Joan White (San Diego Supercomputer Center); Michelle Butler (National Center for Supercomputing Applications)

Recent advances in parallel NFS (pNFS) and GPFS enable pNFS clients to achieve high performance access to IBM’s GPFS. With a 0.75 PB GPFS-WAN located at SDSC exported through pNFS servers over the TeraGrid network, we will demonstrate pNFS clients accessing GPFS servers in a parallel manner. Nodes at SC ’07, NCSA, NCAR and other TeraGrid sites will mount this parallel file system. This new technology allows a diverse set of machines from different manufacturers to easily access GPFS as a high performance global file system.
Panels

The SC07 panels program focuses on several exciting topics that are extremely relevant and timely in today’s world of high performance computing. Of the 22 excellent proposals we received, we selected seven. These panels span a broad spectrum of topics and offer diverse points of view and an opportunity for audience participation. As in previous years, these sessions are expected to engender lively and even passionate debate.

We encourage you to attend the panels to learn, question, offer opinions and engage with the panelists and the rest of the audience for a fun and enriching experience!

Reno Fact

The University of Nevada, Reno is the oldest university in the state of Nevada and Nevada system of Higher Education. In 1886, the state university, previously only a college preparatory school, moved from Elko in remote northeastern Nevada to a site north of downtown Reno, where it became a full-fledged state college. The university’s first building, Morrill Hall, still stands on the historic quad at the campus’ southern end.
Tuesday, Nov. 13

3:30 p.m. - 5:00 p.m.

How to Get a Better Job in Computing—and Keep It!
Room: A3 / A4

Moderator: Roscoe C. Giles (Boston University)
Panelists: Cecilia Aragon (Lawrence Berkeley National Laboratory), Rebecca Austen (IBM), Beverly Clayton (Pittsburgh Supercomputing Center), José Muñoz (National Science Foundation), Kenneth Washington (Lockheed Martin Enterprise Information Systems)

Leaders from industry, academia and national laboratories will share expertise and advice about future directions for the computing workforce, especially in the area of high-performance computing and related technologies. Panelists will each present a two-minute opening statement, encouraging a discussion driven by questions and comments. Panel Chair Dr. Roscoe Giles, Boston University, was General Chair of SC2002 and was named one of the "50 Most Important Blacks in Research Science" in 2004. Panelists include: Dr. Cecilia Aragon, Staff Scientist, Lawrence Berkeley National Laboratory; Rebecca Austen, Director of Deep Computing & Cell Marketing, IBM; Beverly Clayton, Executive Director Emerita, Pittsburgh Supercomputing Center; Dr. José Muñoz, Senior Scientific Advisor, National Science Foundation Office of Cyberinfrastructure; and Dr. Kenneth Washington, Vice President and Chief Technology Officer of Lockheed Martin Enterprise Information Systems. This 90-minute session will close with each panelist offering their single most important piece of advice.

Thursday, Nov. 15

10:30 a.m. - 12:00 p.m.

Fifty Years of Fortran
Room: A3 / A4
Moderator: Frances Allen (IBM Emerita)
Panelists: David Padua (University of Illinois at Urbana-Champaign), Henry M. Tusto (National Center for Atmospheric Research), John Levesque (Cray, Inc), Richard Hanson (Visual Numerics, Inc.)

This special panel session will commemorate the 50-year anniversary of the release of the first Fortran compiler by examining the impact Fortran has had on system software, applications and computer architecture. Included also will be a brief look at the future of the Fortran programming language.

1:30 p.m. - 3:00 p.m.

Progress in Quantum Computing
Room: A3 / A4
Moderator: David DiVincenzo (IBM)
Panelists: Wim Van Dam (University of California, Santa Barbara), Mark Heiligman (Office of the Director of National Intelligence), Geordie Rose (D-Wave Systems), Will Oliver (Lincoln Laboratories), Eli Yablonovich (University of California, Berkeley)

Hardware to perform quantum information processing is being developed on many fronts. Representing points of view from academia, government, and industry, this panel will give an indication of how work is
progressing on quantum computing devices and systems and what the theoretical possibilities and limitations are in this quantum arena.

3:30 p.m. - 5:00 p.m.

**Supercomputer Operating System Kernels: A Weighty Issue**
**Room:** A3 / A4

**Moderator:** Robert Wisniewski (IBM Research)
**Panelists:** Pete Beckman (Argonne National Laboratory), Sean Blanchard (Los Alamos National Laboratory), Bronis R. de Supinski (Lawrence Livermore National Laboratory), Jim Harrell (Cray, Inc.), Barney Maccabe (University of New Mexico), Ron Minnich (Sandia National Laboratories), Jim Sexton (IBM Research)

You just received your shiny new supercomputer. But wait! Your application doesn’t run because kernel support for a system call is missing. You convince the vendor to add the functionality, but now the application runs too slow because the kernel is doing too much. Being the most important customer for your vendor you force them to take out everything but what you need. Ahh, problem solved—or is it? There are different approaches taken for supercomputer kernels. This panel will present system developers’ and application programmers’ perspectives on needed supercomputer kernel support from a system design standpoint and from an application’s requirement point of view. Should we have multiple kernels? Is there a one size fits all? Is it possible to have a single kernel that would be configured differently depending on the requirements of the application? What is the right weight for a supercomputer kernel?

**Friday, Nov. 16**

**8:30 a.m. - 10:00 a.m.**

**Is There an HEC Energy Crisis?**
**Room:** A1 / A6

**Moderator:** Kirk Cameron (Virginia Tech)
**Panelists:** Andrew Fanara (Environmental Protection Agency), Tom Keller (IBM), Satoshi Matsuoka (Tokyo Institute of Technology), Buddy Bland (Oak Ridge National Laboratory), William Tschudi (Lawrence Berkeley National Laboratory)

Now that supercomputers consume as much power as a small city, the power consumption of high-end systems has become disruptive to the "performance at any cost" design paradigm. The question facing this panel is whether or not power will limit the design and implementation of future leadership-class systems. Panelists will present their thoughts and ideas regarding the future impact of power on the SC community from political policy such as the EPA Energy Star program to facility requirements and systems design. We have assembled a team of experts from diverse domains representing data center management, computer architects, power-aware technology researchers and US policy experts to debate the future of power and high-performance computing. Join us for a lively, collegial debate on this timely topic.
Computational scientists are turning their attention to emerging computing technologies (FPGAs, Cell/B.E., GPUs, etc.) as means for continuous application performance improvements while putting limits on the space/power requirements. Even though these chips were not designed with the HPC applications in mind, their performance characteristics and general availability make them very attractive for HPC applications in spite of the relatively high entry cost when considering programmability and architectural limitations. This Panel will look at the range of newly emerging computing architectures from the perspective of HPC systems and applications developers and will attempt to identify the trends, challenges, and opportunities for these technologies to enter the mainstream HPC. Specific questions for the panel include: Can (and under what circumstances) these systems deliver the advertised performance? The challenge of software development and effective use of the technology. What vendors can/should do to satisfy the needs of computational scientists eager to use these architectures?
**Workshops**

This year’s program includes ten independently planned workshops selected to provide interaction and in-depth discussion of stimulating topics of interest to the HPC community. Four of these workshops have not been held at previous SC conferences.

The workshops bring together national and international organizers and presenters from university, laboratory, and commercial arenas, with support from NSF, DOE’s SciDAC, NIH, and other funding agencies.

The formats of the workshop vary; half are planning independent peer-reviewed publications, while several consist exclusively of invited talks. Over half of these are continuations of workshops held at SC06.

**This year’s topics include:**
- Grid Computing Environments
- Manycore and Multicore Computing
- Reconfigurable Computing
- Performance of High-End Computing Systems
- PetaScale Data Storage
- Distributed Virtualization Technologies
- Ultra-scale Visualization
- MATLAB Parallel Computing
- HPC for Nano-science and Technology
- HPC in China

Participation in these workshops is included as part of the Technical Program registration. Note that several workshops are being held in the Atlantis Hotel.

---

**Sunday, Nov. 11**

8:30 a.m. - 5:00 p.m.

**First International Workshop on High-Performance Reconfigurable Computing Technology and Applications (HPRCTA’07)**

**Room: A8**

Volodymyr Kindratenko (National Center for Supercomputing Applications), Tarek El-Ghazawi (George Washington University), Duncan Buell (University of South Carolina), Kris Gaj (George Mason University), Eric Stahlberg (OpenFPGA), Alan D. George (NSF Center for High-Performance Reconfigurable Computing), Martin Herbordt (Boston University), Olaf Storaasli (Oak Ridge National Laboratory)

High-Performance Reconfigurable Computing (HPRC) based on the combination of conventional microprocessors and field-programmable gate arrays (FPGA) is a new and rapidly evolving computing paradigm that offers a potential to accelerate computationally intensive scientific applications beyond what is possible on today’s mainstream HPC systems. The academic community has been actively investigating this technology for the past several years and the technology has proven itself to be practical for a number of HPC applications. Many of the HPC vendors are now offering or are planning to offer various HPRC solutions. The goal of this workshop is to provide a forum for academic researchers and industry to discuss the latest trends and developments in the field, and to set a
research agenda for the upcoming years on
the use of field-programmable gate array
technology in high performance computing.
Workshop’s website: www.ncsa.uiuc.edu/
Conferences/HPRCTA07/

Workshop on Performance
Analysis and Optimization of
High-End Computing Systems
Room: A9
Xian-He Sun (Illinois Institute of Technology /Fermi National Laboratory), Michael Gerndt
(Technische Universitaet Muenchen)

This workshop is a continuation of its suc-
cess at SC06 and is an extension of a series
of SC APART workshops. It is designed to
bring together researchers working on per-
formance analysis and optimization for par-
allel systems and grids. The interests in per-
formance evaluation as well as in perform-
ance analysis tools are high among SC atten-
dees. This workshop will gather experts in
the field and provide a means to meet the
demand. The speakers are invited only.
Last year’s program is available at www.cs.iit.
edu/~scs/sc06workshop/ and the program of
previous APART workshops can be found at
www.fz-juelich.de/apart.

High Performance Computing in
China: Solution Approaches to
Impediments for High
Performance Computing
Room: Atlantis Hotel-Ballroom A
David K. Kahaner (Asian Technology
Information Program), Hai Jin (Huazhong
University of Science and Technology),
Ninghui Sun (Chinese Academy of Sciences),
Yu Zeng (High Performance Computer
Standardization Committee China)

ATIP’s China HPC Workshop consists of a
day-long set of presentations from a delega-
tion of Chinese academic, research laborato-
ry, and industry/vendor experts and will
include individual presentations, posters and
panel discussions addressing a variety of
related topics, including Chinese
Government Plans, University/Institute
Research, and User Applications. In addi-
tion, the workshop will include presenta-
tions by both Chinese and international
HPC vendors to get their perspectives on
the current status of computing in China. A
key aspect of the workshop will be the
unique opportunity for members of the US
research community to interact and have
direct discussions with some of the top
Chinese scientists in this field.

Petascale Data Storage
Workshop
Room: Atlantis Hotel-Ballroom B
Garth A. Gibson (Carnegie Mellon University /
Panasas Inc.), Darrell Long (University of
California, Santa Cruz), Peter Honeyman
(University of Michigan), Gary A. Grider (Los
Alamos National Laboratory), William T.C.
Kramer (Lawrence Berkeley National
Laboratory), Philip C. Roth (Oak Ridge
National Laboratory), Evan J. Felix (Pacific
Northwest National Laboratory), Lee Ward
(Sandia National Laboratories)

Petascale computing infrastructures make
petascale demands on information storage
capacity, performance, concurrency, reliabili-
ty, availability, and manageability. The last
decade has shown that parallel file systems can barely keep pace with high performance computing along these dimensions; this poses a critical challenge when near-future petascale requirements are considered. This recurring one-day workshop focuses on the data storage problems and emerging solutions found in petascale scientific computing environments, with special attention to issues in which community collaboration can be crucial, problem identification, workload capture, solution interoperability, standards with community buy-in, and shared tools. This year the petascale data storage workshop will hold a peer reviewed competitive process for selecting extended abstracts and short papers. The presenter list is the program committee. See www.pdsi-sci-dac.org/SC07 for more details.

**Workshop on Grid Computing Portals and Science Gateways (GCE 2007)**

**Room: C2**

*Marlon E. Pierce (Indiana University)*

Grid computing portals and science gateways are important components of many large-scale Grid computing projects, serving as the primary gateway for extended and diverse user communities. Portals provide well-established mechanisms for providing familiar interfaces to secure grid resources, services, applications, tools, and collaboration services. Furthermore, portals deliver complex grid solutions to users wherever they have access to a web browser and network access without the need to download or install specialized client software. As a result, the science application user is shielded from the complex details and infrastructure needed to operate an application on the grid. This workshop will continue the highly successful GCE workshop series from SC05 and SC06. We seek papers from all aspects of portal development including portal architecture design and standards studies, toolkits developed in support of portals, integration of Web 2.0 techniques, and the development of high-level application portals that utilize these technologies. The workshop web site, with additional information, is http://www.collab-ogce.org/gce07/index.php/Main_Page.

**Manycore and Multicore Computing: Architectures, Applications and Directions I**

**Room: C3**

*Scott Michel (Aerospace Corporation), Dinesh Manocha (University of North Carolina), Naga Govindaraju (Microsoft), Pradeep Dubey (Intel Corporation)*

Manycore processor systems have tremendous potential for high-performance computing and scientific applications, as these processors can be used as accelerators in the design of teraflop or petaflop computers. The significant increase in parallelism within a processor can also lead to other benefits including higher power-efficiency and better memory latency tolerance. Building on the success of last year’s ”General-Purpose GPU Computing: Practice and Experience”, as well as the success of other GPGPU and Edge workshops held over the last few years, this 1 1/2 day workshop will examine the recent trends in these areas as well a range of topic areas through invited talks, poster ses-
sions, and a panel session to discuss many/multicore programmability and overall community research themes.

**Ultra-scale Visualization**

**Room: Atlantis Hotel-Ballroom A**

Kwan-Liu Ma (University of California, Davis), Christopher Johnson (University of Utah)

The output from leading-edge scientific simulations is so voluminous and complex that advanced visualization techniques are necessary to interpret the calculated results. Even though visualization technology has progressed significantly in recent years, we are barely capable of exploiting terascale data to its full extent, and petascale datasets are on the horizon. This workshop aims at addressing this pressing issue by fostering communication between visualization researchers and practitioners, high-performance computing professionals, and application scientists. Attendees will be introduced to the latest and greatest research innovations in large data visualization, and also learn how these innovations impact the scientific supercomputing and discovery process. For more information about this workshop, please visit [http://vis.cs.ucdavis.edu/Ultravis07](http://vis.cs.ucdavis.edu/Ultravis07).

**VTDC 2007: 2nd International Workshop on Virtualization Technologies in Distributed Computing**

**Room: Atlantis Hotel-Ballroom B**

Kate Keahey (Argonne National Laboratory)

The convergence of virtualization technologies and distributed computing is an exciting development and the subject of much research in both academia and industry. Building on a very successful and highly attended VTDC workshop at SC2006, the 2nd VTDC workshop will continue to be a forum for the exchange of ideas and experiences on the use of virtualization technologies in distributed computing, the challenges and opportunities offered by the development of virtual systems themselves, as well as case studies of application of virtualization. The scope of "virtualization technologies" includes techniques and concepts to enable virtual machines, virtual networks, virtual data, virtual storage, virtual applications and virtual instruments. The scope of "distributed computing" includes Grid-computing, cluster computing, peer-to-peer computing and mobile computing. For additional information see [http://workspace.globus.org/vtdc07](http://workspace.globus.org/vtdc07).

**Workshop on Grid Computing Portals and Science Gateways (GCE 2007)**

**Room: C2**

Marlon E. Pierce (Indiana University)

Grid computing portals and science gateways are important components of many large-scale Grid computing projects, serving
as the primary gateway for extended and diverse user communities. Portals provide well-established mechanisms for providing familiar interfaces to secure grid resources, services, applications, tools, and collaboration services. Furthermore, portals deliver complex grid solutions to users wherever they have access to a web browser and network access without the need to download or install specialized client software. As a result, the science application user is shielded from the complex details and infrastructure needed to operate an application on the grid. This workshop will continue the highly successful GCE workshop series from SC05 and SC06. We seek papers from all aspects of portal development including portal architecture design and standards studies, toolkits developed in support of portals, integration of Web 2.0 techniques, and the development of high-level application portals that utilize these technologies. The workshop web site, with additional information, is www.collabogce.org/gce07/index.php/Main_Page.

**Manycore and Multicore Computing: Architectures, Applications and Directions II**

**Room: C3**

*Scott Michel (Aerospace Corporation), Dinesh Manocha (University of North Carolina), Naga Govindaraju (Microsoft), Pradeep Dubey (Intel Corporation)*

Manycore processor systems have tremendous potential for high-performance computing and scientific applications, as these processors can be used as accelerators in the design of teraflop or petaflop computers.

The significant increase in parallelism within a processor can also lead to other benefits including higher power-efficiency and better memory latency tolerance. Building on the success of last year’s "General-Purpose GPU Computing: Practice and Experience", as well as the success of other GPGPU and Edge workshops held over the last few years, this 1-1/2 day workshop will examine the recent trends in these areas as well a range of topic areas through invited talks, poster sessions, and a panel session to discuss many/multicore programmability and overall community research themes.

**1:30 p.m. - 5:00 p.m.**

**Parallel Computing with MATLAB**

**Room: C3**

*Jeremy Kepner (MIT Lincoln Laboratory)*

MATLAB has emerged as one of the languages most commonly used by scientists and engineers for technical computing, with ~1,000,000 users worldwide. The primary benefits of MATLAB are reduced code development time via high levels of abstractions (e.g. first class multi-dimensional arrays and thousands of built in functions), interpretive, interactive programming, and powerful mathematical graphics. The compute intensive nature of technical computing means that many MATLAB users have codes that can significantly benefit from the increased performance offered by parallel computing. There are now a number of parallel programming solutions available for MATLAB. This workshop will bring together the technical leaders in this field and provide a unique opportunity for researchers
and users in the growing community for parallel computing with MATLAB software to interact.

Friday, Nov. 16

8:30 a.m. - 5:00 p.m.

Third International Workshop on High Performance Computing for Nano-science and Nanotechnology (HPCNano07)

Room: A3 / A4

Jun Ni (University of Iowa), Andrew Canning (Lawrence Berkeley National Laboratory)

Nanotechnology is an exciting field with many potential applications. Its impact is already being felt in materials, engineering, electronics, medicine, and other disciplines. Current research in nanotechnology requires multi-disciplinary knowledge, not only in sciences and engineering but also in HPC technology. Many nano-science explorations rely on mature, efficient HPC and computational algorithms, practical and reliable numerical methods, and large-scale computing systems. This workshop offers academic researchers, developers, and practitioners an opportunity to discuss various aspects of HPC-related computational methods and problem solving techniques for nano-science and technology research. The HPCNano05 and HPCNano06 were successful events in SC05 and SC06. This is the third one. We hope to attract people from diverse science and engineering disciplines, nationally and internationally, to attend the workshop, present their research results, share their experiences and ideas, and plan future collaborations. The accepted papers will be published by the IEEE CS Press. Web site: www.hpcnano.org/HPCNano07.
Disruptive Technologies

First introduced at SC06, Exotic Technologies returns in SC07 as Disruptive Technologies.

Generally speaking, a disruptive technology or disruptive innovation is a technological innovation, product, or service that eventually overturns the existing dominant technology or product in the market.

This year we examine those technologies that may significantly change the world of HPC in the next five to fifteen years. Consisting of panel sessions and competitively selected exhibits, Disruptive Technologies showcases technologies not currently on industry’s technology roadmaps. Technologies like reconfigurable computing, heterogeneous multi-core chips, holographic storage, and novel cooling techniques may offer near-term benefits, while quantum computing, chip level optical interconnects, and fundamental material breakthroughs offer potentially paradigm-changing benefits over the long term.

Disruptive Technologies will also host five competitively selected exhibits at SC07. See pages 180-182 for details.

Wednesday, Nov. 14
3:30 p.m. - 5:00 p.m.

Interconnects
Room: A3 / A4
Session Chair: Jeffrey Vetter (Oak Ridge National Laboratory and Georgia Institute of Technology)

Interconnects play a critical role in the architecture of scalable supercomputers. In this panel, we examine technical innovations in the area of interconnects that may dramatically change their performance, cost, or scale. These topics include innovations in optical interconnects, CMOS photonics, switching, routing, and topologies.

Thursday, Nov. 15
3:30 p.m. - 5:00 p.m.

Memory Systems
Room: A10 / A11
Session Chair: Erik P. DeBenedictis (Sandia National Laboratories)

Memory systems play a critical role in the architecture of scalable supercomputers. In this panel, we examine technical innovations in the area of memory systems that may dramatically change their performance, cost, or scale. These topics include innovations in 3D stacking, on-chip photonics, memory materials, and memory device and system architectures.
The SC07 Education Program will engage participants in hands-on activities to assist them in applying computational science, grid computing and high performance computing resources in all fields of study. Participants will gain access to resources they can directly utilize to enhance classroom learning and to excite students in pursuing careers in science, technology, engineering, and mathematics.

**Student Program**
The SC Education Student Program allows undergraduate and graduate students to attend a wide variety of presentations, round tables, and panels led by representatives from colleges, universities, industry, and research institutions. A series of hour long talks will address a broad spectrum of available topics that include education, jobs, internships, careers, and experiences in high performance computing and computational science. Students will learn to optimize grad school applications, obtain internships and tune resumes. Students can choose to attend as many or as few of these talks as they desire. All students attending SC07 are strongly encouraged to attend the Student Program.

**K-12 Program**
The K-12 Teacher’s Day is designed to give teachers a voice, in the current discussions about the ways to create the future of education that will influence the students they reach and the classrooms in which they teach. *The digital divide helps widen an even more alarming divide - the knowledge divide. We will discuss ways to broaden participation in education and to create pathways to STEM careers.* The teachers will participate in a forum that will share with them, the snapshots of the current dialogue from business groups that has reached the congress and that will influence new legislation. Then they will have their say. We will develop a white paper from this effort. The teachers will then be immersed in the offerings of the conference exhibits.

**The Learning & Physical Challenges Education Program**
The Learning & Physical Challenges Education Program will provide K-12 and undergraduate instructors, special education professionals, and students with the knowledge and resources to empower all students to become full participants in utilizing cyberinfrastructure and computational science resources during their education and during their professional careers.
Friday, Nov. 9

5:00 p.m. - 9:00 p.m.
Registration
Room: Nugget Hotel

Saturday, Nov. 10

7:00 a.m. - 8:00 a.m.
Breakfast
Room: D1 / D2 / D3

8:30 a.m. - 8:45 a.m.
Welcome and Introductions
Room: E1 / E2 / E3

8:45 a.m. - 9:15 a.m.
Invited Talk - Research Effectiveness of Computational Science Education
Room: E1 / E2 / E3

9:15 a.m. - 9:45 a.m.
Invited Talk - Computational Science Concept Map
Room: E1 / E2 / E3

9:45 a.m. - 10:00 a.m.
Desktop to Grid Computing--Examples Across Disciplines
Room: E1 / E2 / E3

10:00 a.m. - 10:30 a.m.
Break
Room: E1 / E2 / E3

10:30 a.m. - 12:00 p.m.
Desktop to Grid Parallel Sessions:

CFD
Room: D10

Biology
Room: D4

Physics
Room: D5

Parallel Computing
Room: D6

Environmental and Weather
Room: D7

Mathematics and Algorithms
Room: D9

Chemistry
Room: E1 / E2 / E3

12:00 p.m. - 1:30 p.m.
Lunch
Room: D1 / D2 / D3
1:30 p.m. - 2:30 p.m.
Invited Talk - Computational Science Education for All
Room: E1 / E2 / E3

2:30 p.m. - 3:00 p.m.
Education Program Awards Announcement
Room: E1 / E2 / E3

3:00 p.m. - 3:30 p.m.
Break
Room: E1 / E2 / E3

3:30 p.m. - 4:00 p.m.
Pathways Reports - Inter-disciplinary Impact
Room: E1 / E2 / E3

4:00 p.m. - 5:00 p.m.
Disciplinary Breakout Discussions:

Chemistry
Room: D4

Physics
Room: D5

Biology
Room: D6

Humansities, Arts and Social Sciences
Room: D7

5:00 p.m. - 6:00 p.m.
Birds-of-a-Feather Sessions (BOFs)
BOFs will run in parallel and will be held in Rooms D4, D5, D6 and D7. Check the website for titles and speakers.

6:00 p.m. - 7:00 p.m.
Break

7:00 p.m. - 9:00 p.m.
Education Program Poster Session and Reception
Room: Nugget Hotel

Sunday, Nov. 11

7:00 a.m. - 8:00 a.m.
Breakfast
Room: D1 / D2 / D3

8:30 a.m. - 9:30 a.m.
Invited Talk: Vision for High Performance Computing in Education
Room: E1 / E2 / E3
9:30 a.m. - 10:00 a.m.

**Computational Science Education - Vision, Challenges and Plans**
Room: E1 / E2 / E3

10:00 a.m. - 10:30 a.m.

**Break**
Room: E1 / E2 / E3

10:30 a.m. - 12:00 p.m.

**Parallel Sessions:**

- **nanoHub Science Gateway**
  Room: D10

- **Agent Systems**
  Room: D4

- **Systems Dynamics**
  Room: D5

- **Numerical Systems**
  Room: D6

- **Visualization/GIS**
  Room: D7

- **Grid Computing**
  Room: D9

- **Learning and Physical Challenges Education Program**
  Room: E1 / E2 / E3

3:00 p.m.-3:30 p.m.

**Break**
Room: E1 / E2 / E3
3:30 p.m. - 5:00 p.m.

**Parallel Sessions:**

- **MetaVerses**  
  Room: D10

- **Text Analysis**  
  Room: D4

- **Data Analysis**  
  Room: D5

- **GIS and Mash-ups**  
  Room: D6

- **Easy Java Simulations/OSP**  
  Room: D7

- **Grid Computing**  
  Room: D9

5:00 p.m. - 6:00 p.m.

**Monday, Nov. 12**

7:00 a.m. - 8:00 a.m.

**Breakfast**  
Rooms: D1 / D2 / D3

8:30 a.m. - 12:00 p.m.

**On-Site Student Competition**  
Room: D10

8:30 a.m. - 9:30 a.m.

**Invited Talk - Learning and Physical Challenges Education Program**  
Room: E1 / E2 / E3

9:30 a.m. - 10:00 a.m.

**Plenary Session - SC07 Conference Highlights: Technical Program, Broader Engagement, SCinet**  
Room: E1 / E2 / E3

10:00 a.m. - 10:30 a.m.

**Break**  
Room: E1 / E2 / E3

**BOFs**

BOFs will run in parallel and will be held in Rooms D4, D5, D6 and D7. Check the website for titles and speakers.
10:30 a.m. - 12:00 p.m.

Parallel Sessions:

Biology
Room: D4

Chemistry
Room: D5

Physics
Room: D6

Computer Science
Room: D7

Humanities, Arts and Social Sciences
Room: D9

Mathematics
Room: E1 / E2 / E3

12:00 p.m. - 1:30 p.m.

Lunch
Room: D1 / D2 / D3

1:30 p.m. - 5:00 p.m.

On-Site Student Competition
Room: D10

1:30 p.m. - 3:00 p.m.

Parallel Sessions continued:

Biology
Room: D4

Chemistry
Room: D5

Physics
Room: D6

Computer Science
Room: D7

Humanities, Arts and Social Sciences
Room: D9

Mathematics
Room: E1 / E2 / E3

3:00 p.m. - 3:30 p.m.

Break
Room: D4 / D5

3:30 p.m. - 5:00 p.m.

Parallel Sessions:

Intel/Total View Parallel Programming
Room: D4

Science Gateways
Room: D5
Web Services 2.0  
Room: D6

Matlab Tutorial  
Room: D7

Condor  
Room: D9

Parallel Session - Getting Started Using TeraGrid Resources  
Room: E1 / E2 / E3

**5:00 p.m. - 6:00 p.m.**

**Birds of a Feather Sessions (BOFs)**  
BOFs will run in parallel and will be held in Rooms D4, D5, D6 and D7. Check the website for titles and speakers.

---

**Tuesday, Nov. 13**

**7:00 a.m. - 8:00 a.m.**

Breakfast  
Room: D1 / D2 / D3

**8:30 a.m. - 10:00 a.m.**

General SC Keynote - Education Participants welcome to attend  
Room: Ballroom

---

10:00 a.m. - 10:30 a.m.

Break  
Room: E1 / E2 / E3

10:30 a.m. - 12:00 p.m.

K-12 Program  
Room: D10

10:30 a.m. - 11:00 a.m.

Invited Talk - Empowering students as computational science professionals  
Room: E1 / E2 / E3

11:00 a.m. - 11:30 a.m.

Plenary Session - Student Competition Awards Presentation  
Room: E1 / E2 / E3

11:30 a.m. - 12:00 p.m.

Plenary Session - Next Steps  
Room: E1 / E2 / E3

12:00 p.m. - 1:30 p.m.

Lunch  
Room: D1 / D2 / D3

1:30 p.m. - 3:00 p.m.
Parallel Sessions:

K - 12 Program
Room: D10

HPC University resources, and CSERD -- NSDL resources
Room: D4

Mathematica Tutorial
Room: D5

BCCD
Room: D6

Student Program
Room: D7

Getting started using Open Science Grid resources
Room: D9

Collaboration Tools
Room: E1 / E2 / E3

3:00 p.m. - 3:30 p.m.

Break
Room: D4 / D5

3:30 p.m. - 5:00 p.m.

Parallel Sessions:

K-12 Program
Room: D10

Student Program
Room: D7

Project Reports and Participant Plans
Room: E1 / E2 / E3

Wednesday, Nov. 14

Student Program
Room: D7

The Student Program will run from 8:30 a.m. to 3:00 p.m with breaks where specified during the session.

Thursday, Nov. 15

Student Program
Room: D7

The Student Program will run from 8:30 a.m. to 3:00 p.m with breaks where specified during the session.
Over the years, the SC conference series has sponsored a number of programs to increase participation by groups outside the mainstream of high performance computing and networking. Starting with SC07, these outreach efforts are expanding with the addition of the Broader Engagement initiative, which is committed to broadening the engagement of individuals from groups that have traditionally been under-represented in the field. To achieve this, the Broader Engagement initiative (BE) is providing grants to support participation in the technical program, encouraging technical program submissions, and fostering networking through both a formal mentoring program and informal contacts at SC07.

On-Ramp to SC07

The breadth and depth of the SC conference provide unparalleled opportunities for discovering and exploring innovations in high performance computing, networking, storage, and analysis. But the size and scale of the conference can also be confusing, if not intimidating, to someone who is attending their first SC. To help all attendees, especially those participating for the first time, navigate the myriad conference components, BE will host a kiosk at the Reno-Sparks Convention Center, staffed with volunteers who can answer questions and provide guidance to attendees and help them decide which conference events and resources best fit their needs and interests. Print materials are also available, designed especially for first-time SC attendees.

Reno Fact

The Truckee River serves as Reno’s primary source of drinking water. It supplies Reno with 80,000,000 gallons of water a day during the summer and 40,000,000 gallons of water per day in the winter. As an attempt to save water, golf courses in Reno, like Arrow Creek Golf Course, have been using treated effluent water instead of treated water from one of Reno’s water plants.
**Mentorship Program**

In close collaboration with the Education Program and SC exhibitors, BE has also developed a mentorship program. While general guidance is valuable, a more formal program bringing together people in the early stages of their careers with established professionals can lead to stronger career development, increase awareness of options and build a sense of connection to the community. Through the formation of mentor/protégé pairs, bridges are built to support emerging research and industry careers and to encourage further participation in the SC conference.

**BE With Us at SC07 and Beyond**

Informal networking opportunities and a dynamic Broader Engagement kickoff event will bring together those participating in the BE mentorship program and participation grants, providing an opportunity to develop relationships and collaborations that continue beyond the conference.

Although BE is being launched at SC07, the initiative will to continue through SC08 in Austin and into subsequent years. However, BE is not being created as a once-a-year event, but is being developed as a year-round program to build a broader community. In addition to broadening the engagement of individuals in the SC conference from groups which have traditionally been under-represented in high-performance computing, BE also aims to broaden the engagement of under-represented groups in the field of high-performance computing. The goal is to develop a circular movement, with more people coming into HPC, then participating in the SC conference series, which helps them advance in their careers to the point where they can also serve as mentors and help others become successful.

To do this, BE is expanding the reach of the SC conference by developing relationships with interested partners at the regional and international levels, as well as in specific research disciplines involved in the HPC community. BE committee members are also collaborating across the entire SC conference organization to develop a better understanding of how the conference may engage and serve those from traditionally under-represented groups and emerging fields. This includes working with the Technical Program, the Education Program and other components as appropriate. BE leverages past outreach efforts, such as the successful Minority Serving Institutions program, which introduced instructors from schools serving under-represented students to the SC conference.
To achieve these goals, the BE Initiative is undertaking the following activities:

- Increasing focus on students to contribute to national efforts on recruitment and to help develop a “pipeline” for the next-generation of HPC researchers
- Continuing engagement of faculty, researchers, and other professionals, especially those at a junior level
  - for their own contributions to the conference
  - to build stronger bridges to their students
  - to serve as a bridge for them to build collaborations
  - to serve them in providing more information in areas related to their research
- Developing year-round online activities, creating a community
- Building on past success of MSI Network
- Expanding into new activities
- Annual activities at the SC Conference, including participation grants to enable engagement of individuals who could not otherwise attend
- Cornerstone events which build bridges among BE participants, as well as with others in the HPC user community

So, how will the HPC community know if BE is making a difference? To start with, the BE organizers are conducting surveys of participants to find out the value of the program and identify ways to make it more effective. The committee also plans to increase the visibility of the program with the aim of increasing the number of applications.

Ideally, the success of the program will be evident in the success of BE participants through their career development, the publication of research papers and, ultimately, in their professional growth to the point where they can serve as mentors for the next generation. We would also like to see them close the loop by becoming an integral part of the SC conference itself, contributing to the technical program and serving on the organizing committee, giving input on how the conference can expand its value to all members of the HPC community.
At SC07, industry and research exhibits from the world’s leading companies and organizations are showcased in a dynamic, interactive environment. High performance computing, networking, storage, data management, scientific visualization and collaborative technology are featured.

Industry exhibitors will demonstrate the latest advancements in HPC technology, showcasing new hardware, software, services and innovations. More than 300 industry exhibitors are represented at SC07.

Research exhibits provide an international venue for scientists and engineers to display the latest computational science advances, research and development plans, new concepts, software and other technologies. More than 100 research exhibitors are participating at this year’s conference.

Exhibit Hours

<table>
<thead>
<tr>
<th>Date</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov. 11, 2007</td>
<td>7 p.m.-9 p.m.</td>
</tr>
<tr>
<td>Nov. 12, 2007</td>
<td>10 a.m.-6 p.m.</td>
</tr>
<tr>
<td>Nov. 13, 2007</td>
<td>10 a.m.-6 p.m.</td>
</tr>
<tr>
<td>Nov. 14, 2007</td>
<td>10 a.m.-4 p.m.</td>
</tr>
</tbody>
</table>

Reno Fact

As the mining boom waned early in the twentieth century, Nevada’s centers of political and business activity shifted to the non-mining communities, especially Reno and Las Vegas, and today the former mining metropolises stand as little more than ghost towns. Despite this, Nevada still accounts for over 11% of world gold production.
Disruptive Technologies

This year we examine those technologies that may significantly change the world of HPC in the next five to fifteen years. Consisting of panel sessions (see the Panels & Workshops section) and competitively selected exhibits (described here), Disruptive Technologies showcases technologies not currently on industry’s technology roadmaps. Technologies like reconfigurable computing, heterogeneous multi-core chips, holographic storage, and novel cooling techniques may offer near-term benefits, while quantum computing and optical interconnects and fundamental advances.

Disruptive Technologies exhibits are open during regular exhibit hours.

High Performance Optical Connectivity Based on CMOS Photonics Technology

Marek Tlalka (Luxtera)

Luxtera developed a breakthrough nanophotonic technology that enables manipulation of both photons and electrons on a single semiconductor CMOS die and can be produced in high-volume, low-cost mainstream CMOS processes. This breakthrough silicon photonics technology enables connection of fiberoptic cable directly to a silicon die. It’s a basic toolkit which has the potential to replace nearly any existing optical system with a single chip and in the future will proliferate high-performance optical connectivity at price points traditionally associated with low-performance, legacy copper at the network, system and chip-to-chip level. Enabling a multitude of new semiconductor products and applications in the interconnect space, Blazar is the first commercial application specifically targeted for High Performance Computing. Other applications range from wide-area networking to enterprise networking where a single chip will replace expensive existing optical modules and enable high-performance optical connectivity, rendering emerging 10G BASE-T copper standard obsolete before it enters the market.

Optical Printed Circuit Board Technology and 200+200Gbps Transceiver

Clint Schow, Fuad Doany, Jeffrey Kash, Marc Taubenblatt (IBM)

IBM Research has developed an Optical Printed Circuit Board technology consisting of chip-like optical transceivers (currently supporting 16+16 optical channels at 12.5Gbps each) and polymer waveguides on circuit cards. Our technology is disruptive in that it would replace today’s high cost optical modules based on glass fiber technology with mass manufacturable “optical printed circuit boards.” for short backplane and card level links. Although polymer based waveguides have higher losses than glass fiber technology, the ability to use lithographic processes to mass produce this technology coupled with the use of chip like optical components will allow a low cost solution for this ultra-short interconnect application. We are working to develop a supplier ecosystem to mature this technology in the next 5
to 7 years. As microprocessors become more capable through multi-core architectures, the technology bottleneck for HPC is shifting to the interconnect fabric, requiring immense low power BW at low cost.

**NRAM as a Disruptive Technology**
*Robert Smith, Gerry Taylor (Nantero, Inc.)*

Nantero has developed a new memory technology called NRAM (Non-Volatile Random Access Memory). NRAM is a technology based on the ability to create memory switches using carbon nanotubes and is the first product that enables the scalability and mass production of a real nanotechnology device to the world of complex electronics. The semiconductor industry provides an interesting test case for how the advent of these advanced materials will alter the entire landscape of an entire industry. It is a well known fact that Moore’s law, the principle by which the semiconductor scales down features sizes to improve economics and performance, will cease to provide these advantages in the next fifteen years. Nantero believes that NRAM is a near term product solution and will eventually also change all memory technology as the first “Universal Memory.” All computing devices from laptops to cell phones to HPC will be enhanced by carbon nanotube memory.

**Superconducting Quantum Computing System**
*Geordie Rose (D-Wave Systems)*

We wish to showcase our superconducting quantum computing system, which we believe is a paradigm-changing approach to HPC. Code-named Orion, our current system is designed to solve combinatorial optimization problems using superconducting analog quantum computer processors. These processors (the first of their kind) can, at modest levels of integration, be used to solve problems beyond the scope of traditional HPC systems. In addition, these processors can in principle be operated reversibly---without generation of heat---and in practice are projected to reduce power consumption for solution of hard problems by many orders of magnitude over commodity cluster solutions. Our intent for SC07 is to allow conference attendees to program and use an Orion system directly from our booth.

**Disruptive Technology for Many-core Chip System Software and Logic Co-verification**
*Guang R. Gao (University of Delaware / ET International Inc.), Monty Denneau (IBM Research)*

ETI has developed a disruptive technology for many-core system software and logic co-verification. A complete system may contain many such chips (e.g. 64-bit 160 cores on a chip and many chips in a system in the case of the IBM Cyclops-64 supercomputer). In this exhibit, we demonstrate the co-verification of ETI multithreaded programming model and system software solution on Mrs. Clops - a parallel FPGA based supercomputer for Cyclops-64 emulation. Based on the DIMEs (Delaware Iterative Emulation System) Technology, the Mrs. Clops engine can emulate multiple Cyclops-64 chips at the gate-level at high speed. This allows ETI system software be booted entirely on a
multi-chip C64 system configuration with a large number of parallel programs be executed for co-verification. The innovative ETI TNT (TinyThread) programming model will be demonstrated in addition to legacy SHMEM and OpenMP. Attendees will be able to experience the entire ETI software solution on demonstration in our booth. Programming techniques will be demonstrated at the source code level.

Exhibitor Forum

As at previous SC conferences, a highlight of the Exhibits program is the Exhibitor Forum. These talks showcase the latest advances by our Industry Exhibitors, including new products and upgrades, recent research and development and future plans and roadmaps. Some present case studies by their customers, who have achieved better performance, reduced time-to-market, improved reliability, or enhanced capabilities by using the Exhibitor’s technology. Others give insight into the technology trends driving their strategy, the potential of emerging technologies in their product lines, or the impact of adopting academic research into their development cycle. Whatever the topic, Exhibitor Forum can show you what is available today and tomorrow.

Industrial exhibitors are offered the opportunity to highlight their latest technology breakthroughs in this forum. These sessions run Tuesday-Thursday and are open to all SC07 attendees.

Tuesday, Nov. 13

Petascale
Room: A7

10:30 a.m. - 12:00 p.m.

Production-ready Petascale Computing
Bjorn Anderson (Sun Microsystems)

As high performance computing moves into commercial markets, customers are increasingly demanding "production-ready" HPC. Sun combines open, production-ready supercomputing to "Capability" systems that scale to Petaflops of performance. With a combination of high-density blade computing and the world’s largest InfiniBand switch, the Sun Constellation System brings together innovation, flexibility, and choice to HPC customers seeking competitive advantage. The Network is the Computer.

Innovation beyond Imagination: The Road to PetaFLOPS Computing
Tony Befi (IBM)

As the race to petascale computing heats up, it is increasingly evident that while performance is the ultimate goal, applications are in the driver’s seat with ease-of-use sitting right beside them. The computational vehicles will vary, yet each will need to be built with innovative usability and management features as well as strong reliability and serviceability. Successful PetaFLOPS computing
requires a true systems approach. Chip technology, systems, software, interconnect fabric, system management, usability, reliability, interoperability and security are just some of the key elements requiring innovation and integration. We’ve been on the petaFLOPS road for a while now, taking new routes and detours along the way. This presentation will explore the journey so far, including the bumps along the road as well as some driving factors giving us headlights to the goal.

Beyond Standards: A Look at Innovative HPC Solutions
Scott Misage (Hewlett-Packard)

High performance computing (HPC) is at the forefront of key trends and is the market in which disruptive technologies are created and evolved into broader commercial developments. Scale-out computing and Linux are recent examples. This presentation examines emerging HPC technologies in computation, data management and visualization that are now guiding the future of enterprise computing. Topics include: * Multi-core optimization: Effectively utilize future multi-core processor designs with unprecedented parallelism technologies. * Accelerators: Substantially increase application performance on industry standard platforms through general purpose graphics processing units (GP GPUs) and field programmable gate arrays (FPGAs). * Dense computing: Radically improve performance and operating costs per square meter with high-density blade technology optimized for HPC. * Remote caching: Enable geographically distributed workers to share large amounts of data with low data access latency. * Converged fabrics: Reduce infrastructure and management costs without sacrificing quality of service. * Parallel compositing: Efficiently render huge data sets to greatly enhance the visualization capabilities for complex scientific and engineering applications.

Networking—10 Gb Ethernet
Room: A8
10:30 a.m. - 12:00 p.m.

Scalable, Congestion-free, Low Latency 10 Gigabit Ethernet Fabrics for Compute and Storage Clusters
Bert Tanaka (Woven Systems)

Network traffic congestion is a common reason for performance and scaling degradation in statically routed HPC interconnect fabrics such as InfiniBand, LAG or ECMP routing. Woven Systems’ Active Congestion Management technology solves this problem for multi-stage, multi-path 10 GE fabrics by (1) detecting congestion through one-way latency and jitter measurements, and (2) dynamically re-routing traffic around congestion so as to fully utilize all of the capacity of the fabric. Woven will present results that compare the performance of an InfiniBand and 10 GE cluster using Sandia’s CBENCH HPC benchmarking suite. These results demonstrate the importance of dynamically re-routing traffic to avoid congestion to improve the throughput in large HPC interconnect fabrics.
Dynamic Ethernet Lightpaths: On-demand 10GbE and GbE Connections for Research Networks
Jeff Verrant (Ciena Government Solutions)

The proliferation of optical networking throughout the global research networking community has created the foundation for the creation of Dynamic Ethernet Lightpath provisioning between regional and global research networks. Ciena will provide an overview of dynamic provisioning of Ethernet lightpaths (GbE, 10GbE), fueled by technologies that represent the latest advancements in optical networking. Key capabilities such as OTU/OTN, multi-protocol support, 10GbE optical switching, and automated control plane technologies will be explained in context of providing automated provisioning and management for regional and global research networking requirements. Network deployments will be highlighted that feature these technologies in support of research networks.

Design and Technology for Supercomputers and Grids: Growth of 10 GbE
Jeremy Stieglitz (Force10 Networks)

Networking & supercomputing are continuing to converge and are key elements in the realm of shared cyberinfrastructure and grids. 10 Gigabit Ethernet is continuing to grow as the technology of choice not only for the WAN and grid interconnections, but in the heart of supercomputers. This session will focus on design considerations and new technologies for different 10 Gigabit Ethernet applications, including the use of emerging low latency switching and NICs; density, resiliency & switching economics; and challenges and solutions of security & firewalls with 10 GbE.

Architectures
Room: A7
1:30 p.m. - 3:00 p.m.

Fujitsu’s Solutions and Vision for High Performance Computing
Motoi Okuda (Fujitsu, Ltd.)

Fujitsu provides HPC solutions to meet the widest range of customer requirements. Solutions include PRIMERGY IA cluster systems with high speed interconnects, large scale SMP systems - the Itanium2 PRIMEQUEST and SPARC64 Enterprise servers - as well as application specific systems. Fujitsu is deeply involved in Japan’s next generation supercomputer project, leading to Petascale computing. We are driving the R&D project for a Petascale interconnect together with Kyushu University, while providing the detailed design for a 10 PFlops class supercomputer for RIKEN. In this presentation, we will introduce Fujitsu’s vision for HPC including current products, hardware and software environment, and the challenges of Petascale computing.
**NEC’s HPC Strategy: Consistency and Innovation**

*Rudolf Fischer (NEC HPCE)*

For more than 20 years NEC SX vector supercomputers have been developed with a consistent strategy of innovations. The market leading balance between peak performance and memory bandwidth is the key differentiator. The presentation will focus on the next generation of the SX-series, which is a major breakthrough in providing the world’s fastest processor (>100 GFlops), maintaining a memory bandwidth (256 GByte/s per CPU) to one flat shared memory of up to one TByte. The architecture incorporates many new functionalities which reflect the application-based experience of NEC. Recent successes have proven that value is not only determined in price/performance, but also increasingly more in terms of floor-space, power-consumption, and sustained application performance. Taking these factors into account, SX series delivers very low TCO. NEC is committed to continue the development based on this proven strategy; NEC will continue to provide superior tools to enable breakthroughs in science and engineering.

---

**Cray Advances Adaptive Supercomputing Vision**

*Barry Bolding (Cray, Inc.)*

At SC07 Cray will be introducing the first supercomputer-class hybrid system. This supercomputer builds upon Cray’s highly successful Cray XT product line with new computing technologies for a variety of workloads and workflows as well as extended scalability to 100,000 cores. The hybrid system includes new packaging technologies that increase density while reducing power and cooling costs, significantly lowering total cost of ownership (TCO); and also introduces a new robust, highly scalable Linux OS that allows easy porting of ISV codes to this new platform.

---

**Tools—Scheduling and Debugging**

*Room: A8*

1:30 p.m. - 3:00 p.m.

**Node-level Scheduling Optimization for Multi-core Processors**

*Benoit Marchand (eXludus Technologies)*

The rapid adoption of multi-core processors enables cost-effective processing across a broad range of applications. However, due to increasing core counts per node, increasing cluster scale, and increasingly dynamic job workloads, a scheduling gap exists between the workload manager and the node multi-tasking operating system. Workload managers lay out a static high-level assignment of work to nodes based on ab initio resource requirements, but real-time process scheduling optimization is needed at the node level to maintain maximum throughput. eXludus has developed Grid Optimizer, the industry’s first multi-core scheduling optimizer, that increases application throughput in conjunction with any existing workload manager. We present results demonstrating enhanced throughput of 1.6 to 2.5 times in
application disciplines such as genomic sequencing, protein configuration analysis and mass spec analysis. We also present results obtained running highly heterogeneous workloads.

**What’s In Store: How Organizations Cope as they Mature From Clusters to Adaptive Computing**  
*Michael A. Jackson (Cluster Resources, Inc.)*

Historically, compute resource environments have been modified manually in response to ever-changing needs and increasingly complex technologies. From an HPC perspective, this begins with a simple cluster and evolves as additional resources or clusters are added and sites move through the various cluster and grid stages to ultimately reach utility and adaptive computing, for sites at the more mature end of the cycle. As organizations progress through the cycle, they can either become increasingly overwhelmed or put management responses in place. Moab is a policy-based intelligence layer that integrates with existing middleware for consolidated administrative control and holistic reporting across each step in the complex management evolution. We will present pertinent case studies and focus on management issues and best practice resolutions, offering listeners insight into the issues and resolutions that match their current management levels, and provide them with a view into what is likely to hit them next.

**Debugging for PetaScale**  
*Michael A Rudgyard (Allinea Software)*

One of the greatest challenges for high performance computing users is how to deal with bugs that only occur at high processor counts. Allinea’s Distributed Debugging Tool, DDT, is recognized for its easy-to-use interface but has also been addressing the issues of debugging at scale. In this talk we will discuss these challenges as well as how DDT can help to address them. We will discuss features and supported platforms for both DDT and its companion product the Optimization and Profiling Tool, OPT.

**Network Management**  
*Room: A7*

**3:30 p.m. - 5:00 p.m.**

**The Benefits of Multi-protocol Networking Architecture**  
*Micahel Kagan, Gilad Shainer (Mellanox Technologies)*

With the growing demand for high-performance fabrics, the industry has developed and deployed multiple high-performance fabric infrastructures. A dynamic choice of fabrics facilitates deployment and maintenance of HPC clusters. A Multi-I/O architecture provides seamless connectivity to Ethernet, InfiniBand, Fibre Channel, etc. This architecture provides the IT managers with the flexibility to configure the fabric with the interconnect technology of choice, and easily reconfigure it to adapt to the ever changing requirements on performance,
availability, connectivity, legacy, etc. The Multi-protocol architecture, consolidating 40Gb/s InfiniBand and 10Gb/s Ethernet protocols, with an emphasis for multi-core environments, can leverage the existing InfiniBand protocols to transparently support RDMA channels over both fabrics with the benefits of full transport-offload and achieve accelerated application performance. Furthermore, it brings value-added features for networking, clustering and storage. This session will provide an overview on multi-protocol architecture and the Multi-I/O benefits with the simplicity it brings to HPC clusters.

Managing Network and Storage Infrastructure as an Application Resource
Yaron Haviv (Voltaire)

The expansion of cluster technologies brings benefits such as greater capacity, lower costs and the support of more diverse applications. This trend has also brought additional infrastructure requirements and associated challenges. Next generation HPC architectures will address those complexities by treating infrastructure as a managed resource that can be centrally scheduled and provisioned to provide optimal configuration and resources to meet application requirements. Servers, network topologies and storage elements will be created and torn down to meet application requirements much like compute jobs and processes are scheduled today. This session will examine how clusters can be built around a service oriented infrastructure (SOI) and by leveraging unified fabrics like InfiniBand, network, cluster communication, and storage can be dynamically optimized to meet application requirements. This architecture will be integrated with schedulers and resource management to maximize the value and enable rapid infrastructure deployment, treating hardware as a managed resource pool.

Application of Direct Execution Technology for Analysis of Packets in a High Speed Application Layer Scenario
Kevin Rowett (Mistletoe Technologies)

We describe the implementation of a hardware based parser and direct execution processor tailored for analyzing packets flowing on a LAN. We then describe how this technology provides benefits for analyzing activity on a high performance LAN in a distributed way. Techniques and software have matured in recent times for determining and tracking application layer activity on a LAN (P2P traffic, VoIP, malware propagation). These techniques all depend upon seeing the first 200 bytes of every packet traversing the network. Current techniques require expensive hardware, with custom tuned software, limiting deployment to the core of a network. Using a direct execution technology allows the sensors to be placed at the edge, in large numbers. We show several deployed application scenarios and the benefits.
Networking—Infiniband
Room: A8

High-performance Ethernet and Fibre Channel Connectivity for InfiniBand Clusters
Ariel Cohen (Xsigo Systems)

InfiniBand (IB) provides excellent performance for server-to-server communication within HPC clusters. However, the performance of many HPC applications also depends on server-to-I/O performance. If I/O becomes a bottleneck, the cluster becomes underutilized. I/O is most often provided by non-IB networks: Ethernet and Fibre Channel, so it is critical to be able to bridge the I/O traffic from these non-IB networks to the IB cluster, while maintaining high performance. The Xsigo System provides such a capability. It consists of a chassis with 24 IB ports and 15 I/O slots which can be populated with Ethernet (4x1GE and 1x10GE) NIC and Fibre Channel (2x4Gb/s) HBA modules. IB cluster nodes are provided with virtual NICs and HBAs which provide line-rate connectivity to storage servers and external networks (such as WAN-facing networks, and inter-cluster networks). Each chassis provides 128Gb/s of I/O. This presentation focuses on this novel and cost-effective approach to I/O for IB clusters.

Checklist for Selecting and Deploying Scalable Clusters with InfiniBand Fabrics
Lloyd Dickman (QLogic)

InfiniBand is widely accepted as the interconnect fabric for high-performance cluster computing. Its high-signaling rates, along with implementations that provide high-bandwidth, low communications latencies and high message rates has driven acceptance into many recent cluster deployments. This presentation will provide guidelines for selection and deployment of InfiniBand interconnects based on application and interconnect characteristics, for both host adapters and switching fabrics. We will look at how considerations such as point-to-point microbenchmarks, system microbenchmarks, and application benchmarks contribute to forming a checklist that guides the selection and deployment of the appropriate communications hardware and software components.

Creating and Debugging Low Latency, High Bandwidth InfiniBand Fabrics
Stephen S. Fried (Microway, Inc.)

This talk discusses novel ways to create low latency, high bandwidth InfiniBand Fabrics employing the latest Mellanox ConnectX HCA and InfiniScale III Switches, whose latency is now as low as 1 microsecond. In addition to a discussion of the impact of fabric design on latency, congestion and bandwidth, we will discuss how to debug and control servers connected by InfiniBand Fabrics with Microway’s InfiniScope Real
Time InfiniBand Fabric monitoring tool and "MPI Link Checker" software. Both of these tools work on any brand of InfiniBand switches, servers, and HCAs. One example will show how employing Infiniscope on a system in the field that appeared to be running well, proved it was actually running a factor of three slower than a properly tuned system.

**Wednesday, Nov. 14**

**Data Center Room: A7**

**10:30 a.m. - 12:00 p.m.**

**Innovation for the Data Center of Tomorrow**

*Bill Mannel (SGI)*

The cost of electricity is on the rise, along with the power-density of IT equipment and the cost of its associated infrastructure. Accordingly, today’s IT professionals are faced with an increasingly daunting challenge when planning data centers that are intended to service their organization for 25+ years while housing IT equipment that follows a typical refresh rate of 2 to 5 years. This is balanced against an increasingly important focus on the practices intended to use IT resources in the most efficient way possible. In June of this year, SGI introduced SGI Altix ICE, an innovative new product line designed to address these key issues. This presentation will discuss this new platform from SGI, and specific HPC system design elements intended to reduce complexity and increase the efficiency of large-scale HPC deployments.

---

**Combined Effect of Passive and Low-power Active Equalization: Extend the Life of High-speed Copper Interconnects**

*Eric Gaver (Gore)*

As Double Data Rate (DDR) InfiniBand (5.0Gbps/channel) gains adoption as the preferred interconnect protocol in HPC markets, the maximum reach of copper-based interconnects is limited to approximately 10m, or half the distance required for the largest systems. This reach limitation will be further limited as higher data rates, up to 10Gbps/channel, are introduced over the next few years. Many assume at these higher data rates the only option may be converting to more expensive and less reliable VCSEL-based, parallel optics technology. However, recently introduced to the marketplace is a new low power copper alternative utilizing both passive and active equalization that triples the reach of conventional copper. This new Extended Reach copper alternative offers a solution with the advantages of increased reliability and the lower power consumption of copper, combined with the longer reach and lighter weight advantages of optics.
From the 380V DC Bus to sub-1V Processors: Efficient Power Conversion Solutions
Stephen J. Oliver (V.I Chip)

Without a major architectural review, data centers may consume 100 Billion kWh by 2011. Additionally, inefficient power and cooling techniques may be the downfall of Moore’s law. One efficiency improvement proposal is for the adoption of high voltage DC distribution to enable the AC - 384V stage to be bypassed and the downstream sub-systems (or blades) to be fed directly from the data center distribution buses. This paper represents a follow-on step and proposes high efficiency, high power density power conversion solutions from the HV bus down to processor and low-/medium- and high-power loads. A baseline sub-system is established with a total load of 1320W: 6x processors (1V, 120A), 6x memory (1.5V, 50A) and miscellaneous loads (12V, 12.5A) 380V-12V-1V and 380-48-1V stages are considered using bus converters, synchronous buck converters (VRMs) and factorized power regulators with sine amplitude converters. Efficiency, power density and annual electrical running cost comparisons are presented.

Storage—Analytics and Reliability
Room: A8

The Power of Streaming Analytic Appliances in Supercomputing Environments
John R. Johnson (Lawrence Livermore National Laboratory), Justin Lindsey (Netezza Corporation)

As data volumes continue to grow and the demand for delivering detailed analyses quickly is ever-increasing, organizations are seeking better ways to manage this data and deliver the most powerful insight possible. Lawrence Livermore National Laboratory, is leveraging the streaming analytic architecture that Netezza provides to power their advanced data analyses. By offloading computationally-intensive tasks to the Netezza Performance Server (NPS) streaming analytic appliance, the ability to churn through terabytes of data quickly is now a reality. During this session John R. Johnson of LLNL and Justin Lindsey, CTO, Netezza Corporation will discuss the background behind these advanced analytics applications and their applicability in supercomputing environments.

Advances in Reliability and Data Integrity for High Performance Storage Solutions
Garth Gibson (Panasas)

The constant and relentless growth in disk drive capacity has altered the landscape of reliability and integrity in large scale High Performance storage systems. Panasas leads
the industry in providing solutions that address today and tomorrow’s challenges with capabilities such as de-clustered RAID that dramatically increases the speed of RAID reconstructions. In this presentation you will learn from the co-author of the RAID standard about the latest developments in high-performance RAID. These new developments will increase overall system reliability and data integrity allowing users to take full advantage of increasing disk capacity without any additional burden of risk.

Cyber Crimes Center: A Distributed Heterogeneous Mass Storage Federation for Digital Forensics
Constantin Scheder (Nirvana)

The collaborative digital forensic data grid at the Cyber Crimes Center (C3), a program within the Department of Homeland Security’s ICE division, is designed to provide system security, access control, and data integrity throughout a distributed environment. The system improves field collaboration; enables more efficient cross-site investigations and search; and facilitates case assignment. SRB manages the movement and management of data in the federation throughout its lifecycle, automatically, so users do not have to. SRB helps agents discover, present, share and manage high-value data through familiar interfaces. SRB’s Metadata Catalog (MCAT) provides organization-wide views to authorized C3 agents. The MCAT captures meaningful metadata throughout the lifecycle of evidence which is used to understand data and make decisions about related case handling. In this session Constantin Scheder, Nirvana’s Chief Architect, will demonstrate the benefits of the C3 data federation, and discuss how complex organizations can maximize their existing investments in mass storage.

Storage and HPC—Innovations
Room: A7

Lustre for the Rest of Us: A Fast, Fully Redundant, Fully Configured Lustre Appliance
Larry Genovesi (Terascala)

Lustre has proven itself through many installations as a viable, high performance clustered file system. However, until now, deploying a fully redundant, cost effective, fast installation has required significant investment of time, knowledge and money by end users. Terascala will discuss their unique implementation of a Lustre-based storage appliance that delivers high throughput and scalability in a cost effective blade based solution.

A Paradigm Shift in the Storage Industry: From Monolithic Boxes to Clustered Architectures
Sam Grocott (Isilon Systems)

Tony Asaro, Senior Analyst, Enterprise Strategy Group, said it best in October 2005, when he stated, "Clustered Storage is becoming pervasive and is a major paradigm shift from previous generations of storage
products, much like when CDs made records obsolete." Sam Grocott, Senior Director of Product Management for Isilon Systems and a key visionary behind the development of Isilon’s OneFS(TM) operating system software and product platform, will detail this ongoing paradigm shift in the storage industry from monolithic, refrigerator-type boxes to clustered architectures that parallels the current shift on the server side to clustered computing. The presentation will provide details on why this paradigm shift is occurring now and what storage industry professionals and customers need to know to keep up with - and get ahead of - this industry-transforming trend.

Intel HPC Innovations for the Mainstream and High End
William Magro (Intel Corporation)

Enterprise compute centers are undergoing a transformation, as commercial customers increasingly seek the enhanced capabilities and competitive advantage offered by today's rapidly advancing high performance computing (HPC) offerings. New HPC product designs enable businesses to improve product designs, make better decisions, and shorten time to market, all while lowering costs. But, as HPC moves deeper into the enterprise, mainstream users expect not only the power of HPC, but also the simplicity of packaged systems and a host of compatible applications. At the same time, the appetites of high-end HPC users for processing power now reaches well into the petascale range, driving innovation in both system and component designs. Intel is investing broadly to meet these wide-ranging expectations, and it will share its vision and plans to serve the broad and growing HPC user community by enabling continued performance and usability advancements in all segments.

Multicore Technologies
Room: A8

Data Streaming Compilers for Multi-core CPUs
Michael Wolfe (The Portland Group)

Current multi-core processors increase instruction and computational bandwidth linearly, but don’t increase data bandwidth from main memory. Programming models that use multiple cores like an SMP fail to take this limit into account and fail to deliver the promised performance. Given clock rate constraints, increased performance must come from better core architecture and more efficient use of multiple cores. PGI is developing a multi-core compiler strategy that leverages compute intensity and data streaming. Treating core-local cache as local memory, and inserting inter-core synchronization, we optimize memory bandwidth usage and minimize cache conflicts. Each core loads data into local cache, then hands off control of the memory interface. The core computes at full speed on cache-local data, then stores results the next time it receives control of the memory interface. Often presented as a new model for computing, the Streaming Model, it’s really an incremental modification to classical compiler vectorization.
A New Application Debugging Framework for the Multi-core Age

Chris Gottbrath (TotalView Technologies)

Programmers that are developing applications powered by multi-core and multi-processor technology are finding that they need new tools and techniques in order to realize the full potential of multi-core design. A well-defined, logical, and coherent support structure for organizing and managing the complex debugging process will enable today’s developers to more efficiently write or adapt software to the parallel computing architecture. A comprehensive approach to multi-core debugging includes the following five key technologies: Source Code Debugging; Memory Debugging; Performance Analysis; Data-Centric Debugging; and Active Web Debugging. This presentation will address the implications for application developers as the market evolves to multi-core CPUs. It will include a look at why a new conceptual approach—a framework—for multi-core application debugging is needed, and discuss the elements of such a framework, along with the benefits of taking advantage of this approach.

Multicore: All Eyes on Bottlenecks

Erik Hagersten (Acumem)

The move to multicore architectures implies a smaller fraction of the chip area devoted to caches and that this scarce resource will be fought for by many threads. Limited cache capacity per thread is expected to hamper the potential performance of multicore. Still, more than half the data brought into the cache by many of today’s applications is not used before it gets evicted or invalidated. This more than doubles the frequency of cache misses and indirectly causes more excess misses for other areas of the application because the cache is hogged. This session will unveil Acumem’s technology, which identifies wasteful cache usage of applications. About 20 different types of performance issues related to multithreaded execution and cache usage are identified and fixes suggested at a level of detail allowing a novice programmer to perform performance optimization requiring extreme performance experts today, creating the Virtual Performance Expert.

Distributed Computation

Room: A7

3:30 p.m. - 5:00 p.m.

Cost-effective Grid with EnginFrame

Andrea Rodolico, Fabrizio Magugliani (NICE Srl)

EnginFrame is the industry leading Grid Portal enabling user-friendly and application-oriented HPC job submission, control and monitoring system. It delivers sophisticated data management for all stages of job lifetime and is integrated with all relevant Grid workload management systems. Customers like Airbus, Audi, BMW, Bridgestone, British Gas, Delphi, Ferrari, FIAT, Procter & Gamble, Raytheon,
Schlumberger, Statoil, Toyota, TOTAL, TRW, STMicro, today leverage EnginFrame to interface their engineers, partners and customers with the Grid. EnginFrame can cut costs and improve users’ productivity right across the board from painless roll-out and reduced training to improved use of the Grid / applications. EnginFrame evolves with IT infrastructure and aligns with company’s IT security policies delivering comprehensive transparent data management capabilities and data monitoring, user-friendly job submission and monitoring. In this session participants will hear how EnginFrame can maximize the exploitation of their IT infrastructure through simple, finely tuned, user-friendly and secure access to HPC resources.


deploying a geographically distributed infrastructure to enable global data sharing and continuity of operations

Wayne Karpo (YottaYotta)

Both Government and multi-site organizations must find ways to leverage distributed IT infrastructure, share information efficiently, and provide continuity of operations (COOP) despite system failures or even catastrophic regional disasters. This presentation will outline three use cases in which the Department of Defense (DoD), a Tier 1 Enterprise organization, and multiple agency groups are implementing a geographically distributed infrastructure in order to ensure uninterrupted business operations and provide federated access to data. Attendees will be learn from use cases providing an outline of the storage and network environment, application requirements, and Recovery Time Objectives (RTO) and Recovery Point Objectives (RPO). Attendees will be able to identify key architectural considerations and infrastructure requirements in order to achieve multi-site load balancing, continuous data availability, and COOP.

High Performance Java and .NET Applications for Financial and Other Compute-Intensive Industries

Edward Stewart (Visual Numerics, Inc.)

In many industries today, including financial services, there is a trend towards using Java and .NET for developing computational applications. Both offer an efficient development, testing and deployment platform. At the same time, developers continue to look for new ways to leverage High Performance Computing (HPC) for their compute-intensive applications, especially those with large datasets and/or complex processing requirements. While both of these are popular trends, Java and .NET have historically not been considered as complementary development frameworks for HPC, given the application performance requirements and the perception of lack of speed with Java and .NET. This presentation will challenge the perception of Java and .NET being "too slow" to apply to HPC. We will present several financial-based applications written in Java and the .NET framework for distributed computing, grid computing and shared memory architectures and how, by using optimal computational algorithms, one can significantly improve performance results.
FPGAs—Applications
Room: A8

**Accelerating Key Recovery and Mapping of MD Simulations using FPGAs**

*David Hulton (Pico Computing, Inc.)*

Encryption is simply the act of obfuscating something to the point that it would take too much time or money for an attacker to recover it. Many algorithms have time after time failed due to Moore’s law or large budgets or resources (e.g. distributed.net). There have been many articles published on cracking crypto using specialized hardware, but many were never fully regarded as being practical attacks. Slowly FPGAs (Field Programmable Gate Arrays) have become affordable to consumers and advanced enough to implement some of the conventional software attacks extremely efficiently in hardware. The result is performance up to hundreds of times faster than a modern PC. This presentation will provide a walk through on our experience with using FPGAs for performing key recovery on a variety of crypto ciphers, including Lanman, WPA, WEP, FileVault, WinZip, and Bluetooth. We will also provide live demonstrations of our open source tools from www.openciphers.org.

**60x Faster NCBI BLAST on the Mitrion Virtual Processor**

*Stefan Möhl (Mitrionics), Henrik Abelson (Mitrionics), Göran Sandberg (Mitrionics)*

The Mitrion Virtual Processor is a massively parallel, configurable soft-core processor design that runs in an FPGA. Fine-grain parallel programs running in the Mitrion Virtual Processor achieve the performance benefits of FPGAs without requiring hardware design. In general the Mitrion Virtual Processor achieves 10-30 times higher performance than traditional high-performance processors like Intel Xeon or AMD Opteron. The BLAST algorithm is one of the most important tools within bioinformatics. It allows comparisons between huge databases of genomic information. NCBI BLAST is the industry standard implementation of the algorithm and widely used throughout the world. We will present an open source port of NCBI BLAST for the Mitrion Virtual Processor. The source code is freely available for download on sourceforge.net. BLAST running on the MVP has achieved 60x higher performance than standard BLAST on an Opteron 2.8GHz. The porting process, implementation and performance benchmarking will be presented.
Thursday, Nov. 15

Clusters
Room: A7

10:30 a.m. - 12:00 p.m.

Portable Clusters
Dominic Daninger (Nor-tech)

Organizations within the Supercomputing industry are now requiring high-performance computing that may be easily moved to various locations and have the energy efficiency to operate within standard AC power parameters. REASON together with its parent company, Nor-tech have developed a line of Portable Clusters that are highly ruggedized and extremely mobile. The portable HPC clusters are lightweight and available in various portable rackmount cases including military shock mounted cases that adhere to MIL-SPEC standards. The mobile clusters run cool, save energy and are able to operate up to 96 processor cores on just two 110VAC 20-amp circuits.

Maximum Performance Cluster Design
John L. Gustafson (ClearSpeed Technology)

The typical primary goal in cluster design now is to maximize performance within the constraints of space and power. This contrasts with the primary goal of a decade ago, when the constraint was the budget for parts purchase. We therefore examine how to create a cluster with the highest possible performance using current technology. Specific choices include a pure x86 design versus one with accelerators from ClearSpeed, Nvidia, or IBM. We use mixed integer linear programming arguments to show that the optimum performance is currently achieved with x86 nodes enhanced with ClearSpeed (Advance) accelerators, and that clusters of x86 nodes enhanced with Nvidia (Tesla) or IBM (Cell) accelerators are actually slower than a system with nothing but x86 nodes alone. The reason is that power dissipation becomes the primary limiter.

Measured Performance of CIFS (SMB) as a Global File System for a Moderate Sized Cluster
Frank Chism (Microsoft Corporation)

This presentation will show measured results for both single stream and multiple stream I/O performance for an MPI application run on a 60 node cluster that has multiple CIFS servers with multiple shares and multiple GigE NICs. Aggregate bandwidth appears to be limited by the performance of the local RAID on the servers.

FPGA Programming
Room: A8

Accelerating HPC Applications using C-to-FPGA Techniques
David Pellerin (Impulse Accelerated Technologies)

Software application developers considering FPGAs for high-performance computing are faced with numerous barriers to understand-
ing, due to the radically different architectures of FPGAs when compared to more traditional processors. This presentation describes how recent generations of FPGA-based computing platforms, coupled with advances in compiler tools and programming models, can greatly simplify the FPGA programming process. The presentation includes specific tips on how high levels of algorithm acceleration can be achieved using iterative C-to-hardware design and optimization methods. Real-world examples and acceleration results are presented, and the concepts of system-level and process-level parallelism and hardware/software partitioning and pipelining are explored.

FPGA for High Performance Computing Applications
Shyam Sunder Uma Chander (GiDEL, Inc.)

The objective is to show FPGA’s superior performance compared to high-end processors. A hardware accelerator was developed using Gidel’s reconfigurable hardware to decentralize and to speed up the execution of the algebraic operations of the BLAS library. The idea is to delegate the execution of these operations to the FPGA board; and as it can perform them much more efficiently than a general-purpose processor, the global performance of the system is improved using GiDEL’s hardware reconfigurable systems. Currently, the maximal performances of a processor Intel P4 Dual Core 3 GHz are 3.3 Gflops/s for the operation of matrix multiplication (matrices 1000x1000). This result is obtained by implementing the optimized algorithms of the ATLAS library. For the same operation, the developed prototype is 54% faster with a measured performance of 5.08 Gflops/s.

Programming Tools
Room: A7
1:30 p.m. - 3:00 p.m.

New Technologies in Mathematica
Schoeller Porter (Wolfram Research, Inc.)

gridMathematica’s rich web of algorithms provides straightforward tools for developing and debugging parallel applications and analyzing scientific data. With recent advances in Mathematica’s HPC architecture, a host of new parallel approaches are now possible, including new levels of automation in algorithmic computation, interactive manipulation, and dynamic presentation. This talk will use practical, real-world simulations to demonstrate these new technologies and showcase how they can be used in research for 2D/3D modeling, prototyping, data handling, and deployment on any scale.

Supercomputing Engine for Mathematica
Dean E. Dauger (Dauger Research, Inc.)

The new Supercomputing Engine for Mathematica enables Wolfram Research’s Mathematica to be combined with the programming paradigm of today’s supercomputers. Closely following the industry-standard Message-Passing Interface (MPI), this new toolkit creates a standard way for every
Mathematica kernel in the cluster to communicate with each other directly while performing computations. In contrast to typical grid implementations of Mathematica that are solely master-slave, this solution instead has all kernels communicate with each other directly and collectively the way modern supercomputers do. Besides creating an "all-to-all" communication topology using Mathematica kernels on a cluster, the new technology supports low-level and collective MPI calls and a suite of high-level calls, all within the Mathematica computing environment. We present the technology's API, structure, and supporting technologies with examples of its potential application. http://daugerresearch.com/pooch/mathematica/

**Networking—Performance Room: A8**

**iWARP and Beyond: Performance Enhancements to Ethernet**

*Charles R. Maule (NetEffect)*

The IETF recently approved a set of extensions to the Ethernet standard that provides significant performance enhancements collectively referred to as the iWARP Extensions. This talk will answer the questions: What factors drove the IETF to create extensions to the Ethernet standard?; What is iWARP and why is it important to the future of high-performance computing environments?; How will these extensions help me today and into the future? This talk will also discuss how these enhancements can be effectively applied to non-iWARP situations, such as the acceleration of standard TCP and UDP communications over Ethernet and why they are important for a wide range of high-performance and data center applications.
Scaling I/O Performance with Storage Aggregation Gateways

Nikhil Kelshikar (Cisco Systems, Inc)

The size and performance of highly distributed compute clusters has been growing, resulting from the high bandwidth and low latency delivered by InfiniBand switching. While InfiniBand has optimized interprocess communication between the compute servers, customers are not using this technology to its full potential. Customers are either using local disk on the server for their storage needs, or deploying NAS and Parallel File Systems with 1 Gbps Ethernet technologies, as a matter of convenience, based on the availability of these interfaces as located on the server motherboards. This session will present solutions for scaling NAS and Parallel File System access for HPC clusters using InfiniBand as a primary IPC fabric with the use of an Ethernet and Fibre Channel Gateway. We will discuss the impact the proposed architecture has on application scalability.

Testing Application Performance over a WAN: Using Network Emulation to Find the Weakest Link

Kevin Przybocki (Anue Systems, Inc.)

New applications... Legacy infrastructure... More complexity... Tighter schedules... Deploying mission critical enterprise applications over a WAN is more difficult than ever! These challenges combined with the high-profile nature of enterprise applications increases the pressure to smoothly deploy a system that meets or exceeds users’ expectat-
For seven days, the Reno-Sparks Convention Center will be home to one of the most powerful networks in the world—SCinet. SCinet is the collection of high-performance networks built to support the SC conference. Built each year, SCinet brings to life a highly sophisticated and extreme networking infrastructure that can support the revolutionary applications and network experiments that have become the trademark of the SC conference. SCinet serves as the platform for show exhibitors to demonstrate the advanced computing resources from their home institutions and elsewhere by supporting supercomputing and grid computing applications. SCinet features a high-performance, production-quality network, OpenFabrics InfiniBand and Low Latency Ethernet Networks, and an extremely high-performance experimental network, Xnet.

Volunteers from educational institutions, high-performance computing centers, network equipment vendors, national laboratories, research institutions, research networks and telecommunication carriers work together to design and deliver the SCinet networks. Industry vendors and carriers donate much of the equipment and services needed to build the LAN and WAN infrastructure. Planning begins more than a year in advance of each SC conference and culminates with a high-intensity installation just seven days before the conference begins.

SCinet is providing direct wide area connectivity to many national and worldwide networks through peering relationships with principal networks. Aggregate WAN transport delivered to the industry and research exhibitors is expected to exceed 263 Gigabits/second (263 Gb/s) for SC07.

OpenFabrics and RDMA Services

In 2007, SCinet will offer 10 Gb/s, low latency 20Gb/s InfiniBand (IB), and low latency 10 Gigabit Ethernet (GbE) infrastructures, which will include hardware from many of leading vendors in the IB and the 10GbE industry together with the OpenFabrics software and supporting services. This combination will provide powerful infrastructures for exhibitors within the convention center between booths to experience the advantages InfiniBand and low latency 10 GbE deliver for clustered computing, server-to-server processing, visualization, and file system access to native IB and 10 GbE attached storage. SCinet will also offer limited longer distance point to point IB and routable 10 GbE links on the Wide Area Network (WAN) across the continental US or overseas to support exhibitor requirements.

The InfiniBand architecture is a very high performance, low latency interconnect technology based on an industry-standard approach to Remote Direct Memory Access (RDMA). An InfiniBand fabric is built from
hardware and software that are configured, monitored and operated to deliver a variety of services to users and applications. Characteristics of the technology that differentiate it from comparative interconnects such as Ethernet include: end-to-end reliable delivery, scalable bandwidths from 10 to 60 Gbps available today moving to 120 Gbps in the near future, extremely low latency between devices (less than 1.5 microseconds demonstrated), greatly reduced server CPU utilization for protocol processing, scalability without performance degradation, and efficient I/O channel architecture for network and storage virtualizations.

The OpenFabrics Alliance is an international organization comprised of industry, academic and research groups, along with hardware and software vendors, that have developed a unified core of open source software stacks leveraging RDMA architectures for both the Linux and Windows operating systems over both InfiniBand and Ethernet. The core OpenFabrics software supports all the well known standard upper layer protocols such as MPI, IP, SDP, NFS, SRP, iSER, and RDS.

SCinet encourages exhibitors who are suppliers, users and researchers of these technologies to bring their hardware, software and applications (including switches, host channel adapters, storage systems, gateways to Ethernet and Fibre Channel, WAN extension interfaces) to use this software as a common base for demonstration both within and between their booths and over the WAN at SC07.

SCinet Xnet
SCinet’s Xnet (eXtreme net) provides a venue to showcase emerging, often pre-commercial or precompetitive, developmental networking technologies, protocols, and experimental networking applications. The SCinet Exhibit floor network has evolved into a robust, high-performance, production-quality network that exhibitors and attendees depend on for reliable local area, wide area and commodity network service. Consequently, it has become increasingly difficult for SCinet to showcase cutting-edge and potentially fragile technologies. Simultaneously, OEMs have at times been reticent about showcasing bleeding-edge hardware in SCinet, as it became a mission critical, production network.

Xnet provides the solution to this dichotomy by providing a venue which is by definition leading edge, pre-standard and in which fragility is understood. Xnet thus provides its participants (from both the commercial and research communities) an opportunity to showcase emerging, prototype or experimental network gear or capabilities, prior to their general commercial availability in a fault tolerant (forgiving) environment. Participants in SCinet’s Xnet environment gain insight from the world’s most advanced computing and computer networking resources.

Network Performance Monitoring
The SCinet architecture incorporates a number of features that support network monitoring. Monitoring will be used both to watch the internal network for operational purposes and to characterize the high-per-
formance network applications that traverse SCinet, in particular for the Bandwidth Challenge.

Fine-grained monitoring on the network is particularly challenging because of the large number of high-speed links and packet flows that occur on this specialized network. Utilization and errors for all external links and all major SCinet internal links will be monitored for operational purposes.

Active techniques will be used to monitor reachability over the external links and latency to key sites. One-way delay testing will be done with OWAMP technology developed by Internet2 and partially funded by NSF. Throughput tests will use BWCTL, developed by Internet2, as well as Iperf, developed by NCSA and the University of Illinois at Urbana-Champaign as part of the NSF-funded Network Laboratory for Applied Network Research (NLANR) Distributed Application Support Team.

The passive and active monitoring data will be presented through graphical and web-services interfaces that will be published at http://monitor.sc07.org/. Specifically, Internet2, in conjunction with SCinet, will provide a “weather map” showing current utilization on all SCinet external links, based on the technology used for the Internet2 Network NOC weather map, and developed by the Internet2 Network NOC at Indiana University. Additionally, most of this data will be made available using perfSONAR web services interfaces (http://www.perfsonar.net/) developed by Internet2, the University of Delaware, and other project partners. These web-services interfaces will allow remote visualization and analysis by a variety of network clients.

The Network Diagnostic Tool (NDT) is a Web-based tool (available at http://ndt.sc07.org/ during the show) that allows SC07 users to self-test their end systems and the “last mile” network infrastructure. It detects common infrastructure problems (e.g., duplex mismatch condition) and common performance problems (e.g., incorrect TCP buffer size). The NDT is currently being developed by Internet2, with additional funding from Cisco, Argonne National Laboratory, and the National Library of Medicine.

Flow data (e.g., sFlow, NetFlow, cflow) will be collected from routers and switches. This data will be analyzed and visualized with software from InMon Corporation. This will provide detailed real-time information showing which users and applications are consuming the bandwidth. The information will be used for assessing the Bandwidth Challenge and network operations.

SCinet Network Security Policy

The design characteristics that define the SCinet production networks include high bandwidth, low latency, resiliency, and scalability. SCinet peers with the Internet and with agency and national wide area networks through a series of very high-speed connections.

To maximize performance across these interfaces, there are no firewalls. In this regard, the SCinet network is a logical, albeit temporary, extension of the open Internet.
Exhibitors and attendees are reminded that, in this potentially hostile environment, network security is a collective responsibility. Exhibitors who use insecure communication methods are exposing their networks and systems to compromise. The use of insecure applications including TELNET and FTP is strongly discouraged. These applications are subject to compromise because they send passwords to remote hosts in human-readable, clear text format. Attendees are strongly encouraged to protect their sessions through a mechanism such as Secure Shell (SSH) where all communication is encrypted. SSH implementations are available for little or no cost and are straightforward to install and use.

Each attendee is responsible for ensuring that their communications sessions are protected in accordance with their security requirements. All IEEE 802.11a, 802.11b and 802.11g wireless networks, including those provided by SCinet, are vulnerable by their very nature. The ease of use that makes them attractive is the same feature that is most easily exploited. Wireless networks are open to unauthorized monitoring or snooping by anyone within range of an access point. SCinet will monitor traffic on most external network connections as part of routine network performance monitoring activities. In addition, SCinet has a restricted capability to monitor exhibit floor, wireless network and external network traffic for evidence of security-related activity including compromise or abuse. However, by no means should this coverage be considered a substitute for safe security practices. Please do your part by being cognizant of network security risks and protecting your systems and sessions.

**SCinet Wireless Network Service Policy**

SCinet will deploy IEEE 802.11a, 802.11b and 802.11g wireless networks within the Reno-Sparks Convention Center (RSCC). These wireless networks are part of the production SCinet network, providing access to the Internet and many other national and agency networks. The wireless network will be provided on the exhibit floor, in the Education Program areas, the ballroom and meeting rooms, and in many common areas within the RSCC.

SCinet provides the wireless networks for use by all exhibitors and attendees at no charge. Known wireless network limitations, such as areas of reduced signal strength, limited client capacity or other coverage difficulties may be experienced within certain areas with the center. Please watch for additional signage at appropriate locations throughout the RSCC. Also, as a courtesy to other users, please do not download large files via the wireless network.

Network settings including IP and DNS addresses for wireless clients are provided by SCinet via DHCP. Laptops and other wireless devices configured to request network configuration information via DHCP receive this information automatically upon entering the SCinet wireless coverage area.

SCinet will monitor the health of the wireless networks and maintain this information for exhibitors and attendees. The wireless networks are governed by this policy posted on the SC07 conference Web site. SCinet wants you to have a successful, pleasant experience at SC07. This should include the ability to sit down with your wireless-
equipped laptop or PDA and check e-mail or surf the Web from anywhere in the wireless coverage areas. Please help us achieve this goal by not operating equipment that will interfere with other users. In order to provide the most robust wireless service possible, SCinet must control the entire 2.4GHz and 5.2GHz ISM bands (2.412GHz-2.462GHz) and (5.15GHz to 5.35GHz) within the RSCC where SC07 events are taking place. This has important implications for both exhibitors and attendees:

- Exhibitors and attendees may not operate their own IEEE 802.11 (a,b,g, or other standard) wireless Ethernet access points anywhere within the convention center, including within their own booth.
- Wireless clients may not operate in ad-hoc or peer-to-peer mode due to the potential for interference with other wireless clients.
- Exhibitors and attendees may not operate 2.4GHz or 5.2GHz cordless phones.
- Exhibitors and attendees may not operate 2.4GHz wireless video or security cameras, or any other equipment transmitting in the 2.4GHz or 5.2GHz spectrum
- SCinet reserves the right to disconnect any equipment that interferes with the SCinet wireless networks.

Remember that the SCinet wireless network is a “best effort” network. If you are running demonstrations in your booth that require high availability network access, we advise exhibitors to order a physical network connection.

**SCinet Service Level Policy**

SCinet, provides commodity Internet, research and experimental networks for use by the exhibitors and attendees. While every practical effort shall be made to provide stable and reliable network service on each network, there is no explicit service level agreement for any SCinet network, nor are there any remedies available in the event that network services are lost.

To examine the full SCinet Service Level Policy visit: [http://sc07.supercomputing.org/?pg=servicelevel.html](http://sc07.supercomputing.org/?pg=servicelevel.html)

**SCinet Partner Acknowledgments**

SCinet’s success is due in large part to the support of many vendors and organizations which provide both equipment and engineering expertise. For SC07, SCinet’s premier partners are:

- Ames Laboratory at Iowa State University
- Argonne National Laboratory
- Arizona State University
- CA Labs
- Carnegie Mellon University
- CENIC
- Ciena
- Cisco Systems
- Computer Sciences Corporation
- EMCORE
- ESnet
- Electronic Visualization Laboratory at UIC
- Florida LambdaRail
- Force10 Networks
- Foundry Networks
- Fujitsu
- Fulcrum Microsystems
<table>
<thead>
<tr>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indiana University</td>
</tr>
<tr>
<td>Infinera</td>
</tr>
<tr>
<td>InMon</td>
</tr>
<tr>
<td>Intel</td>
</tr>
<tr>
<td>Internet2</td>
</tr>
<tr>
<td>Juniper Networks</td>
</tr>
<tr>
<td>Lawrence Berkeley National Laboratory</td>
</tr>
<tr>
<td>Level 3</td>
</tr>
<tr>
<td>Los Alamos National Laboratory</td>
</tr>
<tr>
<td>Mellanox</td>
</tr>
<tr>
<td>Mid-Atlantic Crossroads/MAX Gigapop</td>
</tr>
<tr>
<td>MRV</td>
</tr>
<tr>
<td>National Center for Supercomputing Applications</td>
</tr>
<tr>
<td>National Energy Research Scientific Computing Center</td>
</tr>
<tr>
<td>National LambdaRail</td>
</tr>
<tr>
<td>NetEffect</td>
</tr>
<tr>
<td>NetOptics</td>
</tr>
<tr>
<td>Nevada System of Higher Education</td>
</tr>
<tr>
<td>Nortel Networks</td>
</tr>
<tr>
<td>Oak Ridge National Laboratory</td>
</tr>
<tr>
<td>Obsidian Research</td>
</tr>
<tr>
<td>OpenFabrics Alliance</td>
</tr>
<tr>
<td>Pacific Northwest Gigapop/Pacific Wave</td>
</tr>
<tr>
<td>Pacific Northwest National Laboratory</td>
</tr>
<tr>
<td>Purdue University</td>
</tr>
<tr>
<td>QLogic</td>
</tr>
<tr>
<td>Qwest Communications</td>
</tr>
<tr>
<td>Sandia National Laboratories</td>
</tr>
<tr>
<td>San Diego Supercomputer Center</td>
</tr>
<tr>
<td>SARA</td>
</tr>
<tr>
<td>Spirent Communications</td>
</tr>
<tr>
<td>Starlight/Translight</td>
</tr>
<tr>
<td>Trapeze Networks</td>
</tr>
<tr>
<td>United States Air Force Aeronautical Systems Center</td>
</tr>
<tr>
<td>United States Army Research Laboratory</td>
</tr>
<tr>
<td>United States Army Space and Missile Defense Command</td>
</tr>
<tr>
<td>University of Amsterdam</td>
</tr>
<tr>
<td>University of California at San Francisco</td>
</tr>
<tr>
<td>University of Florida</td>
</tr>
<tr>
<td>University of Mannheim</td>
</tr>
<tr>
<td>University of Wisconsin</td>
</tr>
<tr>
<td>Woven Systems</td>
</tr>
</tbody>
</table>
SCinet’s Supporting Organizations
SCinet’s Supporting Organizations

NCSA
NetOptics
Nortel
Oak Ridge National Laboratory
Pacific Northwest National Laboratory
Purdue University
Sandia National Laboratories
sara
SC07
SCinet
SDSC
StarLight
TransLight
Trapeze
Universiteit van Amsterdam
University of Florida
Woven Systems
Planning for each SC conference begins several years in advance and actually kicks into gear two years before the conference opens. The success of SC07 is a reflection of dedicated efforts by hundreds of volunteers. Without the support of the following people, their organizations and their funding sources, SC07 would not be possible. Thank you.

**SC07 COMMITTEE MEMBERS**

**Conference Committee**

**Becky Verastegui**, General Chair
Oak Ridge National Laboratory

**Linda Duncan**, Administrative Assistant
Oak Ridge National Laboratory

**David Cooper**, Vice General Chair
Lawrence Livermore National Laboratory

**Barry V. Hess**, Executive Director
Sandia National Laboratories

**Patricia J. Teller**, Deputy General Chair
University of Texas - El Paso

**Jeff Huskamp**, Local/Regional Participation Liaison
University of Maryland

**Dona Crawford**, Keynote Speaker Chair
Lawrence Livermore National Laboratory

**Society Liaisons**

**Jessica Fried**, ACM

**Donna Cappo**, ACM

**Anne Marie Kelly**, IEEE Computer Society

**Technical Program**

**Harvey J. Wasserman**, Chair
NERSC/Lawrence Berkeley National Laboratory

**Adolfy Hoisie**, Deputy Chair
Los Alamos National Laboratory

**Ricky A. Kendall**, Papers Co-Chair
Oak Ridge National Laboratory

**Josep Torrellas**, Papers Co-Chair
University of Illinois

**Jack Dongarra**, Tutorials Co-Chair
University of Tennessee

**Bruce Loftis**, Tutorials Co-Chair
Purdue University

**Mary Hall**, Workshops Co-Chair
University of Southern California
Acknowledgments

Julia Mullen, Workshops Co-Chair
MIT Lincoln Laboratory

José Moreira, Panels Chair
IBM T. J. Watson Research Center

Tamara K. Grimmett, Posters Co-Chair
Idaho National Laboratory

Charles Koelbel, Posters Co-Chair
Rice University

Fred Johnson, Invited Speakers Chair
US Department of Energy

Stanley Ahalt, Masterworks Co-Chair
Ohio Supercomputer Center

John Grosh, Masterworks Co-Chair
Lawrence Livermore National Laboratory

Steven F. Ashby, Awards Chair
Lawrence Livermore National Laboratory

Jeffrey Vetter, Disruptive Technologies Chair
Oak Ridge National Laboratory and Georgia Institute of Technology

Jeffrey K. Hollingsworth, Doctoral Research Showcase Chair
University of Maryland

Karen Karavanic,
Challenges Coordination Chair
Portland State University

Nancy R. Wilkins-Diehr,
Birds-of-a-Feather Chair
San Diego Supercomputer Center

Wilfred Pinfold, Technical Program Liaison
Intel Corporation

Linklings LLC, Submissions & Infrastructure Database Contractor

Exhibits

Cherri M. Pancake, Chair
NACSE/Oregon State University

Jeffery A. Kuehn, Industry Exhibits Chair
Oak Ridge National Laboratory

Christy Adkinson, Research Exhibits Chair
Cray, Inc.

David Cooper, Exhibitor Contact and Liaison
Lawrence Livermore National Laboratory

L. Eric Greenwade, Exhibitor Forum Chair
Lawrence Livermore National Laboratory

Barbara Horner-Miller, Exhibitor Forum Committee
Arctic Region Supercomputing Center

John R. Sopka, Exhibitor Forum Committee
High Performance System Software

Hall-Erickson, Exhibits Management Contractor

Freeman Companies, Logistics and Decorating Contractor
Acknowledgments

Communications

Kathryn Kelley, Chair
Ohio Supercomputer Center

Vivian Benton, Printed Program Coordinator
Pittsburgh Supercomputing Center

Karen Green, Media Relations Coordinator
Renaissance Computing Institute

Jon Bashor, Media Relations
Lawrence Berkeley National Laboratory

Betsy Riley, Media Room Coordinator
Oak Ridge National Laboratory

Faith Singer-Villalobos, Newsletter Coordinator, Texas Advanced Computing Center

John W. Cobb, Proceedings Chair
Oak Ridge National Laboratory

Kristen Meyer Sunde, Press Releases Coordinator
Louisiana State University

David Hachigian, Press Releases Coordinator
Pacific Northwest National Laboratory

Kay Hunt, Web Content Editor
Purdue University

Wilfred Pinfold, Technical Program Liaison
Intel Corporation

Carlton Bruett Design, Graphics and Web Design Contractor

Moving Tributes, Video and Opening Day Production

Infrastructure

Eric Sills, Co-Chair
North Carolina State University

James W. Ferguson, Co-Chair
National Center for Supercomputing Applications

Barbara Fossum, Housing Chair
Purdue University

Janet Brown, Space Chair
Pittsburgh Supercomputing Center

Gary New, Electrical Chair
National Center for Atmospheric Research

Matt Link, Signage Co-Chair
Indiana University

Mike Madero, Signage Co-Chair
Avetec

Gina Morello, Internet Access Chair
NASA

Tiki L. Suarez, Student Volunteers Co-Chair
Florida A&M University

Laura Arns, Student Volunteers Co-Chair
Purdue University

Jeanine Cook, Deputy Student Volunteer Co-Chair
New Mexico State University
Lorie Liebrock, Deputy Student Volunteer Co-Chair
New Mexico Institute of Mining and Technology

Trish Damkroger, A/V and PC Chair
Lawrence Livermore National Laboratory

Matthew Grove, Deputy Electrical Chair
University of Reading

Jim Costa, Infrastructure Staff
Lawrence Livermore National Laboratory

Linda Duncan, Conference Office Coordinator
Oak Ridge National Laboratory

Cecilia G. Aragon, Conference Office Coordinator
University of New Mexico

Sheryl Hess, Conference Office Coordinator
Sandia National Laboratories

G.M. "Zak" Kozak, Security Chair
Ohio Supercomputer Center

Trey Breckenridge, Deputy Security Chair
Mississippi State University

Jeff Graham, Security Staff
Aeronautical Systems Center

Tim Yeager, Security Staff
Aeronautical Systems Center

Barbara A. Kucera, Photographer
University of Kentucky

Conference Management, Convention Center Catering Contractor

Northstar Event Management, Inc., Local Arrangements Contractor

AV Concepts, AV Contractor

Education

Scott Lathrop, Chair
University of Chicago/Argonne National Laboratory

Robert Panoff, STEM Curriculum Content
Shodor Education Foundation

Rubin H. Landau, STEM Curriculum Content
Oregon State University

Tom Murphy, Parallel/Grid Computing
Contra Costa College

Paul Gray, Parallel/Grid Computing
University of Northern Iowa

Charles Peck, Parallel/Grid Computing
Earlham College

Cathie Dager, Fund Raising
Retired (Stanford Linear Accelerator Center)

Kevin M. Hunter, Student Fellows
Earlham College

Alex Lemann, Student Fellows
Earlham College

Kristina N. Wanous, Student Fellows
University of Northern Iowa
Acknowledgments

Jessica Puls, Student Fellows
University of Northern Iowa

John Hurley, Industry Panel
Boeing

Tom Murphy, Student Program
Contra Costa College

Kristina N. Wanous, Student Program/Web Site
University of Northern Iowa

Mechelle De Craene, K-12 Program
MirandaNet Academy

Bonnie L Bracey-Sutton, K-12 Program
George Lucas Education Foundation

Program/Recruitment
Edee Wiziecki, K-12 Program
National Center for Supercomputing Applications

Paul Gray, Awards and Competitions
University of Northern Iowa

Chuck Swanson, Awards and Competitions
Krell Institute

Susan J. Ragan, Awards and Competitions
Maryland Virtual High School

Masakatsu Watanabe, Posters
University of California, Merced

Chuck Swanson, Mentoring
Krell Institute

Ann Redelfs, General Support
Redelfs LLC

Patty Kobel, General Support
National Center for Supercomputing Applications

Frederick C. Harris, Pathways Engineering
University of Nevada

Tina Garnaat, Logistics
Krell Institute

Michelle King, Logistics
Krell Institute

Learning and Physical Challenges
Lynn Rippe, Lawrence Berkeley National Laboratory

Zaida McCunney, Lawrence Berkeley National Laboratory

Mechelle De Craene, MirandaNet Academy

Broader Engagement

Jennifer Teig von Hoffman, Chair
Boston University

Kenneth Washington, Deputy Chair
Lockheed Martin EIS

Tony Baylis, Mentorship Program Lead
Lawrence Livermore National Laboratory

Dawnette Van Dunk, Kiosk and On-ramp Lead
L-3 Communications, Inc.
Acknowledgments

Cindy Sievers, Kiosk/On-ramp Logistics
Los Alamos National Laboratory

Alson Been, Kiosk/On-ramp Content
Bethune-Cookman College

Bruce Loftis, Kiosk/On-ramp Content
Purdue University

Valerie B. Thomas, Lead Liaison to
Education Program
Department of Defense HPC
Modernization Program

Tim Jones, Liaison to Education Program
Department of Defense HPC
Modernization Program

Liz Bechtel, Liaison to Education Program
Department of Defense HPC
Modernization Program

Stephenie McLean, MSI Outreach
Renaissance Computing Institute

Bonnie L. Bracey-Sutton, K-12 Outreach
George Lucas Education Foundation

Jon Bashor, Liaison to Communications
Lawrence Berkeley National Laboratory

Ariella Rebbi, Recording Secretary
Boston University

Committee Members at Large
(Advisory)
Jan Cuny, National Science Foundation

Roscoe C. Giles, Boston University

John Hurley, Boeing

José Muñoz, National Science Foundation

Zaida McCunney, Lawrence Berkeley
National Laboratory

Tiki L. Suarez, Florida A&M University

Finance

Sandra Huskamp, Finance Chair
University of Maryland

Beverly Clayton, Deputy Finance Chair
Pittsburgh Supercomputing Center

Barbara Horner-Miller, Registration
Co-Chair
Arctic Region Supercomputing Center

Kevin Wohlever, Registration Co-Chair
Ohio Supercomputer Center

Michele Bianchini-Gunn, Onsite
Registration Co-Chair/ Store Chair
Lawrence Livermore National Laboratory

Janet McCord, Onsite Registration
Co-Chair
University of Texas - Austin

Kathy Turnbeaugh, Onsite Registration
Co-Chair
Lawrence Livermore National Laboratory

Spargo, Registration Contractor

Talley Management Group, Finance
Contractor
Challenges

Analytics Challenge
Timothy Leite, Co-Chair, Visual Numerics, Inc.
Paul Fussell, Co-Chair, Boeing
Don Costello, University of Nebraska
Sharan Kalwani, General Motors Corporation
Jill Matzke, SGI
Richard Strelitz, Los Alamos National Laboratory
Danl Pierce, Cray, Inc.
Betsy Riley, Oak Ridge National Laboratory
Michael Schulman, Sun Microsystems

Cluster Challenge
Brent Gorda, Chair, Lawrence Livermore National Laboratory
Bill Boas, System Fabric Works
Ricky A. Kendall, Oak Ridge National Laboratory
Ann Redelfs, Redelfs LLC
George Smith, Lawrence Berkeley National Laboratory
Tom Spelce, Lawrence Livermore National Laboratory
Rich Wolski, University of California, Santa Barbara

Storage Challenge
Raymond L. Paden, Chair, IBM
Mike Knowles, Raytheon
Randy Kreiser, Data Direct Networks
Charles Lively, Texas A&M University
Thomas Ruwart, I/O Performance, Inc.
Chris Semple, Petroleum Geo-Services
John R. Sopka, High Performance System Software
Alan Sussman, University of Maryland
Virginia To, VTK Solutions LLC

Bandwidth Challenge
Debbie Montano, Chair, Force10 Networks
Jim DeLeskie, Force10 Networks
Kevin Walsh, San Diego Supercomputer Center
Anne Richeson, Qwest
Greg Goddard, Spirent Communications
Richard Carlson, Internet2
Stephen Q. Lau, University of California, San Francisco

Technical Program Committee

Applications
Omar Ghattas, Chair
University of Texas - Austin
William L. Barth, University of Texas - Austin
Stephen C. Jardin, Princeton Plasma Physics Laboratory
Theresa L. Windus, Iowa State University
Christian Bischof, Aachen University
Kwan-Liu Ma, University of California - Davis
Brett M. Bode, Ames Laboratory
Charbel Farhat, Stanford University
Tiankai Tu, Carnegie Mellon University
Nikos Chrisochoides, College of William and Mary
Srinivas Aluru, Iowa State University
George Karypis, University of Minnesota
Ulrich J. Ruede, University of Erlangen-Nuremberg
Padma Raghavan, Pennsylvania State University
Gerhard Wellein, Regionalales Rechenzentrum Erlangen
Srinivasan Parthasarathy, Ohio State University
Acknowledgments

David A. Bader, Georgia Institute of Technology
P. Sadayappan, Ohio State University
George Biros, University of Pennsylvania
David R. O’Hallaron, Intel Research
Bart van Bloemen Waanders, Sandia National Laboratories
Rupak Biswas, NASA Ames Research Center
Esmond G. Ng, Lawrence Berkeley National Laboratory

Architecture
John B. Carter, Chair, University of Utah
Ashley Saulsbury, Sun Microsystems Inc.
Viktor K. Prasanna, University of Southern California
Dennis Abts, Cray, Inc
Derek Chiou, University of Texas-Austin
Alex Ramirez, Universitat Politecnica de Catalunya
Mark Heinrich, University of Central Florida
Kei Hiraki, University of Tokyo
Doug Burger, University of Texas - Austin
Xiaowei Shen, IBM China Research
Jose F. Martinez, Cornell University
Lixin Zhang, IBM
Yan Solihin, North Carolina State University

Grids
Anne Trefethen, Co-Chair, Oxford University
Marty A. Humphrey, Co-Chair University of Virginia
Rich Wolski, University of California Santa Barbara
Wolfgang Gentzsch, D-Grid
Beth Plale, Indiana University

Franck Cappello, INRIA
Henrique Andrade, IBM T. J. Watson Research Center
Daniel S. Katz, Louisiana State University
Satoshi Matsuoka, Tokyo Institute of Technology
Jennifer M. Schopf, Argonne National Laboratory
Andrew Grimshaw, University of Virginia
Ann L. Chervenak, USC Information Sciences Institute
Jim Basney, National Center for Supercomputing Applications
Sujoy Basu, Hewlett-Packard
Keith R. Jackson, Lawrence Berkeley National Laboratory
Silvia Figueira, Santa Clara University
Henri E. Bal, Vrije Universiteit
Philip M. Papadopoulos, San Diego Supercomputer Center
Manish Parashar, Rutgers University
David Abramson, Monash University

Networks
Craig B. Stunkel, Chair, IBM T. J. Watson Research Center
Jarek Nieplocha, Pacific Northwest National Laboratory
Steven L. Scott, Cray Inc.
Darius Buntinas, Argonne National Laboratory
Keith Underwood, Intel Corporation
Alan Benner, IBM
Dhabaleswar K. (DK) Panda, Ohio State University
Scott Pakin, Los Alamos National Laboratory
Dongyan Xu, Purdue University
Pete Wyckoff, Ohio Supercomputer Center
Mitchell Gusat, IBM Research
Rami Melhem, University of Pittsburgh
Acknowledgments

Vivek S. Pai, Princeton University
Fabrizio Petrini, IBM TJ Watson Research Center

Performance
Adolfy Hoisie, Chair, Los Alamos National Laboratory
Darren J. Kerbyson, Los Alamos National Laboratory
Karen L. Karavanic, Portland State University
Vladimir Getov, University of Westminster
Jeffrey Vetter, Oak Ridge National Laboratory and Georgia Institute of Technology
Xian-He Sun, Illinois Institute of Technology
Lizy K. John, University of Texas-Austin
Bernd Mohr, Forschungszentrum Juelich
Barton Miller, University of Wisconsin
Philip C. Roth, Oak Ridge National Laboratory
Leonid Oliker, Lawrence Berkeley National Laboratory
Martin Schulz, Lawrence Livermore National Laboratory
Bronis R. de Supinski, Lawrence Livermore National Laboratory
Patrick Haven Worley, Oak Ridge National Laboratory
Allan Snavely, San Diego Supercomputer Center

System Software
Keshav Pingali, Chair, University of Texas
Rudolf Eigenmann, Purdue University
Lawrence Rauchwerger, Texas A&M University
Mary Lou Soffa, University of Virginia
Robert Wisniewski, IBM Research
Frank Mueller, North Carolina State University

Ram Rajamony, IBM Research
Guang R. Gao, University of Delaware
Calin Cascaval, IBM TJ Watson Research Center
Greg Bronevetsky, Lawrence Livermore National Laboratory
Laxmikant Kale, University of Illinois at Urbana-Champaign
Siddhartha Chatterjee, IBM Research
Mary Hall, University of Southern California
Karsten Schwan, Georgia Institute of Technology
Hironori Kasahara, Waseda University
Samuel Midkiff, Purdue University
David Padua, University of Illinois at Urbana-Champaign
Paul Stodghill, USDA Agricultural Research Service
Tony Hey, Microsoft Corporation
Wei Li, Intel Corporation
Calvin Lin, University of Texas - Austin

Tutorials Committee
Jack Dongarra, Co-Chair, University of Tennessee
Bruce Loftis, Co-Chair, Purdue University
Padma Raghavan, Pennsylvania State University
Tom Hacker, Purdue University
Raghu Reddy, Pittsburgh Supercomputing Center
Bryan Embry, National Security Agency
Lauren L. Smith, National Security Agency
Satoshi Matsuoka, Tokyo Institute of Technology
Dieter Kranzlmueller, Johannes Kepler University Linz
Diane Rover, Iowa State University
Frederic Desprez, INRIA
Richard F. Barrett, Oak Ridge National Laboratory
Acknowledgments

Robert B. Ross, Argonne National Laboratory
L. Eric Greenwade, Lawrence Livermore National Laboratory
David William Walker, Cardiff University
José L. Muñoz, National Science Foundation
John Towns, National Center for Supercomputing Applications
Anne C. Elster, Norwegian University of Science and Technology
Glenn R. Luecke, Iowa State University
Fred Johnson, Department of Energy
John W. Cobb, Oak Ridge National Laboratory
Michael M. Resch, High Performance Computing Center Stuttgart
Blaise M. Barney, Lawrence Livermore National Laboratory
Candace Culhane, National Security Agency

Posters Committee
Tamara K. Grimmett, Co-Chair, Idaho National Laboratory
Charles Koelbel, Co-Chair, Rice University
Michela Taufer, University of Delaware
Jeffrey K. Hollingsworth, University of Maryland
Suzanne Shontz, Pennsylvania State University
Randall Bramley, Indiana University
Gabrielle Allen, Louisiana State University
Siddhartha Chatterjee, IBM Research
L. Eric Greenwade, Lawrence Livermore National Laboratory
Robert B. Ross, Argonne National Laboratory
Wu Feng, Virginia Tech
Tia Newhall, Swarthmore College
Jeffrey J. Evans, Purdue University

Kevin James Barker, Los Alamos National Laboratory
Robert A. Ballance, Sandia National Laboratories

Disruptive Technologies Committee
Jeffrey Vetter, Chair, Oak Ridge National Laboratory and Georgia Institute of Technology
Mark Seager, Lawrence Livermore National Laboratory
Richard Linderman, Air Force Research Laboratory
Gary D. Hughes, Department of Defense
Erik P. DeBenedictis, Sandia National Laboratories

Doctoral Research Showcase Committee
Jeffrey K. Hollingsworth, Chair, University of Maryland
David A. Bader, Georgia Institute of Technology
I-Hsin Chung, IBM Research
Martin Schulz, Lawrence Livermore National Laboratory
Karsten Schwan, Georgia Institute of Technology
Allan Snavely, San Diego Supercomputer Center
Erich Strohmaier, Lawrence Berkeley National Laboratory
Chau-Wen Tseng, University of Maryland

Gordon Bell Prize Committee
David H. Bailey, Chair, Lawrence Berkeley National Laboratory
David E. Keyes, Columbia University
James Demmel, University of California, Berkeley
**Acknowledgments**

**Mateo Valero**, Barcelona Supercomputing Center

**Seymour Cray Computer Science and Engineering Award Committee**

**David H. Bailey**, Chair, Lawrence Berkeley National Laboratory

**Mateo Valero**, Barcelona Supercomputing Center

**Steven L. Scott**, Cray, Inc.

**Monty M. Denneau**, IBM

**William J. Dally**, Stanford University

**Sidney Fernbach Memorial Award Committee**

**Steven F. Ashby**, Chair, Lawrence Livermore National Laboratory

**Kwan-Liu Ma**, University of California, Davis

**Edward Seidel**, Louisiana State University

**Sangtae Kim**, Purdue University

**Esmond G. Ng**, Lawrence Berkeley National Laboratory

**Rupak Biswas**, NASA Ames Research Center

**Douglas Kothe**, Oak Ridge National Laboratory

**Ulrich J. Ruede**, University of Erlangen-Nuremberg

**Omar Ghattas**, University of Texas - Austin

**Charles Holland**, DARPA

**External Papers Referees**

**Virat Agarwal**, Georgia Institute of Technology

**Natalia Alexandrov**, NASA Langley Research Center

**Miguel Argaez**, University of Texas - El Paso

**Dorian C. Arnold**, University of Wisconsin

**Ananth Grama**, Purdue University

**Kevin James Barker**, Los Alamos National Laboratory

**Brian Barrett**, Los Alamos National Laboratory

**Richard F. Barrett**, Oak Ridge National Laboratory

**Jonathan Bentz**, Cray, Inc.

**Andrew Bernat**, University of Wisconsin

**Rob H. Bisseling**, Utrecht University

**Michael J. Brim**, University of Wisconsin

**Roy Campbell**, Army Research Laboratory

**Francisco J. Cazorla**, Barcelona Supercomputing Center

**Siddhartha Chatterjee**, IBM Research

**Liqun Cheng**, University of Utah

**Alok Choudhary**, Northwestern University

**Edmond Chow**, D. E. Shaw Research

**Cyrus Umrigar**, Cornell University

**Kei Davis**, Los Alamos National Laboratory

**Larry P. Davis**, DoD High Performance Computing Modernization Program

**David Earl**, University of Pittsburgh

**Michael Deem**, Rice University

**Niels Drost**, Vrije Universiteit

**Iain Duff**, CERFACS

**Denis Zorin**, New York University

**Victor Eijkhout**, University of Texas - Austin

**Dick H. Epema**, Delft University of Technology

**Bard Ermentrout**, University of Pittsburgh

**Hadi Esmaeilzadeh**, University of Texas - Austin

**Blake Fitch**, IBM

**Kelly Gaither**, Texas Advanced Computing Center

**Gregory Byrd**, North Carolina State University

**Dileep George**, Numenta

**Michael Gerndt**, Technische Universität
München Bernard, INRIA
William Douglas Gropp, Argonne National Laboratory
Guy Blelloch, Carnegie Mellon University
Eldad Haber, Emory University
Tsuyoshi Hamada, RIKEN
Robert J. Harrison, University of Tennessee / Oak Ridge National Laboratory
Bruce Hendrickson, Sandia National Laboratories
Chris Henze, NASA
Tony Hey, Microsoft Corporation
Stephen Hodgson, Oak Ridge National Laboratory
Greg Hood, Pittsburgh Supercomputing Center
Robert T. Hood, NASA Ames
Mary Inaba, University of Tokyo
Henry Jin, NASA Ames Research Center
Joel Saltz, Ohio State University
John Salmon, D. E. Shaw Research
John R. Johnson, Lawrence Livermore National Laboratory
Philip Jones, Los Alamos National Laboratory
Ananth Kalyanaraman, Washington State University
Jeremy Kubica, Google, Inc.
Torsten Kuhlen, Aachen University
Karen Willcox, Massachusetts Institute of Technology
Matthew Legendre, University of Wisconsin
Dong Li, University of Texas - Austin
Jian Li, IBM
David Lowenthal, University of Georgia
Jason Maassen, Vrije Universiteit Amsterdam
Dinesh Manocha, University of North Carolina Chapel Hill
Satoshi Matsuoka, Tokyo Institute of Technology
Patrick S. McCormick, NASA Ames Research Center
Jeremy Meredith, Oak Ridge National Laboratory
Klaus Mueller, Stony Brook University
Nigel Goddard, University of Edinburgh
Arifa Nisar, Northwestern University
George Ostrovchov, Oak Ridge National Laboratory
Benno Overeinder, Vrije Universiteit Amsterdam
Yoonho Park, IBM
Miquel Pericas, Barcelona Supercomputing Center
Patrick Moran, NASA
Simon Portegies Zwart, University of Amsterdam
Rizos Sakellariou, University of Manchester
Behnam Robatmili, University of Texas - Austin
Nathan E. Rosenblum, University of Wisconsin
Robert B. Ross, Argonne National Laboratory
Barry Rountree, University of Georgia
Vipin Sachdeva, IBM
Subhash Saini, NASA Ames Research Center
Valentina Salapura, IBM
Nagiza F. Samatova, Oak Ridge National Laboratory
Rob Schreiber, Hewlett-Packard
Philippe Selo, IBM
Mrinal Sen, University of Texas - Austin
Thomas Serre, Massachusetts Institute of Technology
Simha Sethumadhavan, University of Texas - Austin
Luo Si, Purdue University
Chris Simmons, University of Texas - Austin
Acknowledgments

Horst Simon, Lawrence Berkeley National Laboratory
Stuart Johnson, Texas Advanced Computing Center
Brian Smith, IBM
Steven Parker, University of Utah
Eric Stahlberg, Ohio Supercomputer Center
Alexandros P. Stamatakis, Swiss Federal Institute of Technology, Lausanne
Steve McMillan, Drexel University
Craig Stewart, Indiana University
Erich Strohmaier, Lawrence Berkeley National Laboratory
Sai R. Susarla, Network Appliance
Friman Sánchez Castaño, Technical University of Catalonia
Tandy Warnow, University of Texas - Austin
Michela Taufer, University of Texas - El Paso
Rajeev S Thakur, Argonne National Laboratory
John Thornton, Griffith University
Luke Tierney, University of Iowa
Akhilesh Tyagi, Iowa State University
Virginia Torczon, College of William and Mary
Robert A. van de Geijn, University of Texas - Austin
C. van Reeuwijk, Vrije Universiteit Amsterdam
Kees Verstoep, Vrije Univeriteit Amsterdam
Richard W. Vuduc, Lawrence Livermore National Laboratory
Lucas Wilcox, University of Texas - Austin
Tiffani Williams, Texas A&M University
Katherine Yelick, University of California, Berkeley / Lawrence Berkeley National Laboratory
Lexing Ying, University of Texas - Austin
Weikuan Yu, Oak Ridge National Laboratory
ChengXiang Zhai, University of Illinois at Urbana-Champaign
Yi Zhang, University of California, Santa Cruz

SCinet Committee

Jackie Kern, Chair
National Center for Supercomputing Applications
Chuck Fisher, Vice Chair
Oak Ridge National Laboratory
Patrick Dorn, Deputy Chair
National Center for Supercomputing Applications

SCinet Vendors

Michaela Mezo, Juniper Networks
Manjuel Robinson, Fujitsu
Jeff Verrant, Ciena
John W. Maklae, Qwest
Bill Ryan, Foundry Networks

Measurement

Matthew J. Zekauskas, Co-Chair, Internet2
Jeff W. Boote, Co-Chair, Internet2,
Neil H. McKee, InMon Corp.
Martin Swany, University of Delaware
Richard Carlson, Internet2
Sonia Panchen, InMon Corporation
Aaron Brown, Internet2

Architecture

Tracey D. Wilson, Chair, Computer Sciences Corporation
Communications
Lauren Rotman, Liaison, Internet2

Equipment
Jeffery Alan Mauth, Chair, Pacific Northwest National Laboratory

Fiber/Circuits
John S. Long, Chair, Oakridge National Laboratory
Ken Brice, Army Research Laboratory
Annette Hoff, Sandia National Laboratories
Craig Leres, Lawrence Berkeley National Laboratory
Shane Archiquette, Colorado Technical University

Help Desk
Gayle Allen, Help Desk Co-Chair, Sandia National Laboratories
Ann E. Baker, Help Desk Co-Chair, Oak Ridge National Laboratory

IP Services
Rex Duncan, Chair, Oak Ridge National Laboratory
Mike Beckman, US Army and Military Defense Command
Jeffrey Richard Schwab, Purdue University

IT Services
Casey T. Deccio, Co-Chair, Sandia National Laboratories
Davey Wheeler, Co-Chair, National Center for Supercomputing Applications

Logistics
Ralph A. McEldowney, Chair, US Air Force
Bob Williams
John R. Carter, US Air Force

Network Monitoring
Bill Nickless, Chair, Pacific Northwest National Laboratory
Scott Campbell, Lawrence Berkeley National Laboratory
Jason S. Price, MRV Communications
Jason Lee, Lawrence Berkeley National Laboratory

Open Fabrics
Troy Robert Benjegerdes, Chair, DOE Ames Laboratory
Chris Wiggins, Emcore Corporation
Makia Minich, Oak Ridge National Laboratory
Parks Fields, Los Alamos National Laboratory
Cary Whitney, Lawrence Berkeley National Laboratory
Douglas Fuller, Arizona State University

Physical Security
Stephen Q. Lau, Chair, University of California, San Francisco

Power
William R. Wing, Chair, Oak Ridge National Laboratory

Routing
Linda Winkler, Chair, Argonne National Laboratory
Corey Hall, Argonne National Laboratory
Eli Dart, Energy Sciences Network
Brent Sweeney, Indiana University / Internet2
Corby Schmitz, Corby
JP Velders, University of Amsterdam
Thomas Hutton, San Diego Supercomputer Center
Pieter de Boer, SARA
Andrew Lee, Indiana University
Alan Verlo, StarLight / University of Illinois at Chicago

WAN Transport
Chris Costa, Co-Chair, CENIC
Edwin W. Smith, Co-Chair, CENIC
Greg Ebner, Nevada System of Higher Education
Bill Jensen, University of Wisconsin - Madison
Kevin McGrattan, Cisco Systems
Christian Todorov, Internet2

Wireless
Jamie Van Randwyk, Chair, Sandia National Laboratories
Michele Kahn, Sandia National Laboratories
Mark Mitchell, Sandia National Laboratories
Tracey Lamee, Sandia National Laboratories

Xnet
Rod Wilson, Co-Chair, Nortel
E. Paul Love, Co-Chair, Internet Consulting of Vermont
Eric Bernier, Nortel

Steering Committee
ARajkumar Buyya, University of Melbourne
Donna Cappo, ACM
Dona Crawford, Lawrence Livermore National Laboratory
Jack Dongarra, University of Tennessee
John Grosh, Lawrence Livermore National Laboratory
Barbara Horner-Miller, Arctic Region Supercomputing Center
David Kaeli, Northeastern University
Anne Marie Kelly, IEEE Computer Society
William Kramer, Lawrence Berkeley National Laboratory
Fred Johnson, US Department of Energy
Scott Lathrop, University of Chicago/Argonne National Laboratory
Wilfred Pinfold, Intel Corporation
Daniel A. Reed, Renaissance Computing Institute
James H. Rogers, Oak Ridge National Laboratory
Rob Schreiber, Hewlett-Packard
Burton Smith, Microsoft Corporation
Patricia J. Teller, University of Texas-El Paso
Becky Verastegui, Oak Ridge National Laboratory

ACM/IEEE-CS High Performance Computing Ph.D. Fellowship Committee
William Kramer, Chair, Lawrence Berkeley National Laboratory
Charles Koelbel, Rice University
Scott Lathrop, University of Illinois at Urbana-Champaign

Industry Advisory Committee
Frank Baetke, Hewlett-Packard
Mike Bernhardt, Tabor Communications
Rich Brueckner, Sun Microsystems
Denney Cole, The Portland Group, STMICRO
Aaron D’Orlando, Intel CorporationCorpor
Phil Fraher, Visual Numerics, Inc.
George Funk, Dell Inc.
SC07 Student Volunteers
SC07 extends a heartfelt thanks to all the student volunteers whose efforts help to make the conference a success. Student volunteers are the unsung heroes of the conference, taking on any number of tasks and assignments.

In return, these undergraduate and graduate students get first-hand exposure to leaders in the HPC community. The student volunteers in particular provide critical assistance to SCinet, Exhibits, the Education Program, the Technical Program, and Communications, as well as many behind-the-scenes efforts.

Association for Computing Machinery (ACM)
Founded in 1947, ACM is a major force in advancing the skills of information technology professionals and students worldwide. Today, ACM’s 80,000 members and the public turn to ACM for the industry’s leading Portal to Computing Literature, authoritative publications and pioneering conferences, providing leadership for the 21st century. The SC conference series is co-sponsored by the ACM Special Interest Group on Computer Architecture (SIGARCH), which serves a unique community of computer professionals working at the forefront of computer design in both industry and academia. It is ACM’s primary forum for interchange of ideas about tomorrow’s hardware and its interactions with compilers and operating systems.

IEEE Computer Society
The IEEE Computer Society is the world’s leading association of computing professionals with nearly 90,000 members in over 140 countries. Founded in 1946 and today the largest society within the IEEE, this not-for-profit organization is the authoritative provider of technical information and services for computing communities worldwide. In addition to co-sponsoring the SC conference series, IEEE-CS offers a full range of career enhancing products and services through its 214,000-article digital library, 23+ peer-reviewed journals, e-learning courseware, online technical books, 175+ technical conferences, standards development, 40+ technical committees, certification for software professionals, 150 local society chapters, awards and scholarships, and much more.
November 15-21, 2008
Austin, Texas

Conference Dates
November 15-21, 2008

Exhibits Dates
November 17-20, 2008

For more information, go to http://sc08.supercomputing.org

SC08 Sponsors:
ACM SIGARCH/IEEE Computer Society
When SC08 rides into Austin in November 2008, the conference series will celebrate its 20th anniversary as the premier international conference on high performance computing, networking, storage and analysis. SC08 is the forum for demonstrating how these technologies are driving new ideas, new discoveries and new industries. Since 1989, the SC conference has served as an international meeting place, bringing together scientists, engineers, researchers, educators, programmers, system administrators and managers from the global HPC community.

Plan now to be a part of SC08 and its program of trailblazing technical papers, timely tutorials, engaging invited speakers, up-to-the-minute research posters, entertaining panels and thought-provoking birds-of-a-feather sessions. New for 2008 will be two Technology Thrusts: Energy and Biomedical Informatics. Exhibits from industry, academia and government research organizations will demonstrate the latest innovations in computing and networking technology. SC08 promises to be the most compelling and innovative SC conference yet.

In keeping with SC’s history of pushing new frontiers, SC08 will feature several new components.

In the Technical Program, two Technology Thrusts have been added. The Energy Thrust will focus both on the use of HPC in renewable energy and energy efficiency research, as well as the challenge of best practices and technology trends aimed at energy efficient data centers. The Biomedical Informatics Thrust will focus on the use of Grid and HPC technologies to support translational biomedical research, computational biology, large-scale image analysis, personalized medicine and systems biology.

Another new component of the conference, complementing Austin’s title as the Live Music Capital of the World, will be the SC08 Music Initiative. While music and musicians may not readily come to mind when you think of SC conference attendees, it’s no secret that a large number of SC attendees are also composers, musicians and music lovers, as well as scientists, mathematicians and engineers. The SC08 Music Initiative will bring together all these dimensions to make music an exciting and enjoyable aspect of the SC08 conference experience.

Plan now to participate in SC08, the 20th anniversary of the premier international conference on high performance computing, networking, storage and analysis.