



SC12

Salt Lake City, Utah



Conference Program

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

Salt Palace Convention Center
Salt Lake City Utah
November 10-16, 2012
www.sc12.supercomputing.com

[illegible][illegible][illegible]

Table of Contents

3	Welcome from the Chair
4	Governor's Welcome
5	Acknowledgements

7	General Information/SCinet
9	Registration and Conference Store Hours
9	Registration Categories
10	Registration Pass Access
11	Exhibit Hall Hours
11	Information Booth Hours
11	SC13 Preview Booth Hours
11	Social Events
12	Services/Facilities
13	SCinet

17	Map/Daily Schedules
19	Downtown Area Map
21	Daily Schedules

37	Keynote/Invited Talks/Panels
39	Keynote
39	Invited Talks
44	Panels

49	Papers
----	---------------

77	Posters/Scientific Visualization Showcase
79	Research Posters
97	ACM Student Research Competition Posters
100	Scientific Visualization Showcase

105	Tutorials
-----	------------------

117	Workshops
-----	------------------

125	Birds of a Feather
-----	---------------------------

145	Awards
-----	---------------

151	Exhibitor Forum
-----	------------------------

163	Communities
165	HPC Educators
169	Broader Engagement
173	Doctoral Showcase

Table of Contents



SC12

Salt Lake City, Utah



Welcome from the Chair



Welcome to Utah and to the 2012 conference (SC12) on high performance computing, networking, storage and analysis, sponsored by the IEEE Computer Society and the Association for Computing Machinery (ACM).

SC12 once again features a high quality technical program. We had a record 472 paper submissions and accepted 100 papers for presentation during the conference. In addition to papers, we have a great variety of invited talks, panels, posters and birds-of-a-feather sessions (BOFs) throughout the week. Bookending the conference will be tutorials on Sunday and Monday plus workshops on Sunday, Monday, and Friday. Nearly every part of the conference set records for submissions this year. The technical program committee has worked hard to select the best and most diverse program ever.

The focus of SC12 is on **you**—the conference attendee. We are working hard to make the conference more attendee friendly. To start, we have simplified the number of named activities at the conference. For example, the Keynote, plenary speakers and Masterworks are combined into one program called “Invited Talks.” In addition, we are working to lay out the conference space with attendee needs as the highest priority. The Salt Palace Convention Center provides close access from the exhibit hall to the technical program rooms. To allow you to meet with colleagues, we have created three attendee lounges throughout the convention center to provide a place to sit down and recharge both yourself and your electronic devices. Finally, we are trying to make it easier to find activities and meetings. For example, rather than printed cardboard signs, most meeting rooms will have electronic signs which will always be up to date reflecting any last-minute changes.

The Exhibit hall houses a record number of exhibitors from a range of industry, academic, and government research organizations. SC’s exhibit hall provides a unique marketplace for not only commercial hardware and software, but also ways to see the latest science that is enabled by HPC. Connecting the exhibitors to each other and to the world beyond is SCinet, a unique blending of a high performance production network and a bleeding-edge demonstration network.

A distinctive aspect of SC is our commitment to developing the next generation of HPC professionals. This effort is focused in our Communities program. The rebranded HPC Educators program provides a high quality peer-reviewed program that describes how both HPC and scientific computation in general can be taught to students. The Educators program runs throughout the conference and is open to all technical program attendees. New for 2012 is the “Experience HPC for Undergraduates” program that provides an opportunity for talented sophomores and juniors to attend SC12 as well as have special sessions that introduce them to the field.

Thanks for attending and have a great conference!

A handwritten signature in black ink that reads "Jeff Hollingsworth". The signature is written in a cursive, flowing style.

Jeff Hollingsworth
SC12 General Chair



STATE OF UTAH

GARY R. HERBERT
GOVERNOR

OFFICE OF THE GOVERNOR
SALT LAKE CITY, UTAH
84114-2220

GREG BELL
LIEUTENANT GOVERNOR

November 10, 2012

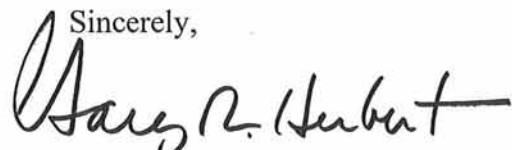
Greetings!

As Governor of the great State of Utah, it is a pleasure to welcome you to SC12, the International Conference for High Performance Computing (HPC), Networking, Storage, and Analysis being held at the Salt Palace Convention Center in Salt Lake City, Utah.

Drawing more than ten thousand research and technology leaders from around the globe, SC12 represents an outstanding opportunity to meet new collaborators and experience the dramatic progress and innovation now underway in fields associated with computational science and engineering. This event provides participants a unique opportunity to explore the ever-changing frontier of computing technology, and to delve further into fields like bio-computing, advanced manufacturing, sustainability, visualization, and the advancing internet.

I also invite you to enjoy our beautiful scenery - in both the Wasatch Mountains and our red rock canyons to the south – as well as the warm and friendly people in Utah. During your stay, I hope you experience some of the closely related technology leadership now taking place in Utah, as well as the favorable environment our state provides for both academic research and corporate innovation.

On behalf of the residents of Utah, best wishes for a successful and memorable event, and may you return soon.

Sincerely,

Gary R. Herbert
Governor

Acknowledgements

No conference this size could be possible without the dedication, commitment and passion of the SC12 committee. The core committee includes more than 100 people who have largely volunteered their time planning for this event for more than three years. Added to that number are the more than 500 people who have helped review submissions and contributed to the planning and preparations. The full list of committee members is posted on the conference website at sc12.supercomputing.org.

SC12 Committee Management

Conference Chair

Jeffrey K. Hollingsworth, University of Maryland

Vice Chair

Wilf Pinfold, Intel

Deputy General Chair

William Douglas Gropp, University of Illinois at Urbana-Champaign

Executive Assistant

Carolyn Peters, Argonne National Laboratory

Communications Co-chairs

Trish Damkroger, Lawrence Livermore National Laboratory
Ian MacConnell, Ohio Supercomputer Center

Communities Chair

John Grosh, Lawrence Livermore National Laboratory

Exhibits Chair

Mary Hall, University of Utah

Finance Chair

Ralph A. McElowney, DOD HPC Modernization Program

Infrastructure Chair

Janet Brown, Pittsburgh Supercomputing Center

SCinet Chair

Linda Winkler, Argonne National Laboratory

Technical Program Chair

Rajeev Thakur, Argonne National Laboratory

Society Liaisons

IEEE-CS

Lynne Harris
Brookes Little
Carmen Saliba

ACM

Donna Cappel





General Information

In this section you'll find information on registration, exhibit hours, conference store hours, descriptions and locations of all conference social events, information booths and their locations, and convention center facilities and services.

General Information

General Information

Registration and Conference Store

The registration area and conference store are located in the South Foyer, Lower Concourse.

Registration and Conference Store Hours

Saturday, November 10	1pm–6pm
Sunday, November 11	7am–6pm
Monday, November 12	7am–9pm
Tuesday, November 13	7:30am–6pm
Wednesday, November 14	7:30am–6pm
Thursday, November 15	7:30am–5pm
Friday, November 15	8am–11am

Registration Categories

Tutorials

Tutorials run for two days, Sunday and Monday. Attendees can purchase a One-Day (Sunday or Monday) Passport or Two-Day (Sunday and Monday) Passport. The fee includes admission to any of the tutorials offered on the selected day for a One-Day Passport or a Two-Day Passport. The fee also includes a set of all of the tutorial notes (provided on a DVD) and lunch on the day(s) of Tutorial registration. Tutorial registration DOES NOT provide access to the keynote and exhibit halls.

Technical Program

Technical Program registration provides access to: the keynote, invited talks, papers, panels, posters (including reception), exhibits, Student Cluster Competition, awards, the Doctoral Research Showcase, Birds of a Feather, Exhibitor Forum, and Scientific Visualization Showcase (including reception), the Monday night Exhibits Opening Gala, Conference Reception on Thursday night, and one copy of the SC12 proceedings (on a DVD). In addition, registrants are admitted to the HPC Educators Program and Broader Engagement Program Tuesday through Thursday.

New for SC12: Workshops are NOT included with the Technical Program registration. If you would like to add workshops to your Technical Program registration, you will need to register for the Workshops Add-On. (See Workshops.)

Exhibitor

Exhibitor registration provides access to: the exhibit hall, the keynote, posters (except the poster reception), plenaries, awards, Exhibitor Forum, Student Cluster Competition, Friday panel sessions, and Scientific Visualization Showcase (but not related reception). The registration also includes access to the Exhibits Gala Opening on Monday evening and participation in the Exhibitor Reception.

Exhibits Only

Exhibits Only registration provides access to: The exhibit halls during regular exhibit hours, Tuesday through Thursday, posters (but not the poster reception), and the Awards ceremony (Thursday). It does NOT provide access to the Monday night Gala Opening or the Sunday evening Exhibitor Reception.

Workshops

New for SC12 are two categories for Workshops—**Workshops Only** and **Workshops Add-On to Technical Program**. This registration provides access to all Workshop sessions on the day(s) of registration. In addition, it provides access to the HPC Educators and Broader Engagement Programs on the day(s) of Workshop registration.

Proceedings

Attendees registered for the Technical Program will receive one copy of the SC12 proceedings on a DVD.

Lost Badge

There is a \$40 processing fee to replace a lost badge.

Age Requirements Policy

- Technical Program attendees must be 16 years of age or older. Age verification is required.
- Exhibits-Only registration is available for children ages 12-16. Age verification is required.
- Children 12 and under are not permitted on the Exhibit Hall other than on Family Day (see below).
- Children under 16 are not allowed in the Exhibit Hall during installation, dismantling or before or after posted exhibit hours. Anyone under 16 must be accompanied by an adult at all times while visiting the exhibition.

Family Day

Family Day is Wednesday, November 14, 4pm-6pm. Adults and children 12 and over are permitted on the floor during these hours when accompanied by a registered conference attendee.

Registration Pass Access

Each registration category provides access to a different set of conference activities, as summarized below.

Type of Event	Tutorials	Technical Program	Technical Program +Workshops	Workshop Only	Exhibitor	Exhibit Hall
Awards (Thursday)		*	*	*	*	*
Birds-of-a-Feather		*	*		*	
Broader Engagement & Educator Sessions (Sun/Mon)			*	*		
Broader Engagement & Educator Sessions (Tue-Thu)		*	*			
Conference Reception (Thursday)		*	*			
Exhibit Floor		*	*		*	*
Exhibitor Forum		*	*		*	
Exhibits Gala Opening (Monday)		*	*		*	
Exhibitor's Reception					*	
Invited Talks (Non-Plenary)		*	*			
Invited Talks (Plenary)		*	*		*	
Keynote (Tuesday)		*	*		*	
Panels (Tue-Thur)		*	*			
Panels (Friday Only)		*	*		*	
Papers		*	*			
Posters		*	*		*	*
Poster Reception (Tuesday)		*	*			
Tutorial Lunch (Sun/Mon ONLY)	*					
Tutorial Sessions	*					
Student Cluster Competition		*	*		*	
Workshops			*	*		

Exhibit Floor Hours

Tuesday, Nov. 13	10am-6pm
Wednesday, Nov. 14	10am-6pm
Thursday, Nov. 15	10am-3pm

SC12 Information Booths

Need up-to-the-minute information about what's happening at the conference. Need to know where to find your next session? What restaurants are close by? Where to get a document printed? These questions and more can be answered by a quick stop at one of the SC Information booths. There are two booth locations for your convenience: one is on the Lower Concourse, South Foyer, just near registration and the conference store; the second booth (satellite) is located on the Upper Concourse Lobby near Room 251.

SC12 Information Booth Hours

	Main Booth	Satellite Booth
Saturday	1pm-6pm	Closed
Sunday	8am-6pm	8am-5pm
Monday	8am-7pm	8am-5pm
Tuesday	7:30am-6pm	7:30am-5:30pm
Wednesday	8am-6pm	8am-5:30pm
Thursday	8am-6pm	8am-Noon
Friday	8:30am-12pm	Closed

SC13 Preview Booth

Members of next year's SC committee will be available in the SC13 booth (located in the South Foyer of the convention center) to offer information and discuss next year's SC conference in Denver. You'll also be greeted by the famous Blue Bear and a representative from the Denver Convention and Visitors Bureau who'll be able to answer questions about local attractions and amenities. Stop by for a picture with the Blue Bear and to pick up your free gift!

The booth will be open during the following hours:

Tuesday, Nov. 12	10am-6pm
Wednesday, Nov. 13	10am-6pm
Thursday, Nov. 14	10am-6pm

Social Events**Exhibitor Reception**

Sunday, November 11
6pm-9pm

SC12 will host an Exhibitor Reception for registered exhibitors. The party is SC12's way of thanking exhibitors for their participation and support of the conference. The reception will be held at The Hotel Elevate, a downtown nightclub located at 155 West 200 South, directly across the street from the South Entrance to the Salt Palace Convention Center. The Hotel Elevate boasts four separate bar areas, with a dance floor in the main bar. Exhibitors will be entertained by live bands, and food and drinks will be served throughout the event.

An Exhibitor badge and government-issued photo ID are required to attend this event, and attendees must be 21 years or older.

Exhibits Gala Opening Reception

Monday, November 12
7pm-9pm

SC12 will host its annual Grand Opening Gala in the Exhibit Hall. This will be your first opportunity to see the latest high performance computing, networking, storage, analysis, and research products, services, and innovations. This event is open to all Technical Program and Exhibitor registrants.

Posters Reception

Tuesday, November 13
5:15pm-7pm

The Posters Reception is an opportunity for attendees to interact with poster presenters. The reception is open to all attendees with Technical Program registration. The Poster Reception is located in the East Lobby.

Scientific Visualization Showcase Reception

Tuesday, November 13
5:15pm-7pm

After you have viewed the posters at the Poster Reception, stop by the Scientific Visualization Showcase Reception for dessert. The reception is open to all attendees with Technical Program registration. The Poster Reception is located in the North Foyer.

Technical Program Conference Reception
Thursday, November 17
6pm-9pm

SC12 will host a conference reception for all Technical Program attendees. Join us for great food, beverages, and entertainment at The Depot (www.depotslc.com). The Depot, a lively nightclub located in an old train station, is only a few blocks from the convention center. There will be quiet rooms to get one last technical conversation in before heading home, as well as live entertainment, including two performances by comedian Ryan Hamilton.

Attendees are required to wear technical program badges throughout the reception, and badges may be checked during the event. In addition, all attendees will be required to present a photo ID (driver’s license or passport) to enter this event and must be 21 years or older to consume alcohol.

Shuttle transportation to the event will run 7pm-10pm from the South Plaza Entrance of the convention center (look for buses with “The Depot” sign in the front window). Transportation also will be available for those on the hotel shuttle routes, with buses running every 15 minutes from the regular pick-up locations (again, look for “The Depot” sign in the front window).

Services/Facilities

ATMs

Two U.S. Bank cash machines are located inside the convention center. You’ll find an ATM on the Upper Concourse (toward Room 254). Another is located in the North Foyer (near the rounded wall).

Business Center

The Salt Palace Business Center is located on the Upper Concourse of the convention center. The center is open most days from 8am to 6pm, but please call (801.534.6305) since, as of this printing, the hours had not been set to accommodate SC12.

Coat and Bag Check

Coat and Bag Check is located in the Lower Concourse, just outside the Ballroom. The hours are:

Saturday, Nov. 10	1pm-6:30pm
Sunday, Nov. 11	7am-10pm
Monday, Nov. 12	7:30am-9:30pm
Tuesday, Nov. 13	7:30am-7:30pm
Wednesday, Nov. 14	7:30am-6pm
Thursday, Nov. 15	7:30am-9:30pm
Friday, Nov. 16	8am-1pm

First-Aid Center

There are two first aid offices in the convention center. One is on the east side, located next to Room 150A; the other is on the west side lobby, outside of Hall 4.

Lost & Found

Lost & Found is located in Room 258.

Prayer and Meditation Room

The Prayer and Meditation Room is located Room 260-B and is open Sunday-Thursday, 9am-5pm.

Restrooms

Restrooms are located conveniently throughout the convention center, as follows:

Lower level:

- Halls A-E (located in the back)
- Hall 1
- Halls 4&5 (west side)
- North and South Foyers
- Outside Room 155

Upper level:

- Across from 254B and near 255
- Upper Mezzanine (on lefthand side)

Visitor’s Center

The Visitor’s Center is located near the East entrance. It is open daily from 9am-5pm.

Wheelchair Rental

Wheelchairs can be acquired through the Business Center.

Where Will SCinet Take You Next?

The era of data intensive science is only today in its infancy. Over the next decade, new large-scale scientific instruments will serve tens of thousands more scientists worldwide. Poised to create petabyte-scale data sets that need to be analyzed and archived, these experiments will need to rely on geographically-dispersed computational and storage resources.

For this reason, the SC conference series has since 1991 created SCinet, a leading edge, high-performance network, assembled each year to enable exhibitors and attendees of the conference to demonstrate HPC innovations in areas that rely on networking for success.

For SC12, SCinet will serve as one of the most powerful networks in the world with nearly 800 Gigabits per second (Gbps) in WAN capacity. Designed and built entirely by volunteers from universities, government and industry, SCinet will link the Salt Palace Convention Center to research networks around the world, such as the Department of Energy's ESnet, Internet2, National LambdaRail, KISTI, SURFnet and others.

SCinet serves as the platform for exhibitors to demonstrate the advanced computing resources of their home institutions and elsewhere by supporting a wide variety of bandwidth-driven applications including supercomputing, cloud computing and data mining. And unlike any commercial network provider, SCinet will utilize advanced virtual circuits and state-of-the-art measurement software that allow attendees and exhibitors to experience peak network performance at all times.

SCinet is also fostering developments in network research that will directly impact data-intensive science. The SCinet Research Sandbox (SRS), now in its third year of the SC conference, allows researchers to showcase "disruptive" network experiments in the unique, live environment of SCinet with access to over 100 Gbps of capacity and an OpenFlow-enabled testbed.

SCinet Research Sandbox Participants

Efficient LHC Data Distribution across 100Gbps Networks

The analysis of data leading to the recent discoveries at the Large Hadron Collider produces data flows of more than 100 Petabytes per year, and increasingly relies on the efficient movement of data sets between the globally distributed computing sites.

The team will demonstrate the state-of-the-art data movement tools, as enabling technology for high-throughput data distribution over 100Gbps WAN circuits. The demo will interconnect 3 major LHC Tier-2 computing sites and the SC12 show floor (booth 809) using 100Gbps technology.

Collaborating organizations: *California Institute of Technology, University of Victoria, University of Michigan, with support from industry partners.*

Demonstration booth: 809

Exploiting Network Parallelism for Improving Data Transfer Performance

The task of scientific bulk data movement, e.g. migrating collected results from the instrumentation to the processing and storage facilities, is hampered by a lack of available network resources. Traditional R&E connectivity can be congested on portions of an end-to-end path causing degradation of overall performance. This SRS project will explore dynamic network control to facilitate efficient bulk data movement, combining opportunistic use of «traditional» networks with dedicated reservations over virtual circuits and OpenFlow enabled resources. The GridFTP application has been instrumented with the eXtensible Session Protocol (XSP), an intelligent system capable of controlling programmable networks. The project intends to show end to end performance improvement between the SC12 conference and campuses involved in the DYNES project, through a combination of regular connectivity, dynamic bandwidth allocations, TCP acceleration, and operations using multiple paths.

Collaborating organizations: *Indiana University, Lawrence Berkeley National Laboratory, Argonne National Laboratory and Internet2*

Demonstration booths: 1042, 1343

Multipathing with MPTCP and OpenFlow

This demo shows several emerging network technologies and how these can be used in big data transfers between data centres. In this demo traffic is sent simultaneously across multiple OpenFlow controlled paths between Geneva and Salt Lake City. The congestion control mechanism of Multipath TCP (MPTCP) favours the least congested paths and ensures that the load balancing across the paths is always optimal.

Collaborating organizations: *SURFnet/SARA, Dutch Research Consortium, iCAIR and California Institute of Technology*

Demonstration booths: 2333, 809, 501

Multi-Science Science DMZ Model with OpenFlow

The emerging era of "Big Science" demands the highest possible network performance. End-to-end circuit automation and workflow-driven customization are two essential capabilities needed for networks to scale to meet this challenge.

This demonstration showcases how combining software-defined networking techniques with virtual circuits capabilities can transform the network into a dynamic, customer-configurable virtual switch. In doing so, users are able to rapidly customize network capabilities to meet their unique workflows with little to no configuration effort. The demo also highlights how the network can be automated to support multiple collaborations in parallel.

Collaborating organizations: *ESnet, Ciena Corporation*
Demonstration booth: 2437

OpenFlow Enabled Hadoop over Local and Wide Area Cluster

The Hadoop Distributed File Systems and Hadoop's implementation of MapReduce is one of the most widely used platforms for data intensive computing. The shuffle and sort phases of a MapReduce computation often saturate network links to nodes and the reduce phase of the computation must wait for data. Our study explores the use of OpenFlow to the control network configuration for different flows and thereby provide different network characteristics for different categories of Hadoop traffic. We demonstrate an OpenFlow enabled version of Hadoop that dynamically modifies the network topology in order to improve the performance of Hadoop.

Collaborating organizations: *University of Chicago*
Demonstration booth: 501

OpenFlow Services for Science: An International Experimental Research Network Demonstrating Multi-Domain Automatic Network Topology Discovery, Direct Dynamic Path Provisioning Using Edge Signaling and Control, Integration With Multipathing Using MPTCP

Large-scale data intensive science requires global collaboration and sophisticated high capacity data management. The emergence of more flexible networking, for example, using techniques based on OpenFlow, provides opportunities to address these issues because these techniques enable a high degree of network customization and dynamic provisioning. These techniques enable large-scale facilities to be created that can be used to prototype new architecture, services, protocols, and technologies. A number of research organizations from several countries have designed and implemented a persistent international experimental research facility that can be used to prototype, investigate, and test network innovations for large-scale global science. For SC12, this international experimental network facility will be extended to from sites across the world to the conference showfloor, and it will be used to support several testbeds and to showcase a series of complementary demonstrations.

Collaborating organizations: *International Center for Advanced Internet Research Northwestern University; National Center for High-Performance Computing, Taiwan; University of Applied Sciences, Taiwan; National Cheng-Kung University, Taiwan; SARA, The Netherlands, California Institute of Technology/CERN; SURFnet. The Netherlands*
Demonstration booths: 2333, 501, 843, 809

Reservoir Labs R-Scope®: Scalable Cyber-Security for Terabit Cloud Computing

Reservoir Labs will demonstrate R-Scope®, a scalable, high-performance network packet inspection technology that forms the core of a new generation of Intrusion Detection Systems enabling the construction and deployment of cyber security infrastructures scaling to terabit per second ingest bandwidths. This scalability is enabled by the use of low- power and high-performance manycore network processors combined with Reservoir's enhancements to Bro, including the incorporation of new sophisticated data structures such as LF- and TED queuing. The innovative R-Scope CSC80 appliance, implemented on a 1U Tiler TILExtreme-Gx platform, will demonstrate the capacity to perform cyber-security analysis at 80Gbps, by combining cyber-security aware front-end network traffic load balancing tightly coupled with the full back-end analytic power of Bro. This fully-programmable platform incorporates the full Bro semantics into the appliance's load-balancing front-end and the back-end analytic nodes.

Collaborating organizations: *Reservoir Labs, SCinet Security Team*

SC12 Wireless Access Policy

In addition to high performance exhibit floor connectivity, SCinet will deploy IEEE 802.11a, 802.11g and 802.11n wireless networks throughout the Salt Palace Convention Center (SPCC) in Salt Lake City. These wireless networks are part of the commodity SCinet network, providing basic access to the Internet. The wireless network will be provided in the meeting rooms, exhibit halls, and common areas of the SPCC. The network can be accessed via SSIDs "SC12" or "eduroam".

eduroam (education roaming) allows users (researchers, teachers, students, staff) from participating institutions to securely access the Internet from any eduroam-enabled institution. The eduroam principle is based on the fact that the user's authentication is done by the user's home institution, whereas the authorization decision granting access to network resources is done by the visited network.

SCinet provides the wireless network for use by all exhibitors and attendees at no charge. Users may experience coverage difficulties in certain areas due to known wireless network limitations, such as areas of reduced signal strength, network congestion, limited client capacity, or other coverage problems. Please watch for additional signage at appropriate locations throughout the convention center. Network settings including IP and DNS addresses for wireless clients are provided by SCinet DHCP services. 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 SCinet wireless networks are governed by this policy posted on the SC12 conference Web site. In summary, while every practical effort shall be made to provide stable reliable network services, there is no explicit service level agreement for any SCinet network, including the wireless networks, nor are there any remedies available in the event that network services are lost.

In order to provide the most robust wireless service possible, SCinet must control the entire 2.4GHz and 5.2GHz ISM bands (2.412GHz to 2.462GHz and 5.15GHz to 5.35GHz) within the SPCC where SC12 conference 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,n 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 or microphones, wireless video or security cameras, or any other equipment transmitting in the 2.4GHz or 5.2GHz spectrum.

SCinet wants everyone to have a successful, pleasant experience at SC12. This should include the ability to sit down with your wireless-equipped laptop or PDA to 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.

SCinet will actively monitor both the 2.4GHz and 5.2GHz frequency spectrums and reserves the right to disconnect any equipment that interferes with the SCinet wireless networks. The SC12 conference reserves the right to deny or remove access to any system in violation of the SCinet acceptable usage policy. Disruptive or illegal activities will not be tolerated.

SCinet reserves the right to revoke access to the wireless network to anyone who uses multicast applications or harms the network in any way, intended or unintended, via computer virus, excessive bandwidth consumption or similar misuse.

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 wired network connection.

SCinet Collaborators

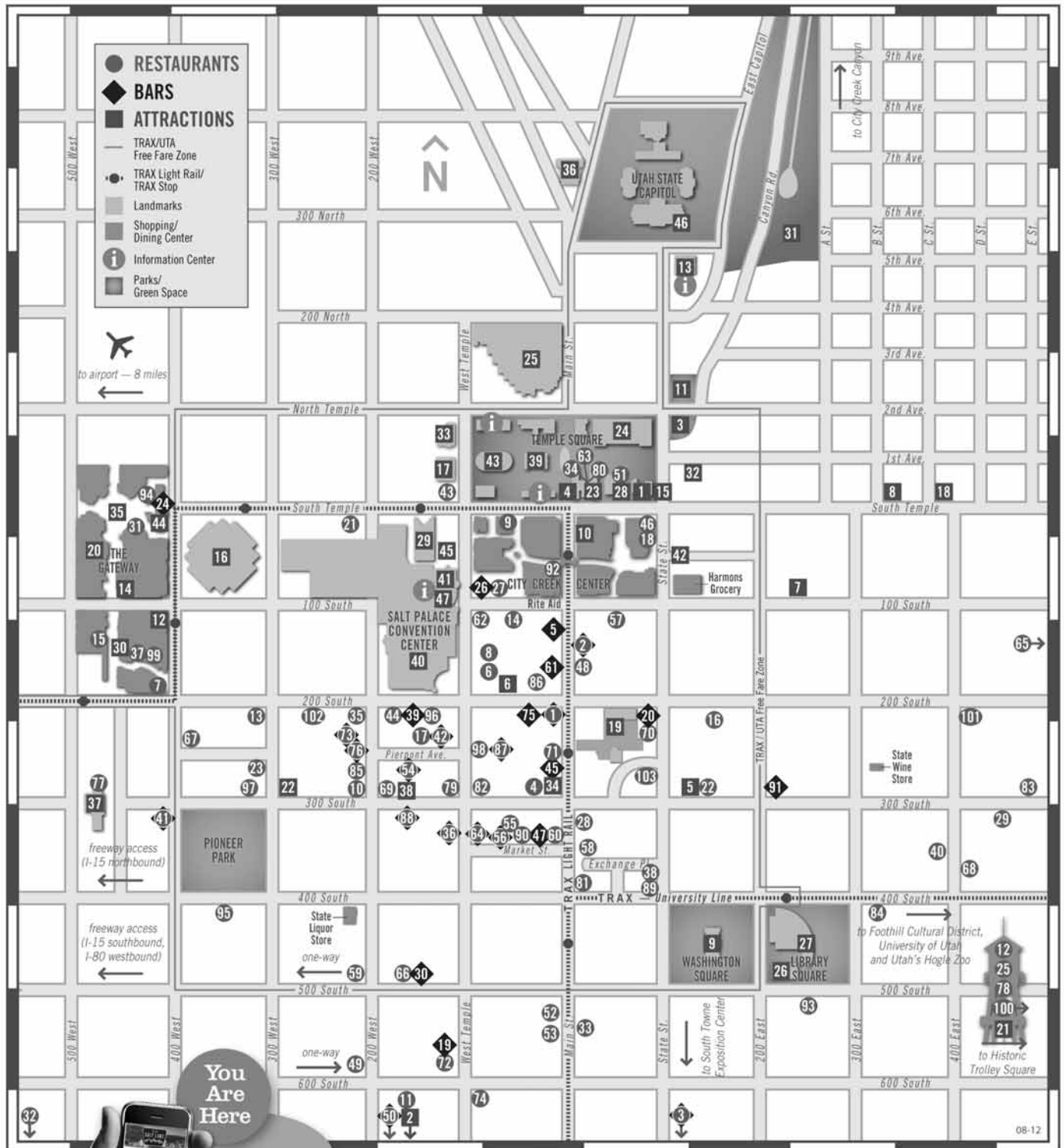
SCinet is the result of the hard work and significant contributions of many government, research, education and corporate collaborators. Collaborators for SC12 include:



Map/Daily Schedules

A schedule of each day's activities (by time/event/location) is provided in this section, along with a map of the Downtown area. A convention center map is located in the back of this booklet.

Salt Lake Downtown RESTAURANTS, BARS & ATTRACTIONS



Different by Nature.

- facebook.com/VisitSaltLake
- twitter.com/VisitSaltLake
- UPCOMING EVENTS
[VisitSaltLake.com/events](https://visitsaltlake.com/events)

Salt Lake Downtown



Visit Salt Lake

90 South West Temple, Salt Lake City, Utah
801-534-4900 info@visitsaltlake.com VisitSaltLake.com

RESTAURANTS & BARS

1 Bambara	801-363-5454	52 Little America Coffee Shop	801-596-5700
2 Bayleaf Bar & Grub	801-359-8490	53 Little America Steak House	801-596-5700
3 The Bayou	801-961-8400	54 Lumpys Downtown	801-938-3070
4 Beehive Tea Room & Wedding Library	801-328-4700	55 Market Street Grill - Downtown	801-322-4668
5 Beerhive Pub	801-364-4268	56 Market Street Oyster Bar - Downtown	801-531-6044
6 Benihana	801-322-2421	57 Martine	801-363-9328
7 Biaggi's Ristorante Italiano	801-596-7222	58 Maxwell's East Coast Eatery	801-328-0304
8 Blue Iguana Restaurant	801-533-8900	59 McDonald's	801-364-1614
9 Blue Lemon	801-328-2583	60 The Melting Pot	801-521-6358
10 Buca di Beppo	801-575-6262	61 Murphy's Bar and Grill	801-359-7271
11 Cafe Olympus	801-521-7373	62 Naked Fish Japanese Bistro	801-595-8888
12 Cafe SuperNatural	801-363-1000	63 Nauvoo Cafe	801-539-3346
13 Cafe Trang	801-359-1638	64 New Yorker	801-363-0166
14 Caffè Molise	801-364-8833	65 Oasis Cafe	801-322-0404
15 Canyon Creek Cafes Food Court		66 Olio Ristorante	801-323-7575
16 Cedars of Lebanon	801-364-4096	67 Pallet Bistro	801-935-4431
17 Christopher's Seafood & Steakhouse	801-519-8515	68 Papa John's Pizza	801-521-7272
18 City Creek Center Food Court		69 P. F. Chang's China Bistro	801-539-0500
19 Clouseau's	801-359-7800	70 Piastra on Gallivan Plaza	801-961-8700
20 Club Piastra	801-961-8700	71 Pier 49 Pizza	801-364-2974
21 Copper Canyon Grill House & Tavern	801-521-7800	72 The Plum Restaurant	801-359-7800
22 Copper Onion	801-355-3282	73 Poplar Street Pub	801-532-2715
23 Cucina Toscana	801-328-3463	74 Raw Bean Coffee House	801-990-2326
24 The Depot	801-456-2800	75 The Red Door	801-363-6030
25 Desert Edge Brewery at the Pub	801-521-8917	76 Red Rock Brewing Company L.C.	801-521-7446
26 Destinations	801-531-0800	77 Rio Grande Cafe	801-364-3302
27 Elevations Restaurant	801-537-6019	78 Rodizio Grill	801-220-0500
28 Eva	801-359-8447	79 Romano's Macaroni Grill	801-521-3133
29 Faustina	801-746-4441	80 The Roof Restaurant	801-539-1911
30 First Press	801-401-2000	81 Royal Eatery	801-532-4301
31 Fleming's	801-355-3704	82 Ruth's Chris Steak House	801-363-2000
Prime Steakhouse & Wine Bar		83 Sage's Cafe	801-322-3790
32 Frida Bistro	801-983-6692	84 Salt Lake Roasting Company	801-363-7572
33 The Garden Cafe	801-258-6708	85 Settebello Pizzeria Napoletana	801-322-3556
34 The Garden Restaurant	801-539-3170	86 Siegfried's Delicatessen, Inc.	801-355-3891
35 Ginza Japanese Cuisine & Sushi Bar	801-322-2224	87 Spencer's For Steaks & Chops	801-238-4748
88 Gracie's	801-819-7565	88 Squatters Pub Brewery	801-363-2739
37 Happy Sumo At Gateway	801-456-7866	89 Star of India Restaurant	801-363-7555
38 Himalayan Kitchen	801-328-2077	90 Takashi	801-519-9595
39 The Hotel Bar & Nightclub	801-487-4310	91 Tavernacle Social Club	801-519-8900
40 Ichiban Sushi & Japanese Cuisine	801-532-7522	92 Texas de Brazil Churrascaria	385-232-8070
41 Iggy's Sports Grill	801-532-9999	93 Thai Lotus Cafe	801-328-4401
42 Inferno Cantina	801-883-8838	94 Thaifoon - Taste of Asia	801-456-8424
43 JB's Family Restaurant	801-328-8344	95 The Tin Angel Cafe	801-328-4155
44 J. Wong's Asian Bistro	801-350-0888	96 Toaster's	801-328-2928
45 Keys On Main	801-363-3638	97 Tony Caputo's Market & Deli	801-328-0222
46 Kneader's Bakery and Cafe	801-428-3051	98 Trofi	801-238-4877
47 Kristauf's Martini Bar	801-366-9490	99 Tucanos Brazilian Grill	801-456-2550
48 Lamb's Grill Cafe	801-364-7166	100 Tucci's Cucina Italiana	801-533-9111
49 Last Samurai Japanese Steakhouse	801-596-2293	101 Vinto	801-539-9999
50 Legends Pub & Grill	801-355-3598	102 Vosen's Bread Paradise	801-322-2424
51 The Lion House Pantry Restaurant*	801-363-5466	103 ZY Restaurant	801-779-4730

ATTRACTIONS

1 Beehive House	801-240-2681
2 Brewvies Cinema Pub	801-355-5500
3 Brigham Young Historic Park	
4 Brigham Young Monument	
5 Broadway Centre Cinemas	801-321-0310
6 Capitol Theatre	801-355-2787
7 Cathedral Church of St. Mark	801-322-3400
8 Cathedral of the Madeleine	801-328-8941
9 City and County Building	801-533-0858
10 City Creek Center — Shopping Area	801-238-5320
11 City Creek Park	
12 Clark Planetarium	801-456-7827
13 Council Hall	801-538-1900
14 Discovery Gateway	801-456-5437
15 Eagle Gate	
16 EnergySolutions Arena	801-325-2000
17 Family History Library	801-240-2584
18 First Presbyterian Church	801-363-3889
19 Gallivan Center	801-535-6110
20 The Gateway — Shopping Area	801-456-0000
21 Historic Trolley Square — Shopping Area	801-521-9877
22 Holy Trinity Cathedral	801-328-9681
Greek Orthodox Church	
23 Joseph Smith Memorial Building, FamilySearch™ Center	801-240-1266
24 LDS Church Office Building	801-240-2190
25 LDS Conference Center	801-240-0075
26 The Leonardo	801-531-9800
27 Library Square	801-524-8200
28 Lion House	801-363-5466
29 Maurice Abravanel Hall	801-355-2787
30 Megaplex 12 at the Gateway	801-304-4553
31 Memory Grove Park	801-972-7800
32 Mormon Pioneer Memorial Monument	
33 Museum of Church History & Art	801-240-3310
34 The Off Broadway Theatre	801-355-4628
35 Olympic Legacy Plaza	
36 Pioneer Memorial Museum	801-532-6479
37 Rio Grande Depot, Utah State Historical Society	801-533-3500
38 Rose Wagner Performing Arts Center	801-323-6800
39 Salt Lake Temple	
40 Salt Palace Convention Center	801-534-4777
41 Simply Salt Lake Gift Shop	801-534-4906
42 Social Hall Heritage Exhibit	801-321-8745
43 Tabernacle	801-240-4872
44 Union Pacific Depot	
45 Utah Museum of Contemporary Art	801-328-4201
46 Utah State Capitol	801-538-3074
47 Visitor Information Center	801-534-4900

Visit Salt Lake Connect Pass attractions

Saturday, November 10

Time	Event	Title	Location
1pm-6pm	Information Booth	Main Booth	South Foyer
1pm-6:30pm	Coat/Bag Check		Lower Concourse

Sunday, November 11

Time	Event	Title	Location
7am-10pm	Coat/Bag Check		Lower Concourse
8am-5pm	Information Booth	Satellite Booth	Upper Concourse
8am-6pm	Information Booth	Main Booth	South Foyer
8:30am-10am	Broader Engagement, HPC Educators	Broader Engagement and Education in the Exascale Era	Ballroom-D
8:30am-12pm	Tutorials	How to Analyze the Performance of Parallel Codes 101	
8:30am-12pm	Tutorials	Hybrid MPI and OpenMP Parallel Programming	
8:30am-12pm	Tutorials	Large Scale Visualization with ParaView	
8:30am-12pm	Tutorials	Parallel Programming with Migratable Objects for Performance and Productivity	
8:30am-12pm	Tutorials	Productive Programming in Chapel: A Language for General, Locality-aware Parallelism	
8:30am-5pm	Tutorials	A Hands-On Introduction to OpenMP	
8:30am-5pm	Tutorials	Debugging MPI and Hybrid-Heterogenous Applications at Scale	
8:30am-5pm	Tutorials	Parallel I/O In Practice	
8:30am-5pm	Tutorials	Productive, Portable Performance on Accelerators Using OpenACC Compilers and Tools	
8:30am-5pm	Tutorials	Scalable Heterogeneous Computing on GPU Clusters	
8:30am-5pm	Tutorials	This Is Not Your Parents' Fortran: Object-Oriented Programming in Modern Fortran	
8:30am-5pm	Tutorials	Using Application Proxies for Co-design of Future HPC Computer Systems and Applications	
9am-5:30pm	Workshops	3rd Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems-Scala	
9am-5:30pm	Workshops	High Performance Computing, Networking and Analytics for the Power Grid	
9am-5:30pm	Workshops	HPCDB 2012-High-Performance Computing Meets Databases	
9am-5:30pm	Workshops	IA3 2012-Second Workshop on Irregular Applications-Architectures and Algorithms	
9am-5:30pm	Workshops	The Second International Workshop on Network-aware Data Management	
9am-5:30pm	Workshops	The Third International Workshop on Data-Intensive Computing in the Clouds-DataCloud	
9am-5:30pm	Workshops	Third Annual Workshop on Energy Efficient High Performance Computing-Redefining System Architecture and Data Centers	
10:30am-11:15am	Broader Engagement	The Importance of Broader Engagement for HPC	355-A
10:30am-12pm	HPC Educators	Supercomputing in Plain English	255-A
11:15am-12pm	Broader Engagement	Programming Exascale Supercomputers	355-A
1:30pm-2:15pm	Broader Engagement	An Unlikely Symbiosis: How the Gaming and Supercomputing Industries are Learning from and Influencing Each Other	355-A
1:30pm-5pm	HPC Educators	A Nifty Way to Introduce Parallelism into the Introductory Programming Sequence	255-A
1:30pm-5pm	Tutorials	An Overview of Fault-Tolerant Techniques for HPC	
1:30pm-5pm	Tutorials	Basics of Supercomputing	
1:30pm-5pm	Tutorials	C++ amP: An introduction to Heterogeneous Programming with C++	
1:30pm-5pm	Tutorials	Developing Scalable Parallel Applications in X10	

1:30pm-5pm	Tutorials	In-Situ Visualization with Catalyst	
1:30pm-5pm	HPC Educators	Introducing Computational Science in the Curriculum	255-D
1:30pm-5pm	Tutorials	Python in HPC	
2:15pm-3pm	Broader Engagement	L33t HPC: How Teh Titan's GPUs Pwned Science	355-A
3:30pm-4:15pm	Broader Engagement	Visual Computing-Making Sense of a Complex World	355-A
4:15pm-5pm	Broader Engagement	XSEDE (Extreme Science and Engineering Discovery Environment)	355-A
5:15pm-6:15pm	Broader Engagement, HPC Educators	SC12 Communities-Conference Orientation	Ballroom-D
6pm-9pm	Reception	Exhibitor Reception	The Hotel Elevate
7pm-10pm	BE, HPC Educators	Broader Engagement/HPC Educators Reception	Ballroom-A

Monday, November 12

Time	Event	Title	Location
7:30am-9:30pm	Coat/Bag Check		Lower Concourse
8am-5pm	Information Booth	Satellite Booth	Upper Concourse
8am-7pm	Information Booth	Main Booth	South Foyer
8:30am-10am	Broader Engagement, HPC Educators	The Fourth Paradigm-Data-Intensive Scientific Discovery	Ballroom-D
8:30am-12pm	Tutorials	InfiniBand and High-speed Ethernet for Dummies	
8:30am-12pm	Tutorials	Introduction to GPU Computing with OpenACC	
8:30am-12pm	Tutorials	Secure Coding Practices for Grid and Cloud Middleware and Services	
8:30am-12pm	Tutorials	Supporting Performance Analysis and Optimization on Extreme-Scale Computer Systems	
8:30am-5pm	Tutorials	A Tutorial Introduction to Big Data	
8:30am-5pm	Tutorials	Advanced MPI	
8:30am-5pm	Tutorials	Advanced OpenMP Tutorial	
8:30am-5pm	Tutorials	Developing and Tuning Parallel Scientific Applications in Eclipse	
8:30am-5pm	Tutorials	Infrastructure Clouds and Elastic Services for Science	
8:30am-5pm	Tutorials	Intro to PGAS-UPC and CAF-and Hybrid for Multicore Programming	
8:30am-5pm	Tutorials	Large Scale Visualization and Data Analysis with VisIt	
8:30am-5pm	Tutorials	Linear Algebra Libraries for High-Performance Computing: Scientific Computing with Multicore and Accelerators	
8:30am-5pm	Tutorials	The Practitioner's Cookbook for Good Parallel Performance on Multi- and Manycore Systems	
8:30am-5pm	Workshops, Broader Engagement	Broadening Engagement Workshop	
9am-5:30pm	Workshops	3rd International Workshop on Performance Modeling, Benchmarking and Simulation of High Performance Computing Systems	
9am-5:30pm	Workshops	3rd SC Workshop on Petascale Data Analytics: Challenges and Opportunities	
9am-5:30pm	Workshops	5th Workshop on High Performance Computational Finance	
9am-5:30pm	Workshops	7th Parallel Data Storage Workshop	
9am-5:30pm	Workshops	Climate Knowledge Discovery Workshop	
9am-5:30pm	Workshops	The 5th Workshop on Many-Task Computing on Grids and Supercomputers (MTAGS) 2012	
9am-5:30pm	Workshops	The 7th Workshop on Ultrascale Visualization	
9am-5:30pm	Workshops	The Seventh Workshop on Workflows in Support of Large-Scale Science-WORKS12	

10am-6pm	Exhibits	Exhibit Hall	
10am-6pm	Preview Booth	SC13 Preview Booth	South Foyer
10:30am-11:15am	Broader Engagement	The Sequoia System and Facilities Integration Story	355-A
10:30am-12pm	HPC Educators	Python for Parallelism in Introductory Computer Science Education	255-D
10:30am-5pm	HPC Educators	LittleFe Buildout Workshop	255-A
11:15am-12pm	Broader Engagement	Using Power Efficient ARM-Based Servers for Data Intensive HPC	355-A
1:30pm-2:15pm	Broader Engagement	OpenMP: The “Easy” Path to Shared Memory Computing	355-A
1:30pm-5pm	Tutorials	Advanced GPU Computing with OpenACC	
1:30pm-5pm	Tutorials	Asynchronous Hybrid and Heterogeneous Parallel Programming with MPI/OmpSs for Exascale Systems	
1:30pm-5pm	Tutorials	Designing High-End Computing Systems with InfiniBand and High-Speed Ethernet	
1:30pm-5pm	HPC Educators	Going Parallel with C++11	255-D
1:30pm-5pm	Tutorials	The Global Arrays Toolkit-A Comprehensive, Production-Level, Application-Tested Parallel Programming Environment	
2:15pm-3pm	Broader Engagement	OpenACC, An Effective Standard for Developing Performance Portable Applications for Future Hybrid Systems	355-A
3:30pm-4:15pm	Broader Engagement	The Growing Power Struggle in HPC	355-A
4:15pm-5pm	Broader Engagement	Heading Towards Exascale-Techniques to Improve Application Performance and Energy Consumption Using Application-Level Tools	355-A
5pm-7pm	Broader Engagement	Mentor-Protégé Mixer	Ballroom-A
6pm-7pm	Other Event	Experiencing HPC for Undergraduates-Welcome and Orientation	250-AB
7pm-9pm	Reception	Opening Gala	Exhibit Hall

Tuesday, November 13

Time	Event	Title	Location
7:30am-5:30pm	Information Booth	Satellite Booth	Lower Concourse
7:30am-6pm	Information Booth	Main Booth	South Foyer
7:30am-7:30pm	Coat/Bag Check		South Lobby
8:30am-10am	Keynote	Keynote: Physics of the Future Dr. Michio Kaku, City University of New York	Ballroom-CDEFGH
10:30am-11am	Paper	Demonstrating Lustre over a 100Gbps Wide Area Network of 3,500km	355-EF
10:30am-11am	Paper	Portable Section-Level Tuning of Compiler Parallelized Applications	355-D
10:30am-11am	Paper (Best Student Paper Finalist)	Direction-Optimizing Breadth-First Search	255-EF
10:30am-11am	Paper	Hybridizing S3D into an Exascale Application using OpenACC	255-BC
10:30am-11am	Exhibitor Forum	Taking HPC to the Cloud-Overcoming Complexity and Accelerating Time-to-Results with Unlimited Compute	155-C
10:30am-11am	Exhibitor Forum	PCI Express as a Data Center Fabric	155-B
10:30am-11:15am	Invited Talk	The Sequoia System and Facilities Integration Story	Ballroom-EFGH
10:30am-12pm	Other Event	Experiencing HPC for Undergraduates-Introduction to HPC	250-AB
10:30am-12pm	Panels	HPC’s Role In The Future of American Manufacturing	355-BC
10:30am-12pm	HPC Educators	Invited Talk: TCPP Parallel and Distributed Curriculum Initiative	255-D
10:30am-12pm	Broader Engagement	Mentoring: Building Functional Professional Relationships	355-A
10:30am-12pm	HPC Educators	Unveiling Parallelization Strategies at Undergraduate Level	255-A
11am-11:30am	Paper	A Study on Data Deduplication in HPC Storage Systems	355-EF
11am-11:30am	Paper	A Multi-Objective Auto-Tuning Framework for Parallel Codes	355-D

Tuesday, November 13

Time	Event	Title	Location
11am-11:30am	Paper	Breaking the Speed and Scalability Barriers for Graph Exploration on Distributed-Memory Machines	255-EF
11am-11:30am	Paper	High Throughput Software for Direct Numerical Simulations of Compressible Two-Phase Flows	255-BC
11am-11:30am	Exhibitor Forum	HPC Cloud ROI and Opportunities Cloud Brings to HPC	155-C
11am-11:30am	Exhibitor Forum	Mellanox Technologies – Paving the Road to Exascale Computing	155-B
11:15am-12pm	Invited Talk	Titan-Early Experience with the Titan System at Oak Ridge National Laboratory	Ballroom-EFGH
11:30am-12pm	Paper (Best Student Paper Finalist, Best Paper Finalist)	Characterizing Output Bottlenecks in a Supercomputer	355-EF
11:30am-12pm	Paper	PATUS for Convenient High-Performance Stencils: Evaluation in Earthquake Simulations	355-D
11:30am-12pm	Paper	Large-Scale Energy-Efficient Graph Traversal-A Path to Efficient Data-Intensive Supercomputing	255-EF
11:30am-12pm	Exhibitor Forum	The Technical Cloud: When Remote 3D Visualization Meets HPC	155-C
11:30am-12pm	Exhibitor Forum	Affordable Shared Memory for Big Data	155-B
12:15pm-1:15pm	Birds of a Feather	ACM SIGHPC First Annual Members Meeting	155-E
12:15pm-1:15pm	Birds of a Feather	Collaborative Opportunities with the Open Science Data Cloud	250-DE
12:15pm-1:15pm	Birds of a Feather	Data and Software Preservation for Big-Data Science Collaborations	255-A
12:15pm-1:15pm	Birds of a Feather	Exascale IO Initiative: Progress Status	155-F
12:15pm-1:15pm	Birds of a Feather	Fifth Graph500 List	255-BC
12:15pm-1:15pm	Birds of a Feather	Genomics Research Computing: The Engine that Drives Personalized Medicine Forward	251-A
12:15pm-1:15pm	Birds of a Feather	HDF5: State of the Union	250-C
12:15pm-1:15pm	Birds of a Feather	How the Government can enable HPC and emerging technologies	355-D
12:15pm-1:15pm	Birds of a Feather	Implementing Parallel Environments: Training and Education	251-D
12:15pm-1:15pm	Birds of a Feather	Interoperability in Scientific Cloud Federations	250-AB
12:15pm-1:15pm	Birds of a Feather	MPICH: A High-Performance Open-Source MPI Implementation	155-B
12:15pm-1:15pm	Birds of a Feather	Network Measurement	255-EF
12:15pm-1:15pm	Birds of a Feather	Obtaining Bitwise Reproducible Results-Perspectives and Latest Advances	251-E
12:15pm-1:15pm	Birds of a Feather	OpenACC API Status and Future	255-D
12:15pm-1:15pm	Birds of a Feather	Parallel and Accelerated Computing Experiences for Successful Industry Careers in High-Performance Computing	251-F
12:15pm-1:15pm	Birds of a Feather	Python for High Performance and Scientific Computing	155-C
12:15pm-1:15pm	Birds of a Feather	Scalable Adaptive Graphics Environment (SAGE) for Global Collaboration	251-C
12:15pm-1:15pm	Birds of a Feather	System wide Programming Models for Exascale	355-BC
12:15pm-1:15pm	Birds of a Feather	The 2012 HPC Challenge Awards	355-A
1:30pm-2pm	ACM Gordon Bell Finalists	Billion-Particle SIMD-Friendly Two-Point Correlation on Large-Scale HPC Cluster Systems	155-E
1:30pm-2pm	Paper (Best Student Paper Finalist)	McrEngine-A Scalable Checkpointing System Using Data-Aware Aggregation and Compression	255-EF
1:30pm-2pm	Paper	Scalia: An Adaptive Scheme for Efficient Multi-Cloud Storage	355-D
1:30pm-2pm	Paper	Early Evaluation of Directive-Based GPU Programming Models for Productive Exascale Computing	355-EF
1:30pm-2pm	Exhibitor Forum	Transforming HPC Yet Again with NVIDIA Kepler GPUs	155-B
1:30pm-2pm	Paper	Unleashing the High Performance and Low Power of Multi-Core DSPs for General-Purpose HPC	255-BC
1:30pm-2pm	Exhibitor Forum	Faster, Better, Easier Tools: The Shortcut to Results	155-C

Tuesday, November 13

Time	Event	Title	Location
1:30pm-2:15pm	Invited Talk	Pushing Water Up Mountains: Green HPC and Other Energy Oddities	Ballroom-EFGH
1:30pm-3pm	Broader Engagement	Impact of the BE Program-Lessons Learned and Broad Applicability	355-A
1:30pm-5pm	HPC Educators	GPU Computing as a Pathway to System-conscious Programmers	255-A
1:30pm-5pm	HPC Educators	Test-Driven Development for HPC Computational Science & Engineering	255-D
2pm-2:30pm	ACM Gordon Bell Finalist	Toward Real-Time Modeling of Human Heart Ventricles at Cellular Resolution: Simulation of Drug-Induced Arrhythmias	155-E
2pm-2:30pm	Paper	Alleviating Scalability Issues of Checkpointing Protocols	255-EF
2pm-2:30pm	Paper	Host Load Prediction in a Google Compute Cloud with a Bayesian Model	355-D
2pm-2:30pm	Paper	Automatic Generation of Software Pipelines for Heterogeneous Parallel Systems	355-EF
2pm-2:30pm	Exhibitor Forum	Addressing Big Data Challenges with Hybrid-Core Computing	155-B
2pm-2:30pm	Paper	A Scalable, Numerically Stable, High-Performance Tridiagonal Solver Using GPUs	255-BC
2pm-2:30pm	Exhibitor Forum	Scalable Debugging with TotalView for Xeon Phi, BlueGene/Q, and more	155-C
2:15pm-3pm	Invited Talk	The Costs of HPC-Based Science in the Exascale Era	Ballroom-EFGH
2:30pm-3pm	ACM Gordon Bell Finalist	Extreme-Scale UQ for Bayesian Inverse Problems Governed by PDEs	155-E
2:30pm-3pm	Paper	Design and Modeling of a Non-Blocking Checkpointing System	255-EF
2:30pm-3pm	Paper	Cost- and Deadline-Constrained Provisioning for Scientific Workflow Ensembles in IaaS Clouds	355-D
2:30pm-3pm	Paper	Accelerating MapReduce on a Coupled CPU-GPU Architecture	355-EF
2:30pm-3pm	Exhibitor Forum	Flash Memory and GPGPU Supercomputing: A Winning Combination	155-B
2:30pm-3pm	Paper (Best Paper Finalist)	Efficient Backprojection-Based Synthetic Aperture Radar Computation with Many-Core Processors	255-BC
2:30pm-3pm	Exhibitor Forum	Advanced Programming of Many-Core Systems Using CAPS OpenACC Compiler	155-C
3:30pm-4pm	Paper	Parametric Flows-Automated Behavior Equivalencing for Symbolic Analysis of Races in CUDA Programs	255-BC
3:30pm-4pm	Paper	RamZzz: Rank-Aware DRam Power Management with Dynamic Migrations and Demotions	355-D
3:30pm-4pm	Paper	Protocols for Wide-Area Data-Intensive Applications-Design and Performance Issues	255-EF
3:30pm-4pm	Exhibitor Forum	How Memory and SSDs can Optimize Data Center Operations	155-C
3:30pm-4pm	Exhibitor Forum	Integrating ZFS RAID with Lustre Today	155-B
3:30pm-4pm	Paper (Best Student Paper Finalist)	A Divide and Conquer Strategy for Scaling Weather Simulations with Multiple Regions of Interest	355-EF
3:30pm-4:15pm	Invited Talk	Communication-Avoiding Algorithms for Linear Algebra and Beyond	Ballroom-EFGH
3:30pm-5pm	Panel	NSF-TCPP Curriculum Initiative on Parallel and Distributed Computing-Core Topics for Undergraduates	355-BC
4pm-4:30pm	Paper (Best Paper Finalist)	MPI Runtime Error Detection with MUST-Advances in Deadlock Detection	255-BC
4pm-4:30pm	Paper	MAGE-Adaptive Granularity and ECC for Resilient and Power Efficient Memory Systems	355-D
4pm-4:30pm	Paper	High Performance RDMA-Based Design of HDFS over InfiniBand	255-EF
4pm-4:30pm	Exhibitor Forum	Beyond von Neumann With a 1 Million Element Massively Parallel Cognitive Memory	155-C
4pm-4:30pm	Exhibitor Forum	The End of Latency: A New Storage Architecture	155-B
4pm-4:30pm	Paper	Forward and Adjoint Simulations of Seismic Wave Propagation on Emerging Large-Scale GPU Architectures	355-EF
4:15pm-5pm	Invited Talk	Stochastic Simulation Service-Towards an Integrated Development Environment for Modeling and Simulation of Stochastic Biochemical Systems	Ballroom-EFGH
4:30pm-5pm	Paper	Novel Views of Performance Data to Analyze Large-Scale Adaptive Applications	255-BC

Tuesday, November 13

Time	Event	Title	Location
4:30pm-5pm	Paper (Best Student Paper Finalist)	Efficient and Reliable Network Tomography in Heterogeneous Networks Using BitTorrent Broadcasts and Clustering Algorithms	255-EF
4:30pm-5pm	Exhibitor Forum	Hybrid Memory Cube (HMC): A New Paradigm for System Architecture Design	155-C
4:30pm-5pm	Exhibitor Forum	The Expanding Role of Solid State Technology in HPC Storage Applications	155-B
5:15pm-7pm	Receptions	Research Exhibits, Poster Exhibits, ACM SRC Receptions	East Entrance
5:15pm-7pm	Reception	Scientific Visualization Showcase Reception	North Foyer
5:15pm-7pm	ACM SRC Competition	On the Cost of a General GPU Framework-The Strange Case of CUDA 4.0 vs. CUDA 5.0	East Entrance
5:15pm-7pm	ACM SRC Competition	High Quality Real-Time Image-to-Mesh Conversion for Finite Element Simulations	East Entrance
5:15pm-7pm	ACM SRC Competition	Optimus: A Parallel Optimization Framework With Topology Aware PSO and Applications	East Entrance
5:15pm-7pm	ACM SRC Competition	An MPI Library Implementing Direct Communication for Many-Core Based Accelerators	East Entrance
5:15pm-7pm	ACM SRC Competition	Massively Parallel Model of Evolutionary Game Dynamics	East Entrance
5:15pm-7pm	ACM SRC Competition	Scalable Cooperative Caching with RDMA-Based Directory Management for Large-Scale Data Processing	East Entrance
5:15pm-7pm	ACM SRC Competition	An Ultra-Fast Computing Pipeline for Metagenome Analysis with Next-Generation DNA Sequencers	East Entrance
5:15pm-7pm	ACM SRC Competition	Reducing the Migration Times of Multiple VMs on WANS	East Entrance
5:15pm-7pm	ACM SRC Competition	Performing Cloud Computation on a Parallel File System	East Entrance
5:15pm-7pm	ACM SRC Competition	Crayons: An Azure Cloud Based Parallel System for GIS Overlay Operations	East Entrance
5:15pm-7pm	ACM SRC Competition	Pay as You Go in the Cloud: One Watt at a Time	East Entrance
5:15pm-7pm	ACM SRC Competition	Optimizing pF3D using Model-Based, Dynamic Parallelism	East Entrance
5:15pm-7pm	ACM SRC Competition	Norm-Coarsened Ordering for Parallel Incomplete Cholesky Preconditioning	East Entrance
5:15pm-7pm	ACM SRC Competition	Neural Circuit Simulation of Hodgkin-Huxley Event Neurons Toward Peta Scale Computers	East Entrance
5:15pm-7pm	Poster	Matrices Over Runtime Systems at Exascale	East Entrance
5:15pm-7pm	Poster	Assessing the Predictive Capabilities of Mini-applications	East Entrance
5:15pm-7pm	Poster	Towards Highly Accurate Large-Scale Ab Initio Calculations Using Fragment Molecular Orbital Method in GamESS	East Entrance
5:15pm-7pm	Poster	Acceleration of the BLAST Hydro Code on GPU	East Entrance
5:15pm-7pm	Poster	A Novel Hybrid CPU-GPU Generalized Eigensolver for Electronic Structure Calculations Based on Fine Grained Memory Aware Tasks	East Entrance
5:15pm-7pm	Poster	HTCaaS: A Large-Scale High-Throughput Computing by Leveraging Grids, Supercomputers and Cloud	East Entrance
5:15pm-7pm	Poster	Evaluation of Magneto-Hydro-Dynamic Simulation on Three Events of Scalar	East Entrance
5:15pm-7pm	Poster	Three Steps to Model Power-Performance Efficiency for Emergent GPU-Based Parallel Systems	East Entrance
5:15pm-7pm	Poster	Impact of Integer Instructions in Floating Point Applications	East Entrance
5:15pm-7pm	Poster	Operating System Assisted Hierarchical Memory Management for Heterogeneous Architectures	East Entrance
5:15pm-7pm	Poster	The MPACK-Arbitrary Accurate Version of BLAS and LAPACK	East Entrance
5:15pm-7pm	Poster	Scalable Direct Eigenvalue Solver ELPA for Symmetric Matrices	East Entrance
5:15pm-7pm	Poster	Hybrid Breadth First Search Implementation for Hybrid-Core Computers	East Entrance
5:15pm-7pm	Poster	Interface for Performance Environment Autoconfiguration Framework	East Entrance
5:15pm-7pm	Poster	Imaging Through Cluttered Media Using Electromagnetic Interferometry on a Hardware-Accelerated High-Performance Cluster	East Entrance
5:15pm-7pm	Poster	Memory-Conscious Collective IO for Extreme-Scale HPC Systems	East Entrance

Tuesday, November 13

Time	Event	Title	Location
5:15pm-7pm	Poster	Visualization Tool for Development of Topology-Aware Network Communication Algorithm	East Entrance
5:15pm-7pm	Poster	Multi-GPU-Based Calculation of Percolation Problem on the TSUBamE 2.0 Supercomputer	East Entrance
5:15pm-7pm	Poster	Beating MKL and ScaLAPACK at Rectangular Matrix Multiplication Using the BFS/DFS Approach	East Entrance
5:15pm-7pm	Poster	Evaluating Topology Mapping via Graph Partitioning	East Entrance
5:15pm-7pm	Poster	Communication Overlap Techniques for Improved Strong Scaling of Gyrokinetic Eulerian Code Beyond 100k Cores on K-Computer	East Entrance
5:15pm-7pm	Poster	Polarization Energy On a Cluster of Multicores	East Entrance
5:15pm-7pm	Poster	Exploring Performance Data with Boxfish	East Entrance
5:15pm-7pm	Poster	Reservation-Based I/O Performance Guarantee for MPI-IO Applications using Shared Storage Systems	East Entrance
5:15pm-7pm	Poster	Visualizing and Mining Large Scale Scientific Data Provenance	East Entrance
5:15pm-7pm	Poster	Using Active Storage Concept for Seismic Data Processing	East Entrance
5:15pm-7pm	Poster	Slack-Conscious Lightweight Loop Scheduling for Scaling Past the Noise amplification Problem	East Entrance
5:15pm-7pm	Poster	Solving the Schroedinger and Dirac Equations of Atoms and Molecules with Massively Parallel Supercomputer	East Entrance
5:15pm-7pm	Poster	Leveraging PEPPIER Technology for Performance Portable Supercomputing	East Entrance
5:15pm-7pm	Poster	Networking Research Activities at Fermilab for Big Data Analysis	East Entrance
5:15pm-7pm	Poster	Collective Tuning: Novel Extensible Methodology, Framework and Public Repository to Collaboratively Address Exascale Challenges	East Entrance
5:15pm-7pm	Poster	High-Speed Decision Making on Live Petabyte Data Streams	East Entrance
5:15pm-7pm	Poster	Gossip-Based Distributed Matrix Computations	East Entrance
5:15pm-7pm	Poster	Scalable Fast Multipole Methods for Vortex Element Methods	East Entrance
5:15pm-7pm	Poster	PLFS/HDFS: HPC Applications on Cloud Storage	East Entrance
5:15pm-7pm	Poster	High Performance GPU Accelerated TSP Solver	East Entrance
5:15pm-7pm	Poster	Speeding-Up Memory Intensive Applications Through Adaptive Hardware Accelerators	East Entrance
5:15pm-7pm	Poster	FusedOS: A Hybrid Approach to Exascale Operating Systems	East Entrance
5:15pm-7pm	Poster	Using Provenance to Visualize Data from Large-Scale Experiments	East Entrance
5:15pm-7pm	Poster	Cascaded TCP: Big Throughput for Big Data Applications in Distributed HPC	East Entrance
5:15pm-7pm	Poster	Automatically Adapting Programs for Mixed-Precision Floating-Point Computation	East Entrance
5:15pm-7pm	Poster	MAAPED: A Predictive Dynamic Analysis Tool for MPI Applications	East Entrance
5:15pm-7pm	Poster	Memory and Parallelism Exploration using the LULESH Proxy Application	East Entrance
5:15pm-7pm	Poster	Auto-Tuning of Parallel I/O Parameters for HDFS Applications	East Entrance
5:15pm-7pm	Poster	Uintah Framework Hybrid Task-Based Parallelism Algorithm	East Entrance
5:15pm-7pm	Poster	Programming Model Extensions for Resilience in Extreme Scale Computing	East Entrance
5:15pm-7pm	Poster	Seismic Imaging on Blue Gene/Q	East Entrance
5:15pm-7pm	Poster	Using Business Workflows to Improve Quality of Experiments in Distributed Systems Research	East Entrance
5:15pm-7pm	Poster	Build to Order Linear Algebra Kernels	East Entrance
5:15pm-7pm	Poster	Distributed Metadata Management for Exascale Parallel File System	East Entrance

Tuesday, November 13

Time	Event	Title	Location
5:15pm-7pm	Poster	Advances in Gyrokinetic Particle-in-Cell Simulation for Fusion Plasmas to Extreme Scale	East Entrance
5:15pm-7pm	Poster	The Hashed Oct-Tree N-Body Algorithm at a Petaflop	East Entrance
5:15pm-7pm	Poster	Asynchronous Computing for Partial Differential Equations at Extreme Scales	East Entrance
5:15pm-7pm	Poster	GPU Accelerated Ultrasonic Tomography Using Propagation and Backpropagation Method	East Entrance
5:15pm-7pm	Poster	Application Restructuring for Vectorization and Parallelization: A Case Study	East Entrance
5:15pm-7pm	Poster	Parallel Algorithms for Counting Triangles and Computing Clustering Coefficients	East Entrance
5:15pm-7pm	Poster	Improved OpenCL Programmability with clUtil	East Entrance
5:15pm-7pm	Poster	Hadoop's Adolescence: A Comparative Workload Analysis from Three Research Clusters	East Entrance
5:15pm-7pm	Poster	Preliminary Report for a High Precision Distributed Memory Parallel Eigenvalue Solver	East Entrance
5:15pm-7pm	Poster	Analyzing Patterns in Large-Scale Graphs Using MapReduce in Hadoop	East Entrance
5:15pm-7pm	Poster	Digitization and Search: A Non-Traditional Use of HPC	East Entrance
5:15pm-7pm	Poster	An Exascale Workload Study	East Entrance
5:15pm-7pm	Poster	Visualization for High-Resolution Ocean General Circulation Model via Multi-Dimensional Transfer Function and Multivariate Analysis	East Entrance
5:15pm-7pm	Poster	Portals 4 Network Programming Interface	East Entrance
5:15pm-7pm	Poster	Quantum Mechanical Simulations of Crystalline Helium Using High Performance Architectures	East Entrance
5:15pm-7pm	Poster	Multiple Pairwise Sequence Alignments with the Needleman-Wunsch Algorithm on GPU	East Entrance
5:15pm-7pm	Poster	GenASIS: An Object-Oriented Approach to High Performance Multiphysics Code with Fortran 2003	East Entrance
5:15pm-7pm	Poster	Exploring Design Space of a 3D Stacked Vector Cache	East Entrance
5:15pm-7pm	Poster	A Disc-Based Decomposition Algorithm with Optimal Load Balancing for N-body Simulations	East Entrance
5:15pm-7pm	Poster	Remote Visualization for Large-Scale Simulation using Particle-Based Volume Rendering	East Entrance
5:15pm-7pm	Poster	Tracking and Visualizing Evolution of the Universe: In Situ Parallel Dark Matter Halo Merger Trees	East Entrance
5:15pm-7pm	Poster	Autonomic Modeling of Data-Driven Application Behavior	East Entrance
5:15pm-7pm	Poster	Automated Mapping Streaming Applications onto GPUs	East Entrance
5:15pm-7pm	Poster	Planewave-Based First-Principles MD Calculation on 80,000-Node K-computer	East Entrance
5:15pm-7pm	Poster	Bringing Task- and Data-Parallelism to Analysis of Climate Model Output	East Entrance
5:15pm-7pm	Poster	Evaluating Asynchrony in Gibraltar RAID's GPU Reed-Solomon Coding Library	East Entrance
5:15pm-7pm	Poster	Matrix Decomposition Based Conjugate Gradient Solver for Poisson Equation	East Entrance
5:15pm-7pm	Poster	Evaluating the Error Resilience of GPGPU Applications	East Entrance
5:15pm-7pm	Poster	Comparing GPU and Increment-Based Checkpoint Compression	East Entrance
5:15pm-7pm	Poster	The Magic Determination of the Magic Constants by ttgLib Autotuner	East Entrance
5:15pm-7pm	Poster	MemzNet: Memory-Mapped Zero-copy Network Channel for Moving Large Datasets over 100Gbps Networks	East Entrance
5:15pm-7pm	Poster	Evaluating Communication Performance in Supercomputers BlueGene/Q and Cray XE6	East Entrance
5:15pm-7pm	Poster	Statistical Power and Energy Modeling of Multi-GPU kernels	East Entrance
5:15pm-7pm	Poster	Virtual Machine Packing Algorithms for Lower Power Consumption	East Entrance
5:15pm-7pm	Poster	PanDA: Next Generation Workload Management and Analysis System for Big Data	East Entrance
5:15pm-7pm	Poster	Numerical Studies of the Klein-Gordon Equation in a Periodic Setting	East Entrance
5:15pm-7pm	SciViz Showcase	Computing the Universe-From Big Bang to Stars	North Foyer

Tuesday, November 13

Time	Event	Title	Location
5:15pm-7pm	SciViz Showcase	Investigation of Turbulence in the Early Stages of a High Resolution Supernova Simulation	North Foyer
5:15pm-7pm	SciViz Showcase	Two Fluids Level Set: High Performance Simulation and Post Processing	North Foyer
5:15pm-7pm	SciViz Showcase	SiO ₂ Fissure in Molecular Dynamics	North Foyer
5:15pm-7pm	SciViz Showcase	Direct Numerical Simulations of Cosmological Reionization: Field Comparison: Density	North Foyer
5:15pm-7pm	SciViz Showcase	Direct Numerical Simulations of Cosmological Reionization: Field Comparison: Ionization Fraction	North Foyer
5:15pm-7pm	SciViz Showcase	Direct Numerical Simulation of Flow in Engine-Like Geometries	North Foyer
5:15pm-7pm	SciViz Showcase	Cosmology on the Blue Waters Early Science System	North Foyer
5:15pm-7pm	SciViz Showcase	Explosive Charge Blowing a Hole in a Steel Plate Animation	North Foyer
5:15pm-7pm	SciViz Showcase	Computational Fluid Dynamics and Visualization	North Foyer
5:15pm-7pm	SciViz Showcase	Effect of Installation Geometry on Turbulent Mixing Noise from Jet Engine Exhaust	North Foyer
5:15pm-7pm	SciViz Showcase	Virtual Rheoscopic Fluid for Large Dynamics Visualization	North Foyer
5:15pm-7pm	SciViz Showcase	Inside Views of a Rapidly Spinning Star	North Foyer
5:15pm-7pm	SciViz Showcase	A Dynamic Portrait of Global Aerosols	North Foyer
5:15pm-7pm	SciViz Showcase	Probing the Effect of Conformational Constraints on Binding	North Foyer
5:15pm-7pm	SciViz Showcase	In-Situ Feature Tracking and Visualization of a Temporal Mixing Layer	North Foyer
5:30pm-7pm	Birds of a Feather	Computing Research Testbeds as a Service: Supporting Large-scale Experiments and Testing	251-E
5:30pm-7pm	Birds of a Feather	Critically Missing Pieces in Heterogeneous Accelerator Computing	155-A
5:30pm-7pm	Birds of a Feather	Cyber Security's Big Data, Graphs, and Signatures	250-AB
5:30pm-7pm	Birds of a Feather	Energy Efficient High Performance Computing	155-C
5:30pm-7pm	Birds of a Feather	Exascale Research – The European Approach	255-A
5:30pm-7pm	Birds of a Feather	High Performance Computing Programming Techniques For Big Data Hadoop	251-F
5:30pm-7pm	Birds of a Feather	High Productivity Languages for High Performance Computing	251-B
5:30pm-7pm	Birds of a Feather	High-level Programming Models for Computing Using Accelerators	250-DE
5:30pm-7pm	Birds of a Feather	HPC Cloud: Can Infrastructure Clouds Provide a Viable Platform for HPC?	355-BC
5:30pm-7pm	Birds of a Feather	HPC Runtime System Software	255-EF
5:30pm-7pm	Birds of a Feather	Hybrid Programming with Task-based Models	251-C
5:30pm-7pm	Birds of a Feather	Large-Scale Reconfigurable Supercomputing	255-D
5:30pm-7pm	Birds of a Feather	Managing Big Data: Best Practices from Industry and Academia	255-BC
5:30pm-7pm	Birds of a Feather	OCI-led Activities at NSF	155-E
5:30pm-7pm	Birds of a Feather	OpenMP: Next Release and Beyond	355-A
5:30pm-7pm	Birds of a Feather	Policies and Practices to Promote a Diverse Workforce	253
5:30pm-7pm	Birds of a Feather	Scientific Application Performance in Heterogeneous Supercomputing Clusters	155-F
5:30pm-7pm	Birds of a Feather	SPEC HPG Benchmarks For Next Generation Systems	155-B
5:30pm-7pm	Birds of a Feather	The Apache Software Foundation, Cyberinfrastructure, and Scientific Software: Beyond Open Source	251-A
5:30pm-7pm	Birds of a Feather	The Eclipse Parallel Tools Platform	250-C
5:30pm-7pm	Birds of a Feather	TOP500 Supercomputers	Ballroom-EFGH
5:30pm-7pm	Birds of a Feather	TORQUE; Rpm, Cray and MIC	251-D
5:30pm-7pm	Birds of a Feather	XSEDE User Meeting	355-D

Wednesday, November 14

Time	Event	Title	Location
7:30am-6pm	Coat/Bag Check		Lower Concourse
8am-5:30pm	Information Booth	Satellite Booth	Upper Concourse
8am-6pm	Information Booth	Main Booth	South Foyer
8:30am-9:15am	Invited Talk	Simulating the Human Brain-An Extreme Challenge for Computing	Ballroom-EFGH
9:15am-10am	Invited Talk	The K Computer-Toward Its Productive Application to Our Life	Ballroom-EFGH
10am-3pm	Broader Engagement	Student Job/Opportunity Fair	251-ABCDEF
10am-6pm	Exhibits	Exhibit Hall	
10am-6pm	Preview Booth	SC13 Preview Booth	South Foyer
10:30am-10:45am	Doctoral Showcase	Analyzing and Reducing Silent Data Corruptions Caused By Soft-Errors	155-F
10:30am-11am	Paper	Bamboo-Translating MPI Applications to a Latency-Tolerant, Data-Driven Form	255-EF
10:30am-11am	Paper, Best Paper Finalists	A Framework for Low-Communication 1-D FFT	255-BC
10:30am-11am	Award	Kennedy Award Recipient: Mary Lou Soffa	155-E
10:30am-11am	Paper	Petascale Lattice Quantum Chromodynamics on a Blue Gene/Q Supercomputer	355-EF
10:30am-11am	Exhibitor Forum	Hybrid Solutions with a Vector-Architecture for Efficiency	155-B
10:30am-11am	Paper	Byte-Precision Level of Detail Processing for Variable Precision Analytics	355-D
10:30am-11am	Exhibitor Forum	Create Flexible Systems As Your Workload Requires	155-C
10:30am-11:15am	Invited Talk	The Long Term Impact of Codesign	Ballroom-EFGH
10:30am-12pm	Panel	Boosting European HPC Value Chain: the vision of ETP4HPC, the European Technology Platform for High Performance Computing	355-BC
10:30am-12pm	HPC Educators	Computational Examples for Physics (and Other) Classes Featuring Python, Mathematica, an eTextBook and More	255-A
10:30am-12pm	Other Event	Experiencing HPC for Undergraduates-Graduate Student Perspective	250-AB
10:30am-5pm	HPC Educators	An Educators's Toolbox for CUDA	255-D
10:45am-11am	Doctoral Showcase	Fast Multipole Methods on Heterogeneous Architectures	155-F
11am-11:15am	Doctoral Showcase	Algorithmic Approaches to Building Robust Applications for HPC Systems of the Future	155-F
11am-11:30am	Paper	Tiling Stencil Computations to Maximize Parallelism	255-EF
11am-11:30am	Paper	Parallel Geometric-Algebraic Multigrid on Unstructured Forests of Octrees	255-BC
11am-11:30am	Award	Seymour Cray Award Recipient: Peter Kogge	155-E
11am-11:30am	Paper	Massively Parallel X-Ray Scattering Simulations	355-EF
11am-11:30am	Exhibitor Forum	Appro's Next Generation Xtreme-X Supercomputer	155-B
11am-11:30am	Paper	Combining In-Situ and In-Transit Processing to Enable Extreme-Scale Scientific Analysis	355-D
11am-11:30am	Exhibitor Forum	The OpenOnload User-level Network Stack	155-C
11:15am-11:30am	Doctoral Showcase	Total Energy Optimization for High Performance Computing Data Centers	155-F
11:15am-12pm	Invited Talk	High-Performance Techniques for Big Data Computing in Internet Services	Ballroom-EFGH
11:30am-11:45am	Doctoral Showcase	Parallel Algorithms for Bayesian Networks Structure Learning with Applications to Gene Networks	155-F
11:30am-12pm	Paper (Best Student Paper Finalist)	Compiler-Directed File Layout Optimization for Hierarchical Storage Systems	255-EF
11:30am-12pm	Paper	Scalable Multi-GPU 3-D FFT for TSUBamE 2.0 Supercomputer	255-BC
11:30am-12pm	Award	Sidney Fernbach Award Recipients: Laxmikant Kale and Klaus Schulten	155-E
11:30am-12pm	Paper	High Performance Radiation Transport Simulations-Preparing for TITAN	355-EF
11:30am-12pm	Paper	Efficient Data Restructuring and Aggregation for I/O Acceleration in PIDX	355-D

Wednesday, November 14

Time	Event	Title	Location
11:30am-12pm	Exhibitor Forum	Innovation and HPC Transformation	155-B
11:30am-12pm	Exhibitor Forum	Runtimes and Applications for Extreme-Scale Computing	155-C
11:45am-12pm	Doctoral Showcase	Exploring Multiple Levels of Heterogeneous Performance Modeling	155-F
12:15pm-1:15pm	Birds of a Feather	Architecture and Systems Simulators	155-A
12:15pm-1:15pm	Birds of a Feather	Building an Open Community Runtime (OCR) Framework for Exascale Systems	255-EF
12:15pm-1:15pm	Birds of a Feather	Chapel Lightning Talk 2012	255-A
12:15pm-1:15pm	Birds of a Feather	Early Experiences Debugging on the Blue Gene/Q	155-E
12:15pm-1:15pm	Birds of a Feather	International Collaboration on System Software Development for Post-petascale Computing	355-BC
12:15pm-1:15pm	Birds of a Feather	Open MPI State of the Union	155-B
12:15pm-1:15pm	Birds of a Feather	PGAS: The Partitioned Global Address Space Programming Model	355-EF
12:15pm-1:15pm	Birds of a Feather	PRObE: A 1000 Node Facility for Systems Infrastructure Researchers	255-BC
12:15pm-1:15pm	Birds of a Feather	Science-as-a-Service: Exploring Clouds for Computational and Data-Enabled Science and Engineering	155-C
12:15pm-1:15pm	Birds of a Feather	Setting Trends for Energy Efficiency	250-AB
12:15pm-1:15pm	Birds of a Feather	The Way Forward: Addressing the data challenges for Exascale Computing	355-A
12:15pm-1:15pm	Birds of a Feather	Unistack: Interoperable Community Runtime Environment for Exascale Systems	355-D
12:15pm-1:15pm	Birds of a Feather	XSEDE Metrics on Demand (XDMoD) Technology Auditing Framework	250-C
1:30pm-1:45pm	Doctoral Showcase	Automatic Selection of Compiler Optimizations Using Program Characterization and Machine Learning	155-F
1:30pm-2pm	ACM Gordon Bell Finalist	The Universe at Extreme Scale-Multi-Petaflop Sky Simulation on the BG/Q	155-E
1:30pm-2pm	Paper	Measuring Interference Between Live Datacenter Applications	255-BC
1:30pm-2pm	Exhibitor Forum	Cray's Adaptive Supercomputing Vision	155-B
1:30pm-2pm	Paper (Best Paper Finalist)	Compass-A Scalable Simulator for an Architecture for Cognitive Computing	355-EF
1:30pm-2pm	Paper	Classifying Soft Error Vulnerabilities in Extreme-Scale Scientific Applications Using a Binary Instrumentation Tool	255-EF
1:30pm-2pm	Exhibitor Forum	A Plague of Petabytes	155-C
1:30pm-2pm	Paper	Parallel IO, Analysis, and Visualization of a Trillion Particle Simulation	355-D
1:30pm-2:15pm	Invited Talk	Design, Implementation and Evolution of High Level Accelerator Programming	Ballroom-EFGH
1:30pm-3pm	ACM SRC Competition	ACM SRC Competition Semi-Finals (Session 1)	250-C
1:30pm-3pm	Panel	Exascale and Big Data IO-Which Will Drive Future IO Architectures, Standards and Protocols-Should They be Open or Proprietary	355-BC
1:30pm-5pm	HPC Educators	Cyber-Physical Systems	255-A
1:30pm-5pm	Broader Engagement, HPC Educators	HPC: Suddenly Relevant to Mainstream CS Education?	355-A
1:45pm-2pm	Doctoral Showcase	High Performance Non-Blocking and Power-Aware Collective Communication for Next Generation InfiniBand Clusters	155-F
2pm-2:15pm	Doctoral Showcase	Virtualization of Accelerators in High Performance Clusters	155-F
2pm-2:30pm	ACM Gordon Bell Finalist	4.45 Pflops Astrophysical N-Body Simulation on K Computer-The Gravitational Trillion-Body Problem	155-E
2pm-2:30pm	Paper	T*-A Data-Centric Cooling Energy Costs Reduction Approach for Big Data Analytics Cloud	255-BC
2pm-2:30pm	Exhibitor Forum	Findings From Real Petascale Computer Systems and Fujitsu's Challenges in Moving Towards Exascale Computing	155-B
2pm-2:30pm	Paper	Optimizing Fine-Grained Communication in a Biomolecular Simulation Application on Cray XK6	355-EF

Wednesday, November 14

Time	Event	Title	Location
2pm-2:30pm	Paper (Best Student Paper Finalist)	Containment Domains-A Scalable, Efficient, and Flexible Resiliency Scheme for Exascale Systems	255-EF
2pm-2:30pm	Exhibitor Forum	Big Data, Big Opportunity: Maximizing the Value of Data in HPC Environments	155-C
2pm-2:30pm	Paper	Data-Intensive Spatial Filtering in Large Numerical Simulation Datasets	355-D
2:15pm-2:30pm	Doctoral Showcase	Heterogeneous Scheduling for Performance and Programmability	155-F
2:15pm-3pm	Invited Talk	Dealing with Portability and Performance on Heterogeneous Systems with Directive-Based Programming Approaches	Ballroom-EFGH
2:30pm-2:45pm	Doctoral Showcase	Integrated Parallelization of Computation and Visualization for Large-Scale Weather Applications	155-F
2:30pm-3pm	Paper	ValuePack-Value-Based Scheduling Framework for CPU-GPU Clusters	255-BC
2:30pm-3pm	Exhibitor Forum	On Solution-Oriented HPC	155-B
2:30pm-3pm	Paper	Heuristic Static Load-Balancing Algorithm Applied to the Fragment Molecular Orbital Method	355-EF
2:30pm-3pm	Exhibitor Forum	FhGFS-Parallel Filesystem Performance at the Maximum	155-C
2:30pm-3pm	Paper	Parallel Particle Advection and FTLE Computation for Time-Varying Flow Fields	355-D
2:45pm-3pm	Doctoral Showcase	Programming High Performance Heterogeneous Computing Systems: Paradigms, Models and Metrics	155-F
3:30pm-4pm	Paper	A New Scalable Parallel DBSCAN Algorithm Using the Disjoint-Set Data Structure	355-EF
3:30pm-4pm	Exhibitor Forum	The HPC Advisory Council Outreach and Research Activities	155-B
3:30pm-4pm	Exhibitor Forum	Something New in HPC? EXTOLL: A Scalable Interconnect for Accelerators	155-C
3:30pm-4pm	Paper (Best Student Paper Finalist)	Characterizing and Mitigating Work Time Inflation in Task Parallel Programs	255-EF
3:30pm-4pm	Paper	Design and Implementation of an Intelligent End-to-End Network QoS System	255-BC
3:30pm-4pm	Paper	Critical Lock Analysis-Diagnosing Critical Section Bottlenecks in Multithreaded Applications	355-D
3:30pm-4:15pm	Invited Talk	Modelling the Earth's Climate System-Data and Computing Challenges	Ballroom-EFGH
3:30pm-5pm	ACM SRC Competition	ACM SRC Competition Semi-Finals (Session 2)	250-C
3:30pm-5pm	Doctoral Showcase	A Cloud Architecture for Reducing Costs in Local Parallel and Distributed Virtualized Environments	155-F
3:30pm-5pm	Doctoral Showcase	Towards the Support for Many-Task Computing on Many-Core Computing Platforms	155-F
3:30pm-5pm	Doctoral Showcase	Software Support for Regular and Irregular Applications in Parallel Computing	155-F
3:30pm-5pm	Doctoral Showcase	Towards Scalable and Efficient Scientific Cloud Computing	155-F
3:30pm-5pm	Doctoral Showcase	On Bandwidth Reservation for Optimal Resource Utilization in High-Performance Networks	155-F
3:30pm-5pm	Doctoral Showcase	Distributed File Systems for Exascale Computing	155-F
3:30pm-5pm	Doctoral Showcase	Dynamic Load-Balancing for Multicores	155-F
3:30pm-5pm	Doctoral Showcase	Numerical Experimentations in the Dynamics of Particle-Laden Supersonic Impinging Jet Flow	155-F
3:30pm-5pm	Doctoral Showcase	An Efficient Runtime Technology for Many-Core Device Virtualization in Clusters	155-F
3:30pm-5pm	Doctoral Showcase	Simulating Forced Evaporative Cooling Utilising a Parallel Direct Simulation Monte Carlo Algorithm	155-F
3:30pm-5pm	Doctoral Showcase	Autonomic Modeling of Data-Driven Application Behavior	155-F
3:30pm-5pm	Doctoral Showcase	Metrics, Workloads and Methodologies for Energy Efficient Parallel Computing	155-F
3:30pm-5pm	Doctoral Showcase	Adaptive, Resilient Cloud Platform for Dynamic, Mission-Critical Dataflow Applications	155-F
3:30pm-5pm	Doctoral Showcase	Using Computational Fluid Dynamics and High Performance Computing to Model a Micro-Helicopter Operating Near a Flat Vertical Wall	155-F

Wednesday, November 14

Time	Event	Title	Location
3:30pm-5pm	Doctoral Showcase	Paving the Road to Exascale with Many-Task Computing	155-F
3:30pm-5pm	Doctoral Showcase	High Performance Computing in Simulating Carbon Dioxide Geologic Sequestration	155-F
3:30pm-5pm	Doctoral Showcase	Uncovering New Parallel Program Features with Parallel Block Vectors and Harmony	155-F
3:30pm-5pm	Doctoral Showcase	New Insights into the Colonisation of Australia Through the Analysis of the Mitochondrial Genome	155-F
3:30pm-5pm	Doctoral Showcase	A Meshfree Particle Based Model for Microscale Shrinkage Mechanisms of Food Materials in Drying Conditions	155-F
3:30pm-5pm	Doctoral Showcase	Reproducibility and Scalability in Experimentation through Cloud Computing Technologies	155-F
3:30pm-5pm	Doctoral Showcase	Programming and Runtime Support for Enabling In-Situ/In-Transit Scientific Data Processing	155-F
3:30pm-5pm	Doctoral Showcase	Ensemble-Based Virtual Screening to Expand the Chemical Diversity of LSD1 Inhibitors	155-F
3:30pm-5pm	Panel	Visualization Frameworks for Multi-Core and Many-Core Architectures	355-BC
4pm-4:30pm	Paper	Parallel Bayesian Network Structure Learning with Application to Gene Networks	355-EF
4pm-4:30pm	Exhibitor Forum	Differences Between Cold and Hot Water Cooling on CPU and Hybrid Supercomputers	155-B
4pm-4:30pm	Exhibitor Forum	QFabric Technology: Revolutionizing the HPC Interconnect Architecture	155-C
4pm-4:30pm	Paper	Legion-Expressing Locality and Independence with Logical Regions	255-EF
4pm-4:30pm	Paper	Looking Under the Hood of the IBM Blue Gene/Q Network	255-BC
4pm-4:30pm	Paper	Code Generation for Parallel Execution of a Class of Irregular Loops on Distributed Memory Systems	355-D
4pm-6pm	Family Day	Exhibit Halls	
4:15pm-5pm	Invited Talk	Achieving Design Targets by Stochastic Car Crash Simulations-The Relation between Bifurcation of Deformation and Quality of FE Models	Ballroom-EFGH
4:30pm-5pm	Paper	A Multithreaded Algorithm for Network Alignment via Approximate Matching	355-EF
4:30pm-5pm	Exhibitor Forum	100% Server Heat Recapture in Data Centers is Now a Reality	155-B
4:30pm-5pm	Exhibitor Forum	Deploying 40 and 100GbE: Optics and Fiber Media	155-C
4:30pm-5pm	Paper	Designing a Unified Programming Model for Heterogeneous Machines	255-EF
4:30pm-5pm	Paper (Best Paper Finalist, Best Student Paper Finalist)	Design of a Scalable InfiniBand Topology Service to Enable Network-Topology-Aware Placement of Processes	255-BC
5:30pm-7pm	Birds of a Feather	Application Grand Challenges in the Heterogeneous Accelerator Era	355-BC
5:30pm-7pm	Birds of a Feather	Co-design Architecture and Co-design Efforts for Exascale: Status and Next Steps	355-A
5:30pm-7pm	Birds of a Feather	Common Practices for Managing Small HPC Clusters	355-D
5:30pm-7pm	Birds of a Feather	Cool Supercomputing: Achieving Energy Efficiency at the Extreme Scales	155-A
5:30pm-7pm	Birds of a Feather	Cyberinfrastructure services for long tail research	253
5:30pm-7pm	Birds of a Feather	DARPA's High Productivity Computing Systems Program: A Final Report	255-D
5:30pm-7pm	Birds of a Feather	Exploiting Domain Semantics and High-Level Abstractions in Computational Science	155-B
5:30pm-7pm	Birds of a Feather	HPC Centers	155-E
5:30pm-7pm	Birds of a Feather	Intel MIC Processors and the Stampede Petascale Computing System	255-A
5:30pm-7pm	Birds of a Feather	OpenCL: Supporting Mainstream Heterogeneous Computing	Ballroom-A
5:30pm-7pm	Birds of a Feather	Operating Systems and Runtime Technical Council	355-EF
5:30pm-7pm	Birds of a Feather	Power and Energy Measurement and Modeling on the Path to Exascale	255-EF
5:30pm-7pm	Birds of a Feather	PRACE Future Technologies Evaluation Results	250-AB
5:30pm-7pm	Birds of a Feather	The Green500 List	255-BC
5:30pm-7pm	Birds of a Feather	The Ground is Moving Again in Paradise: Supporting Legacy Codes in the New Heterogeneous Age	155-F

5:30pm-7pm	Birds of a Feather	Using Application Proxies for Exascale Preparation	250-C
5:30pm-7pm	Birds of a Feather	What Next for On-Node Parallelism? Is OpenMP the Best We Can Hope For?	155-C

Thursday, November 15

Time	Event	Title	Location
7:30am-9:30pm	Coat/Bag Check		Lower Concourse
8am-12pm	Information Booth	Satellite Booth	Upper Concourse
8am-6pm	Information Booth	Main	South Foyer
8:30am-9:15am	Invited Talk	A Journey to Exascale Computing	Ballroom-EFGH
9:15am-10am	Invited Talk	The Evolution of GPU Accelerated Computing	Ballroom-EFGH
10am-3pm	Exhibits	Exhibit Hall	
10am-6pm	Preview Booth	SC13 Preview Booth	South Lobby
10:30am-11am	Paper	First-Ever Full Observable Universe Simulation	255-EF
10:30am-11am	Paper	A Study of DRam Failures in the Field	255-BC
10:30am-11am	Other Event	Efficient and Scalable Runtime for GAS Programming Models on Petascale Systems	155-E
10:30am-11am	Paper	ATLAS Grid Workload on NDGF Resources: Analysis, Modeling, and Workload Generation	355-D
10:30am-11am	Exhibitor Forum	An OpenCL Application for FPGAs	155-B
10:30am-11am	Exhibitor Forum	Addressing HPC Compute Center Challenges with Innovative Solutions	155-C
10:30am-11am	Paper	Dataflow-Driven GPU Performance Projection for Multi-Kernel Transformations	355-EF
10:30am-11:15am	Invited Talk	Application Development for Titan-A Multi-Petaflop Hybrid-Multicore MPP System	Ballroom-EFGH
10:30am-12pm	HPC Educators	Computational Examples for Physics (and Other) Classes Featuring Python, Mathematica, an eTextBook and More	255-A
10:30am-12pm	Panels	Current Status of HPC Progress in China	355-BC
10:30am-12pm	Other Event	Experiencing HPC for Undergraduates—Careers in HPC	250-AB
10:30am-12pm	HPC Educators	Teaching Parallel Computing through Parallel Prefix	255-D
11am-11:30am	Paper	Optimizing the Computation of N-Point Correlations on Large-Scale Astronomical Data	255-EF
11am-11:30am	Paper	Fault Prediction Under the Microscope-A Closer Look Into HPC Systems	255-BC
11am-11:30am	Paper	On the Effectiveness of Application-Aware Self-Management for Scientific Discovery in Volunteer Computing Systems	355-D
11am-11:30am	Exhibitor Forum	MIC and GPU: Application, Porting and Optimization	155-B
11am-11:30am	Exhibitor Forum	Achieving Cost-Effective HPC with Process Virtualization	155-C
11am-11:30am	Paper	A Practical Method for Estimating Performance Degradation on Multicore Processors and its Application to HPC Workloads	355-EF
11am-12pm	Other Event	Planning the Future of the SC Student Cluster Competition	155-E
11:15am-12pm	Invited Talk	Application Performance Characterization and Analysis on Blue Gene/Q	Ballroom-EFGH
11:30am-12pm	Paper	Hierarchical Task Mapping of Cell-Based amR Cosmology Simulations	255-EF
11:30am-12pm	Paper	Detection and Correction of Silent Data Corruption for Large-Scale High-Performance Computing	255-BC
11:30am-12pm	Paper	On Using Virtual Circuits for GridFTP Transfers	355-D
11:30am-12pm	Exhibitor Forum	Acceleration of ncRNA Discovery via Reconfigurable Computing Platforms	155-B
11:30am-12pm	Exhibitor Forum	Reducing First Costs and Improving Future Flexibility in the Construction of High Performance Computing Facilities	155-C
11:30am-12pm	Paper	Aspen-A Domain Specific Language for Performance Modeling	355-EF
12:15pm-1:15pm	Birds of a Feather	Charm++: Adaptive Runtime-Assisted Parallel Programming	255-A
12:15pm-1:15pm	Birds of a Feather	Data Analysis through Computation and 3D Stereo Visualization	355-EF

Thursday, November 15

Time	Event	Title	Location
12:15pm-1:15pm	Birds of a Feather	Discussing Biomedical Data Management as a Service	250-C
12:15pm-1:15pm	Birds of a Feather	Graph Analytics in Big Data	255-EF
12:15pm-1:15pm	Birds of a Feather	HPC Advisory Council University Award Ceremony	155-B
12:15pm-1:15pm	Birds of a Feather	In-silico Bioscience: Advances in the Complex, Dynamic Range of Life Sciences Applications	155-F
12:15pm-1:15pm	Birds of a Feather	New Developments in the Global Arrays Programming Model	155-E
12:15pm-1:15pm	Birds of a Feather	OpenSHMEM: A standardized SHMEM for the PGAS community	155-C
12:15pm-1:15pm	Birds of a Feather	Petascade Systems Management	355-BC
12:15pm-1:15pm	Birds of a Feather	Resilience for Extreme-scale High Performance Computing	255-BC
12:15pm-1:15pm	Birds of a Feather	SLURM User Group Meeting	155-A
12:15pm-1:15pm	Birds of a Feather	The MPI 3.0 Standard	355-A
12:15pm-1:15pm	Birds of a Feather	The UDT Forum: A Community for UDT Developers and Users	255-D
12:30pm-1:30pm	Awards	SC12 Conference Award Presentations	Ballroom-EFGH
1:30pm-2pm	Paper	Design and Analysis of Data Management in Scalable Parallel Scripting	255-EF
1:30pm-2pm	Exhibitor Forum	HPC Cloud and Big Data Analytics-Transforming High Performance Technical Computing	155-B
1:30pm-2pm	Paper	Application Data Prefetching on the IBM Blue Gene/Q Supercomputer	355-D
1:30pm-2pm	Paper	A Parallel Two-Level Preconditioner for Cosmic Microwave Background Map-Making	355-EF
1:30pm-2pm	Paper	Extending the BT NAS Parallel Benchmark to Exascale Computing	255-BC
1:30pm-2pm	Exhibitor Forum	SET-Supercomputing Engine Technology	155-C
1:30pm-5pm	HPC Educators	CSinParallel: An incremental approach to adding PDC throughout the CS curriculum	255-A
1:30pm-5pm	HPC Educators	High-Level Parallel Programming using Chapel	255-D
2pm-2:30pm	Paper	Usage Behavior of a Large-Scale Scientific Archive	255-EF
2pm-2:30pm	Exhibitor Forum	Gompute Software and Remote Visualization for a Globalized Market	155-B
2pm-2:30pm	Paper	Hardware-Software Coherence Protocol for the Coexistence of Caches and Local Memories	355-D
2pm-2:30pm	Paper	A Massively Space-Time Parallel N-Body Solver	355-EF
2pm-2:30pm	Paper	NUMA-Aware Graph Mining Techniques for Performance and Energy Efficiency	255-BC
2pm-2:30pm	Exhibitor Forum	The Future of OpenMP	155-C
2:30pm-3pm	Paper	On Distributed File Tree Walk of Parallel File Systems	255-EF
2:30pm-3pm	Exhibitor Forum	Windows HPC in the Cloud	155-B
2:30pm-3pm	Paper	What Scientific Applications Can Benefit from Hardware Transactional Memory	355-D
2:30pm-3pm	Paper	High-Performance General Solver for Extremely Large-Scale Semidefinite Programming Problems	355-EF
2:30pm-3pm	Paper	Optimization of Geometric Multigrid for Emerging Multi- and Manycore Processors	255-BC
2:30pm-3pm	Exhibitor Forum	Let The Chips Fall where they May-PGI Compilers & Tools for Heterogeneous HPC Systems	155-C
3:30pm-4pm	Paper	Mapping Applications with Collectives over Sub-Communicators on Torus Networks	255-EF
3:30pm-4pm	Paper	Communication Avoiding and Overlapping for Numerical Linear Algebra	355-D
3:30pm-4pm	Paper	Cray Cascade-A Scalable HPC System Based on a Dragonfly Network	255-BC
3:30pm-4:15pm	Invited Talk	Low Mach Number Models in Computational Astrophysics	Ballroom-EFGH
3:30pm-5pm	Broader Engagement	Wrap-up Session-Program Evaluation and Lessons Learned	355-A
3:30pm-5pm	Panel	Big Data and Data Integrity-How Do We Handle Reliability, Reproducibility, and Accessibility	355-BC
3:30pm-5pm	Exhibitor Forum	OrangeFS Drop-In	155-B

Thursday, November 15

Time	Event	Title	Location
4pm-4:30pm	Paper	Optimization Principles for Collective Neighborhood Communications	255-EF
4pm-4:30pm	Paper	Communication-Avoiding Parallel Strassen-Implementation and Performance	355-D
4pm-4:30pm	Paper	GRAPE-8-An Accelerator for Gravitational N-Body Simulation with 20.5GFLOPS/W Performance	255-BC
4:15pm-5pm	Invited Talk	Very Large-Scale Fluid-Flow and Aeroacoustical Simulations for Engineering Applications Performed on Supercomputer "K"	Ballroom-EFGH
4:30pm-5pm	Paper	Optimizing Overlay-Based Virtual Networking Through Optimistic Interrupts and Cut-Through Forwarding	255-EF
4:30pm-5pm	Paper	Managing Data-Movement for Effective Shared-Memory Parallelization of Out-of-Core Sparse Solvers	355-D
4:30pm-5pm	Paper	SGI UV2-A Fused Computation and Data Analysis Machine	255-BC
6pm-9pm	Reception	Technical Program Conference Reception	*The Depot

Friday, November 16

Time	Event	Title	Location
8am-1pm	Coat/Bag Check		Lower Concourse
8:30am-10am	Panel	Applying High Performance Computing at a National Laboratory to the Needs of Industry Case Study-hpc4energy Incubator	355-EF
8:30am-12pm	Information Booth	Main (Satellite Booth closed)	South Lobby
8:30am-12:30pm	Workshop	Extreme-Scale Performance Tools	
8:30am-12:30pm	Workshop	Multi-Core Computing Systems-MuCoCoS-Performance Portability and Tuning	
8:30am-12:30pm	Workshop	Preparing Applications for Exascale Through Co-design	
8:30am-12:30pm	Workshop	Python for High Performance and Scientific Computing	
8:30am-12:30pm	Workshop	Sustainable HPC Cloud Computing 2012	
8:30am-12:30pm	Workshop	The First International Workshop on Data Intensive Scalable Computing Systems-DISCS	
8:30am-12:30pm	Workshop	Workshop on Domain-Specific Languages and High-Level Frameworks for High-Performance Computing	
10:30am-12pm	Panel	Is the Cloud a Game Changer for the Film Industry	355-EF
10:30am-12pm	Panel	Questions: Rhetorically Volleying the Terabit Network Ball on the HPC Court	355-BC

*Bussing service for the Technical Program Conference Reception is provided at the South Entrance.

Keynote/Invited Talks/Panels

In the past, SC conferences have featured a variety of invited talks under various names such as Masterworks, plenary talks, and state of the field. To reduce confusion this year, we grouped all talks under a single banner: Invited Talks. These talks feature leaders who detail innovative work in the area of high performance, networking, storage, analysis, and their application to the world's most challenging problems. You will hear about the latest innovations in computing and how they are fueling new approaches to addressing the toughest and most complex questions of our time.

Panels at SC12 will be, as in past years, among the most important and heavily attended events of the conference. Panels provide a unique forum for engagement and interaction of the community for exchange of information, ideas, and opinions about a number of hot topics spanning the field of high performance computing and related domains. Panels will bring together the key thinkers and producers in the field to consider in a lively and rapid-fire context some of the key questions challenging HPC this year.

Please plan on attending one or more of the panel offerings. We look forward to your help, through your participation, in making this year's panels a major success and lots of fun.

Keynote/Invited Talks

Tuesday, November 13

Keynote

Chair: Jeffrey K. Hollingsworth (University of Maryland)

8:30am-10am

Room: Ballroom-CDEFGH

Keynote: Physics of the Future

Michio Kaku (City University of New York)

Bio: Dr. Michio Kaku holds the Henry Semat Chair in Theoretical Physics at the City University of New York (CUNY). He graduated from Harvard University summa cum laude and first in his physics class. He received his Ph.D. in physics from the University of California at Berkeley, and been a professor at CUNY for almost 30 years. He has taught at Harvard and Princeton as well. Dr. Kaku's goal is to complete Einstein's dream of a "theory of everything," to derive an equation, perhaps no more than one inch long, which will summarize all the physical laws of the universe. He is the co-founder of string field theory, a major branch of string theory, which is the leading candidate today for the theory of everything.

Our community has entered a phase of radical change as we address the challenges of reaching exascale computation and the opportunities that big data will bring to science. Building on ideas presented in his most recent book *Physics of the Future: How Science will Change Daily Life by 2100*, Dr. Kaku will open the SC12 technical program. Based on interviews with over 300 of the world's top scientists Dr Kaku presents the revolutionary developments in medicine, computers, quantum physics, and space travel that will forever change our way of life and alter the course of civilization itself.

Fielding Large Scale Systems

Chair: Bronis R. de Supinski

(Lawrence Livermore National Laboratory)

10:30am-12pm

Room: Ballroom-EFGH

The Sequoia System and Facilities Integration Story

Kimberly Cupps (Lawrence Livermore National Laboratory)

Sequoia, a 20PF/s Blue Gene/Q system, will serve National Nuclear Security Administration's Advanced Simulation and Computing (ASC) program to fulfill stockpile stewardship requirements through simulation science. Problems at the highest end of this computational spectrum are a principal ASC driver as highly predictive codes are developed. Sequoia is an Uncertainty Quantification focused system at Lawrence Livermore National Laboratory (LLNL). Sequoia will simultaneously run integrated design code and science materials calculations

enabling sustained performance of 24 times ASC's Purple calculations and 20 times ASC's Blue Gene/L calculations. LLNL prepared for Sequoia's delivery for over three years. During the past year we have been consumed with the integration challenges of siting the system and its facilities and infrastructure. Sequoia Integration continues, acceptance testing begins in September and production level computing is expected in March 2013. This talk gives an overview of Sequoia and its facilities and system integration victories and challenges.

Titan—Early Experience with the Titan System at Oak Ridge National Laboratory

Arthur S. Bland (Oak Ridge National Laboratory)

In 2011, Oak Ridge National Laboratory began a two-phase upgrade to convert the Cray XT5 Jaguar system into a Cray XK6 system named Titan. The first phase, completed in early 2012, replaced all XT5 node boards with XK6 boards, including the AMD Opteron 6274 16-core processors, 600 terabytes of system memory, Cray's new Gemini network, and 960 NVIDIA X2090 "Fermi" processors. The second phase will add 20 petaflops of NVIDIA's next generation K20 "Kepler" processor. The most important aspect of the Titan project has been developing a programming strategy to allow the applications to run efficiently and effectively on the accelerators, while maintaining compatibility with other architectures. ORNL's Center for Accelerated Applications Readiness has worked for over two years to implement this strategy on several key Department of Energy applications. This talk describes Titan, the upgrade process, and the challenges of developing its programming strategy and programming environment.

Green Supercomputing

Chair: David Lowenthal (University of Arizona)

1:30pm-3pm

Room: Ballroom-EFGH

Pushing Water Up Mountains: Green HPC and Other Energy Oddities

Kirk W. Cameron (Virginia Tech)

Green HPC is an oxymoron. How can something be "green" when it consumes over 10 megawatts of power? Utility companies pay customers to use less power. Seriously, energy use per capita continues to increase worldwide yet most agree new power production facilities should not be built in their backyards. HPC cannot operate in a vacuum. Whether we like it or not, we are part of a large multi-market ecosystem at the intersection of the commodity markets for advanced computer hardware and the energy markets for power. This talk will provide a historical view of the Green HPC movement including some of my own power-aware software successes and failures. I'll discuss the challenges facing computer energy efficiency research and how market forces will likely affect big changes in the future of HPC.

The Costs of HPC-Based Science in the Exascale Era

Thomas Ludwig (*German Climate Computing Center*)

Many science fields base their knowledge-gaining process on high performance computing. Constant exponential increase in performance allows in particular natural sciences to run more and more sophisticated numerical simulations. However, one may wonder, does the quality of results correlate to the increase in costs? In particular with the advent of the Exascale Era and with Big Data we are confronted with possibly prohibitive energy costs. In addition, our installations grow in size and we typically replace them every 4-6 years. The talk will analyze the cost-benefit ratio of HPC-based science and consider economic and ecological aspects. We will have a closer look into different science fields and evaluate the impact of their research results on society.

Algorithmic Innovations for Large-Scale Computing

Chair: Bernd Mohr (*Juelich Supercomputing Centre*)

3:30pm-5pm

Room: Ballroom-EFGH

Communication-Avoiding Algorithms for Linear Algebra and Beyond

James Demmel (*University of California, Berkeley*)

Algorithms have two costs: arithmetic and communication, i.e., moving data between levels of a memory hierarchy or processors over a network. Communication costs (measured in time or energy per operation) already greatly exceed arithmetic costs, and the gap is growing over time following technological trends. Thus, our goal is to design algorithms that minimize communication. We present algorithms that attain provable lower bounds on communication and show large speedups compared to their conventional counterparts. These algorithms are for direct and iterative linear algebra, for dense and sparse matrices, as well as direct n -body simulations. Several of these algorithms exhibit perfect strong scaling, in both time and energy: run time (resp. energy) for a fixed problem size drops proportionally to p (resp. is independent of p). Finally, we describe extensions to algorithms involving arbitrary loop nests and array accesses, assuming only that array subscripts are linear functions of the loop indices.

Stochastic Simulation Service—Towards an Integrated Development Environment for Modeling and Simulation of Stochastic Biochemical Systems

Linda Petzold (*University of California, Santa Barbara*)

In recent years it has become increasingly clear that stochasticity plays an important role in many biological processes. Examples include bistable genetic switches, noise enhanced robustness of oscillations, and fluctuation enhanced sensitivity or “stochastic focusing.” In many cellular systems, local low

species populations can create stochastic effects even if total cellular levels are high. Numerous cellular systems, including development, polarization and chemotaxis, rely on spatial stochastic noise for robust performance. In this talk we report on our progress in developing next-generation algorithms and software for modeling and simulation of stochastic biochemical systems, and in building an integrated development environment that will enable researchers to build a such a model and scale it up to increasing levels of complexity.

Wednesday, November 14

Current Large-Scale Computing Activities

Chair: Wilfred Pinfold (*Intel Corporation*)

8:30am-10am

Room: Ballroom-EFGH

Simulating the Human Brain - An Extreme Challenge for Computing

Henry Markram (*EPFL*)

Knowledge of the brain is highly fragmented---neuroscientists are locked into their subspecialties---and while it is obvious that we need much more data and new theories in order to understand the brain, we have no way to assemble and make sense of what we know today or to prioritize the vast number of experiments still required to obtain an integrated view of the brain. It is time for a radically new strategy of collaboration, where scientists of many disciplines can come around the same table and begin reassembling the pieces that we have, find out what data and knowledge is really missing, what gaps can be filled using statistical models, and what parts require new experiments. In the Blue Brain Project, we have been building a technology platform to catalyze this collaboration and integrate our collective knowledge into unifying computer models of the brain. The platform enables supercomputer simulations of brain models, developed to account for all we know, to predict what we cannot measure, and to test all we can hypothesize about how the brain works. To develop this platform for the human brain is an extreme undertaking with a far larger and more multidisciplinary consortium of scientists - a Human Brain Project. To account for all genes, proteins, cells, circuits, brain regions, the whole brain all the way to cognition and behavior, we need to build on relevant information from all of biology and everything we have discovered about the brains of animals. Building and simulating brain models across its spatial (nine orders of magnitude) and temporal (twelve orders of magnitude) scales will demand extreme solutions for data management, cloud-based computing, internet-based interactivity, visualization and supercomputing. We believe the effort will transform supercomputing by driving the hardware and software innovations required to turn supercomputers

into visually interactive scientific instruments, and that it will spur a new era of computing technology combining the best of von Neumann and neuromorphic computing.

The K Computer - Toward Its Productive Application to Our Life

Mitsuo Yokokawa (RIKEN)

No one doubts that computer simulations are now indispensable techniques to elucidate natural phenomena and to design artificial structures with the help of the growing power of supercomputers. Many countries are committed to have supercomputers as a fundamental tool for their national competitiveness.

The Next-Generation Supercomputer Development Project was started in 2006 as a seven-year project under direction of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) in Japan. Its objectives are to develop the world's most advanced and high-performance supercomputer (later it was named "Kei" in Japanese or "K computer"), to develop and to deploy its usage technologies including application software in various science and engineering fields, and to establish a center of excellence for computational sciences as one of key technologies of national importance designated in the Japanese Third Science and Technology Basic Plan.

The K computer, located in the RIKEN Advanced Institute for Computational Science (AICS) at Kobe, Japan, broke a 10PFLOPS wall for the first time in the world with high efficiency in LINPACK benchmark in November, 2011. It has also achieved more than petaflops sustained performance in real applications. This powerful and stable computing capability will bring us a big step toward the realization of the sustainable human society and break-through in research and development we have never had.

In this talk, the development history of the K computer will be introduced and the overall features of the K computer and some results obtained by the early access will be given.

Innovations in Computer Architecture

Chair: Jack Dongarra (University of Tennessee, Knoxville)

10:30am-12pm

Room: Ballroom-EFGH

The Long Term Impact of Codesign

Alan Gara (Intel Corporation)

Supercomputers of the future will utilize new technologies in an effort to provide exponential improvements in cost and energy efficiency. While today's supercomputing applications concern domains that have been of focus for decades, the complexity of the physics and the scale is ever increasing resulting in greater system demands. In this talk, I will discuss the long term (10 to 15 years) technology trends as well as

how these trends will likely shape system architectures and the algorithms that will exploit these architectures. Optimal utilization of new technologies is truly a co-design effort. Both algorithms and system design must be taken into account to reap optimal performance from new technologies.

High-Performance Techniques for Big Data Computing in Internet Services

Zhiwei Xu (Chinese Academy of Sciences)

Internet services directly impact billions of users and have a much larger market size than traditional high-performance computing. However, these two fields share common technical challenges. Exploiting locality and providing efficient communication are common research issues. Internet services increasingly feature big data computing, involving petabytes of data and billions of records. This talk focuses on three problems that recur frequently in Internet services systems: the data placement, data indexing, and data communication problems, which are essential in enhancing performance and reducing energy consumption. After presenting a formulation of the relationship between performance and power consumption, I will provide examples of high-performance techniques developed to address these problems, including a data placement method that significantly reduces storage space needs, a data indexing method that enhances throughput by orders of magnitude, and a key-value data communication model with its high-performance library. Applications examples include Facebook in the USA and Taobao in China.

Compiling for Accelerators

Chair: Robert F. Lucas (Information Sciences Institute)

1:30pm-3pm

Room: Ballroom-EFGH

Design, Implementation and Evolution of High Level Accelerator Programming

Michael Wolfe (The Portland Group, Inc.)

In 2008, PGI designed the PGI Accelerator programming model and began work on an implementation to target heterogeneous X64 host + NVIDIA GPU systems. In November 2011, Cray, NVIDIA and CAPS Enterprise joined with PGI to refine and standardize directive-based GPU and accelerator programming with the introduction of the OpenACC API. This presentation will discuss three aspects of this language design evolution. We will describe how the programming model changed over time to take advantage of the features of current accelerators, while trying to avoid various performance cliffs. We describe advantages and problems associated with committee-designed languages and specifications. Finally, we describe several specific challenges related to the implementation of OpenACC for the current generation of targets, and how we solved them in the PGI Accelerator compilers.

Dealing with Portability and Performance on Heterogeneous Systems with Directive-Based Programming Approaches

François Bodin (CAPS)

Directive-based programming is a very promising technology for dealing with heterogeneous many-core architectures. Emerging standards such as OpenACC and other initiatives such as OpenHMPP provide a solid ground for users to invest in such paradigm. On one side portability is required to ensure long software lifetime and to reduce maintenance cost. On the other-hand, obtaining efficient code requires to have a tight mapping between the code and the target architecture. In this presentation we describe the challenges in building programming tools based on directives. We show how OpenACC and OpenHMPP directives offer an incremental development for various heterogeneous architectures ranging from AMD, Intel, Nvidia to ARM. We explain why source-to-source compilers are particularly adequate when dealing with heterogeneity. Finally, we propose an auto-tuning framework for achieving better performance portability. On this later topic we advocate for a standard API to be included into current standardization initiatives.

HPC Applications and Society

Chair: Yoshio Oyanagi (Kobe University)

3:30pm-5pm

Room: Ballroom-EFGH

Modelling the Earth's Climate System—Data and Computing Challenges

Sylvie Joussaume (CNRS)

Climate models are used to assess mitigation and adaptation strategies for climate change. The international community has just completed an unprecedented coordinated set of experiments, the Coupled Modeling Intercomparison Project (CMIP5), to which the European Network for Earth System Modelling (ENES) has contributed with seven global climate models. These experiments have triggered a new way to manage the Petabyte distributed datasets produced and widely used to study climate change and its impacts. The European IS-ENES infrastructure contributes to this international challenge. The future of climate modeling highly depends on available computing power: ensemble of prediction experiments, increase of resolution to better represent small scale processes, complexity of the Earth's climate system, duration of experiments to investigate climate stability, are all limited by computing power. Massively parallel computing starts to address resolution increase and ensemble runs but still raises a number of issues as emphasized by the ENES infrastructure strategy.

Achieving Design Targets by Stochastic Car Crash Simulations - The Relation between Bifurcation of Deformation and Quality of FE Models

Tsuyoshi Yasuki (Toyota Motor Corporation)

Accuracy of car crash simulation is one of key issues for car development of automotive industries. A stochastic car crash simulation was performed using a very detailed car crash FE model. Car deformations of these simulations indicated bifurcations of buckling modes of frontal side rails, while rough car crash FE models did not indicate it.

Thursday, November 15

Thinking about the Future of Large-Scale Computing

Chair: Bronis R. de Supinski (Lawrence Livermore National Laboratory)

8:30am-10am

Room: Ballroom-EFGH

A Journey to Exascale Computing

William J. Harrod (DOE Office of Advanced Scientific Computing Research)

Exascale computing is a shared international pursuit aimed at creating a new class of high performance computing systems that can achieve a thousand times the sustained performance of today's petascale computers while limiting growth in space and power requirements. Although the primary goal of this pursuit is to develop leading edge computing assets for new scientific discovery, medical science, climate modeling, and other compute- and data-intensive applications, the resulting technologies will have a profound impact on all future computing systems down to laptops and handheld devices.

Computing is now at a critical crossroads. We can no longer proceed down the path of steady but incremental progress to which we have become accustomed. Thus, exascale computing is not simply an effort to provide the next level of computational power by creative scaling up of current petascale computing systems.

New architectures will be required to achieve the exascale computing goals. Although there are many daunting challenges, which have been identified and extensively examined in numerous previous studies, past and ongoing pilot projects have indicated the feasibility of achieving the exascale goals. However, development of exascale technology is not just a hardware problem. A significant investment in system software, programming models, and applications algorithms and codes is required as well.

While there are different interpretations of the specific system details for an exascale computer, there is fundamental agreement concerning the challenges and general design features. The pioneering generation of exascale computers will likely consist of heterogeneous processors that have thousands of computing elements per processor. Data movement can no longer be considered a “free” operation, as it drives power consumption across the system. Resiliency will also be a significant concern. The potential complexity of the system could be a significant challenge for achieving highly programmable computers.

Industry will not be able to achieve these goals without substantial governmental investment. The realization of exascale computing systems and technology rests on partnerships among academia, industry, and government institutions, and on international collaboration.

This presentation will focus on the strategy and plans for developing and deploying energy efficient, highly programmable exascale computers by the early 2020s. Various challenges will be discussed, including technical, programmatic, and policy issues.

The Evolution of GPU Accelerated Computing

Steve Scott (NVIDIA)

GPUs were invented in 1999, replacing fixed-function graphics pipelines with fully programmable processors, and ushering in the era of computational graphics. By 2003, early pioneers were using graphics APIs to perform general purpose scientific calculations on GPUs, and by 2007, NVIDIA responded by creating the CUDA architecture and programming language. GPU accelerated computing has evolved rapidly from there. This talk will briefly recount the history of GPU computing, assess where we are today, and explore how GPU computing will evolve over the coming decade as we pursue Exascale computing.

Optimizing Applications for New Systems

Chair: Satoshi Matsuoka (Tokyo Institute of Technology)

10:30am-12pm

Room: Ballroom-EFGH

Application Development for Titan - A Multi-Petaflop Hybrid-Multicore MPP System

John Levesque (Cray Inc.)

Oak Ridge National Laboratory has installed the Oak Ridge Leadership Computing Facility - 3, a large multi-Petaflop XK6 consisting of 10s of thousands of Nvidia GPUs. The system is called Titan. Two years ago, five major applications were identified for porting to the Titan system so that break-through science could be run as soon as the system was installed.

This talk will discuss the process over the two years, the various approaches taken including CUDA, CUDA Fortran, C++ meta-programming templates and OpenACC directives targeting the accelerator. Going forward, a large group of users will be faced with the challenges for moving their applications to the new Titan system. After reviewing the strategies and successes of porting the five applications to the system, recommendations will be made outlining an approach for moving to the architecture that promises to deliver over an exaflop in performance in the future.

Application Performance Characterization and Analysis on Blue Gene/Q

Bob Walkup (IBM)

The Blue Gene/Q system presents a significant challenge to understanding application performance because of the degree of concurrency supported both within a node and across the whole system. These challenges are representative of the application characterization problem for future exascale systems. In this talk, we present tools developed at IBM research to collect and analyze performance data from applications running at scale on Blue Gene/Q. Using these tools we characterize performance for a range of applications running at scale on Blue Gene/Q.

Large-Scale Simulations

Chair: Robert F. Lucas (Information Sciences Institute)

3:30pm-5pm

Room: Ballroom-EFGH

Low Mach Number Models in Computational Astrophysics

Ann Almgren (Lawrence Berkeley National Laboratory)

A number of astrophysical phenomena are characterized by low Mach number flows. Examples include the convective phase of a white dwarf prior to ignition as a Type Ia supernova, X-ray bursts, and convection in massive stars. Low Mach number models analytically remove acoustic wave propagation while retaining compressibility effects resulting from nuclear reactions and ambient stratification. This enables time steps controlled by advective time scales, resulting in significant reduction in computational costs. We will discuss the derivation and implementation of low Mach number models for astrophysical applications, particularly in the context of structured grid AMR, and will present computational results from three-dimensional adaptive mesh simulations.

Very Large-Scale Fluid-Flow and Aeroacoustical Simulations for Engineering Applications Performed on the K Supercomputer

Chisachi Kato (University of Tokyo)

With tremendous speed-up of high-end computers, applications of fully-resolved Large Eddy Simulation (FRLES) is becoming feasible for engineering problems. FRLES directly computes all eddies responsible for the production of turbulence and thus, is expected to give as accurate results as Direct Numerical Simulation (DNS) does with by a factor of several tens smaller computational cost than that of DNS. The authors have developed flow and acoustical solvers capable of performing very large-scale engineering applications. We have already achieved a sustained performance of 8 percent to the theoretical processor performance and a parallel scalability more than 90 percent for 524,288 cores. The first results of very large-scale fluid-flow and aeroacoustical simulations of industrial problems using several tens of billion grids will be presented.

Panels

Tuesday, November 13

HPC's Role In The Future of American Manufacturing

10:30am-12pm

Room: 355-BC

Moderator: C. William Booher (Council on Competitiveness)

Panelists: Thomas Lange (The Procter and Gamble Company), Richard Arthur (G.E. Global Research), Sridhar Kota (Office of Science and Technology Policy), Brian Rosenboom (Rosenboom Machine & Tool Inc.)

Given the pressures of a global economy, American manufacturing competitiveness and, in fact, the future of our manufacturing sector, hinges on our ability to become highly innovative, adaptable, agile and efficient. Fortune 50 manufacturers utilize HPC for modeling, simulation, design, process and supply chain optimization, and on the manufacturing floor to accomplish these goals. But adoption of the digital manufacturing lifecycle paradigm is only just now taking root in this country. Many speak about the potential for a manufacturing renaissance in the U. S. based in large part on our technological and computational prowess. Digital manufacturing is front and center in the national agenda. It is a key point in almost every report, address and discussion on the future of the U. S. economy. What is the role of HPC in this vision and how will this play out? How can we succeed in bootstrapping an American manufacturing renaissance?

NSF-TCPP Curriculum Initiative on Parallel and Distributed Computing - Core Topics for Undergraduates

3:30pm-5pm

Room: 355-BC

Moderator: Sushil Prasad (Georgia State University)

Panelists: Alan Sussman (University of Maryland), Andrew Lumsdaine (Indiana University), Almadena Chtchelkanova (National Science Foundation), David Padua (University of Illinois at Urbana-Champaign), Krishna Kant (Intel Corporation), Manish Parashar (Rutgers University), Arny Rosenberg (Northeastern University), Jie Wu (Temple University)

A working group from IEEE Technical Committee on Parallel Processing (TCPP), NSF, and the sister communities, including ACM, has taken up proposing a parallel and distributed computing (PDC) curriculum for computer science (CS) and computer engineering (CE) undergraduates. The premise of the working group is that every CS/CE undergraduate should achieve a specified skill level in PDC-related topics from the required coursework. We released a preliminary report in

Dec 2010, and have had about four dozen early adopter institutions worldwide try the proposed curriculum. We have assimilated feedback from these early adopters, stakeholders and the community at large, and are preparing to release its first formal version. The proposed panel will introduce this curriculum and its rationale to the SC community. Similar panels at HiPC-10, SIGCSE-11, and IPDPS-11 were very successful, frequently arousing audience with passionate arguments and disagreements, but all agreeing to the urgent need.

Wednesday, November 14

Boosting European HPC Value Chain: The Vision of ETP4HPC, the European Technology Platform for High Performance Computing

10:30am – 12pm

Room: 355-BC

Moderator: Jean-Francois Lavignon (Bull)

Panelists: Giampietro Tecchioli (Eurotech), Hugo Falter (ParTec), David Lecomber (Allinea), Francesc Subirada (Barcelona Supercomputing Center), François Bodin (CAPS Enterprise), Jean Gonnord (CEA), Guy Lonsdale (Fraunhofer), Thomas Lippert (FZI), Andreas Pflieger (IBM), Bernadette Andrietti (Intel Corporation), Arndt Bode (Leibniz Supercomputing Centre), Catherine Rivière (GENCI), Ken Claffey (Xyratex)

ETP4HPC has been created to define a Strategic Research Agenda in the area of HPC technology supply, and discuss with the European Commission on implementing HPC research programs within such frameworks as HORIZON 2020. ETP4HPC is an industry-led forum with both industrial and academic members, aiming at improving the competitiveness of European HPC industry, which can benefit the entire European economy. During this session, panelists from ETP4HPC and other stakeholders of the European or international HPC ecosystem will present and discuss several aspects of the ETP4HPC mission and vision: Explain the motivation for creating ETP4HPC and demonstrate the benefits of being an ETP4HPC member; Present current status of the Strategic Research Agenda and its main research priorities; Discuss the implementation of these research tasks; and Discuss how to strengthen the international cooperation in HPC.

Exascale and Big Data I/O: Which Will Drive Future I/O Architectures, Standards and Protocols? Should they be Open or Proprietary?

1:30pm-3pm

Room: 355-BC

Moderator: Bill Boas (InfiniBand Trade Association)

Panelists: Peter Braam (Xyratex), Sorin Fabish (EMC), Ronald Luijten (IBM Zurich Research Laboratory), Duncan Roweth (Cray Inc.), Michael Kagan (Mellanox Technologies), Paul Grun (Cray Inc.), Moray McLaren (HP), Manoj Wadekar (QLogic)

The requirements for Exascale and Big Data I/O are driving research, architectural deliberation, technology selection and product evolution throughout HPC and Enterprise/Cloud/Web computing, networking and storage systems. Analysis of the Top500 interconnect families over the past decade reveals an era when standard commodity I/O technologies have come to dominate, almost completely. Speeds have gone from 1 Gigabit to over 50 Gigabits, latencies have decreased 10X to below a microsecond and software has evolved towards one software stack—OpenFabrics/OFED. The Enterprise is now adopting these same technologies at a rapid rate. From the perspective of the major OEM suppliers worldwide of computer systems fabrics and storage, what does the future hold for the next generation of interconnects and system I/O architectures as the fabric integrates into the industry standard processing chips.

Visualization Frameworks for Multi-Core and Many-Core Architectures

3:30pm-5pm

Room: 355-BC

Moderator: Hank Childs (Lawrence Berkeley National Laboratory)

Panelists: Jeremy Meredith (Oak Ridge National Laboratory), Patrick McCormick, Christopher Sewell (Los Alamos National Laboratory), Kenneth Moreland (Sandia National Laboratories)

Multi-core and many-core nodes, already prevalent today, are essential to address the power constraints for achieving greater compute levels. Today's visualization software packages have been slow to keep pace; they often employ only distributed memory parallel techniques even when running on hardware where hybrid parallelism would provide substantial benefit. Worse, power costs and relative disk performance will mandate in situ visualization in the future; visualization software will be required to run effectively on multi-core and many-core nodes. Fortunately, visualization software is emerging for these environments. In this panel, developers of DAX, EAVL, PISTON, as well as a developer of a DSL for visualization, will describe their frameworks. The panel format will have

each panelist answer the same questions, to inform the audience about: - their approaches to exascale issues, such as massive concurrency, memory overhead, fault tolerance, etc, - the long-term result for this effort (Production software? Research prototype?)

Thursday, November 15

Current Status of HPC Progress in China

10:30am-12pm

Room: 355-BC

Moderator: David Kahaner (Asian Technology Information Program)

Panelists: Kai Lu (National University of Defense Technology), Zeyao Mo (Institute of Applied Physics and Computational Mathematics), Zhiwei Xu (Institute of Computing Technology), Jingshan Pan (National Supercomputing Center in Jinan), Depei Qian (Beihang University), Yunquan Zhang (Chinese Academy of Sciences), James Lin (Shanghai Jiao Tong University)

In recent years, China has made remarkable progress in the field of HPC. A full-day workshop held by ATIP and NSF at SC10 showcased Chinese developments at exactly the time China's Tianhe 1A became #1 on the Top500 List. Rapid technical progress has continued since then, and China's participation in the SC conference has grown. This year features nine separate exhibits from China (PRC). ATIP's Panel of Chinese HPC Experts will provide an opportunity for conference participants to hear from representatives of the key organizations building computers and developing HPC software applications in China.

Big Data and Data Integrity - How Do We Handle Reliability, Reproducibility, and Accessibility?

3:30pm-5pm

Room: 355-BC

Moderator: Tracey D. Wilson (Computer Sciences Corporation)

Panelists: Ron Bewtra (NOAA), Daniel Duffy (NASA), James Rogers (Oak Ridge National Laboratory), Lee Ward (Sandia National Laboratories), Keith Michaels (Boeing)

As HPC data becomes larger and more complex, so does our need to maintain its integrity. Threats to our data's integrity come at different levels. The HPC community has experience these already in transfers of large data over the wide area, reproduction of data from large complex systems, silent corruption of data in our compute and file systems, and the requirements and expectations the community expects from HPC

compute and storage clouds. This panel of experts will discuss their organizations views on the various issues and explain potential solutions to these pressing issues.

Friday, November 16

Applying High Performance Computing at a National Laboratory to the Needs of Industry Case Study - hpc4energy Incubator

8:30am-10am

Room: 355-EF

Moderator: John Grosh (Lawrence Livermore National Laboratory)

Panelists: Devin Van Zandt (GE Energy Consulting), Madhusudan Pai (GE Global Research), Eric Doran (Robert Bosch LLC), Tom Wideman (Potter Drilling Inc.), Robert LaBarre (United Technologies Research Center), Eugene Litvinov (ISO New England)

Hpc4energy—using HPC to reduce development time and costs in the energy industry. This panel will discuss the different methods used to access HPC using the hpc4energy incubator as a case study. Representatives from companies will discuss the barriers to adoption of HPC, the benefits and opportunities made available by the use of HPC, and future directions for public-private partnerships solidifying methods to access HPC. Lawrence Livermore has begun to work with companies of all sizes in the energy industry to produce more efficient and advanced energy technologies and strengthen U.S. industrial competitiveness in the hpc4energy incubator. Viewing this initiative as part of its mission to protect national security, Lawrence Livermore teamed computer scientists and domain experts, in addition to computational power with energy experts in industry to demonstrate the ability of high performance computing (HPC) to catalyze rapid advancement of energy technologies.

Is the Cloud a Game Changer for the Film Industry

10:30am-12pm

Room: 355-EF

Moderator: Scott Houston (GreenButton)

Panelists: Chris Ford (Pixar), Doug Hauger (Microsoft Corporation)

Pixar and GreenButton have recently launched a cloud-based rendering service aimed at the film industry. The service is based on the Windows Azure platform and enables studios to use the cloud to render large scale motion picture productions. This is a game changer for the industry as it enables smaller studios to compete with major production companies without the need to invest capital in expensive infrastructure

and technical support. The panel discusses the particular technical challenges to support this business model and how long before a visual effects Oscar will be won by a company that has no major technical infrastructure and artists that are spread across the globe.

Questions: Rhetorically Volleying the Terabit Network Ball on the HPC Court

10:30am-12pm

Room: 355-BC

Moderator: Dan Gunter (Lawrence Berkeley National Laboratory)

Panelists: Ezra Kissel (University of Delaware), Raj Kettimuthu (Argonne National Laboratory), Jason Zurawski (Internet2)

The landscape of questions presented by the intersection of high-performance computing, high-performance storage, and high-performance networking is always full of questions. This panel attempts to raise at least as many questions as it answers by engaging the panelists and the audience in a rapid-fire dialogue with no holds barred. Questions such as “What would we do with terabit networks if we had them”? Will software-defined networking change everything? Is TCP dead for bulk data transfer? Should we all start using RDMA? Will solid-state disks outstrip networks? Are parallel file systems really important? Whatever happened to P2P? Will users continue using thumb drives? How can we stop them? Should we stop them? Does SaaS solve this problem? Will our data start to live “in the cloud”? How does that drive our plans for the future? Strap on your shoes, grab a rhetorical racquet and join the conversation!



Papers

The SC12 Technical Papers program received 472 submissions covering a wide variety of research topics in high performance computing. We followed a rigorous peer review process with a newly introduced author rebuttal period, careful management of conflicts, and four reviews per submission (in most cases). At a two-day face-to-face committee meeting June 25-26 in Salt Lake City, over 100 technical committee members discussed every paper and finalized the selections. At the conclusion of the meeting, the committee accepted 100 papers, reflecting an acceptance rate of 21 percent. Additionally, 13 of the 100 accepted papers have been selected as finalists for the Best Paper and Best Student Paper awards. Winners will be announced during the Awards Session on Thursday, November 15.

Papers

Tuesday, November 13

Analysis of I/O and Storage

Chair: Robert B. Ross (Argonne National Laboratory)

10:30am-12pm

Room: 355-EF

Demonstrating Lustre over a 100Gbps Wide Area Network of 3,500km

Authors: Robert Henschel, Stephen Simms, David Hancock, Scott Michael, Tom Johnson, Nathan Heald (Indiana University), Thomas William (Technical University Dresden), Donald Berry, Matt Allen, Richard Knepper, Matthew Davy, Matthew Link, Craig Stewart (Indiana University)

As part of the SCinet Research Sandbox at SC11, Indiana University demonstrated use of the Lustre high performance parallel file system over a dedicated 100 Gbps wide area network (WAN) spanning more than 3,500 km (2,175 mi). This demonstration functioned as a proof of concept and provided an opportunity to study Lustre's performance over a 100 Gbps WAN. To characterize the performance of the network and file system, low level iperf network tests, file system tests with the IOR benchmark, and a suite of real-world applications reading and writing to the file system were run over a latency of 50.5 ms. In this article we describe the configuration and constraints of the demonstration and outline key findings.

A Study on Data Deduplication in HPC Storage Systems

Authors: Dirk Meister, Jürgen Kaiser, Andre Brinkmann (Johannes Gutenberg University Mainz), Toni Cortes (Barcelona Supercomputing Center), Michael Kuhn, Julian Kunkel (University of Hamburg)

Deduplication is a storage-saving technique that is successful in backup environments. On a file system a single data block might be stored multiple times across different files; for example, multiple versions of a file might exist that are mostly identical. With deduplication this data replication is localized and redundancy is removed. This paper presents the first study on the potential of data deduplication in HPC centers, which belongs to the most demanding storage producers. We have quantitatively assessed this potential for capacity reduction for four data centers. We have analyzed over 1212 TB of file system data. The evaluation shows that typically 20% to 30% of this online data could be removed by applying data deduplication techniques, peaking up to 70% for some data sets. Interestingly, this reduction can only be achieved by a subfile deduplication approach, while approaches based on whole-file comparisons only lead to small capacity savings.

Characterizing Output Bottlenecks in a Supercomputer

Authors: Bing Xie, Jeff Chase (Duke University), David Dillow (Oak Ridge National Laboratory), Oleg Drokin (Whamcloud, Inc.), Scott Klasky, Sarp Oral, Norbert Podhorszki (Oak Ridge National Laboratory)

Supercomputer I/O loads are often dominated by writes. HPC file systems are designed to absorb these bursty outputs at high bandwidth through massive parallelism. However, the delivered write bandwidth often falls well below the peak. This paper characterizes the data absorption behavior of a center-wide shared Lustre parallel file system on the Jaguar supercomputer. We use a statistical methodology to address the challenges of accurately measuring a shared machine under production load and to obtain the distribution of bandwidth across samples of compute nodes, storage targets, and time intervals. We observe and quantify limitations from competing traffic, contention on storage servers and I/O routers, concurrency limitations in the client compute node operating systems, and the impact of variance (stragglers) on coupled output such as striping. We then examine the implications of our results for application performance and the design of I/O middleware systems on shared supercomputers.

Finalist: Best Student Paper Award, Best Paper Award

Autotuning and Search-Based Optimization

Chair: Francois Bodin (CAPS)

10:30am-12pm

Room: 355-D

Portable Section-Level Tuning of Compiler Parallelized Applications

Authors: Dheya Mustafa, Rudolf Eigenmann (Purdue University)

Automatic parallelization of sequential programs combined with tuning is an alternative to manual parallelization. This method has the potential to substantially increase productivity and is thus of critical importance for exploiting the increased computational power of today's multicores. A key difficulty is that parallelizing compilers are generally unable to estimate the performance impact of an optimization on a whole program or a program section at compile time; hence, the ultimate performance decision today rests with the developer. Building an autotuning system to remedy this situation is not a trivial task. This work presents a portable empirical autotuning system that operates at program-section granularity and partitions the compiler options into groups that can be tuned independently. To our knowledge, this is the first approach delivering an autoparallelization system that ensures performance improvements for nearly all programs, eliminating the users' need to "experiment" with such tools to strive for highest application performance.

A Multi-Objective Auto-Tuning Framework for Parallel Codes

Authors: Herbert Jordan, Peter Thoman, Juan J. Durillo, Simone Pellegrini, Philipp Gschwandtner, Thomas Fahringer, Hans Moritsch (University of Innsbruck)

In this paper we introduce a multi-objective auto-tuning framework comprising compiler and runtime components. Focusing on individual code regions, our compiler uses a novel search technique to compute a set of optimal solutions, which are encoded into a multi-versioned executable. This enables the runtime system to choose specifically tuned code versions when dynamically adjusting to changing circumstances. We demonstrate our method by tuning loop tiling in cache-sensitive parallel programs, optimizing for both runtime and efficiency. Our static optimizer finds solutions matching or surpassing those determined by exhaustively sampling the search space on a regular grid, while using less than 4% of the computational effort on average. Additionally, we show that parallelism-aware multi-versioning approaches like our own gain a performance improvement of up to 70% over solutions tuned for only one specific number of threads.

PATUS for Convenient High-Performance Stencils: Evaluation in Earthquake Simulations

Authors: Matthias Christen, Olaf Schenk (USI Lugano), Yifeng Cui (San Diego Supercomputer Center)

Patus is a code generation and auto-tuning framework for the class of stencil computations targeted at modern multi- and many-core processors. The goals of the framework are productivity, portability, and achieving a high performance on the target platform. Its stencil specification DSL allows the programmer to express the computation in a concise way independently of hardware architecture-specific details. Thus, it increases the programmer productivity by disburdening her or him of low level programming model issues. We illustrate the impact of the stencil code generation in seismic applications. The challenges in computational seismology are to harness these stencil-code generation techniques for wave propagation problems, for which weak and strong scaling are important. We evaluate the performance showing two examples: (1) focusing on a scalable discretization of the wave equation, and (2) testing complex simulation types of the AWP-ODC code to enable petascale 3D earthquake calculations and aims on aggressive parallel efficiency.

Breadth-First Search

Chair: Umit Catalyurek (Ohio State University)

10:30am-12pm

Room: 255-EF

Direction-Optimizing Breadth-First Search

Authors: Scott Beamer, Krste Asanović, David Patterson (University of California, Berkeley)

Breadth-First Search is an important kernel used by many graph-processing applications. In many of these emerging applications of BFS, such as analyzing social networks, the input graphs are low-diameter and scale-free. We present an efficient breadth-first search algorithm that is advantageous for low-diameter graphs. We adopt a hybrid approach, combining a conventional top-down algorithm along with a novel bottom-up algorithm. The bottom-up algorithm can dramatically reduce the number of edges examined, which in turn accelerates the search as a whole. On a multi-socket server, our hybrid approach demonstrates speedups of 3.3-7.8 on a range of standard synthetic graphs and speedups of 1.4-3.8 on graphs from real social networks compared to a strong baseline. We also show speedups of greater than 2.3 over the state-of-the-art multicore implementation when using the same hardware and input graphs.

Finalist: Best Student Paper Award

Breaking the Speed and Scalability Barriers for Graph Exploration on Distributed-Memory Machines

Authors: Fabio Checconi, Fabrizio Petrini (IBM T.J. Watson Research Center), Jeremiah Willcock, Andrew Lumsdaine (Indiana University), Yogish Sabharwal, Anamitra Choudhury (IBM India)

In this paper, we describe the challenges involved in designing a family of highly-efficient Breadth-First Search (BFS) algorithms and in optimizing these algorithms on the latest two generations of Blue Gene machines, Blue Gene/P and Blue Gene/Q. With our recent winning Graph 500 submissions in November 2010, June 2011, and November 2011, we have achieved unprecedented scalability results in both space and size. On Blue Gene/P, we have been able to parallelize the largest BFS search presented in the literature, running a scale 38 problem with 238 vertices and 242 edges on 131,072 processing cores. Using only four racks of an experimental configuration of Blue Gene/Q, we have achieved the fastest processing rate reported to date on a BFS search, 254 billion edges per second on 65,536 processing cores. This paper describes the algorithmic design and the main classes of optimizations that we have used to achieve these results.

Large-Scale Energy-Efficient Graph Traversal - A Path to Efficient Data-Intensive Supercomputing

Authors: Nadathur Satish (Intel Corporation), Changkyu Kim (Intel Corporation), Jatin Chhugani (Intel Corporation), Pradeep Dubey (Intel Corporation)

Graph-traversal is used in many fields including social-networks, bioinformatics and HPC. The push for HPC machines to be rated in "GigaTEPS" (billions-of-traversed-edges-per-second) has led to the Graph500 benchmark. Graph-traversal is well-optimized for single-node CPUs. However, current cluster implementations suffer from high-latency and large-volume inter-node communication, with low performance and energy-efficiency. In this work, we use novel low-overhead data-compression techniques to reduce communication-volumes along with new latency-hiding techniques. Keeping the same optimized single-node algorithm, we obtain 6.6X performance improvement and order-of-magnitude energy savings over state-of-the-art techniques. Our Graph500 implementation achieves 115 GigaTEPS on 320-node Intel-Endeavor cluster with E5-2700 Sandy Bridge nodes, matching the second-ranked result in the November-2011 Graph500 list with 5.6X fewer nodes. Our per-node performance only drops 1.8X over optimized single-node implementations, and is highest in the top 10 of the list. We obtain near-linear scaling with node count. On 1024 Westmere-nodes of the NASA-Pleiadas system, we obtain 195 GigaTEPS.

Direct Numerical Simulations

Chair: Martin Berzins (University of Utah)

10:30am-12pm

Room: 255-BC

Hybridizing S3D into an Exascale Application using OpenACC

Authors: John Michael Levesque (Cray Inc.), Grout Ray (National Renewable Energy Laboratory), Ramanan Sankaran (Oak Ridge National Laboratory)

Hybridization is the process of converting an application with a single level of parallelism to an application with multiple levels of parallelism. Over the past 15 years a majority of the applications that run on HPC systems have employed MPI for all of the parallelism within the application. In the peta-exascale computing regime, effective utilization of the hardware requires multiple levels of parallelism matched to the macro architecture of the system to achieve good performance. A hybridized code base is performance portable when sufficient parallelism is expressed in a architecture agnostic form to achieve good performance on available systems. The hybridized S3D code is performance portable across today's leading many core and GPU accelerated systems. The OpenACC framework allows a unified code base to be deployed for either (Manycore CPU or Manycore CPU+GPU) while permitting architecture specific optimizations to expose new dimensions of parallelism to be utilized.

High Throughput Software for Direct Numerical Simulations of Compressible Two-Phase Flows

Authors: Babak Hejiazalhosseini, Diego Rossinelli, Christian Conti, Petros Koumoutsakos (ETH Zurich)

We present an open source, object-oriented software for high throughput Direct Numerical Simulations of compressible, two-phase flows. The Navier-Stokes equations are discretized on uniform grids using high order finite volume methods. The software exploits recent CPU micro-architectures by explicit vectorization and adopts NUMA-aware techniques as well as data and computation reordering. We report a compressible flow solver with unprecedented fractions of peak performance: 45% of the peak for a single node (nominal performance of 840 GFLOP/s) and 30% for a cluster of 47'000 cores (nominal performance of 0.8 PFLOP/s). We suggest that the present work may serve as a performance upper bound, regarding achievable GFLOP/s, for two-phase flow solvers using adaptive mesh refinement. The software enables 3D simulations of shock-bubble interaction including, for the first time, effects of diffusion and surface tension, by efficiently employing two hundred billion computational elements.

Checkpointing

Chair: Frank Mueller (North Carolina State University)

1:30pm-3pm

Room: 255-EF

mcrEngine: A Scalable Checkpointing System Using Data-Aware Aggregation and Compression

Authors: Tanzima Z. Islam (Purdue University), Kathryn Mohror (Lawrence Livermore National Laboratory), Saurabh Bagchi (Purdue University), Adam Moody, Bronis R. de Supinski (Lawrence Livermore National Laboratory), Rudolf Eigenmann (Purdue University)

HPC systems use checkpoint and restart for tolerating failures. Typically, applications store their states in checkpoints on a parallel file system (PFS). As applications scale up, checkpoint-restart incurs high overheads due to contention for overloaded PFS resources. The high overheads force large-scale applications to reduce checkpoint frequency, which means more compute time is lost in the event of a failure. To alleviate these problems, we demonstrate a scalable checkpoint-restart system, mcrEngine. mcrEngine aggregates checkpoints from multiple application processes with knowledge of the data semantics available through widely-used I/O libraries, e.g., HDF5 and netCDF and compresses them. Our novel scheme improves compressibility of checkpoints up to 115% over simple concatenation and compression. Experimental results with large-scale application checkpoints show that mcrEngine reduces checkpointing overhead by up to 87% and recovery overhead by up to 62% over a baseline with no aggregation or compression.

Finalist: Best Student Paper Award

Alleviating Scalability Issues of Checkpointing Protocols

Authors: Rolf Riesen (IBM), Kurt Ferreira (Sandia National Laboratories), Dilma Da Silva, Pierre Lemarinier (IBM), Dorian Arnold, Patrick G. Bridges (University of New Mexico)

Current fault-tolerance protocols are not sufficiently scalable for the exascale era. The most widely-used method, coordinated checkpointing, places enormous demands on the I/O subsystem and imposes frequent synchronizations. Uncoordinated protocols use message logging which introduces message rate limitations or undesired memory and storage requirements to hold payload and event logs. In this paper, we propose a combination of several techniques, namely coordinated checkpointing, optimistic message logging, and a protocol that glues them together. This combination eliminates some of the drawbacks of each individual approach and proves to be an alternative for many types of exascale applications. We evaluate performance and scaling characteristics of this combination using simulation and a partial implementation. While not a universal solution, the combined protocol is suitable for a large range of existing and future applications that use coordinated checkpointing and enhances their scalability.

Design and Modeling of a Non-Blocking Checkpointing System

Authors: Kento Sato (Tokyo Institute of Technology), Adam Moody, Kathryn Mohror, Todd Gamblin, Bronis R. de Supinski (Lawrence Livermore National Laboratory), Naoya Maruyama (RIKEN), Satoshi Matsuoka (Tokyo Institute of Technology)

As the capability and component count of PFS systems increase, the MTBF correspondingly decreases. Typically, applications tolerate failures with checkpoint/restart using a PFS. While simple, this approach suffers from high overhead due to contention for PFS resources. A promising solution to this problem is multi-level checkpointing. However, while multi-level checkpointing is successful on today's machines, it is not expected to be sufficient for exascale class machines, where the total memory sizes and failure rates are predicted to be orders of magnitude higher. Our solution to this problem is a system that combines the benefits of non-blocking and multi-level checkpointing. In this paper, we present the design of our system and a model describing its performance. Our experiments show that our system can improve efficiency by 1.1 to $2.0 \times$ on future machines. Additionally, applications using our checkpointing system can achieve high efficiency even when using a PFS with lower bandwidth.

Cloud Computing

Chair: Marty A. Humphrey (University of Virginia)

1:30pm-3pm

Room: 355-D

Scalia: An Adaptive Scheme for Efficient Multi-Cloud Storage

Authors: Thanasis G. Papaioannou, Nicolas Bonvin, Karl Aberer (EPFL)

A growing amount of data is produced daily resulting in a growing demand for storage solutions. While cloud storage providers offer a virtually infinite storage capacity, data owners seek geographical and provider diversity in data placement, in order to avoid vendor lock-in and to increase availability and durability. Moreover, depending on the customer data access pattern, a certain cloud provider may be cheaper than another. In this paper, we introduce Scalia, a cloud storage brokerage solution that continuously adapts the placement of data based on its access pattern and subject to optimization objectives, such as storage costs. Scalia cleverly considers re-positioning of only selected objects that may significantly lower the storage cost. By extensive simulation experiments, we prove the cost effectiveness of Scalia against static placements and its proximity to the ideal data placement in various scenarios of data access patterns, of available cloud storage solutions and of failures.

Host Load Prediction in a Google Compute Cloud with a Bayesian Model

Authors: Sheng Di, Derrick Kondo (INRIA), Walfredo Cirne (Google)

Prediction of host load in Cloud systems is critical for achieving service-level agreements. However, accurate prediction of host load in Clouds is extremely challenging because it fluctuates drastically at small timescales. We design a prediction method based on Bayes model to predict the mean load over a long-term time interval, as well as the mean load in consecutive future time intervals. We identify novel predictive features of host load that capture the expectation, predictability, trends and patterns of host load. We also determine the most effective combinations of these features for prediction. We evaluate our method using a detailed one-month trace of a Google data center with thousands of machines. Experiments show that the Bayes method achieves high accuracy with a mean squared error of 0.0014. Moreover, the Bayes method improves the load prediction accuracy by 5.6-50% compared to other state-of-the-art methods based on moving averages, auto-regression, and/or noise filters.

Cost and Deadline-Constrained Provisioning for Scientific Workflow Ensembles in IaaS Clouds

Authors: Maciej Malawski (AGH University of Science and Technology), Gideon Juve, Ewa Deelman (University of Southern California), Jarek Nabrzyski (University of Notre Dame)

Large-scale applications expressed as scientific workflows are often grouped into ensembles of inter-related workflows. In this paper, we address a new and important problem concerning the efficient management of such ensembles under budget and deadline constraints on Infrastructure-as-a-Service (IaaS) clouds. We discuss, develop, and assess algorithms based on static and dynamic strategies for both task scheduling and resource provisioning. We perform the evaluation via simulation using a set of scientific workflow ensembles with a broad range of budget and deadline parameters, taking into account uncertainties in task runtime estimations, provisioning delays, and failures. We find that the key factor determining the performance of an algorithm is its ability to decide which workflows in an ensemble to admit or reject for execution. Our results show that an admission procedure based on workflow structure and estimates of task runtimes can significantly improve the quality of solutions.

GPU Programming Models and Patterns

Chair: Michael A. Heroux (Sandia National Laboratories)

1:30pm-3pm

Room: 355-EF

Early Evaluation of Directive-Based GPU Programming Models for Productive Exascale Computing

Authors: Seyong Lee, Jeffrey S. Vetter (Oak Ridge National Laboratory)

Graphics Processing Unit (GPU)-based parallel computer architectures have shown increased popularity as a building block for high performance computing, and possibly for future exascale computing. However, their programming complexity remains as a major hurdle for their widespread adoption. To provide better abstractions for programming GPU architectures, researchers and vendors have proposed several directive-based GPU programming models. These directive-based models provide different levels of abstraction and required different levels of programming effort to port and optimize applications. Understanding these differences among these new models provides valuable insights on their applicability and performance potential. In this paper, we evaluate existing directive-based models by porting thirteen application kernels from various scientific domains to use CUDA GPUs, which, in turn, allows us to identify important issues in the functionality, scalability, tunability, and debuggability of the existing models. Our evaluation shows that directive-based models can achieve reasonable performance, compared to hand-written GPU codes.

Automatic Generation of Software Pipelines for Heterogeneous Parallel Systems

Authors: Jacques A. Pienaar, Anand Raghunathan (Purdue University), Srimat Chakradhar (NEC Laboratories America)

Pipelining is a well-known approach to increasing parallelism and performance. We address the problem of software pipelining for heterogeneous parallel platforms that consist of different multi-core and many-core processing units. In this context, pipelining involves two key steps---partitioning an application into stages and mapping and scheduling the stages onto the processing units of the heterogeneous platform. We show that the inter-dependency between these steps is a critical challenge that must be addressed in order to achieve high performance. We propose an Automatic Heterogeneous Pipelining framework (AHP) that automatically generates an optimized pipelined implementation of a program from an annotated unpipelined specification. Across three complex applications (image classification, object detection, and document retrieval) and two heterogeneous platforms (Intel Xeon multi-core CPUs with Intel MIC and NVIDIA GPGPU accelerators), AHP achieves a throughput improvement of up to 1.53x (1.37x on average) over a heterogeneous baseline that exploits data and task parallelism.

Accelerating MapReduce on a Coupled CPU-GPU Architecture

Authors: Linchuan Chen, Xin Huo, Gagan Agrawal (Ohio State University)

The work presented here is driven by two observations. First, heterogeneous architectures that integrate the CPU and the GPU on the same chip are emerging, and hold much promise for supporting power-efficient and scalable high performance computing. Secondly, MapReduce has emerged as a suitable framework for simplified parallel application development for many classes of applications, including data mining and machine learning applications that benefit from accelerators. This paper focuses on the challenge of scaling a MapReduce application using the CPU and GPU together in an integrated architecture. We use different methods for dividing the work: a map-dividing scheme, which divides map tasks on both devices, and a pipelining scheme, which pipelines the map and the reduce stages on different devices. We develop dynamic work distribution schemes for both the approaches. To achieve high performance, we use a runtime tuning method to adjust task block sizes.

Maximizing Performance on Multi-Core and Many-Core Architectures

Chair: Atsushi Hori (RIKEN)

1:30pm-3pm

Room: 255-BC

Unleashing the High Performance and Low Power of Multi-Core DSPs for General-Purpose HPC

Authors: Francisco D. Igual (Texas Advanced Computing Center), Murtaza Ali, Arnon Friedmann, Eric Stotzer (Texas Instruments), Timothy Wentz (University of Illinois at Urbana-Champaign), Robert A. van de Geijn (University of Texas at Austin)

Take a multicore Digital Signal Processor (DSP) chip designed for cellular base stations and radio network controllers, add floating-point capabilities to support 4G networks, and out of thin air a HPC engine is born. The potential for HPC is clear: it promises 128 GFLOPS (single precision) for 10 Watts; it is used in millions of network related devices and hence benefits from economies of scale; it should be simpler to program than a GPU. Simply put, it is fast, green, and cheap. But is it easy to use? In this paper, we show how this potential can be applied to general-purpose high performance computing, more specifically to dense matrix computations, without major changes in existing codes and methodologies and with excellent performance and power consumption numbers.

A Scalable, Numerically Stable, High-Performance Tridiagonal Solver Using GPUs

Authors: Li-Wen Chang, John A. Stratton, Hee-Seok Kim, Wen-Mei W. Hwu (University of Illinois at Urbana-Champaign)

In this paper, we present a scalable, numerically stable, high-performance tridiagonal solver. The solver is based on the SPIKE algorithm, a method for partitioning a large matrix into small independent matrices, which can be solved in parallel. For each small matrix, our solver applies a general 1-by-1 or 2-by-2 diagonal pivoting algorithm, which is known to be numerically stable. Our paper makes two major contributions. First, our solver is the first numerically stable tridiagonal solver for GPUs. Our solver provides comparable quality of stable solutions to Intel MKL and Matlab, at a speed comparable to the GPU tridiagonal solvers in existing packages like NVIDIA CUSPARSE. It is also scalable to multiple GPUs and CPUs. Second, we present and analyze two key optimization strategies for our solver: a high-throughput data layout transformation for memory efficiency, and a dynamic tiling approach for reducing the memory access footprint caused by branch divergence.

Efficient Backprojection-Based Synthetic Aperture Radar Computation with Many-Core Processors

Authors: Jongsoo Park, Ping Tak Peter Tang, Mikhail Smelyanskiy, Daehyun Kim (Intel Corporation), Thomas Benson (Georgia Institute of Technology)

Tackling computationally challenging problems with high efficiency often requires the combination of algorithmic innovation, advanced architecture and thorough exploitation of parallelism. We demonstrate this synergy through synthetic aperture radar (SAR) via backprojection, an image reconstruction method that can require hundreds of TFLOPS. Computation cost is significantly reduced by our new algorithm of approximate strength reduction; data movement cost is economized by software locality optimizations facilitated by advanced architecture supports; parallelism is fully harnessed in various patterns and granularities. We deliver over 35 billion backprojections per second throughput per compute node on a Sandy Bridge-based cluster, equipped with Intel Knights Corner coprocessors. This corresponds to processing a 3K×3K image within a second using a single node. Our study can be extended to other settings: backprojection is applicable elsewhere including medical imaging, approximate strength reduction is a general code transformation technique, and many-core processors are emerging as a solution to energy-efficient computing.

Finalist: Best Paper Award

Auto-Diagnosis of Correctness and Performance Issues

Chair: Kenjiro Taura (University of Tokyo)

3:30pm-5pm

Room: 255-BC

Parametric Flows - Automated Behavior Equivalencing for Symbolic Analysis of Races in CUDA Programs

Authors: Peng Li (University of Utah), Guodong Li (Fujitsu Laboratories of America), Ganesh Gopalakrishnan (University of Utah)

The growing scale of concurrency requires automated abstraction techniques to cut down the effort in concurrent system analysis. In this paper, we show that the high degree of behavioral symmetry present in GPU programs allows CUDA race detection to be dramatically simplified through abstraction. Our abstraction techniques is one of automatically creating parametric flows—control-flow equivalence classes of threads that diverge in the same manner—and checking for data races only across a pair of threads per parametric flow. We have implemented this approach as an extension of our recently proposed GKLEE symbolic analysis framework and show that all our previous results are dramatically improved in that (1) the parametric flow-based analysis takes far less time, and (2)

because of the much higher scalability of the analysis, we can detect even more data race situations that were previously missed by GKLEE because it was forced to downscale examples to limit analysis complexity.

MPI Runtime Error Detection with MUST - Advances in Deadlock Detection

Authors: Tobias Hilbrich, Joachim Protze (Technical University Dresden), Martin Schulz, Bronis R. de Supinski (Lawrence Livermore National Laboratory), Matthias S. Mueller (Technical University Dresden)

The widely used Message Passing Interface (MPI) is complex and rich. As a result, application developers require automated tools to avoid and to detect MPI programming errors. We present the Marmot Umpire Scalable Tool (MUST) that detects such errors with a significantly increased scalability. We present improvements to our graph-based deadlock detection approach for MPI, which cover complex MPI constructs, as well as future MPI extensions. Further, our enhancements check complex MPI constructs that no previous graph-based detection approach handled correctly. Finally, we present optimizations for the processing of MPI operations that reduce runtime deadlock detection overheads. Existing approaches often require $O(p)$ analysis time per MPI operation, for p processes. We empirically observe that our improvements lead to sub-linear or better analysis time per operation for a wide range of real world applications. We use two major benchmark suites with up to 1024 cores for this evaluation.

Finalist: Best Paper Award

Novel Views of Performance Data to Analyze Large-Scale Adaptive Applications

Authors: Abhinav Bhatele, Todd Gamblin (Lawrence Livermore National Laboratory), Katherine E. Isaacs (University of California, Davis), Brian T. N. Gunney, Martin Schulz, Peer-Timo Bremer (Lawrence Livermore National Laboratory), Bernd Hamann (University of California, Davis)

Performance analysis of parallel scientific codes is becoming increasingly difficult due to the rapidly growing complexity of applications and architectures. Existing tools fall short in providing intuitive views that facilitate the process of performance debugging and tuning. In this paper, we extend recent ideas of projecting and visualizing performance data for faster, more intuitive analysis of applications. We collect detailed per-level and per-phase measurements in a dynamically load-balanced, structured AMR library and relate the information back to the application's communication structure. We show how our projections and visualizations lead to a simple diagnosis of and mitigation strategy for a previously elusive scaling bottleneck in the library that is hard to detect using conventional tools. Our new insights have resulted in a 22% performance improvement for a 65,536-core run on an IBM Blue Gene/P system.

DRAM Power and Resiliency Management

Chair: Nathan Debardeleben (Los Alamos National Laboratory)

3:30pm-5pm

Room: 355-D

RAMZzz: Rank-Aware DRAM Power Management with Dynamic Migrations and Demotions

Authors: Donghong Wu, Bingsheng He, Xueyan Tang (Nanyang Technological University), Jianliang Xu (Hong Kong Baptist University), Minyi Guo (Shanghai Jiao Tong University)

Main memory is a significant energy consumer which may contribute to over 40% of the total system power, and will become more significant for server machines with more main memory. In this paper, we propose a novel memory system design named RAMZzz with rank-aware energy-saving optimizations. Specifically, we rely on a memory controller to monitor the memory access locality, and group the pages with similar access locality into the same rank. We further develop dynamic page migrations to adapt to data access patterns and a prediction model to estimate the demotion time for accurate control on power state transitions. We experimentally compare our algorithm with other energy-saving policies with cycle-accurate simulation. Experiments with benchmark workloads show that RAMZzz achieves significant improvement on energy-delay and energy consumption over other power-saving techniques.

MAGE - Adaptive Granularity and ECC for Resilient and Power Efficient Memory Systems

Authors: Sheng Li, Doe Hyun Yoon (Hewlett-Packard), Ke Chen (University of Notre Dame), Jishen Zhao (Pennsylvania State University), Jung Ho Ahn (Seoul National University), Jay Brockman (University of Notre Dame), Yuan Xie (Pennsylvania State University), Norman Jouppi (Hewlett-Packard)

Resiliency is one of the toughest challenges in high performance computing, and memory accounts for a significant fraction of errors. Providing strong error tolerance in memory usually requires a wide memory channel that incurs a large access granularity (hence, a large cache line). Unfortunately, applications with limited spatial locality waste memory power and bandwidth on systems with a large access granularity. Thus, careful design considerations must be made to balance memory system performance, power efficiency, and resiliency. In this paper, we propose MAGE, a Memory system with Adaptive Granularity and ECC, to achieve high performance, power efficiency, and resiliency. MAGE enables adaptive selection of appropriate granularities and ECC schemes for applications with different memory behaviors. Our experiments show that MAGE achieves more than a 28% energy-delay product improvement compared to the best existing systems with static granularity and ECC.

Grids/Clouds Networking

Chair: Michael Lang (Los Alamos National Laboratory)

3:30pm-5pm

Room: 255-EF

Protocols for Wide-Area Data-Intensive Applications - Design and Performance Issues

Authors: Yufei Ren, Tan Li (Stony Brook University), Dantong Yu (Brookhaven National Laboratory), Shudong Jin, Thomas Robertazzi (Stony Brook University), Brian Tierney, Eric Pouyoul (Lawrence Berkeley National Laboratory)

Providing high-speed data transfer is vital to various data-intensive applications. While there have been remarkable technology advances to provide ultra-high-speed network bandwidth, existing protocols and applications may not be able to fully utilize the bare-metal bandwidth due to their inefficient design. We identify the same problem remains in the field of Remote Direct Memory Access (RDMA) networks. RDMA offloads TCP/IP protocols to hardware devices. However, its benefits have not been fully exploited due to the lack of efficient software and application protocols, in particular in wide-area networks. In this paper, we address the design choices to develop such protocols. We describe a protocol implemented as part of a communication middleware. The protocol has its flow control, connection management, and task synchronization. It maximizes the parallelism of RDMA operations. We demonstrate its performance benefit on various local and wide-area testbeds, including the DOE ANI testbed with RoCE links and InfiniBand links.

High Performance RDMA-Based Design of HDFS over InfiniBand

Authors: Nusrat S. Islam, Md W. Rahman, Jithin Jose, Raghu-nath Rajachandrasekar, Hao Wang, Hari Subramoni (Ohio State University), Chet Murthy (IBM T.J. Watson Research Center), Dhableswar K. Panda (Ohio State University)

Hadoop Distributed File System (HDFS) acts as primary storage of Hadoop and has been adopted by reputed organizations (Facebook, Yahoo! etc.) due to its portability and fault tolerance. The existing implementation of HDFS uses Java-socket interface for communication which delivers suboptimal performance in terms of latency and throughput. For data-intensive applications, network performance becomes a key component as the amount of data being stored and replicated to HDFS increases. In this paper, we present a novel design of HDFS using Remote Direct Memory Access (RDMA) over InfiniBand via JNI interfaces. Experimental results show that, for 5GB HDFS file writes, the new design reduces the communication time by 87% and 30% over 1Gigabit Ethernet (1GigE) and IP-over-InfiniBand (IPoIB), respectively, on QDR platform (32Gbps). For HBase, the Put operation performance is improved by 20% with our new design. To the best of our knowledge, this is the first design of HDFS over InfiniBand networks.

Efficient and Reliable Network Tomography in Heterogeneous Networks Using BitTorrent Broadcasts and Clustering Algorithms

Authors: Kiril Dichev, Fergal Reid, Alexey Lastovetsky (University College Dublin)

In the area of network performance and discovery, network tomography focuses on reconstructing network properties using only end-to-end measurements at the application layer. One challenging problem in network tomography is reconstructing available bandwidth along all links during multiple source/multiple destination transmissions. The traditional measurement procedures used for bandwidth tomography are extremely time consuming. We propose a novel solution to this problem. Our method counts the fragments exchanged during a BitTorrent broadcast. While this measurement has a high level of randomness, it can be obtained very efficiently, and aggregated into a reliable metric. This data is then analyzed with state-of-the-art algorithms, which reliably reconstruct logical clusters of nodes inter-connected by high bandwidth, as well as bottlenecks between these logical clusters. Our experiments demonstrate that the proposed two-phase approach efficiently solves the presented problem for a number of settings on a complex grid infrastructure.

Finalist: Best Student Paper Award

Weather and Seismic Simulations

Chair: Amik St-Cyr (Royal Dutch Shell)

3:30pm-5pm

Room: 355-EF

A Divide and Conquer Strategy for Scaling Weather Simulations with Multiple Regions of Interest

Authors: Preeti Malakar (Indian Institute of Science), Thomas George (IBM India Research Lab), Sameer Kumar (IBM T.J. Watson Research Center), Rashmi Mittal (IBM India Research Lab), Vijay Natarajan (Indian Institute of Science), Yogish Sabharwal (IBM India Research Lab), Vaibhav Saxena (IBM India Research Lab), Sathish S. Vadhiyar (Indian Institute of Science)

Accurate and timely prediction of weather phenomena, such as hurricanes and flash floods, require high-fidelity compute-intensive simulations of multiple finer regions of interest within a coarse simulation domain. Current weather applications execute these nested simulations sequentially using all the available processors, which is sub-optimal due to their sublinear scalability. In this work, we present a strategy for parallel execution of multiple nested domain simulations based on partitioning the 2-D processor grid into disjoint rectangular regions associated with each domain. We propose a novel combination of performance prediction, processor allocation methods and topology-aware mapping of the regions on torus interconnects. Experiments on IBM Blue Gene systems using WRF show that the proposed strategies result in performance improvement of up to 33% with topology-oblivious mapping and up to additional 7% with topology-aware mapping over the default sequential strategy.

Finalist: Best Student Paper Award

Forward and Adjoint Simulations of Seismic Wave Propagation on Emerging Large-Scale GPU Architectures

Authors: Max Rietmann (USI Lugano), Peter Messmer (NVIDIA), Tarje Nissen-Meyer (ETH Zurich), Daniel Peter (Princeton University), Piero Basini (ETH Zurich), Dimitri Komatitsch (CNRS), Olaf Schenk (USI Lugano), Jeroen Tromp (Princeton University), Lapo Boschi (ETH Zurich), Domenico Giardini (ETH Zurich)

SPECFEM3D is a widely used community code which simulates seismic wave propagation in earth-science applications. It can be run either on multi-core CPUs only or together with many-core GPU devices on large GPU clusters. The new implementation is optimally fine tuned and achieves excellent performance results. Mesh coloring enables an efficient accumulation of border nodes in the assembly process over an unstructured mesh on the GPU and asynchronous GPU-CPU memory transfers and non-blocking MPI are used to overlap communication and computation, effectively hiding synchronizations. To demonstrate the performance of the inversion, we present two case studies run on the Cray XE6 and XK6 architectures up to 896 nodes: (1) focusing on most commonly used forward simulations, we simulate wave propagation generated by earthquakes in Turkey, and (2) testing the most complex simulation type of the package, we use ambient seismic noise to image 3D crust and mantle structure beneath western Europe.

Wednesday, November 14

Compiler-Based Analysis and Optimization

Chair: Xipeng Shen (College of William & Mary)

10:30am-12pm

Room: 255-EF

Bamboo - Translating MPI Applications to a Latency-Tolerant, Data-Driven Form

Authors: Tan Nguyen, Pietro Cicotti (University of California, San Diego), Eric Bylaska (Pacific Northwest National Laboratory), Dan Quinlan (Lawrence Livermore National Laboratory), Scott Baden (University of California, San Diego)

We present Bamboo, a custom source-to-source translator that transforms MPI C source into a data-driven form that automatically overlaps communication with available computation. Running on up to 98304 processors of NERSC's Hopper system, we observe that Bamboo's overlap capability speeds up MPI implementations of a 3D Jacobi iterative solver and Cannon's matrix multiplication. Bamboo's generated code meets or exceeds the performance of hand optimized MPI, which includes split-phase coding, the method classically employed to hide communication. We achieved our results

with only modest amounts of programmer annotation and no intrusive reprogramming of the original application source.

Tiling Stencil Computations to Maximize Parallelism

Authors: Vinayaka Bandishti (Indian Institute of Science), Irshad Pananilath (Indian Institute of Science), Uday Bondhugula (Indian Institute of Science)

Most stencil computations allow tile-wise concurrent start, i.e., there always exists a face of the iteration space and a set of tiling hyperplanes such that all tiles along that face can be started concurrently. This provides load balance and maximizes parallelism. However, existing automatic tiling frameworks often choose hyperplanes that lead to pipelined start-up and load imbalance. We address this issue with a new tiling technique that ensures concurrent start-up as well as perfect load-balance whenever possible. We first provide necessary and sufficient conditions on tiling hyperplanes to enable concurrent start for programs with affine data accesses. We then provide an approach to find such hyperplanes. Experimental evaluation on a 12-core Intel Westmere shows that our code is able to outperform a tuned domain-specific stencil code generator by 4 to 20 percent, and previous compiler techniques by a factor of 2x to 10.14x.

Compiler-Directed File Layout Optimization for Hierarchical Storage Systems

Authors: Wei Ding (Pennsylvania State University), Yuanrui Zhang (Intel Corporation), Mahmut Kandemir (Pennsylvania State University), Seung Woo Son (Northwestern University)

File layout of array data is a critical factor that effects the behavior of storage caches, and has so far taken not much attention in the context of hierarchical storage systems. The main contribution of this paper is a compiler-driven file layout optimization scheme for hierarchical storage caches. This approach, fully automated within an optimizing compiler, analyzes a multi-threaded application code and determines a file layout for each disk-resident array referenced by the code, such that the performance of the target storage cache hierarchy is maximized. We tested our approach using 16 I/O intensive application programs and compared its performance against two previously proposed approaches under different cache space management schemes. Our experimental results show that the proposed approach improves the execution time of these parallel applications by 23.7% on average.

Finalist: Best Student Paper Award

Fast Algorithms

Chair: *Torsten Hoefer (ETH Zurich)*

10:30am-12pm

Room: 255-BC

A Framework for Low-Communication 1-D FFT

Authors: Ping Tak, Peter Tang, Jongsoo Park, Daehyun Kim, Vladimir Petrov (Intel Corporation)

In high performance computing on distributed-memory systems, communication often represents a significant part of the overall execution time. The relative cost of communication will certainly continue to rise as compute density growth follows the current technology and industry trends. Design of lower-communication alternatives to fundamental computational algorithms has become an important field of research. For distributed 1-D FFT, communication cost has hitherto remained high as all industry-standard implementations perform three all-to-all internode data exchanges (also called global transpose). These communication steps indeed dominate execution time. In this paper, we present a mathematical framework from which many single-all-to-all and easy-to-implement 1-D FFT algorithms can be derived. For large-scale problems, our implementation can be twice as fast as leading FFT libraries on state-of-the-art computer clusters. Moreover, our framework allows tradeoff between accuracy and performance, further boosting performance if reduced accuracy is acceptable.

Finalist: Best Paper Award

Parallel Geometric-Algebraic Multigrid on Unstructured Forests of Octrees

Authors: Hari Sundar, George Biros, Carsten Burstedde, Johann Rudi, Omar Ghattas, Georg Stadler (University of Texas at Austin)

We present a parallel multigrid method for solving variable-coefficient elliptic partial differential equations on arbitrary geometries using highly-adapted meshes. Our method is designed for meshes that are built from an unstructured hexahedral macro mesh, in which each macro element is adaptively refined as an octree. This forest-of-octrees approach enables us to generate meshes for complex geometries with arbitrary levels of local refinement. We use geometric multigrid (GMG) for each of the octrees and algebraic multigrid (AMG) as the coarse grid solver. We designed our GMG sweeps to entirely avoid collectives, thus minimizing communication cost. We present weak and strong scaling results for the 3D variable-coefficient Poisson problem that demonstrates high parallel scalability. As a highlight, the largest problem we solve is on a non-uniform mesh with 100 billion unknowns on 262,144 cores of NCCS's Cray XK6 "Jaguar"; in this solve we sustain 272 TFlops/s.

Scalable Multi-GPU 3-D FFT for TSUBAME 2.0 Supercomputer

Authors: Akira Nukada, Kento Sato, Satoshi Matsuoka (Tokyo Institute of Technology)

For scalable 3-D FFT computation using multiple GPUs, efficient all-to-all communication between GPUs is the most important factor in good performance. Implementations with point-to-point MPI library functions and CUDA memory copy APIs typically exhibit very large overheads especially for small message sizes in all-to-all communications between many nodes. We propose several schemes to minimize the overheads, including employment of lower-level API of InfiniBand to effectively overlap intra- and inter-node communication, as well as auto-tuning strategies to control scheduling and determine rail assignments. As a result we achieve very good strong scalability as well as good performance, up to 4.8 TFLOPS using 256 nodes of TSUBAME 2.0 Supercomputer (768 GPUs) in double precision.

Massively Parallel Simulations

10:30am-12pm

Chair: *Kamesh Madduri (Pennsylvania State University)*

Room: 355-EF

Petascale Lattice Quantum Chromodynamics on a Blue Gene/Q Supercomputer

Authors: Jun Doi (IBM Research, Tokyo)

Lattice Quantum Chromodynamics (QCD) is one of the most challenging applications running on massively parallel supercomputers. To reproduce these physical phenomena on a supercomputer, a precise simulation is demanded requiring well optimized and scalable code. We have optimized lattice QCD programs on Blue Gene family supercomputers and show the strength in lattice QCD simulation. Here we optimized on the third generation Blue Gene/Q supercomputer (1) by changing the data layout, (2) by exploiting new SIMD instruction sets, and (3) by pipelining boundary data exchange to overlap communication and calculation. The optimized lattice QCD program shows excellent weak scalability on the large scale Blue Gene/Q system, and with 16 racks we sustained 1.08 Pflops, 32.1% of the theoretical peak performance, including the conjugate gradient solver routines.

Massively Parallel X-Ray Scattering Simulations

Authors: Abhinav Sarje, Xiaoye S. Li, Slim Chourou, Elaine R. Chan, Alexander Hexemer (Lawrence Berkeley National Laboratory)

Although present X-ray scattering techniques can provide tremendous information on the nano-structural properties of materials that are valuable in the design and fabrication of energy-relevant nano-devices, a primary challenge remains in the analyses of such data. In this paper we describe a high-performance, flexible, and scalable Grazing Incidence Small Angle X-ray Scattering simulation algorithm and codes that

we have developed on multi-core/CPU and many-core/GPU clusters. We discuss in detail our implementation, optimization and performance on these platforms. Our results show speedups of ~125x on a Fermi-GPU and ~20x on a Cray-XE6 24-core node, compared to a sequential CPU code, with near linear scaling on multi-node clusters. To our knowledge, this is the first GISAXS simulation code that is flexible to compute scattered light intensities in all spatial directions allowing full reconstruction of GISAXS patterns for any complex structures and with high-resolutions while reducing simulation times from months to minutes.

High Performance Radiation Transport Simulations - Preparing for TITAN

Authors: Christopher Baker, Gregory Davidson, Thomas Evans, Steven Hamilton, Joshua Jarrell, Wayne Joubert (Oak Ridge National Laboratory)

In this paper, we describe the Denovo code system. Denovo solves the seven-dimensional linear Boltzmann transport equation, of central importance to nuclear technology applications such as reactor core analysis (neutronics), radiation shielding, nuclear forensics and radiation detection. The code features multiple spatial differencing schemes, state-of-the-art linear solvers, the Koch-Baker-Alcouffe (KBA) parallel wavefront sweep algorithm for modeling radiation flux, a new multilevel energy decomposition method scaling to hundreds of thousands of processing cores, and a modern, novel code architecture that supports straightforward integration of new features. In this paper we discuss the port of Denovo to the 20+ petaflop ORNL GPU-based system, Titan. We describe algorithms and techniques used to exploit the capabilities of Titan's heterogeneous compute node architecture and the challenges of obtaining good parallel performance for this sparse hyperbolic PDE solver containing inherently recursive computations. Numerical results demonstrating Denovo performance on representative early hardware are presented.

Optimizing I/O For Analytics

Chair: Dean Hildebrand (IBM Almaden Research Center)

10:30am-12pm

Room: 355-D

Byte-Precision Level of Detail Processing for Variable Precision Analytics

Authors: John Jenkins, Eric Schendel, Sriram Lakshminarasimhan, Terry Rogers, David A. Boyuka (North Carolina State University), Stephane Ethier (Princeton Plasma Physics Laboratory), Robert Ross (Argonne National Laboratory), Scott Klasky (Oak Ridge National Laboratory), Nagiza F. Samatova (North Carolina State University)

I/O bottlenecks in HPC applications are becoming a more pressing problem as compute capabilities continue to outpace

I/O capabilities. While double-precision simulation data often must be stored losslessly, the loss of some of the fractional component may introduce acceptably small errors to many types of scientific analyses. Given this observation, we develop a precision level of detail (APLOD) library, which partitions double-precision datasets along user-defined byte boundaries. APLOD parameterizes the analysis accuracy-I/O performance tradeoff, bounds maximum relative error, maintains I/O access patterns compared to full precision, and operates with low overhead. Using ADIOS as an I/O use-case, we show proportional reduction in disk access time to the degree of precision. Finally, we show the effects of partial precision analysis on accuracy for operations such as k-means and Fourier analysis, finding a strong applicability for the use of varying degrees of precision to reduce the cost of analyzing extreme-scale data.

Combining In-Situ and In-Transit Processing to Enable Extreme-Scale Scientific Analysis

Authors: Janine C. Bennett (Sandia National Laboratories), Hasan Abbasi (Oak Ridge National Laboratory), Peer-Timo Bremer (Lawrence Livermore National Laboratory), Ray W. Grout (National Renewable Energy Laboratory), Attila Gyulassy (University of Utah), Tong Jin (Rutgers University), Scott Klasky (Oak Ridge National Laboratory), Hemanth Kolla (Sandia National Laboratories), Manish Parashar (Rutgers University), Valerio Pascucci (University of Utah), Philippe Pébay (Kitware, Inc.), David Thompson (Sandia National Laboratories), Hongfeng Yu (Sandia National Laboratories), Fan Zhang (Rutgers University), Jacqueline Chen (Sandia National Laboratories)

With the onset of extreme-scale computing, scientists are increasingly unable to save sufficient raw simulation data to persistent storage. Consequently, the community is shifting away from a post-process centric data analysis pipeline to a combination of analysis performed in-situ (on primary compute resources) and in-transit (on secondary resources using asynchronous data transfers). In this paper we summarize algorithmic developments for three common analysis techniques: topological analysis, descriptive statistics, and visualization. We describe a resource scheduling system that supports various analysis workflows, and discuss our use of the DataSpaces and ADIOS frameworks to transfer data between in-situ and in-transit computations. We demonstrate the efficiency of our lightweight, flexible framework on the Jaguar XK6, analyzing data generated by S3D, a massively parallel turbulent combustion code. Our framework allows scientists dealing with the data deluge at extreme-scale to perform analyses at increased temporal resolutions, mitigate I/O costs, and significantly improve time to insight.

Efficient Data Restructuring and Aggregation for IO Acceleration in PIDX

Authors: Sidharth Kumar (University of Utah), Venkatram Vishwanath, Philip Carns (Argonne National Laboratory), Joshua A. Levine (University of Utah), Robert Latham (Argonne National Laboratory), Giorgio Scorzelli (University of Utah), Hemanth Kolla (Sandia National Laboratories), Ray Grout (National Renewable Energy Laboratory), Jacqueline Chen (Sandia National Laboratories), Robert Ross, Michael E. Papka (Argonne National Laboratory), Valerio Pascucci (University of Utah)

Hierarchical, multi-resolution data representations enable interactive analysis and visualization of large-scale simulations. One promising application of these techniques is to store HPC simulation output in a hierarchical Z (HZ) ordering that translates data from a Cartesian coordinate scheme to a one dimensional array ordered by locality at different resolution levels. When the dimensions of the simulation data are not an even power of two, however, parallel HZ-ordering produces sparse memory and network access patterns that inhibit I/O performance. This work presents a new technique for parallel HZ-ordering of simulation datasets that restructures simulation data into large power of two blocks to facilitate efficient I/O aggregation. We perform both weak and strong scaling experiments using the S3D combustion application on both Cray-XE6 (65536 cores) and IBM BlueGene/P (131072 cores) platforms. We demonstrate that data can be written in hierarchical, multiresolution format with performance competitive to that of native data ordering methods.

Datacenter Technologies

Chair: Rolf Riesen (IBM)

1:30pm-3pm

Room: 255-BC

Measuring Interference Between Live Datacenter Applications

Authors: Melanie Kambadur (Columbia University), Tipp Moseley, Rick Hank (Google), Martha A. Kim (Columbia University)

Application interference is prevalent in datacenters due to contention over shared hardware resources. Unfortunately, understanding interference in live datacenters is more difficult than in controlled environments or on simpler architectures. Most approaches to mitigating interference rely on data that cannot be collected efficiently in a production environment. This work exposes eight specific complexities of live datacenters that constrain measurement of interference. It then introduces new, generic measurement techniques for analyzing interference in the face of these challenges and restrictions. We use the measurement techniques to conduct the first large-scale study of application interference in live production datacenter workloads. Data is measured across 1000 12-core

Google servers observed to be running 1102 unique applications. Finally, our work identifies several opportunities to improve performance that use only the available data; these opportunities are applicable to any datacenter.

T* - A Data-Centric Cooling Energy Costs Reduction Approach for Big Data Analytics Cloud

Authors: Rini Kaushik, Klara Nahrstedt (University of Illinois at Urbana-Champaign)

Explosion in Big Data has led to a surge in extremely large-scale Big Data analytics platforms, resulting in burgeoning energy costs. T* takes a novel, data-centric approach to reduce cooling energy costs and to ensure thermal-reliability of the servers. T* is cognizant of the difference in thermal-profile and thermal-reliability-driven load threshold of the servers, and the difference in the computational jobs arrival rate, size, and evolution life spans of the Big Data placed in the cluster. Based on this knowledge, and coupled with its predictive file models and insights, T* does proactive, thermal-aware file placement, which implicitly results in thermal-aware job placement in the Big Data analytics compute model. T* evaluation results with one-month long real-world Big Data analytics production traces from Yahoo! show up to 42% reduction in the cooling energy costs, lower and more uniform thermal-profile, and 9x better performance than the state-of-the-art data-agnostic, job-placement-centric cooling techniques.

ValuePack - Value-Based Scheduling Framework for CPU-GPU Clusters

Authors: Vignesh T. Ravi (Ohio State University), Michela Becchi (University of Missouri), Gagan Agrawal (Ohio State University), Srimat Chakradhar (NEC Laboratories America)

Heterogeneous computing nodes are becoming commonplace today, and recent trends strongly indicate that clusters, supercomputers, and cloud environments will increasingly host more heterogeneous resources, with some being massively parallel (e.g., GPU, MIC). With such heterogeneous environments becoming common, it is important to revisit scheduling problems for clusters and cloud environments. In this paper, we formulate and address the problem of value-driven scheduling of independent jobs on heterogeneous clusters, which captures both the urgency and relative priority of jobs. Our overall scheduling goal is to maximize the aggregate value or yield of all jobs. Exploiting the portability available from the underlying programming model, we propose four novel scheduling schemes that can automatically and dynamically map jobs onto heterogeneous resources. Additionally, to improve the utilization of massively parallel resources, we also propose heuristics to automatically decide when and which jobs can share a single resource.

Optimizing Application Performance

Chair: Martin Schulz (Lawrence Livermore National Laboratory)

1:30pm-3pm

Room: 355-EF

Compass - A Scalable Simulator for an Architecture for Cognitive Computing

Authors: Robert Preissl, Theodore M. Wong, Pallab Datta, Raghav Singh, Steven Esser, William Risk (IBM Research), Horst Simon (Lawrence Berkeley National Laboratory), Myron Flickner, Dharmendra Modha (IBM Research)

Inspired by the function, power, and volume of the organic brain, we are developing TrueNorth, a novel modular, non-von Neumann, ultra-low power, compact architecture. TrueNorth consists of a scalable network of neurosynaptic cores, with each core containing neurons, dendrites, synapses, and axons. To set sail for TrueNorth, we have developed Compass, a multi-threaded, massively parallel functional simulator and a parallel compiler that maps a network of long-distance pathways in the Macaque monkey brain to TrueNorth. We demonstrate near-perfect weak scaling on a 16 rack IBM Blue Gene/Q (262144 CPUs, 256 TB memory), achieving an unprecedented scale of 256 million neurosynaptic cores containing 65 billion neurons and 16 trillion synapses running only 388× slower than real-time with an average spiking rate of 8.1 Hz. By using emerging PGAS communication primitives, we also demonstrate 2× better real-time performance over MPI primitives on a 4 rack Blue Gene/P (16384 CPUs, 16 TB memory).

Finalist: Best Paper Award

Optimizing Fine-Grained Communication in a Biomolecular Simulation Application on Cray XK6

Authors: Yanhua Sun, Gengbin Zheng, Chao Mei, Eric J. Bohm, Laxmikant V. Kale, James C. Phillips (University of Illinois at Urbana-Champaign), Terry R. Jones (Oak Ridge National Laboratory)

Achieving good scaling for fine-grained communication intensive applications on modern supercomputers remains challenging. In our previous work, we have shown that such an application --- NAMD --- scales well on the full Jaguar XT5 without long-range interactions; yet, with them, the speedup falters beyond 64K cores. Although the new Gemini interconnect on Cray XK6 has improved network performance, the challenges remain, and are likely to remain for other such networks as well. We analyze communication bottlenecks in NAMD and its CHARM++ runtime, using the Projections performance analysis tool. Based on the analysis, we optimize the runtime, built on the uGNI library for Gemini. We present several techniques to improve the fine-grained communication. Consequently, the performance of running 92224-atom ApoA1 on GPUs is improved by 36%. For 100-million-atom STMV, we improve upon the prior Jaguar XT5 result of 26 ms/step to 13 ms/step using 298,992 cores on Titan XK6.

Heuristic Static Load-Balancing Algorithm Applied to the Fragment Molecular Orbital Method

Authors: Yuri Alexeev, Ashutosh Mahajan, Sven Leyffer, Graham Fletcher (Argonne National Laboratory), Dmitri Fedorov (National Institute of Advanced Industrial Science and Technology)

In the era of petascale supercomputing, the importance of load balancing is crucial. Although dynamic load balancing is widespread, it is increasingly difficult to implement effectively with thousands of processors or more, prompting a second look at static load-balancing techniques even though the optimal allocation of tasks to processors is an NP-hard problem. We propose a heuristic static load-balancing algorithm, employing fitted benchmarking data, as an alternative to dynamic load-balancing. The problem of allocating CPU cores to tasks is formulated as a mixed-integer nonlinear optimization problem, which is solved by using an optimization solver. On 163,840 cores of Blue Gene/P, we achieved a parallel efficiency of 80% for an execution of the fragment molecular orbital method applied to model protein-ligand complexes quantum-mechanically. The obtained allocation is shown to outperform dynamic load balancing by at least a factor of 2, thus motivating the use of this approach on other coarse-grained applications.

Resilience

Chair: Bronis R. de Supinski (Lawrence Livermore National Laboratory)

1:30pm-3pm

Room: 255-EF

Classifying Soft Error Vulnerabilities in Extreme-Scale Scientific Applications Using a Binary Instrumentation Tool

Authors: Dong Li, Jeffrey Vetter (Oak Ridge National Laboratory), Weikuan Yu (Auburn University)

Extreme-scale scientific applications are at a significant risk of being hit by soft errors on future supercomputers. To better understand soft error vulnerabilities in scientific applications, we have built an empirical fault injection and consequence analysis tool - BIFIT - to evaluate how soft errors impact applications. BIFIT is designed with capability to inject faults at specific targets: execution point and data structure. We apply BIFIT to three scientific applications and investigate their vulnerability to soft errors. We classify each application's individual data structures in terms of their vulnerabilities, and generalize these classifications. Our study reveals that these scientific applications have a wide range of sensitivities to both the time and the location of a soft error. Yet, we are able to identify relationships between vulnerabilities and classes of data structures. These classifications can be used to apply appropriate resiliency solutions to each data structure within an application.

Containment Domains - A Scalable, Efficient, and Flexible Resiliency Scheme for Exascale Systems

Authors: Jinsuk Chung, Ikhwan Lee, Michael Sullivan, Jee Ho Ryoo, Dongwan Kim (University of Texas at Austin), Doe Hyun Yoon (Hewlett-Packard), Larry Kaplan (Cray Inc.), Mattan Erez (University of Texas at Austin)

This paper describes and evaluates a scalable and efficient resiliency scheme based on the concept of containment domains. Containment domains are a programming construct that enables applications to express resiliency needs and interact with the system to tune and specialize error detection, state preservation and restoration, and recovery schemes. Containment domains have weak transactional semantics and are nested to take advantage of the machine hierarchy and to enable distributed and hierarchical state preservation, restoration, and recovery as an alternative to non-scalable and inefficient checkpoint-restart. We evaluate the scalability and efficiency of containment domains using generalized trace-driven simulation and analytical analysis and show that containment domains are superior to both checkpoint restart and redundant execution approaches.

Finalist: Best Paper Award

Visualization and Analysis of Massive Data Sets

Chair: Hank Childs (Lawrence Berkeley National Laboratory)

1:30pm-3pm

Room: 355-D

Parallel I/O, Analysis, and Visualization of a Trillion Particle Simulation

Authors: Surendra Byna (Lawrence Berkeley National Laboratory), Jerry Chou (Tsinghua University), Oliver Ruebel, Mr Prabhat (Lawrence Berkeley National Laboratory), Homa Karimabadi (University of California, San Diego), William Daughton (Los Alamos National Laboratory), Vadim Roytershteyn (University of California, San Diego), Wes Bethel (Lawrence Berkeley National Laboratory), Mark Howison (Brown University), Ke-Jou Hsu, Kuan-Wu Lin (Tsinghua University), Arie Shoshani, Andrew Uzelton, Kesheng Wu (Lawrence Berkeley National Laboratory)

Petascale plasma physics simulations have recently entered the regime of simulating trillions of particles. These unprecedented simulations generate massive amounts of data, posing significant challenges in storage, analysis, and visualization. In this paper, we present parallel I/O, analysis, and visualization results from a VPIC trillion particle simulation running on 120,000 cores, which produces ~30TB of data for a single timestep. We demonstrate the successful application of H5Part, a particle data extension of parallel HDF5, for writing the dataset at a significant fraction of system peak I/O rates. To enable efficient analysis, we develop hybrid parallel FastQuery to index and query data using multi-core CPUs on distributed memory hardware. We show good scalability results for the

FastQuery implementation using up to 10,000 cores. Finally, we apply this indexing/query-driven approach to facilitate the first-ever analysis and visualization of the trillion-particle dataset.

Data-Intensive Spatial Filtering in Large Numerical Simulation Datasets

Authors: Kalin Kanov, Randal Burns, Greg Eyink, Charles Meneveau, Alexander Szalay (Johns Hopkins University)

We present a query processing framework for the efficient evaluation of spatial filters on large numerical simulation datasets stored in a data-intensive cluster. Previously, filtering of large numerical simulations stored in scientific databases has been impractical owing to the immense data requirements. Rather, filtering is done during simulation or by loading snapshots into the aggregate memory of an HPC cluster. Our system performs filtering within the database and supports large filter widths. We present two complementary methods of execution: I/O streaming computes a batch filter query in a single sequential pass using incremental evaluation of decomposable kernels, summed volumes generates an intermediate data set and evaluates each filtered value by accessing only eight points in this dataset. We dynamically choose between these methods depending upon workload characteristics. The system allows us to perform filters against large data sets with little overhead: query performance scales with the cluster's aggregate I/O throughput.

Parallel Particle Advection and FTLE Computation for Time-Varying Flow Fields

Authors: Boonthanome Nouanesengsy, Teng-Yok Lee, Kewei Lu, Han-Wei Shen (Ohio State University), Tom Peterka (Argonne National Laboratory)

Flow fields are an important product of scientific simulations. One popular flow-visualization technique is particle advection, in which seeds are traced through the flow field. One use of these traces is to compute a powerful analysis tool called the Finite-Time Lyapunov Exponent (FTLE) field, but no existing particle tracing algorithms scale to the particle injection frequency required for high-resolution FTLE analysis. In this paper, a framework to trace the massive number of particles necessary for FTLE computation is presented. A new approach is explored, in which processes are divided into groups, and are responsible for mutually exclusive spans of time. This pipelining over time intervals reduces overall idle time of processes and decreases I/O overhead. Our parallel FTLE framework is capable of advecting hundreds of millions of particles at once, with performance scaling up to tens of thousands of processes.

Graph Algorithms

Chair: Esmond G. Ng (Lawrence Berkeley National Laboratory)
3:30pm-5pm
Room: 355-EF

A New Scalable Parallel DBSCAN Algorithm Using the Disjoint-Set Data Structure

Authors: Md. Mostofa Ali Patwary, Diana Palsetia, Ankit Agrawal, Wei-keng Liao (Northwestern University), Fredrik Manne (University of Bergen), Alok Choudhary (Northwestern University)

DBSCAN is a well-known density-based clustering algorithm capable of discovering arbitrary shaped clusters and eliminating noise data. However, parallelization of DBSCAN is challenging as it exhibits an inherent sequential data access order. Moreover, existing parallel implementations adopt a master-slave strategy which can easily cause an unbalanced workload resulting in low parallel efficiency. We present a new parallel DBSCAN algorithm (PDSDBSCAN) using graph algorithmic concepts. More specifically, we employ the disjoint-set data structure to break the access sequentiality of DBSCAN. In addition, we use a tree-based bottom-up approach to construct the clusters. This yields a better-balanced workload distribution. We implement the algorithm both for shared and for distributed memory. Using data sets containing several hundred million high-dimensional points, we show that PDSDBSCAN significantly outperforms the master-slave approach, achieving speedups up to 30.3 using 40 cores on shared memory architecture, and speedups up to 5,765 using 8,192 cores on distributed memory architecture.

Parallel Bayesian Network Structure Learning with Application to Gene Networks

Authors: Olga Nikolova, Srinivas Aluru (Iowa State University)

Bayesian networks (BN) are probabilistic graphical models which are widely utilized in various research areas, including modeling complex biological interactions in the cell. Learning the structure of a BN is an NP-hard problem and exact solutions are limited to a few tens of variables. In this work, we present a parallel BN structure learning algorithm that combines principles of both heuristic and exact approaches and facilitates learning of larger networks. We demonstrate the applicability of our approach by an implementation on a Cray AMD cluster, and present experimental results for the problem of inferring gene networks. Our approach is work-optimal and achieves nearly perfect scaling.

A Multithreaded Algorithm for Network Alignment via Approximate Matching

Authors: Arif Khan, David Gleich (Purdue University), Mahantesh Halappanavar (Pacific Northwest National Laboratory), Alex Pothen (Purdue University)

Network alignment is an optimization problem to find the best one-to-one map between the vertices of a pair of graphs that overlaps in as many edges as possible. It is a relaxation of the graph isomorphism problem and is closely related to the sub-graph isomorphism problem. The best current approaches are entirely heuristic and iterative in nature. They generate real-valued heuristic weights that must be rounded to find integer solutions. This rounding requires solving a bipartite maximum weight matching problem at each iteration in order to avoid missing high quality solutions. We investigate substituting a parallel, half-approximation for maximum weight matching instead of an exact computation. Our experiments show that the resulting difference in solution quality is negligible. We demonstrate almost a 20-fold speedup using 40 threads on an 8 processor Intel Xeon E7-8870 system and now solve real-world problems in 36 seconds instead of 10 minutes.

Locality in Programming Models and Runtimes

Chair: Milind Kulkarni (Purdue University)

3:30pm-5pm

Room: 255-EF

Characterizing and Mitigating Work Time Inflation in Task Parallel Programs

Authors: Stephen L. Olivier (University of North Carolina at Chapel Hill), Bronis R. de Supinski, Martin Schulz (Lawrence Livermore National Laboratory), Jan F. Prins (University of North Carolina at Chapel Hill)

Task parallelism raises the level of abstraction in shared-memory parallel programming to simplify the development of complex applications. However, task parallel applications can exhibit poor performance due to thread idleness, scheduling overheads, and work time inflation -- additional time spent by threads in a multithreaded computation beyond the time required to perform the same work in a sequential computation. We identify the contributions of each factor to lost efficiency in various task parallel OpenMP applications and diagnose the causes of work time inflation in those applications. A major cause of work time inflation in NUMA systems is increased latency to access data for computations. To mitigate this source of work time inflation in some applications, we propose a locality framework for task parallel OpenMP programs. As implemented in our extensions to the Qthreads library, locality-aware scheduling demonstrates up to 3X improvement compared to the Intel OpenMP task scheduler.

Finalist: Best Student Paper Award

Legion - Expressing Locality and Independence with Logical Regions

Authors: Michael Baue, Sean Treichler, Elliott Slaughter, Alex Aiken (Stanford University)

Modern parallel architectures have both heterogeneous processors and deep, complex memory hierarchies. We present Legion, a programming model and runtime system for achieving high performance on these machines. Legion is organized around logical regions, which express both locality and independence of program data, and tasks that perform computations on regions. We describe a runtime system that dynamically extracts parallelism from Legion programs using a distributed, parallel scheduling algorithm that identifies both independent tasks and nested parallelism. Legion also enables explicit, programmer controlled movement of data through the memory hierarchy and placement of tasks based on locality information via a novel mapping interface. We evaluate our Legion implementation on three applications: fluid-flow on a regular grid, a three-level AMR code solving a heat diffusion equation, and a circuit simulation.

Designing a Unified Programming Model for Heterogeneous Machines

Authors: Michael Garland, Manjunath Kudlur (NVIDIA), Yili Zheng (Lawrence Berkeley National Laboratory)

While high-efficiency machines are increasingly embracing heterogeneous architectures and massive multithreading, contemporary mainstream programming languages reflect a mental model in which processing elements are homogeneous, concurrency is limited, and memory is a flat undifferentiated pool of storage. Moreover, the current state of the art in programming heterogeneous machines tends towards using separate programming models, such as OpenMP and CUDA, for different portions of the machine. Both of these factors make programming emerging heterogeneous machines unnecessarily difficult. We describe the design of the Phalanx programming model, which seeks to provide a unified programming model for heterogeneous machines. It provides constructs for bulk parallelism, synchronization, and data placement which operate across the entire machine. Our prototype implementation is able to launch and coordinate work on both CPU and GPU processors within a single node, and by leveraging the GASNet runtime, is able to run across all the nodes of a distributed-memory machine.

Networks

Chair: Sadaf R. Alam (Swiss National Supercomputing Centre)

3:30pm-5pm

Room: 255-BC

Design and Implementation of an Intelligent End-to-End Network QoS System

Authors: Sushant Sharma, Dimitrios Katramatos, Dantong Yu (Brookhaven National Laboratory), Li Shi (Stony Brook University)

End-to-End guaranteed network QoS is a requirement for predictable data transfers between geographically distant end-hosts. Existing QoS systems, however, do not have the capability/intelligence to decide what resources to reserve and which paths to choose when there are multiple and flexible resource reservation requests. In this paper, we design and implement an intelligent system that can guarantee end-to-end network QoS for multiple flexible reservation requests. At the heart of this system is a polynomial time algorithm called resource reservation and path construction (RRPC). The RRPC algorithm schedules multiple flexible end-to-end data transfer requests by jointly optimizing the path construction and bandwidth reservation along these paths. We show that constructing such schedules is NP-hard. We implement our intelligent QoS system, and present the results of deployment on real world production networks (ESnet and Internet2). Our implementation does not require modifications or new software to be deployed on the routers within network.

Looking Under the Hood of the IBM Blue Gene/Q Network

Authors: Dong Chen, Anamitra Choudhury, Noel Easley, Philip Heidelberger, Sameer Kumar, Amith Mamidala, Jeff Parker, Fabrizio Petrini, Yogish Sabharwal, Robert Senger, Swati Singhal, Steinmacher-Burow Burkhard, Yutaka Sugawara, Robert Walkup (IBM)

This paper explores the performance and optimization of the IBM Blue Gene/Q (BG/Q) five dimensional torus network on up to 16K nodes. The BG/Q hardware supports multiple dynamic routing algorithms and different traffic patterns may require different algorithms to achieve best performance. Between 85% to 95% of peak network performance is achieved for all-to-all traffic, while over 85% of peak is obtained for challenging bisection pairings. A new software-controlled hardware algorithm is developed for bisection traffic that achieves better performance than any individual hardware algorithm. To evaluate memory and network performance, the HPC Random Access benchmark was tuned for BG/Q and achieved 858 Giga Updates per Second (GUPS) on 16K nodes. To further accelerate message processing, the message libraries on BG/Q enable the offloading of messaging overhead onto dedicated communication threads. Several applications, including Algebraic Multigrid (AMG), exhibit from 3 to 20% gain using communication threads.

Design of a Scalable InfiniBand Topology Service to Enable Network-Topology-Aware Placement of Processes

Authors: Hari Subramoni, Sreeram Potluri, Krishna Kandalla (Ohio State University), Bill Barth (University of Texas at Austin), Jerome Vienne (Ohio State University), Jeff Keasler (Lawrence Livermore National Laboratory), Karen Tomko (Ohio Supercomputer Center), Karl Schulz (University of Texas at Austin), Adam Moody (Lawrence Livermore National Laboratory), Dhabaleswar Panda (Ohio State University)

Over the last decade, InfiniBand has become an increasingly popular interconnect for deploying modern supercomputing systems. However, there exists no detection service that can discover the underlying network topology in a scalable manner and expose this information to runtime libraries and users of the high performance computing systems in a convenient way. In this paper, we design a novel and scalable method to detect the InfiniBand network topology by using Neighbor-Joining techniques (NJ). To the best of our knowledge, this is the first instance where the neighbor joining algorithm has been applied to solve the problem of detecting InfiniBand network topology. We also design a network-topology-aware MPI library that takes advantage of the network topology service. The library places processes taking part in the MPI job in a network-topology-aware manner with the dual aim of increasing intra-node communication and reducing the long distance inter-node communication across the InfiniBand fabric.

Finalist: Best Paper Award, Best Student Paper Award

Runtime-Based Analysis and Optimization

Chair: Siegfried Benkner (University of Vienna)

3:30pm-5pm

Room: 355-D

Critical Lock Analysis - Diagnosing Critical Section Bottlenecks in Multithreaded Applications

Authors: Guancheng Chen (IBM Research - China), Per Stenstrom (Chalmers University of Technology)

Critical sections are well known potential performance bottlenecks in multithreaded applications and identifying the ones that inhibit scalability are important for performance optimizations. While previous approaches use idle time as a key measure, we show such a measure is not reliable. The reason is that idleness does not necessarily mean the critical section is on the critical path. We introduce critical lock analysis, a new method for diagnosing critical section bottlenecks in multithreaded applications. Our method firstly identifies the critical sections appearing on the critical path, and then quantifies the impact of such critical sections on the overall performance by using quantitative performance metrics. Case studies show that our method can successfully identify critical sections that are most beneficial for improving overall performance as well as quantify their performance impact on the critical path, which results in a more reliable establishment of the inherent critical section bottlenecks than previous approaches.

Code Generation for Parallel Execution of a Class of Irregular Loops on Distributed Memory Systems

Authors: Mahesh Ravishankar, John Eisenlohr, Louis-Noel Pouchet (Ohio State University), J. Ramanujam (Louisiana State University), Atanas Rountev, P. Sadayappan (Ohio State University)

Parallelization and locality optimization of affine loop nests has been successfully addressed for shared-memory machines. However, many large-scale simulation applications must be executed in a distributed environment, and use irregular/sparse computations where the control-flow and array-access patterns are data-dependent. In this paper, we propose an approach for effective parallel execution of a class of irregular loop computations in a distributed memory environment, using a combination of static and run-time analysis. We discuss algorithms that analyze sequential code to generate an inspector and an executor. The inspector captures the data-dependent behavior of the computation in parallel and without requiring replication of any of the data structures used in the original computation. The executor performs the computation in parallel. The effectiveness of the framework is demonstrated on several benchmarks and a climate modeling application.

Thursday, November 15

Cosmology Applications

Chair: Subhash Saini (NASA Ames Research Center)

10:30am-12pm

Room: 255-EF

First-Ever Full Observable Universe Simulation

Authors: Jean-Michel Alimi, Vincent Bouillot (Paris Observatory), Yann Rasera (Paris Diderot University), Vincent Reverdy, Pier-Stefano Corasaniti, Irène Balmès (Paris Observatory), Stéphane Requena (GENCI), Xavier Delaruelle (CEA), Jean-Noël Richet (CEA)

We performed a massive N-body simulation of the full observable universe. This has evolved 550 billion particles on an Adaptive Mesh Refinement grid with more than two trillion computing points along the entire evolutionary history of the Universe, and across 6 orders of magnitudes length scales, from the size of the Milky Way to the whole observable Universe. To date, this is the largest and most advanced cosmological simulation ever run. It will have a major scientific impact and provide an exceptional support to future observational programs dedicated to mapping the distribution of matter and galaxies in the Universe. The simulation has run on 4752 (of 5040) thin nodes of BULL supercomputer CURIE, using 300 TB of memory for 10 million hours of computing time. 50 PBytes of rough data were generated throughout the run, reduced to a useful amount of 500 TBytes using an advanced and innovative reduction workflow.

Optimizing the Computation of N-Point Correlations on Large-Scale Astronomical Data

Authors: William B. March, Kenneth Czechowski, Marat Dukhan, Thomas Benson, Dongryeol Lee (Georgia Institute of Technology), Andrew J. Connolly (University of Washington), Richard Vuduc, Edmond Chow, Alexander G. Gray (Georgia Institute of Technology)

The n -point correlation functions (npcf) are powerful statistics that are widely used for data analyses in astronomy and other fields. These statistics have played a crucial role in fundamental physical breakthroughs, including the discovery of dark energy. Unfortunately, directly computing the npcf at a single value requires $\mathcal{O}(N^n)$ time for N points and values of n of 2, 3, 4, or even larger. Astronomical data sets can contain billions of points, and the next generation of surveys will generate terabytes of data per night. To meet these computational demands, we present a highly-tuned npcf computation code that shows an order-of-magnitude speedup over current state-of-the-art. This enables a much larger 3-point correlation computation on the galaxy distribution than was previously possible. We show a detailed performance evaluation on many different architectures.

Hierarchical Task Mapping of Cell-Based AMR Cosmology Simulations

Authors: Jingjin Wu, Zhiling Lan, Xuanxing Xiong (Illinois Institute of Technology), Nickolay Y. Gnedin (Fermi National Laboratory), Andrey V. Kravtsov (University of Chicago)

Cosmology simulations are highly communication-intensive; thus, it is critical to exploit topology-aware task mapping techniques for performance optimization. To exploit the architectural properties of multiprocessor clusters (the performance gap between inter-node and intra-node communication as well as the gap between inter-socket and intra-socket communication), we design and develop a hierarchical task-mapping scheme for cell-based AMR (Adaptive Mesh Refinement) cosmology simulations, in particular, the ART application. Our scheme consists of two parts: (1) an inter-node mapping to map application processes onto nodes with the objective of minimizing network traffic among nodes and (2) an intra-node mapping within each node to minimize the maximum size of messages transmitted between CPU sockets. Experiments on production supercomputers with 3D torus and fat-tree topologies show that our scheme can significantly reduce application communication cost by up to 50%. More importantly, our scheme is generic and can be extended to many other applications.

Fault Detection and Analysis

Chair: Pedro C. Diniz (University of Southern California)

10:30am-12pm

Room: 255-BC

A Study of DRAM Failures in the Field

Authors: Vilas Sridharan, Dean Liberty (AMD)

Most modern computer systems use dynamic random access memory (DRAM) as a main memory store. Recent publications have confirmed that DRAM is a common source of failures in the field. Therefore, further attention to the faults experienced by DRAM is warranted. We present a study of 11 months of DRAM errors in a large high-performance computing cluster. Our goal is to understand the failure modes, rates, and fault types experienced by DRAM in production settings. We draw several conclusions from our study. First, DRAM failure modes appear dominated by permanent faults. Second, DRAMs are susceptible to large multi-bit failures, such as failures that affect an entire DRAM row or column. Third, some DRAM failures can affect shared board-level circuitry, disrupting accesses to other DRAM devices that share the same circuitry. Finally, we find that chipkill functionality is extremely effective, reducing the node failure rate from DRAM errors by over 36x.

Fault Prediction Under the Microscope - A Closer Look Into HPC Systems

Authors: Ana Gainaru (University of Illinois at Urbana-Champaign), Franck Cappello (INRIA), William Kramer (National Center for Supercomputing Applications), Marc Snir (University of Illinois at Urbana-Champaign)

A large percentage of computing capacity in today's large HPC systems is wasted due to failures. As a consequence, current research is focusing on providing fault tolerance strategies that aim to minimize fault's effects on applications. A complement to this approach is failure avoidance, where the occurrence of a fault is predicted and preventive measures are taken. For this, monitoring systems require a reliable prediction system to give information on what will be generated and at what location. In this paper, we merge signal analysis concepts with data mining techniques to extend the ELSA toolkit to offer an adaptive and overall more efficient prediction module. To this end, a large part of the paper is focused on a detailed analysis of the prediction method, by applying it to two large-scale systems. Furthermore, we analyze the prediction's impact on current checkpointing strategies and highlight future improvements and directions.

Detection and Correction of Silent Data Corruption for Large-Scale High-Performance Computing

Authors: David Fiala, Frank Mueller (North Carolina State University), Christian Engelmann (Oak Ridge National Laboratory), Rolf Riesen (IBM Ireland), Kurt Ferreira, Ron Brightwell (Sandia National Laboratories)

Faults have become the norm rather than the exception for high-end computing clusters. Exacerbating this situation, some of these faults remain undetected, manifesting themselves as silent errors that allow applications to compute incorrect results. This paper studies the potential for redundancy to detect and correct soft errors in MPI message-passing applications while investigating the challenges inherent to detecting soft errors within MPI applications by providing transparent MPI redundancy. By assuming a model wherein corruption in application data manifests itself by producing differing MPI messages between replicas, we study the best suited protocols for detecting and correcting corrupted MPI messages. Using our fault injector, we observe that even a single error can have profound effects on applications by causing a cascading pattern of corruption which, in most cases, spreads to all other processes. Results indicate that our consistency protocols can successfully protect applications experiencing even high rates of silent data corruption.

Grid Computing

Chair: Manish Parashar (Rutgers University)

10:30am-12pm

Room: 355-D

ATLAS Grid Workload on NDGF Resources: Analysis, Modeling, and Workload Generation

Authors: Dmytro Karpenko, Roman Vitenberg, Alexander Lincoln Read (University of Oslo)

Evaluating new ideas for job scheduling or data transfer algorithms in large-scale grid systems is known to be notoriously challenging. Existing grid simulators expect to receive a realistic workload as an input. Such input is difficult to provide in absence of an in-depth study of representative grid workloads. In this work, we analyze the ATLAS workload processed on the resources of NDG Facility. ATLAS is one of the biggest grid technology users, with extreme demands for CPU power and bandwidth. The analysis is based on the data sample with ~1.6 million jobs, 1723TB of data transfer, and 873 years of processor time. Our additional contributions are (a) scalable workload models that can be used to generate a synthetic workload for a given number of jobs, (b) an open-source workload generator software integrated with existing grid simulators, and (c) suggestions for grid system designers based on the insights of data analysis.

On the Effectiveness of Application-Aware Self-Management for Scientific Discovery in Volunteer Computing Systems

Authors: Trilce Estrada, Michela Taufer (University of Delaware)

An important challenge faced by high-throughput, multiscale applications is that human intervention has a central role in driving their success. However, manual intervention is inefficient, error-prone and promotes resource wasting. This paper presents an application-aware modular framework that provides self-management for computational multiscale applications in volunteer computing (VC). Our framework consists of a learning engine and three modules that can be easily adapted to different distributed systems. The learning engine of this framework is based on our novel tree-like structure called KOTree. KOTree is a fully automatic method that organizes statistical information in a multi-dimensional structure that can be efficiently searched and updated at runtime. Our empirical evaluation shows that our framework can effectively provide application-aware self-management in VC systems. Additionally, we observed that our algorithm is able to predict accurately the expected length of new jobs, resulting in an average of 85% increased throughput with respect to other algorithms.

On Using Virtual Circuits for GridFTP Transfers

Authors: Zhengyang Liu, Malathi Veeraraghavan, Zhenzhen Yan (University of Virginia), Chris Tracy (Energy Sciences Network), Jing Tie, Ian Foster (University of Chicago), John Dennis (National Center for Atmospheric Research), Jason Hick (Lawrence Berkeley National Laboratory), Yee-Ting Li (SLAC National Accelerator Laboratory), Wei Yang (SLAC National Accelerator Laboratory)

GridFTP transfer logs obtained from NERSC, SLAC, and NCAR, were analyzed. The goal of the analyses is to characterize these transfers and determine the suitability of dynamic virtual circuit (VC) service for these transfers instead of the currently used IP-routed service. Given VC setup overhead, the first analysis of the GridFTP transfer logs characterizes the duration of sessions. Of the NCAR-NICS sessions analyzed, 56% of sessions would have been long enough to be served with dynamic VC service. An analysis of transfer throughput across four paths, NCAR-NICS, SLAC-BNL, NERSC-ORNL and NERSC-ANL, shows significant variance. An analysis of the potential causes of this variance shows that server-related factors are more important than network-related factors. This is because most of the network links are lightly loaded, which implies that throughput variance is likely to remain unchanged with virtual circuits.

Performance Modeling

Chair: Dimitris Nikolopoulos (Queen's University, Belfast)

10:30am-12pm

Room: 355-EF

Dataflow-Driven GPU Performance Projection for Multi-Kernel Transformations

Authors: Jiayuan Meng, Vitali Morozov, Venkatram Vishwanath, Kalyan Kumaran (Argonne National Laboratory)

Applications often have a sequence of parallel operations to be offloaded to graphics processors; each operation can become an individual GPU kernel. Developers typically explore different transformations for each kernel. It is well known that efficient data management is critical in achieving high GPU performance and that “fusing” multiple kernels into one may greatly improve data locality. Doing so, however, requires transformations across multiple, potentially nested, parallel loops; at the same time, the original code semantics must be preserved. Since each kernel may have distinct data access patterns, their combined dataflow can be nontrivial. As a result, the complexity of multi-kernel transformations often leads to significant effort with no guarantee of performance benefits. This paper proposes a dataflow-driven analytical framework to project GPU performance for a sequence of parallel operations without implementing GPU code or using physical hardware. The framework also suggests multi-kernel transformations that can achieve the projected performance.

A Practical Method for Estimating Performance Degradation on Multicore Processors and its Application to HPC Workloads

Authors: Tyler Dwyer, Alexandra Fedorova, Sergey Blagodurov, Mark Roth, Fabien Gaud, Jian Pei (Simon Fraser University)

When multiple threads or processes run on a multicore CPU they compete for shared resources, such as caches and memory controllers, and can suffer performance degradation as high as 200%. We design and evaluate a new machine learning model that estimates this degradation online, on previously unseen workloads, and without perturbing the execution. Our motivation is to help data center and HPC cluster operators effectively use workload consolidation. Consolidation places many runnable entities on the same server to maximize hardware utilization, but may sacrifice performance as threads compete for resources. Our model helps determine when consolidation is overly harmful to performance. Our work is the first to apply machine learning to this problem domain, and we report on our experience reaping the advantages of machine learning while navigating around its limitations. We demonstrate how the model can be used to improve performance fidelity and save power for HPC workloads.

Aspen - A Domain Specific Language for Performance Modeling

Authors: Kyle L. Spafford, Jeffrey S. Vetter (Oak Ridge National Laboratory)

We present a new approach to analytical performance modeling using Aspen (Abstract Scalable Performance Engineering Notation), a domain specific language. Aspen fills an important gap in existing performance modeling techniques and is designed to enable rapid exploration of new algorithms and architectures. It includes a formal specification of an application's performance behavior and an abstract machine model. We provide an overview of Aspen's features and demonstrate how it can be used to express a performance model for a three dimensional Fast Fourier Transform. We then demonstrate the composability and modularity of Aspen by importing and reusing the FFT model in a molecular dynamics model. We have also created a number of tools that allow scientists to balance application and system factors quickly and accurately.

Big Data

Chair: Dennis Gannon (Microsoft Corporation)

1:30pm-3pm

Room: 255-EF

Design and Analysis of Data Management in Scalable Parallel Scripting

Authors: Zhao Zhang, Daniel S. Katz (University of Chicago), Justin M. Wozniak (Argonne National Laboratory), Allan Espinosa, Ian Foster (University of Chicago)

We seek to enable efficient large-scale parallel execution of applications in which a shared filesystem abstraction is used to couple many tasks. Such parallel scripting (Many-Task-Computing) applications suffer poor performance and utilization on large parallel computers due to the volume of filesystem I/O and a lack of appropriate optimizations in the shared filesystem. Thus, we design and implement a scalable MTC data management system that uses aggregated compute node local storage for more efficient data movement strategies. We co-design the data management system with the data-aware scheduler to enable dataflow pattern identification and automatic optimization. The framework reduces the time-to-solution of parallel stages of an astronomy data analysis application, Montage, by 83.2% on 512 cores, decreases time-to-solution of a seismology application, CyberShake, by 7.9% on 2,048 cores, and delivers BLAST performance better than mpiBLAST at various scales up to 32,768 cores, while preserving the flexibility of the original BLAST application.

Usage Behavior of a Large-Scale Scientific Archive

Authors: Ian F. Adams, Brian A. Madden, Joel C. Frank (University of California, Santa Cruz), Mark W. Storer (NetApp), Ethan L. Miller (University of California, Santa Cruz), Gene Harano (National Center for Atmospheric Research)

Archival storage systems for scientific data have been growing in both size and relevance over the past two decades, yet researchers and system designers alike must rely on limited and obsolete knowledge to guide archival management and design. To address this issue, we analyzed three years of file-level activities from the NCAR mass storage system, providing valuable insight into a large-scale scientific archive with over 1600 users, tens of millions of files, and petabytes of data. Our examination of system usage showed that, while a subset of users were responsible for most of the activity, this activity was widely distributed at the file level. We also show that the physical grouping of files and directories on media can improve archival storage system performance. Based on our observations, we provide suggestions and guidance for both future scientific archival system designs as well as improved tracing of archival activity.

On Distributed File Tree Walk of Parallel File Systems

Authors: Jharrod LaFon (Los Alamos National Laboratory), Satyajayant Misra (New Mexico State University), Jon Bringham (Los Alamos National Laboratory)

Supercomputers generate vast amounts of data, typically organized into large directory hierarchies on parallel file systems. While the supercomputing applications are parallel, the tools used to process them requiring complete directory traversals, are typically serial. We present an algorithm framework and three fully distributed algorithms for traversing large parallel file systems, and performing file operations in parallel. The first algorithm introduces a randomized work-stealing scheduler; the second improves the first with topology-awareness; and the third improves upon the second by using a hybrid approach. We have tested our implementation on Cielo, a 1.37 petaflop supercomputer at the Los Alamos National Laboratory and its 7 petabyte file system. Test results show that our algorithms execute orders of magnitude faster than state-of-the-art algorithms while achieving ideal load balancing and low communication cost. We present performance insights from the use of our algorithms in production systems at LANL, performing daily file system operations.

Memory Systems

Chair: Jaejin Lee (Seoul National University)

1:30pm-3pm

Room: 355-D

Application Data Prefetching on the IBM Blue Gene/Q Supercomputer

Authors: I-Hsin Chung, Changhoan Kim, Hui-Fang Wen, Guojing Cong (IBM T.J. Watson Research Center)

Memory access latency is often a crucial performance limitation for high performance computing. Prefetching is one of the strategies used by system designers to bridge the processor-memory gap. This paper describes a new innovative list prefetching feature introduced in the IBM Blue Gene/Q supercomputer. The list prefetcher records the L1 cache miss addresses and prefetches them in the next iteration. The evaluation shows this list prefetching mechanism reduces L1 cache misses and improves the performance for high performance computing applications with repeating non-uniform memory access patterns. Its performance is compatible with classic stream prefetcher when properly configured.

Hardware-Software Coherence Protocol for the Coexistence of Caches and Local Memories

Authors: Lluc Alvarez, Lluís Vilanova, Marc Gonzalez, Xavier Martorell, Nacho Navarro, Eduard Ayguade (Barcelona Supercomputing Center)

Cache coherence protocols limit the scalability of chip multiprocessors. One solution is to introduce a local memory alongside the cache hierarchy, forming a hybrid memory system. Local memories are more power-efficient than caches and they do not generate coherence traffic but they suffer from poor programmability. When non-predictable memory access patterns are found compilers do not succeed in generating code because of the incoherency between the two storages. This paper proposes a coherence protocol for hybrid memory systems that allows the compiler to generate code even in the presence of memory aliasing problems. Coherency is ensured by a simple software/hardware co-design where the compiler identifies potentially incoherent memory accesses and the hardware diverts them to the correct copy of the data. The coherence protocol introduces overheads of 0.24% in execution time and of 1.06% in energy consumption to enable the usage of the hybrid memory system.

What Scientific Applications Can Benefit from Hardware Transactional Memory

Authors: Martin Schindewolf (Karlsruhe Institute of Technology), Martin Schulz, John Gyllenhaal, Barna Bihari (Lawrence Livermore National Laboratory), Amy Whang (IBM Toronto Lab), Wolfgang Karl (Karlsruhe Institute of Technology)

Achieving efficient and correct synchronization of multiple threads is a difficult and error-prone task at small scale and, as we march towards extreme scale computing, will be even more challenging when the resulting application is supposed to utilize millions of cores efficiently. Transactional Memory (TM) is a promising technique to ease the burden on the programmer, but only recently has become available on commercial hardware in the new Blue Gene/Q system and hence the real benefit for scientific applications has not been studied yet. This paper presents the first performance results of TM embedded into OpenMP on a prototype system of BG/Q and characterizes code properties that will likely lead to benefits when augmented with TM primitives. Finally, we condense our findings into a set of best practices and apply them to a Monte Carlo Benchmark and a Smoothed Particle Hydrodynamics method to optimize the performance.

Numerical Algorithms

1:30pm-3pm

Room: 355-EF

A Parallel Two-Level Preconditioner for Cosmic Microwave Background Map-Making

Authors: Laura Grigori (INRIA), Radek Stompor (Paris Diderot University), Mikolaj Szydlarski (INRIA)

In this work we study performance of two-level preconditioners in the context of iterative solvers of the generalized least square systems, where the weights are assumed to be described by a block-diagonal matrix with Toeplitz blocks. Such cases are physically well motivated and arise whenever the instrumental noise displays a piece-wise stationary behavior. Our iterative algorithm is based on a conjugate gradient method with a parallel two-level preconditioner (2lvl-PCG) for which we construct its coarse space from a limited number of sparse vectors estimated solely from coefficients of the initial linear system. Our prototypical application is the map-making problem in the Cosmic Microwave Background observations. We show experimentally that our parallel implementation of 2lvl-PCG outperforms by as much as a factor of 5 the standard one-level PCG in terms of both the convergence rate and the time to solution.

A Massively Space-Time Parallel N-Body Solver

Authors: Robert Speck, Daniel Ruprecht, Rolf Krause (Università della Svizzera italiana), Matthew Emmett (Lawrence Berkeley National Laboratory), Michael Minion (Stanford University), Mathias Winkel, Paul Gibbon (Forschungszentrum Juelich)

We present a novel space-time parallel version of the Barnes-Hut tree code PEPC using PFASST, the Parallel Full Approximation Scheme in Space and Time. The naive use of increasingly more processors for a fixed-size N-body problem is prone to saturate as soon as the number of unknowns per core becomes too small. To overcome this intrinsic strong-scaling limit, we introduce temporal parallelism on top of PEPC's existing hybrid MPI/PThreads spatial decomposition. Here, we use PFASST which is based on a combination of the iterations of the parallel-in-time algorithm parareal with the sweeps of spectral deferred correction (SDC) schemes. By combining these sweeps with multiple space-time discretization levels, PFASST relaxes the theoretical bound on parallel efficiency in parareal. We present results from runs on up to 262,144 cores on the IBM Blue Gene/P installation JUGENE, demonstrating that the space-time parallel code provides speedup beyond the saturation of the purely space-parallel approach.

High Performance General Solver for Extremely Large-Scale Semidefinite Programming Problems

Authors: Katsuki Fujisawa (Chuo University), Toshio Endo, Hitoshi Sato, Makoto Yamashita, Satoshi Matsuoka (Tokyo Institute of Technology), Maho Nakata (RIKEN)

Semidefinite Programming (SDP) is one of the most important optimization problems, which covers a wide range of applications such as combinatorial optimization, control theory, quantum chemistry, truss topology design, etc. Solving extremely large-scale SDP problems has significant importance for the current and future applications of SDPs. We have developed SDPA aimed for solving large-scale SDP problems with numerical stability. SDPARA is a parallel version of SDPA, which replaces two major bottleneck parts (the generation of the Schur complement matrix and its Cholesky factorization) of SDPA by their parallel implementation. In particular, it has been successfully applied on combinatorial optimization and truss topology optimization, new SDPARA(7.5.0-G) on a large-scale supercomputer called TSUBAME2.0 has succeeded to solve the largest SDP problem which has over 1.48 million constraints and make a new world record. Our implementation has also achieved 533 TFlops for the large-scale Cholesky factorization using 2,720 CPUs and 4,080 GPUs.

Performance Optimization

Chair: Padma Raghavan (Pennsylvania State University)

1:30pm-3pm

Room: 255-BC

Extending the BT NAS Parallel Benchmark to Exascale Computing

Authors: Rob F. Van Der Wijngaart, Srinivas Sridharan, Victor W. Lee (Intel Corporation)

The NAS Parallel Benchmarks (NPB) are a well-known suite of benchmarks that proxy scientific computing applications. They specify several problem sizes that represent how such applications may run on different sizes of HPC systems. However, even the largest problem (Class F) is still far too small to exercise properly a Petascale supercomputer. Our work shows how one may scale the Block Tridiagonal (BT) NPB from today's size to Petascale and Exascale computing systems. In this paper we discuss the pros and cons of various ways of scaling. We discuss how scaling BT would impact computation, memory access and communications, and highlight the expected bottleneck, which turns out to be not memory or communication bandwidth, but latency. Two complementary ways are presented to overcome latency obstacles. We also describe a practical method to gather approximate performance data for BT at exascale on actual hardware, without requiring an exascale system.

NUMA-Aware Graph Mining Techniques for Performance and Energy Efficiency

Authors: Michael R. Frasca, Kamesh Madduri, Padma Raghavan (Pennsylvania State University)

We investigate dynamic methods to improve power and performance performance profiles of large irregular applications on modern multi-core systems. In this context, we study a large sparse graph application, Betweenness Centrality, and focus on memory behavior as core count scales. We introduce new techniques to efficiently map the computational demands onto non-uniform memory architectures (NUMA). Our dynamic design adapts to hardware topology and dramatically improves both energy and performance. These gains are more significant at higher core counts. We implement a scheme for adaptive data layout, which reorganizes the graph after observing parallel access patterns, and a dynamic task scheduler that encourages shared data between neighboring cores. We measure performance and energy consumption on a modern multi-core machine and observe that mean execution time is reduced by 51.2% and energy is reduced by 52.4%.

Optimization of Geometric Multigrid for Emerging Multi- and Manycore Processors

Authors: Samuel W. Williams (Lawrence Berkeley National Laboratory), Dhiraj D. Kalamkar (Intel Corporation), Amik Singh (University of California, Berkeley), Anand M. Deshpande (Intel Corporation), Brian Van Straalen (Lawrence Berkeley National Laboratory), Mikhail Smelyanskiy (Intel Corporation), Ann Almgren (Lawrence Berkeley National Laboratory), Pradeep Dubey (Intel Corporation), John Shalf, Leonid Oliker (Lawrence Berkeley National Laboratory)

Multigrid methods are widely used to accelerate the convergence of iterative solvers for linear systems. We explore optimization techniques for geometric multigrid on existing and emerging multicore systems including the Cray XE6, Intel SandyBridge and Nehalem-based Infiniband clusters, as well as Intel's forthcoming Knights Corner (KNC) Coprocessor. Our work examines a variety of techniques including communication-aggregation, threaded wavefront-based DRAM communication-avoiding, dynamic threading decisions, SIMDization, and fusion of operators. We quantify performance through each phase of the V-cycle for both single-node and distributed-memory experiments and provide detailed analysis for each class of optimization. Results show our optimizations yield significant speedups across a variety of subdomain sizes while demonstrating the potential of multi- and manycore processors to dramatically accelerate single-node performance. Our analysis also indicates that improvements in networks and communication will be essential to reap the potential of many-core processors in large-scale multigrid simulations.

Communication Optimization

Chair: Ron Brightwell (Sandia National Laboratories)

3:30pm-5pm

Room: 255-EF

Mapping Applications with Collectives over Sub-Communicators on Torus Networks

Authors: Abhinav Bhatele, Todd Gamblin, Steven H. Langer, Peer-Timo Bremer (Lawrence Livermore National Laboratory), Erik W. Draeger (Lawrence Livermore National Laboratory), Bernd Hamann, Katherine E. Isaacs (University of California, Davis), Aaditya G. Landge, Joshua A. Levine, Valerio Pascucci (University of Utah), Martin Schulz, Charles H. Still (Lawrence Livermore National Laboratory)

The placement of tasks in a parallel application on specific nodes of a supercomputer can significantly impact performance. Traditionally, task mapping has focused on reducing the distance between communicating processes on the physical network. However, for applications that use collectives over sub-communicators, this strategy may not be optimal. Many collectives can benefit from an increase in bandwidth even at the cost of an increase in hop count, especially when

sending large messages. We have developed a tool, Rubik, that provides a simple API to create a wide variety of mappings for structured communication patterns. Rubik supports several operations that can be combined into a large number of unique patterns. Each mapping can be applied to disjoint groups of MPI processes involved in collectives to increase the effective bandwidth. We demonstrate the use of these techniques for improving performance of two parallel codes, pF3D and Qbox, which use collectives over sub-communicators.

Optimization Principles for Collective Neighborhood Communications

Authors: Torsten Hoefler, Timo Schneider (University of Illinois at Urbana-Champaign)

Many scientific applications work in a bulk-synchronous mode of iterative communication and computation steps. Even though the communication steps happen at the same time, important patterns such as stencil computations cannot be expressed as collective communications in MPI. Neighborhood collective operations allow to specify arbitrary collective communication relations during runtime and enable optimizations similar to traditional collective calls. We show a number of optimization opportunities and algorithms for different communication scenarios. We also show how users can assert additional constraints that provide new optimization opportunities in a portable way. Our communication and protocol optimizations result in a performance improvement of up to a factor of two for stencil communications. We found that our optimization heuristics can automatically generate communication schedules that are comparable to hand-tuned collectives. With those optimizations, we are able to accelerate arbitrary collective communication patterns, such as regular and irregular stencils with optimization methods for collective communications.

Optimizing Overlay-Based Virtual Networking Through Optimistic Interrupts and Cut-Through Forwarding

Authors: Zheng Cui (University of New Mexico), Lei Xia (Northwestern University), Patrick G. Bridges (University of New Mexico), Peter A. Dinda (Northwestern University), John R. Lange (University of Pittsburgh)

Overlay-based virtual networking provides a powerful model for realizing virtual distributed and parallel computing systems with strong isolation, portability, and recoverability properties. However, in extremely high throughput and low latency networks, such overlays can suffer from bandwidth and latency limitations, which is of particular concern if we want to apply the model in HPC environments. Through careful study of an existing very high performance overlay-based virtual network system, we have identified two core issues limiting performance: delayed and/or excessive virtual interrupt delivery into guests, and copies between host and guest data buffers done during encapsulation. We respond with two novel

optimizations: optimistic, timer-free virtual interrupt injection, and zero-copy cut-through data forwarding. These optimizations improve the latency and bandwidth of the overlay network on 10 Gbps interconnects, resulting in near-native performance for a wide range of microbenchmarks and MPI application benchmarks.

Linear Algebra Algorithms

Chair: X. Sherry Li (Lawrence Berkeley National Laboratory)
3:30pm-5pm
Room: 355-D

Communication Avoiding and Overlapping for Numerical Linear Algebra

Authors: Evangelos Georganas (University of California, Berkeley), Jorge González-Domínguez (University of A Coruña), Edgar Solomonik (University of California, Berkeley), Yili Zheng (Lawrence Berkeley National Laboratory), Juan Touriño (University of A Coruña), Katherine Yelick (Lawrence Berkeley National Laboratory)

To efficiently scale dense linear algebra problems to future exascale systems, communication cost must be avoided or overlapped. Communication-avoiding 2.5D algorithms improve scalability by reducing inter-processor data transfer volume at the cost of extra memory usage. Communication overlap attempts to hide messaging latency by pipelining messages and overlapping with computational work. We study the interaction and compatibility of these two techniques for two matrix multiplication algorithms (Cannon and SUMMA), triangular solve, and Cholesky factorization. For each algorithm, we construct a detailed performance model which considers both critical path dependencies and idle time. We give novel implementations of 2.5D algorithms with overlap for each of these problems. Our software employs UPC, a partitioned global address space (PGAS) language that provides fast one-sided communication. We show communication avoidance and overlap provide a cumulative benefit as core counts scale, including results using over 24K cores of a Cray XE6 system.

Communication-Avoiding Parallel Strassen - Implementation and Performance

Authors: Benjamin Lipshitz, Grey Ballard, Oded Schwartz, James Demmel (University of California, Berkeley)

Matrix multiplication is a fundamental kernel of many high performance and scientific computing applications. Most parallel implementations use classical $O(n^3)$ matrix multiplication, even though there exist algorithms with lower arithmetic complexity. We recently presented a new Communication-Avoiding Parallel Strassen algorithm (CAPS), based on Strassen's fast matrix multiplication, that minimizes communication (SPAA '12). It communicates asymptotically less than all classical and all previous Strassen-based algorithms, and it attains theoretical lower bounds. In this paper we show that

CAPS is also faster in practice. We benchmark and compare its performance to previous algorithms on Hopper (Cray XE6), Intrepid (IBM BG/P), and Franklin (Cray XT4). We demonstrate significant speedups over previous algorithms both for large matrices and for small matrices on large numbers of processors. We model and analyze the performance of CAPS and predict its performance on future exascale platforms.

Managing Data-Movement for Effective Shared-Memory Parallelization of Out-of-Core Sparse Solvers

Authors: Haim Avron, Anshul Gupta (IBM T.J. Watson Research Center)

Direct methods for solving sparse linear systems are robust and typically exhibit good performance, but often require large amounts of memory due to fill-in. Many industrial applications use out-of-core techniques to mitigate this problem. However, parallelizing sparse out-of-core solvers poses some unique challenges because accessing secondary storage introduces serialization and I/O overhead. We analyze the data-movement costs and memory versus parallelism trade-offs in a shared-memory parallel out-of-core linear solver for sparse symmetric systems. We propose an algorithm that uses a novel memory management scheme and adaptive task parallelism to reduce the data-movement costs. We present experiments to show that our solver is faster than existing out-of-core sparse solvers on a single core, and is more scalable than the only other known shared-memory parallel out-of-core solver. This work is also directly applicable at the node level in a distributed-memory parallel scenario.

New Computer Systems

Chair: Jeffrey Vetter (Oak Ridge National Laboratory)

3:30pm-5pm

Room: 255-BC

Cray Cascade - A Scalable HPC System Based on a Dragonfly Network

Authors: Gregory Faanes, Abdulla Bataineh, Duncan Roweth, Tom Court, Edwin Froese, Bob Alverson, Tim Johnson, Joe Kopnick, Michael Higgins, James Reinhard (Cray Inc.)

Higher global bandwidth requirement for many applications and lower network cost have motivated the use of the Dragonfly network topology for high performance computing systems. In this paper we present the architecture of the Cray Cascade system, a distributed memory system based on the Dragonfly network topology. We describe the structure of the system, its Dragonfly network the routing algorithms, and a set of advanced features supporting both mainstream high performance computing applications and emerging global address space programming models. With a combination of performance results from prototype systems and simulation data for large

systems, we demonstrate the value of the Dragonfly topology and the benefits obtained through extensive use of adaptive routing.

GRAPE-8: An Accelerator for Gravitational N-Body Simulation with 20.5GFLOPS/W Performance

Authors: Junichiro Makino (Tokyo Institute of Technology), Hiroshi Daisaka (Hitotsubashi University)

In this paper, we describe the design and performance of GRAPE-8 accelerator processor for gravitational N-body simulations. It is designed to evaluate gravitational interaction with cutoff between particles. The cutoff function is useful for schemes like TreePM or Particle-Particle Particle-Tree, in which gravitational force is divided to short-range and long-range components. A single GRAPE-8 processor chip integrates 48 pipeline processors. The effective number of floating-point operations per interaction is around 40. Thus the peak performance of a single GRAPE-8 processor chip is 480 Gflops. A GRAPE-8 processor card houses two GRAPE-8 chips and one FPGA chip for PCI-Express interface. The total power consumption of the board is 46W. Thus, theoretical peak performance per wattage is 20.5 Gflops/W. The effective performance of the total system, including the host computer, is around 5Gflops/W. This is more than a factor of two higher than the highest number in the current Green500 list.

SGI UV2 - A Fused Computation and Data Analysis Machine

Authors: Gregory M. Thorson (SGI), Michael Woodacre (SGI)

UV2 is SGI's 2nd generation Data Fusion system. UV2 was designed to meet the latest challenges facing users in computation and data analysis. Its unique ability to perform both functions on a single platform enables efficient, easy to manage workflows.

This platform has a hybrid infrastructure, leveraging the latest Intel EP processors to provide industry leading computation. Due to its high bandwidth, extremely low latency NumaLink6 interconnect, plus vectorized synchronization and data movement, UV2 provides industry leading data intensive capability. It supports a single operating system (OS) image up to 64TB and 4K threads. Multiple OS images can be deployed on a single NL6 fabric, which has a single flat address space up to 8PB and 256K threads. These capabilities allow for extreme performance on a broad range of programming models and languages including: OpenMP, MPI, UPC, CAF, and SHMEM.



Posters/Scientific Visualization Showcase

Posters provide an excellent opportunity for short presentations and informal discussions with conference attendees. Posters display cutting-edge, interesting research in high performance computing, storage, networking and analytics. Posters will be prominently displayed for the duration of the conference, giving presenters a chance to showcase their latest results and innovations.

The Scientific Visualization Showcase is back for a second year. We received 26 submissions, which showcased a wide variety of research topics in HPC. Of those, we selected 16 for presentation this year. Selected entries are being displayed live in a museum/art gallery format so attendees can experience and enjoy the latest in science and engineering HPC results expressed through state-of-the-art visualization technologies.

Posters/
Scientific
Visualization
Showcase

Posters

Tuesday, November 13

Reception & Exhibit

5:15pm-7pm

Chair: Torsten Hoefer (ETH Zurich)

Room: East Entrance

Research Posters

ACM Student Research Competition Posters

Electronic Posters

Wednesday, November 14 -

Thursday, November 15

Exhibit

8:30am-5pm

Room: East Entrance

Research Posters

Matrices Over Runtime Systems at Exascale

Emmanuel Agullo (INRIA), George Bosilca (University of Tennessee, Knoxville), Cédric Castagnède (INRIA), Jack Dongarra (University of Tennessee, Knoxville), Hatem Ltaief (King Abdulah University of Science & Technology), Stan Tomov (University of Tennessee, Knoxville)

The goal of the Matrices Over Runtime Systems at Exascale (MORSE) project is to design dense and sparse linear algebra methods that achieve the fastest possible time to an accurate solution on large-scale multicore systems with GPU accelerators, using all the processing power that future high end systems can make available. In this poster, we propose a framework for describing linear algebra algorithms at a high level of abstraction and delegating the actual execution to a runtime system in order to design software whose performance is portable across architectures. We illustrate our methodology on three classes of problems: dense linear algebra, sparse direct methods and fast multipole methods. The resulting codes have been incorporated into Magma, Pastix and ScalFMM solvers, respectively.

Assessing the Predictive Capabilities of Mini-applications

Richard Barrett, Paul Crozier, Douglas Doerfler, Simon Hammond, Michael Heroux, Paul Lin, Timothy Trucano, Courtenay Vaughan, Alan Williams (Sandia National Laboratories)

The push to exascale computing is informed by the assumption that the architecture, regardless of the specific design, will be fundamentally different from petascale computers. The

Mantevo project has been established to produce a set of proxies, or “miniapps,” which enable rapid exploration of key performance issues that impact a broad set of scientific applications programs of interest to ASC and the broader HPC community. Understanding the conditions under which a miniapp can be confidently used as predictive of an applications’ behavior must be clearly elucidated. Toward this end, we have developed a methodology for assessing the predictive capabilities of application proxies. Adhering to the spirit of experimental validation, our approach provides a framework for examining data from the application with that provided by their proxies. In this poster we present this methodology, and apply it to three miniapps developed by the Mantevo project.

Towards Highly Accurate Large-Scale Ab Initio Calculations Using Fragment Molecular Orbital Method in GAMESS

Maricris L. Mayes (Argonne National Laboratory), Graham D. Fletcher (Argonne National Laboratory), Mark S. Gordon (Iowa State University)

One of the major challenges of modern quantum chemistry (QC) is to apply it to large systems with thousands of correlated electrons and basis functions. The availability of supercomputers and development of novel methods are necessary to realize this challenge. In particular, we employ the linear scaling Fragment Molecular Orbital (FMO) method which decomposes the large system into smaller, localized fragments which can be treated with a high-level QC method like MP2. FMO is inherently scalable since the individual fragment calculations can be carried out simultaneously on separate processor groups. It is implemented in GAMESS, a popular ab-initio QC program. We present the scalability and performance of FMO on Intrepid (Blue Gene/P) and Blue Gene/Q systems at Argonne Leadership Computing Facility. We also describe our work on multithreading the integral kernels in GAMESS to effectively use the enormous number of cores and threads of new generation supercomputers.

Acceleration of the BLAST Hydro Code on GPU

Tingxing Dong (University of Tennessee, Knoxville), Tzanio Kolev, Robert Rieben, Veselin Dobrev (Lawrence Livermore National Laboratory), Stanimire Tomov, Jack Dongarra (University of Tennessee, Knoxville)

The BLAST code implements a high-order numerical algorithm that solves the equations of compressible hydrodynamics using the Finite Element Method in a moving Lagrangian frame. BLAST is coded in C++ and parallelized by MPI. We accelerated the most computational intensive parts (80%-95%) of BLAST on NVIDIA GPUs with the CUDA programming model. Several 2D and 3D problems were tested and achieved an overall speedup of 4.3 on 1 M2050.

A Novel Hybrid CPU-GPU Generalized Eigensolver for Electronic Structure Calculations Based on Fine Grained Memory Aware Tasks

Raffaele Solcà (ETH Zurich), Azzam Haidar, Stanimire Tomov (University of Tennessee, Knoxville), Thomas C. Schulthess (ETH Zurich), Jack Dongarra (University of Tennessee, Knoxville)

The adoption of hybrid GPU-CPU nodes in traditional supercomputing platforms such as the Cray-XK6 opens acceleration opportunities for electronic structure calculations in materials science and chemistry applications, where medium-sized generalized eigenvalue problems must be solved many times. These eigenvalue problems are too small to scale on distributed systems, but can benefit from the massive compute performance concentrated on a single node, hybrid GPU-CPU system. However, hybrid systems call for the development of new algorithms that efficiently exploit heterogeneity and massive parallelism of not just GPUs, but of multi/many-core CPUs as well. Addressing these demands, we developed a novel algorithm featuring innovative: Fine grained memory aware tasks; Hybrid execution/scheduling, and Increased computational intensity. The resulting eigensolvers are state-of-the-art in HPC, significantly outperforming existing libraries. We describe the algorithm and analyze its performance impact on applications of interest when different fractions of eigenvectors are needed by the host electronic structure code.

HTCaaS: A Large-Scale High-Throughput Computing by Leveraging Grids, Supercomputers and Cloud

Seungwoo Rho, Seoyoung Kim, Sangwan Kim, Seokkyoo Kim, Jik-Soo Kim, Soonwook Hwang (Korea Institute of Science and Technology Information)

With the growing number of jobs and complexity in HTC problems, it is inevitable to utilize as many computing resources (such as grids, supercomputers and cloud) as possible. However, it is challenging for researchers to effectively utilize available resources that are under control by independent resource providers as the number of jobs (that should be submitted at once) increase dramatically (as in parameter sweeps or N-body calculations). We designed a HTCaaS (HTC as a Service) system which aims to provide researchers with ease of exploring a large-scale and complex HTC problems by leveraging grids, supercomputers and cloud. Our contributions include meta-job based automatic job split and submission, intelligent resource selection algorithm that automatically selects more responsive and effective resources, pluggable resource interface to heterogeneous computing environment and client application independency through a simple and uniform WS-Interface. We demonstrate the architecture of our system and how it works with several examples.

Evaluation of Magneto-Hydro-Dynamic Simulation on Three Types of Scalar Systems

Keiichiro Fukazawa, Takeshi Nanri, Toshiya Takami (RIIT, Kyuhsu Univ.)

The massively parallel computational performance of magneto-hydro-dynamic (MHD) code is evaluated on three scalar-type supercomputer systems. We have made performance tuning of a three-dimensional MHD code for planetary magnetosphere simulation on the Fujitsu PRIMEHPC FX10 (SPARC64IVfx) 4800 nodes, PRIMERGY CX400 S6 (Sandy Bridge Xeon) 1476 nodes Cray XE6 (Interlagos Opteron) 256 nodes. To adapt the exascale computing, we use two- and three-dimensional domain decomposition methods in the parallelization. As a result, we obtained the computation efficiency around 20% and good scalability on each system. The suitable optimization, however, is different among them. The cache tuning is important for the FX10, and the vectorization is more effective to the CX400 and XE6. In this study we also compare the evaluation results with those of other computer systems (Hitachi HA8000, SR16000/L2, Fujitsu FX1, RX200 S6 and NEC SX-9).

Three Steps to Model Power-Performance Efficiency for Emergent GPU-Based Parallel Systems

Shuaiwen Song (Virginia Tech), Chun-Yi Su (Virginia Tech), Barry Rountree (Lawrence Livermore National Laboratory), Kirk Cameron (Virginia Tech)

Massive parallelism combined with complex memory hierarchies form a barrier to efficient application and architecture design. These challenges are exacerbated with GPUs as parallelism increases an order of magnitude and power consumption can easily double. Models have been proposed to isolate power and performance bottlenecks and identify their root causes. However, no current models combine usability, accuracy, and support for emergent GPU architectures (e.g. NVIDIA Fermi). We combine hardware performance counter data with machine learning and advanced analytics to create a power-performance efficiency model for modern GPU-based systems. Our performance counter based approach is general and does not require detailed understanding of the underlying architecture. The resulting model is accurate for predicting power (within 2.1%) and performance (within 6.7%) for application kernels on modern GPUs. Our model can identify power-performance bottlenecks and their root causes for various complex computation and memory access patterns (e.g. global, shared, texture).

Impact of Integer Instructions in Floating Point Applications

Hisanobu Tomari (University of Tokyo), Kei Hiraki (University of Tokyo)

The performance of floating-point oriented applications is determined not only by the performance of floating-point instructions, but also by the speed of integer instruction execution. Dynamic instruction trace of NAS Parallel Benchmarks (NPB) workloads show that integer instructions are often executed more than floating-point instructions in the floating-point application benchmark. Some vendors are taking the SIMD-only strategy where integer performance stays the same as generations-old ones, while floating-point application performance is increased using SIMD instructions. We show that there is a limit for this approach and that the slow integer execution has a huge impact on the per-socket NPB scores. When these performances are compared to other historic processors, we found that some of the latest processors can be improved by using the known techniques to accelerate the integer performance.

Operating System Assisted Hierarchical Memory Management for Heterogeneous Architectures

Balazs Gerofi, Akio Shimada, Atsushi Hori (RIKEN), Yutaka Ishikawa (University of Tokyo)

Heterogeneous architectures, where a multicore processor—which is optimized for fast single-thread performance—is accompanied with a large number of simpler, but more power-efficient cores optimized for parallel workloads, are receiving a lot attention recently. Currently, these co-processors, such as Intel's Many Integrated Core (MIC) software development platform, come with a limited on-board RAM, which requires partitioning computational problems manually into pieces that can fit into the device's memory, and at the same time, efficiently overlapping computation and data movement between the host and the device. In this poster we present an operating system (OS) assisted hierarchical memory management system. We are aiming at transparent data movement between the device and the host memory, as well as tight integration with other OS services, such as file and network I/O.

MPACK - Arbitrary Accurate Version of BLAS and LAPACK

Maho Nakata (RIKEN)

We are interested in the accuracy of linear algebra operations, accuracy of the solution of linear equation, eigenvalue and eigenvectors of some matrices, etc. For this reason we have been developing the MPACK. MPACK consists of MBLAS and MLAPACK, multiple precision versions of BLAS and LAPACK, respectively. Features of MPACK are: (1) based on LAPACK 3.1.1; (2) provides a reference implementation and/or API; (3) written in C++, rewrite from FORTRAN77; (4) supports GMP and QD as multiple-precision arithmetic library; and (5) is

portable. The current version of MPACK is 0.7.0 and it supports 76 MBLAS routines and 100 MLAPACK routines. Some routines are accelerated via GPU or via OpenMP. These software packages are available at <http://mplapack.sourceforge.net/>.

Scalable Direct Eigenvalue Solver ELPA for Symmetric Matrices

Hermann Lederer (RZG, Max Planck Society), Andreas Marek (RZG, Max Planck Society)

ELPA is a new efficient distributed parallel direct eigenvalue solver for symmetric matrices. It contains both an improved one-step ScaLAPACK type solver (ELPA1) and the two-step solver ELPA2 [1,2]. ELPA has demonstrated good scalability for large matrices up to 294,000 cores of a BlueGene/P system [3]. ELPA is especially beneficial when a significant part, but not all eigenvectors are needed. For a quantification of this statement, matrix sizes of 10,000, 20,000, and 50,000 have been solved with ELPA1, ELPA2 and ScaLAPACK routines from Intel MKL 10.3 for real and complex matrices with eigenvector fractions of 10%, 25%, 50% and 100% on 1024 cores of an Intel Sandy Bridge based Linux cluster with FDR10 Infiniband interconnect.

Hybrid Breadth First Search Implementation for Hybrid-Core Computers

Kevin Wadleigh (Convey Computer)

The Graph500 benchmark is designed to evaluate the suitability of supercomputing systems on graph algorithms, which are increasingly important in HPC. The timed Graph500 kernel, Breadth First Search, exhibits memory access patterns typical of these types of applications, with poor spatial locality and synchronization between multiple streams of execution. The Graph500 benchmark was ported to a Convey HC-2ex, a hybrid-core computer with an Intel host system and a coprocessor incorporating four reprogrammable Xilinx FPGAs. The computer incorporates a unique memory system designed to sustain high bandwidth for random memory accesses. The BFS kernel was implemented as a hybrid algorithm with concurrent processing on both the host and coprocessor. The early steps use a top-down algorithm on the host with results copied to coprocessor memory for use in a bottom-up algorithm. The coprocessor uses thousands of threads to traverse the graph. The resulting implementation runs at over 11 billion TEPS.

Interface for Performance Environment Autoconfiguration Framework

Liang Men (University of Arkansas), Bilel Hadri, Haihang You (University of Tennessee, Knoxville)

Performance Environment Autoconfiguration framework (PEAK) is presented to help developers and users of scientific applications find the optimal configurations for their application on a given platform with rich computational resources and complicated options. The choices to be made include the compiler with its settings of compiling options, the numerical libraries and settings of library parameters, and settings of other environment variables to take advantage of the NUMA systems. A website-based interface is developed for users' convenience of choosing the optimal configuration to get a significant speedup for some scientific applications executed on different systems.

Imaging Through Cluttered Media Using Electromagnetic Interferometry on a Hardware-Accelerated High-Performance Cluster

Esam El-Araby, Ozlem Kilic, Vinh Dang (Catholic University of America)

Detecting concealed objects, such as weapons behind cluttered media, is essential for security applications. Terahertz frequencies are usually employed for imaging in security checkpoints due to their safe non-ionizing properties for humans. Interferometric images are constructed based on the complex correlation function of the received electric fields from the medium of interest. Interferometric imaging, however, is computationally intensive, which makes it impractical for real time requirements. It is essential, therefore, to have efficient implementations of the algorithm using HPC platforms. In this paper, we exploit the capabilities of a 13-node GPU-accelerated cluster using CUDA and MVAPICH2 environments for electromagnetic terahertz interferometric imaging through cluttered media. With efficient load balancing, the experimental results demonstrate the performance gain achieved in comparison to conventional platforms. The results also show potential scalability characteristics for larger HPC systems. This work will be presented in the poster session in the format of a short PowerPoint presentation.

Memory-Conscious Collective IO for Extreme-Scale HPC Systems

Yin Lu, Yong Chen (Texas Tech University), Rajeev Thakur (Argonne National Laboratory), Yu Zhuang (Texas Tech University)

The continuing decrease in memory capacity per core and the increasing disparity between core count and off-chip memory bandwidth create significant challenges for I/O operations in exascale systems. The exascale challenges require rethinking collective I/O for the effective exploitation of the cor-

relation among I/O accesses in the exascale system. In this study, considering the major constraint of the memory space, we introduce a Memory-Conscious collective I/O. Given the importance of I/O aggregator in improving the performance of collective I/O, the new collective I/O strategy restricts aggregation data traffic within disjointed subgroups, coordinates I/O accesses in intra-node and inter-node layer and determines I/O aggregators at run time considering data distribution and memory consumption among processes. The preliminary results have demonstrated that the new collective I/O strategy holds promise in substantially reducing the amount of memory pressure, alleviating contention for memory bandwidth and improving the I/O performance for extreme-scale systems.

Visualization Tool for Development of Topology-Aware Network Communication Algorithm

Ryohei Suzuki, Hiroaki Ishihata (Tokyo University of Technology)

We develop a visualization tool for designing a topology-aware communication algorithm. This tool visualizes the communication behavior from the logs of a network simulator or an existing parallel computer. Using multiple views, filtering functions, and an animation function, the tool affords users an intuitive understanding of communication behavior and provides statistical information. The topology view represents the spatial load distribution of the network topology in 3D space. The user can analyze the communication behavior on a specific network topology. A distinction between the behaviors of the new all-to-all communication algorithm and the conventional one is drawn clearly by the tool. In addition to the poster presentation, we are going to present a communication algorithm behavior on PC using our visualization tool.

Multi-GPU-Based Calculation of Percolation Problem on the TSUBAME 2.0 Supercomputer

Yukihiro Komura, Yutaka Okabe (Tokyo Metropolitan university)

We present the multi-GPU-based calculation of percolation problem on the 2D regular lattice for the multiple GPUs on the large-scale open science supercomputer TSUBAME 2.0. Recently, we presented the multiple GPU computing with the common unified device architecture (CUDA) for the cluster labeling. We adapt this cluster labeling algorithm to the percolation problem. In addition, we modify this cluster labeling algorithm in order to simplify the analysis for the percolation. As a result, we realized the large scale and rapid calculations without a decay of computational speed on the analysis for the percolation, and the calculation time for the 2D bond percolation with $L=65536$ is obtained as 180 milliseconds per a single realization.

Beating MKL and ScaLAPACK at Rectangular Matrix Multiplication Using the BFS/DFS Approach

James Demmel, David Elichu, Armando Fox, Shoaib Kamil, Benjamin Lipshitz, Oded Schwartz, Omer Spillinger (University of California, Berkeley)

We present CARMA, the first implementation of a communication-avoiding parallel rectangular matrix multiplication algorithm, attaining significant speedups over both MKL and ScaLAPACK. Combining the recursive BFS/DFS approach of Ballard, Demmel, Holtz, Lipshitz and Schwartz (SPAA '12) with the dimension splitting technique of Frigo, Leiserson, Prokop and Ramachandron (FOCS '99), CARMA is communication-optimal, cache- and network-oblivious, and simple to implement (60 lines of code for the shared-memory version). Since CARMA minimizes communication across the network, between NUMA domains, and between levels of cache, it performs well on both shared- and distributed-memory machines.

Evaluating Topology Mapping via Graph Partitioning

Anshu Arya (University of Illinois at Urbana-Champaign), Todd Gamblin, Bronis R. de Supinski (Lawrence Livermore National Laboratory), Laxmikant V. Kale (University of Illinois at Urbana-Champaign)

Intelligently mapping applications to machine network topologies has been shown to improve performance, but considerable developer effort is required to find good mappings. Techniques from graph partitioning have the potential to automate topology mapping and relieve the developer burden. Graph partitioning is already used for load balancing parallel applications, but can be applied to topology mapping as well. We show performance gains by using a topology-targeting graph partitioner to map sparse matrix-vector and volumetric 3-D FFT kernels onto a 3-D torus network.

Communication Overlap Techniques for Improved Strong Scaling of Gyrokinetic Eulerian Code Beyond 100k Cores on K-Computer

Yasuhiro Idomura, Motoki Nakata, Susumu Yamada (Japan Atomic Energy Agency), Toshiyuki Imamura (RIKEN), Tomohiko Watanabe (National Institute for Fusion Science), Masahiko Machida (Japan Atomic Energy Agency), Masanori Nunami (National Institute for Fusion Science), Hikaru Inoue, Shigenobu Tsutsumi, Ikuo Miyoshi, Naoyuki Shida (Fujitsu)

A plasma turbulence research based on 5D gyrokinetic simulations is one of the most critical and demanding issues in fusion science. To pioneer new physics regimes both in problem sizes and in time scales, an improvement of strong scaling is essential. Overlap of computations and communications is a promising approach, but it often fails on practical applications with conventional MPI libraries. In this work, this classical issue is clarified, and resolved by developing communication overlap

techniques with mpi_test and communication threads, which work even on conventional MPI libraries and hardware. These techniques dramatically improve the parallel efficiency of a gyrokinetic Eulerian code GT5D on K and Helios, which adopt dedicated and commodity networks. On K, excellent strong scaling is confirmed beyond 10⁵ cores with keeping the peak ratio of 10% (307 TFlops at 196,608 cores), and simulations for ITER-size fusion devices are significantly accelerated.

Polarization Energy On a Cluster of Multicores

Jesmin Jahan Tithi (Stony Brook University)

We have implemented distributed-memory and distributed-shared-memory parallel octree-based algorithms for approximating polarization energy of protein molecules by extending prior work of Chowdhury et al. (2010) for shared-memory architectures. This is an octree-based hierarchical algorithm, based on Greengard-Rokhlin type near and far decomposition of data points (i.e., atoms and points sampled from the molecular surface) which calculates the polarization energy of protein molecules using the r^6 approximation of Generalized Born radius of atoms. We have shown that our implementations outperform some state-of-the-art polarization energy implementations such as Amber and GBr6. Using approximations and efficient load-balancing scheme, we have achieved a speedup factor of about 10k w.r.t. the naïve exact algorithm with less than 1% error using as few as 144 cores (i.e., 12 compute nodes with 12 cores each) for molecules with millions of atoms.

Exploring Performance Data with Boxfish

Katherine E. Isaacs (University of California, Davis), Aaditya G. Landge (University of Utah), Todd Gamblin (Lawrence Livermore National Laboratory), Peer-Timo Bremer (Lawrence Livermore National Laboratory), Valerio Pascucci (University of Utah), Bernd Hamann (University of California, Davis)

The growth in size and complexity of scaling applications and the systems on which they run pose challenges in analyzing and improving their overall performance. To aid the process of exploration and understanding, we announce the initial release of Boxfish, an extensible tool for manipulating and visualizing data pertaining to application behavior. Combining and visually presenting data and knowledge from multiple domains, such as the application's communication patterns and hardware's network configuration and routing policies, can yield the insight necessary to discover the underlying causes of observed behavior. Boxfish allows users to query, filter and project data across these domains to create interactive, linked visualizations.

Reservation-Based I/O Performance Guarantee for MPI-IO Applications using Shared Storage Systems

Yusuke Tanimura (National Institute of Advanced Industrial Science & Technology), Rosa Filgueira (University of Edinburgh), Isao Kojima (National Institute of Advanced Industrial Science & Technology), Malcolm Atkinson (University of Edinburgh)

While optimized collective I/O methods are proposed for MPI-IO implementations, a problem in concurrent use of the shared storage system is raised. In order to prevent performance degradation of parallel I/O due to such I/O conflict, we propose an advance reservation approach, including possible integration with existing batch scheduler on HPC clusters. In this work, we use Dynamic-CoMPI as a MPI-IO implementation and Papio as a shared storage system which implements parallel I/O and performance reservation. Then we have been developing the ADIO layer to connect these systems and to evaluate the benefits of the reservation-based performance isolation. Our prototype implementation, Dynamic-CoMPI/Papio, was evaluated using the MPI-IO Test benchmark and the BISP3D application. While total execution time increased 3~12% with the Dynamic-CoMPI/PVFS2 under the situation where additional workload affects on MPI execution, there was no obvious time increase with Dynamic-CoMPI/Papio.

Visualizing and Mining Large Scale Scientific Data Provenance

Peng Chen (Indiana University), Beth Plale (Indiana University)

The provenance of digital scientific data is an important piece of the metadata of a data object. It can help increase the understanding and thus the acceptance of scientific result by showing all factors that contribute to the result. Provenance of scientific data from HPC experiments, however, can grow voluminous quickly because of its larger amount of (intermediate) data and ever-increasing complexity. While previous research focuses on small/medium size of provenance data, we have designed two new approaches for large scale provenance: we developed a provenance visualization component that enables the scientists to interactively navigate, manipulate, and analyze large-scale data provenance; we proposed a representation of provenance based on logical time that reduces the feature space and preserves interesting features so that data mining on the representation yields provenance-useful information. We demonstrate provenance visualizations from different types of experiments, and the evaluation result of mining a 10GB provenance database.

Using Active Storage Concept for Seismic Data Processing

Ekaterina Tyutlyaeva (PSI of RAS), Alexander Moskovsky (RSK SKIF), Sergey Konyuhov (RSK SKIF), Evgeny Kurin (GEOLAB)

This poster presents an approach to distributed seismic data processing using Seismic Un*x Software and the Active Storage system based on TSim C++ template library and the Lustre file system. Active Storage concept implies the use of distributed system architecture, where each node has processing and storage capabilities. Data are distributed across these nodes, computational tasks are submitted to the most suitable nodes, to reduce network traffic. The main benefits of our approach are: (a) Effective storing and processing of seismic data with minimal changes in CWP/SU modules; (2) Performance testing shows that the Active Storage system is effective; and (c) Meta-programming with C++ templates permits a flexible implementation of scheduling algorithms. The study analyzes performance results of the developed system as well as the usability of Active Storage and Seismic Unix integration. In the nearest future, we will get results for a 1.2 TB data processing.

Slack-Conscious Lightweight Loop Scheduling for Scaling Past the Noise Amplification Problem

Vivek Kale (University of Illinois at Urbana-Champaign), Todd Gamblin (Lawrence Livermore National Laboratory), Torsten Hoefler (University of Illinois at Urbana-Champaign), Bronis de Supinski (Lawrence Livermore National Laboratory), William Gropp (University of Illinois at Urbana-Champaign)

The amount of overhead that noise amplification causes can increase dramatically as we scale the application to a very large numbers of processes (10,000 or more). In prior work, we have introduced lightweight scheduling, which combines dynamic and static task scheduling to reduce the total number of dequeue operations while still absorbing noise on a node. In this work, we exploit a priori knowledge of per-process MPI slack to reduce the static fraction for those MPI processes that are known not to be on the critical path and thus likely not to amplify noise. This technique gives a 11% performance gain over the original lightweight scheduling (17% gain over static scheduling) when we run an AMG application on up to 16,384 process runs (1024 nodes) of a NUMA cluster, and are able to project further performance gains on machines with node counts beyond 10,000. (More details on poster in dynHybSummary.pdf)

Solving the Schroedinger and Dirac Equations of Atoms and Molecules with Massively Parallel Super-Computer

Hiroyuki Nakashima, Atsushi Ishikawa, Yusaku Kurokawa, Hiroshi Nakatsuji (Quantum Chemistry Research Institute)

Schroedinger and relativistic Dirac equations are the most fundamental equations in quantum mechanics and they govern most of phenomena in molecular material science. In spite of that importance, however, their exact solutions have not been able to be solved for over 80 years. Recently, one of the authors was successful to propose a new general theory of solving these equations. In addition, the method proposed for general atoms and molecules is very suitable for massively parallel computing since the sampling procedure is used for solving the local Schroedinger equation. In the present presentation, we will show some practical applications of our method to general atoms and molecules. Our final purpose is to create quantum chemistry as a predictive science with the solution of the Schroedinger and relativistic Dirac equations and the massively parallel super computing should be a help for that purpose.

Leveraging PEPPER Technology for Performance Portable Supercomputing

Raymond Namyst (INRIA), Christoph Kessler, Usman Dastgeer, Mudassar Majeed (Linköping University), Siegfried Benkner, Sabri Pillana (University of Vienna), Jesper Larsson Traff (Vienna University of Technology)

PEPPER is a 3-year EU FP7 project that develops a novel approach and framework for performance portability and programmability of heterogeneous multi-core systems. Its primary target is single-node heterogeneous systems, where several CPU cores are supported by accelerators such as GPUs. With this poster we give a short survey of selected parts of the PEPPER framework for single-node systems and then elaborate the perspectives for leveraging the PEPPER approach to generate performance-portable code for heterogeneous multinode systems.

Networking Research Activities at Fermilab for Big Data Analysis

Phil Demar, David Dykstra, Gabriele Garzoglio, Parag Mhashilkar, Anupam Rajendran (Fermi National Laboratory), Wenji Wu (Fermi National Laboratory)

Exascale science translates to big data. In the case of the Large Hadron Collider (LHC), the data is not only immense, it is also globally distributed. Fermilab is host to the LHC Compact Muon Solenoid (CMS) experiments US Tier-1 Center, the largest of the LHC Tier-1s. The Laboratory must deal with both scaling and wide-area distribution challenges in processing its CMS data. Fortunately, evolving technologies in the form of 100Gigabit ethernet, multi-core architectures, and GPU processing provide tools to help meet these challenges.

Current Fermilab R&D efforts in these areas include optimization of network I/O handling in multi-core systems, modification of middleware to improve application performance in 100GE network environments, and network path reconfiguration and analysis for effective use of high bandwidth networks. This poster describes the ongoing network-related R&D activities at Fermilab as a mosaic of efforts that combine to facilitate big data processing and movement.

Collective Tuning: Novel Extensible Methodology, Framework and Public Repository to Collaboratively Address Exascale Challenges

Grigori Fursin (INRIA)

Designing and optimizing novel computing systems became intolerably complex, ad-hoc, costly and error prone due to an unprecedented number of available tuning choices and complex interactions between all software and hardware components. In this poster, we present a novel methodology, extensible infrastructure and public repository to overcome the rising complexity of computer systems by distributing their characterization and optimization among multiple users. Our technology effectively combines auto-tuning, run-time adaptation, data mining and predictive modeling to collaboratively analyze thousands of codelets and datasets, explore large optimization spaces and detect abnormal behavior. It extrapolates collected knowledge to suggest program optimizations, run-time adaptation scenarios or architecture designs to balance performance, power consumption and other characteristics. This technology has been recently successfully validated and extended in several academic and industrial projects with NCAR, Intel Exascale Lab, Google, IBM and CAPS Enterprise, and we believe that it will be vital for developing future Exascale systems.

High-Speed Decision Making on Live Petabyte Data Streams

Jim Kowalkowski, Kurt Biery, Chris Green, Marc Paterno, Rob Roser, William F. Badgett Jr. (Fermi National Laboratory)

High Energy Physics has a long history of coping with cutting-edge data rates in its efforts to extract meaning from experimental data. The quantity of data from planned future experiments that must be analyzed practically in real-time to enable efficient filtering and storage of the scientifically interesting data has driven the development of sophisticated techniques which leverage technologies such as MPI, OpenMP and Intel TBB. We show the evolution of data collection, triggering and filtering from the 1990s with Tevatron experiments into the future of Intensity Frontier and Cosmic Frontier experiments and show how the requirements of upcoming experiments lead us to the development of high-performance streaming triggerless DAQ systems.

Gossip-Based Distributed Matrix Computations

Hana Strakova, Wilfried N. Gansterer (University of Vienna)
Two gossip-based algorithms for loosely coupled distributed networks with potentially unreliable components are discussed – a distributed QR factorization algorithm and a distributed eigensolver. Due to their randomized communication restricted only to direct neighbors they are very flexible. They can operate on arbitrary topologies and they can be made resilient against dynamic changes in the network, against message loss or node failures, and against asynchrony between compute nodes. Moreover, their overall cost can be reduced by accuracy-communication tradeoffs. On this poster, first results about the two algorithms with respect to numerical accuracy, convergence speed, communication cost and resilience against message loss or node failures are reviewed and extended, and they are compared to the state-of-the-art. Due to the growth in the number of nodes for future extreme-scale HPC systems and the anticipated decrease in reliability, some properties of gossip-based distributed algorithms are expected to become very important in the future.

Scalable Fast Multipole Methods for Vortex Element Methods

Qi Hu, Nail A. Gumerov (University of Maryland), Rio Yokota (King Abdullah University of Science & Technology), Lorena Barba (Boston University), Ramani Duraiswami (University of Maryland)

We use a particle-based method to simulate incompressible flows, where the Fast Multipole Method (FMM) is used to accelerate the calculation of particle interactions. The most time-consuming kernels—the Biot-Savart equation and stretching term of the vorticity equation—are mathematically reformulated so that only two Laplace scalar potentials are used instead of six, while automatically ensuring divergence-free far-field computation. Based on this formulation, and on our previous work for a scalar heterogeneous FMM algorithm, we develop a new FMM-based vortex method capable of simulating general flows including turbulence on heterogeneous architectures. Our work for this poster focuses on the computation perspective and our implementation can perform one time step of the velocity+stretching for one billion particles on 32 nodes in 55.9 seconds, which yields 49.12 Tflop/s.

PLFS/HDFS: HPC Applications on Cloud Storage

Chuck Cranor, Milo Polte, Garth Gibson (Carnegie Mellon University)

Long running large-scale HPC applications protect themselves from failures by periodically checkpointing their state to a single file stored in a distributed network filesystem. These filesystems commonly provide a POSIX-style interface for reading and writing files. HDFS is a filesystem used in cloud computing by Apache Hadoop. HDFS is optimized for Hadoop jobs that do not require full POSIX I/O semantics. Only one process may write to an HDFS file, and all writes are appends.

Our work enables multiple HPC processes to checkpoint their state into an HDFS file using PLFS. PLFS is a middleware filesystem that converts random I/O into log-based I/O. We added a new I/O store layer to PLFS that allows it to use non-POSIX filesystems like HDFS as backing store. HPC applications can now checkpoint to HDFS, allowing HPC and cloud to share the same storage systems and work with each other's data.

High Performance GPU Accelerated TSP Solver

Kamil Rocki, Reiji Suda (University of Tokyo)

We are presenting a high performance GPU accelerated implementation of the Iterated Local Search algorithm using 2-opt local search to solve the Traveling Salesman Problem (TSP). GPU usage greatly decreases the time needed to optimize the route and requires a well-tuned implementation. Our results show that at least 90% of the time during Iterated Local Search is spent on the local search; therefore, GPU is used to accelerate this part of the algorithm. The main contribution of this work is the problem division scheme which allows us to solve arbitrarily big problem instances using GPU. We tested our algorithm using different TSPLIB instances on a GTX 680 GPU, and we achieved very high performance of over 700 GFLOPS during calculation of the distances. Compared to the CPU implementation, GPU is able to perform local optimization approximately 150 times faster allowing us to solve very large problem instances on a single machine.

Speeding-Up Memory Intensive Applications Through Adaptive Hardware Accelerators

Vito Giovanni Castellana, Fabrizio Ferrandi (Politecnico di Milano, Dip. di Elettronica e Informazione)

Heterogeneous architectures are becoming an increasingly relevant component for HPC: they combine the computational power of multi-core processors with the flexibility of reconfigurable co-processor boards. Such boards are often composed of a set of standard Field Programmable Gate Arrays (FPGAs), coupled with a distributed memory architecture. This allows the concurrent execution of memory accesses. Nevertheless, since the execution latency of these operations may be unknown at compile time, the synthesis of such parallelizing accelerators becomes a complex task. In fact, standard approaches require the construction of Finite State Machines whose complexity, in terms of number of states and transitions, increases exponentially with respect to the number of unbounded operations that may execute concurrently. We propose an adaptive architecture for such accelerators which overcome this limitation, while exploiting the available parallelism. The proposed design methodology is compared with FSM-based approaches by means of a motivational example.

FusedOS: A Hybrid Approach to Exascale Operating Systems

Yoonho Park, Eric Van Hensbergen, Marius Hillenbrand, Todd Inglett, Bryan Rosenburg, Kyung Ryu (IBM), Robert Wisniewski (Intel Corporation)

Historically, both Light-Weight and Full-Weight Kernel (LWK and FWK) approaches have been taken in providing an operating environment for HPC. The implications of these approaches impact the mechanisms providing resource management for heterogeneous cores on a single chip. Rather than starting with an LWK or an FWK, we combined the two into an operating environment called “FusedOS.” In particular, we combine Linux and CNK environments on a single node by providing an infrastructure capable of partitioning resources of a manycore heterogeneous system. Our contributions are threefold. We present an architectural description of a novel manner for combining an FWK and LWK, retaining the best of each showing we can manage cores a traditional kernel cannot. We describe a prototype environment running on current hardware. We present performance results demonstrating low noise, and show micro-benchmarks running with performance commensurate with CNK.

Using Provenance to Visualize Data from Large-Scale Experiments

Felipe Horta, Jonas Dias, Kary Ocaña, Daniel Oliveira (Federal University of Rio de Janeiro), Eduardo Ogasawara (Federal Centers of Technological Education - Rio de Janeiro), Marta Mattoso (Federal University of Rio de Janeiro)

Large-scale scientific computations are often organized as a composition of many computational tasks linked through data flow. The data that flows along this many-task computing often moves from a desktop to a high-performance environment and to a visualization environment. Keeping track of this data flow is a challenge to provenance support in high-performance Scientific Workflow Management Systems. After the completion of a computational scientific experiment, a scientist has to manually select and analyze its staged-out data, for instance, by checking inputs and outputs along computational tasks that were part of the experiment. In this paper, we present a provenance management system that describes the production and consumption relationships between data artifacts, such as files, and the computational tasks that compose the experiment. We propose a query interface that allows for scientists to browse provenance data and select the output they want to visualize using browsers or a high-resolution tiled display.

Cascaded TCP: Big Throughput for Big Data Applications in Distributed HPC

Umar Kalim, Mark Gardner, Eric Brown, Wu-chun Feng (Virginia Tech)

Saturating high capacity and high latency paths is a challenge with vanilla TCP implementations. This is primarily due to congestion-control algorithms which adapt window sizes when acknowledgements are received. With large latencies, the congestion-control algorithms have to wait longer to respond to network conditions (e.g., congestion), and thus result in less aggregate throughput. We argue that throughput can be improved if we reduce the impact of large end-to-end latencies by introducing layer-4 relays along the path. Such relays would enable a cascade of TCP connections, each with lower latency, resulting in better aggregate throughput. This would directly benefit typical applications as well as Big Data applications in distributed HPC. We present empirical results supporting our hypothesis.

Automatically Adapting Programs for Mixed-Precision Floating-Point Computation

Michael O. Lam (University of Maryland), Bronis R. de Supinski (Lawrence Livermore National Laboratory), Matthew P. LeGendre (Lawrence Livermore National Laboratory), Jeffrey K. Hollingsworth (University of Maryland)

As scientific computation continues to scale, it is crucial to use floating-point arithmetic processors as efficiently as possible. Lower precision allows streaming architectures to perform more operations per second and can reduce memory bandwidth pressure on all architectures. However, using a precision that is too low for a given algorithm and data set will result in inaccurate results. In this poster, we present a framework that uses binary instrumentation and modification to build mixed-precision configurations of existing binaries that were originally developed to use only double-precision. This allows developers to easily experiment with mixed-precision configurations without modifying their source code, and it permits auto-tuning of floating-point precision. We also implemented a simple search algorithm to automatically identify which code regions can use lower precision. We include results for several benchmarks that show both the efficacy and overhead of our tool.

MAAPED: A Predictive Dynamic Analysis Tool for MPI Applications

Subodh Sharma (University of Oxford), Ganesh Gopalakrishnan (University of Utah), Greg Bronevetsky (Lawrence Livermore National Laboratory)

Formal dynamic analysis of MPI programs is critically important since conventional testing tools for message passing programs don't cover the space of possible non-deterministic communication matches, thus may miss bugs in the unexam-

ined execution scenarios. While modern dynamic verification techniques guarantee the coverage of non-deterministic communication matches, they do so indiscriminately, inviting exponential interleaving explosion. Though the general problem is difficult to solve, we show that a specialized dynamic analysis method can be developed for dramatically reducing the number of interleavings when looking for certain safety properties such as deadlocks. Our MAAPED (Messaging Application Analysis with Predictive Error Discovery) tool collects a single program trace and predicts deadlock presence in other (unexplored) traces of an MPI program. MAAPED hinges on initially computing the potential alternate matches for non-deterministic communication operations and then analyzes such matches which may lead to a deadlock. The results collected are encouraging.

Memory and Parallelism Exploration using the LULESH Proxy Application

Ian Karlin, Jim McGraw (Lawrence Livermore National Laboratory), Esthela Gallardo (University of Texas at El Paso), Jeff Keasler, Edgar A. Leon, Bert Still (Lawrence Livermore National Laboratory)

Current and planned computer systems present challenges for scientific programming. Memory capacity and bandwidth are limiting performance as floating point capability increases due to more cores per processor and wider vector units. Effectively using hardware requires finding greater parallelism in programs while using relatively less memory. In this poster, we present how we tuned the Livermore Unstructured Lagrange Explicit Shock Hydrodynamics proxy application for on-node performance resulting in 62% fewer memory reads, a 19% smaller memory footprint, 770% more floating point operations vectorizing and less than 0.1% serial section runtime. Tests show serial runtime decreases of up to 57% and parallel runtime reductions of up to 75%. We are also applying these optimizations to GPUs and a subset of ALE3D, from which the proxy application was derived. So far we achieve up to a 1.9x speedup on GPUs, and a 13% runtime reduction in the application for the same problem.

Auto-Tuning of Parallel IO Parameters for HDF5 Applications

Babak Behzad (University of Illinois at Urbana-Champaign), Joey Huchette (Lawrence Berkeley National Laboratory), Huong Luu (University of Illinois at Urbana-Champaign), Ruth Aydt, Quincey Koziol (HDF Group), Mr Prabhat, Suren Byna (Lawrence Berkeley National Laboratory), Mohamad Chaarawi (HDF Group), Yushu Yao (Lawrence Berkeley National Laboratory)

I/O is often a limiting factor for HPC applications. Although well-tuned codes have shown good I/O throughput compared to the theoretical maximum, the majority of applications use default parallel I/O parameter values and achieve poor

performance. We have built an extensible framework for benchmark-guided auto-tuning of HDF5, MPI-IO, and Lustre parameters. The framework includes three main components. H5AutoTuner uses a control file to adjust I/O parameters without changing or recompiling the application. H5Perf-Capture records performance metrics for HDF5 and MPI-IO. H5Evolve uses genetic algorithms to explore the parameter search space until well-performing values are identified. Early results for three HDF5 application-based I/O benchmarks on two different HPC systems have shown 1.3x–6.8x speedup using auto-tuned parameters compared to default values. Our auto-tuning framework can improve I/O performance without hands-on optimization and also provides a general platform for exploring parallel I/O behavior. The printed poster details framework architecture and experimental results.

Uintah Framework Hybrid Task-Based Parallelism Algorithm

Qingyu Meng, Martin Berzins (University of Utah)

Uintah is a software framework that provides an environment for solving large-scale science and engineering problems involving the solution of partial differential equations. Uintah uses a combination of fluid-flow solvers and particle-based methods for solids, together with adaptive meshing and asynchronous task-based approach with automated load balancing. When applying Uintah to fluid-structure interaction problems, the combination of adaptive meshing and the movement of structures through space present a formidable challenge in terms of achieving scalability on large-scale parallel computers. Adopting a model that uses MPI to communicate between nodes and a shared memory model on-node is one approach to achieve scalability on large-scale systems. This scalability challenge is addressed here for Uintah, by the development of new hybrid runtime and scheduling algorithms combined with novel lock-free data structures, making it possible for Uintah to achieve excellent scalability for a challenging fluid-structure problem with mesh refinement on as many as 256K cores.

Programming Model Extensions for Resilience in Extreme Scale Computing

Saurabh Hukerikar, Pedro C. Diniz, Robert F. Lucas (University of Southern California)

System resilience is a key challenge to building extreme scale systems. A large number of HPC applications are inherently resilient, but application programmers lack mechanisms to convey their fault tolerance knowledge to the system. We present a cross-layer approach to resilience in which we propose a set of programming model extensions and develop a runtime inference framework that can reason about the context and significance of faults, as they occur, to the application programmer's fault tolerance expectations. We demonstrate using a set accelerated fault injection experiments the validity of our approach with a set of real scientific and engineering

codes. Our experiments show that a cross-layer approach that explicitly engages the programmer in expressing fault tolerance knowledge which is then leveraged across the layers of system abstraction can significantly improve the dependability of long running HPC applications.

Seismic Imaging on Blue Gene/Q

Ligang Lu, James Sexton, Michael Perrone, Karen Magerlein, Robert Walkup (IBM T.J. Watson Research Center)

Blue Gene/Q (BG/Q) is an early representative of increasing scale and thread count that will characterize future HPC systems. This work helps to address two questions important to future HPC system development: how HPC systems with high levels of scale and thread count will perform in applications; and how systems with many degrees of freedom in choosing the number of nodes, cores, and threads can be calibrated to achieve optimal performance. Our investigation of Reverse Time Migration (RTM) seismic imaging on BG/Q helps to answer such questions and provides an example of how HPC systems like BG/Q can accelerate applications to unprecedented levels of performance. Our analyses of various levels and aspects of optimization also provide valuable experience and insights into how BG/Q's architecture and hardware features can be utilized to facilitate the advance of seismic imaging technologies. Our BG/Q RTM solution achieved a 14.93x speedup over the BG/P implementation.

Using Business Workflows to Improve Quality of Experiments in Distributed Systems Research

Tomasz Buchert (INRIA), Lucas Nussbaum (Lorraine Research Laboratory in Computer Science and its Applications)

Distributed systems pose many difficult problems to researchers. Due to their large-scale complexity, their numerous constituents (e.g., computing nodes, network links) tend to fail in unpredictable ways. This particular fragility of experiment execution threatens reproducibility, often considered to be a foundation of experimental science. We present a new approach to description and execution of experiments involving large-scale computer installations. The main idea consists in describing the experiment as workflow and using achievements of Business Workflow Management to reliably and efficiently execute it. Moreover, to facilitate the design process, the framework provides abstractions that hide unnecessary complexity from the user. The implementation of an experiment engine that fulfills our goals is underway. During the poster presentation I will discuss motivation for my work and explain why we chose our approach. Moreover, I will show demonstration of currently developed experiment engine.

Build to Order Linear Algebra Kernels

Christopher Gropp (Rose-Hulman Institute of Technology), Geoffrey Belter, Elizabeth Jessup, Thomas Nelson (University of Colorado Boulder), Boyana Norris (Argonne National Laboratory), Jeremy Siek (University of Colorado Boulder)

Tuning linear algebra kernels for specific machines is a time-consuming and difficult process. The Build-to-Order (BTO) compiler takes MATLAB-like kernels and generates fully tuned C code for the machine on which it runs. BTO allows users to create optimized linear algebra kernels customized to their needs and machines. Previous work has shown that BTO kernels perform competitively with or better than hand-tuned code. We now test BTO on a full application. We have selected bidiagonalization, a non-trivial matrix operation useful in computing singular value decompositions. We use BTO to tune the core kernels of the algorithm, rather than the entire application. This poster shows the comparative performance of four implementations of bidiagonalization: the LAPACK routine DGBD2, a BLAS 2.5 based algorithm developed by Howell et al., and those same making use of BTO kernels.

Distributed Metadata Management for Exascale Parallel File System

Keiji Yamamoto, Atsushi Hori (RIKEN), Yutaka Ishikawa (University of Tokyo)

In today's supercomputer, most file systems provide scalable I/O bandwidth. But these systems cannot concurrent creation of millions to billions of files in a single directory. Toward exascale-era, we propose scalable metadata management method by using many metadata servers. Directory entry and inode are distributed by using consistent hashing. Directory entry and inode are stored as same MDS as possible. In many case, two requests of lookup dentry and inode are merged into one request. In evaluation, our file system throughput is more than 490,000 file creates per second on 64 metadata servers and 2048 clients. Concurrent file creation time of 40 million is 42 second in a single directory.

Advances in Gyrokinetic Particle-in-Cell Simulation for Fusion Plasmas to Extreme Scale

Bei Wang (Princeton University), Stephane Eithier (Princeton Plasma Physics Laboratory), William Tang (Princeton University), Khaled Ibrahim, Kamesh Madduri, Sam Williams (Lawrence Berkeley National Laboratory)

The Gyrokinetic Particle-in-cell (PIC) method has been successfully applied in studies of low-frequency microturbulence in magnetic fusion plasmas. While the excellent scaling of PIC codes on modern computing platforms is well established, significant challenges remain in achieving high on-chip concurrency for the new path to exascale systems. In addressing associated issues, it is necessary to deal with the basic gather-scatter operation and the relatively low computational

intensity in the PIC method. Significant advancements have been achieved in optimizing gather-scatter operations in the gyrokinetic PIC method for next-generation multi-core CPU and GPU architectures. In particular, we will report on new techniques that improve locality, reduce memory conflict, and efficiently utilize shared memory on GPU's. Performance benchmarks on two high-end computing platforms—the IBM BlueGene/Q (Mira) system at the Argonne Leadership Computing Facility and the Cray XK6 (Titan Dev) with the latest GPU at Oak Ridge Leadership Computing Facility—will be presented.

The Hashed Oct-Tree N-Body Algorithm at a Petaflop

*Michael S. Warren (Los Alamos National Laboratory),
Ben Bergen (Los Alamos National Laboratory)*

Cosmological simulations are the cornerstone of theoretical analysis of large-scale structure. During the next few years, observational projects will measure the spatial distribution of large-scale structure in enormous volumes of space across billions of years of cosmic evolution. Advances in modeling must keep pace with observational advances if we are to understand the Universe which led to these observations. We have recently demonstrated our hashed octree N-body code (HOT) scaling to 256k processors on Jaguar at Oak Ridge National Laboratory with a performance of 1.79 Petaflops (single precision) using 2 trillion particles. We have additionally performed preliminary studies with NVIDIA Fermi GPUs, achieving single GPU performance on our hexadecapole inner loop of 1 Tflop (single precision) and application performance speedup of 2x by offloading the most computationally intensive part of the code to the GPU.

Asynchronous Computing for Partial Differential Equations at Extreme Scales

Aditya Konduri, Diego A. Donzis (Texas A&M University)

Advances in computing technology have made numerical simulations an indispensable research tool in the pursuit of understanding real life problems. Due to their complexity, these simulations demand massive computations with extreme levels of parallelism. At extreme scales, communication between processors could take up a substantial amount of time. This results in substantial waste in computing cycles, as processors remain idle for most of the time. We investigate a novel approach based on widely used finite-difference schemes in which computations are carried out in an asynchronous fashion—synchronization among cores is not enforced and computations proceed regardless of the status of messages. This drastically reduces idle times resulting in much larger computation rates and scalability. However, stability, consistency and accuracy have to be shown in order for these schemes to be viable. This is done through mathematical theory and numerical simulations. Results are used to design new numerical schemes robust to asynchronicity.

GPU Accelerated Ultrasonic Tomography Using Propagation and Back Propagation Method

**Pedro Bello Maldonado (Florida International University),
Yuanwei Jin (University of Maryland Eastern Shore), Enyue Lu (Salisbury University)**

This paper develops implementation strategy and method to accelerate the propagation and back propagation (PBP) tomographic imaging algorithm using Graphic Processing Units (GPUs). The Compute Unified Device Architecture (CUDA) programming model is used to develop our parallelized algorithm since the CUDA model allows the user to interact with the GPU resources more efficiently than traditional Shader methods. The results show an improvement of more than 80x when compared to the C/C++ version of the algorithm, and 515x when compared to the MATLAB version while achieving high quality imaging for both cases. We test different CUDA kernel configurations in order to measure changes in the processing-time of our application. By examining the acceleration rate and the image quality, we develop an optimal kernel configuration that maximizes the throughput of CUDA implementation for the PBP method.

Application Restructuring for Vectorization and Parallelization: A Case Study

Karthik Raj Saanthalingam, David Hudak, John Eisenlohr (Ohio Supercomputer Center), P. Sadayappan (Ohio State University)

Clock rates remain flat while transistor density increases, so microprocessor designers are providing more parallelism on a chip by increasing vector length and core count. For example, the Intel Westmere architecture has a vector length of four floats (128 bits) and six cores compared to eight floats (256 bits) and eight cores on the Intel Sandy Bridge. Applications must get good vector and shared-memory performance in order to leverage these hardware advances. Dissipative Particle Dynamics (DPD) is analogous to traditional molecular dynamics techniques applied to mesoscale simulations. We analyzed and restructured an existing DPD implementation to improve vector and OpenMP performance for the Intel Xeon and MIC architectures. We designed an efficient partitioned global address space (PGAS) implementation using the Global Arrays Toolkit using this experience. We present performance results on representative architectures.

Parallel Algorithms for Counting Triangles and Computing Clustering Coefficients

S. M. Arifuzzaman, Maleq Khan, Madhav V. Marathe (Virginia Tech)

We present MPI-based parallel algorithms for counting triangles and computing clustering coefficients in massive networks. Counting triangles is important in the analysis of various networks, e.g., social, biological, web etc. Emerging

massive networks do not fit in the main memory of a single machine and are very challenging to work with. Our distributed-memory parallel algorithm allows us to deal with such massive networks in a time- and space-efficient manner. We were able to count triangles in a graph with 2 billions of nodes and 50 billions of edges in 10 minutes. Our parallel algorithm for computing clustering coefficients uses efficient external memory aggregation. We also show how edge sparsification technique can be used with our parallel algorithm to find approximate number of triangles without sacrificing the accuracy of estimation. In addition, we propose a simple modification of a state-of-the-art sequential algorithm that improves both runtime and space requirement.

Improved OpenCL Programmability with clUtil

Rick Weber, Gregory D. Peterson (University of Tennessee, Knoxville)

CUDA was the first GPGPU programming environment to achieve widespread adoption and interest. This API owes much of its success to its highly productive abstraction model while still exposing enough hardware details to achieve high performance. OpenCL sacrifices much of the programmability in its front-end API for portability; while a less productive API than CUDA, it supports many more devices. In this poster, we present clUtil, which aims to reunite OpenCL's portability and CUDA's ease of use via C++11 language features. Furthermore, clUtil supports high-level parallelism motifs, namely a parallel-for loop that can automatically load balance applications onto heterogeneous OpenCL devices.

Hadoop's Adolescence: A Comparative Workload Analysis from Three Research Clusters

Kai Ren, Garth Gibson (Carnegie Mellon University), YongChul Kwon, Magdalena Balazinska, Bill Howe (University of Washington)

We analyze Hadoop workloads from three different research clusters from an application-level perspective, with two goals: (1) explore new issues in application patterns and user behavior and (2) understand key performance challenges related to IO. Our analysis suggests that Hadoop usage is still in its adolescence. We see underuse of Hadoop features, extensions, and tools as well as significant opportunities for optimization. We see significant diversity in application styles, including some "interactive" workloads, motivating new tools in the ecosystem. We find that some conventional approaches to improving performance are not especially effective and suggest some alternatives. Overall, we find significant opportunity for simplifying the use and optimization of Hadoop.

Preliminary Report for a High Precision Distributed Memory Parallel Eigenvalue Solver

Toshiyuki Imamura (RIKEN), Susumu Yamada, Masahiko Machida (Japan Atomic Energy Agency)

This study covers design and implementation of a DD (double-double) extended parallel eigenvalue solver, namely DD-Eigen. We extended most of underlying numerical software layers from BLAS, LAPACK, and ScaLAPACK as well as MPI. Preliminary results shows that DD-Eigen performs on several platforms, and shows good accuracy and parallel efficiency. We can conclude that the DD format is reasonable data format instead of real(16) format from the viewpoint of programming and performance.

Analyzing Patterns in Large-Scale Graphs Using MapReduce in Hadoop

Joshua Schultz (Salisbury University), Jonathan Vierya (California State Polytechnic University, Pomona), Enyue Lu (Salisbury University)

Analyzing patterns in large-scale graphs, such as social networks (e.g. Facebook, LinkedIn, Twitter) has many applications including community identification, blog analysis, intrusion and spamming detections. Currently, it is impossible to process information in large-scale graphs with millions even billions of edges with a single computer. In this paper, we take advantage of MapReduce, a programming model for processing large datasets, to detect important graph patterns using open source Hadoop on Amazon EC2. The aim of this paper is to show how MapReduce cloud computing with the application of graph pattern detection scales on real world data. We implement Cohen's MapReduce graph algorithms to enumerate patterns including triangles, rectangles, trusses and barycentric clusters using real world data taken from Snap Stanford. In addition, we create a visualization algorithm to visualize the detected graph patterns. The performance of MapReduce graph algorithms has been discussed too.

Digitization and Search: A Non-Traditional Use of HPC

Liana Diesendruck, Luigi Marini, Rob Kooper, Mayank Kejriwal, Kenton McHenry (University of Illinois at Urbana-Champaign)

We describe our efforts to provide a form of automated search of handwritten content for digitized document archives. To carry out the search we use a computer vision technique called word spotting. A form of content-based image retrieval, it avoids the still difficult task of directly recognizing text by allowing a user to search using a query image containing handwritten text and ranking a database of images in terms of those that contain more similar looking content. In order to make this search capability available on an archive three computationally expensive pre-processing steps are required. We augment this automated portion of the process with a passive

crowd sourcing element that mines queries from the systems users in order to then improve the results of future queries. We benchmark the proposed framework on 1930s Census data, a collection of roughly 3.6 million forms and 7 billion individual units of information.

An Exascale Workload Study

Rinku Gupta, Prasanna Balaprakash, Darius Buntinas, Anthony Chan (Argonne National Laboratory), Apala Guha (University of Chicago), Sri Hari Krishna Narayanan (Argonne National Laboratory), Andrew Chien (University of Chicago), Paul Hovland, Boyana Norris (Argonne National Laboratory)

While Amdahl's 90-10 approach has been used in the past to drive supercomputing speedup in the last few decades, increasing heterogeneous architectures combined with power and energy limitations dictates a need for a new paradigm. In this poster, we describe our 10x10 paradigm, which identifies top ten distinct dominant characteristics in a set of applications. One could then exploit customized architectures (accelerators), best suited to optimize each dominant characteristic. Every application will typically have multiple characteristics and thus will use several customized accelerators/tools during its various execution phases. The goal is to ensure that the application runs efficiently and that the architecture is used in an energy-efficient manner. In this poster, we describe our efforts in three directions (1) understanding applications characterization, (2) developing extrapolation statistical models to understand application characteristics at exascale level, (3) evaluating extrapolated applications with technologies that might potentially be available during the exascale era.

Visualization for High-Resolution Ocean General Circulation Model via Multi-Dimensional Transfer Function and Multi-variate Analysis

Daisuke Matsuoka, Fumiaki Araki, Shinichiro Kida, Hideharu Sasaki, Bunmei Taguchi (Japan Agency for Marine-Earth Science and Technology)

Ocean currents and vortices play an important role in transferring heat, salt or carbon as well as atmospheric circulation. With advances in supercomputing technology, high-resolution large-scale simulation study has been focused in the field of ocean science. However, it is difficult to intuitively understand characteristic features defined as multivariable hiding in the high-resolution dataset. In order to obtain scientific knowledge from large-scale simulation data, it is important to effectively extract and to efficiently express the characteristic feature. The aim of this study is how to efficiently extract and how to effectively visualize ocean currents which affect the heat transportation. In this research, new multi-dimensional transfer function to emphasis the ocean currents and vortices is proposed. Furthermore, multivariate analyses to extract such features are developed. This presentation describes the

methodologies and experimental results of these methods. Evaluation of visualization results and feedback to the parameter optimization will be also reported.

Portals 4 Network Programming Interface

Brian Barrett, Ron Brightwell (Sandia National Laboratories), Keith Underwood (Intel Corporation), K. Scott Hemmert (Sandia National Laboratories)

Portals 4 is an advanced network programming interface which allows for the development of a rich set of upper layer protocols. By careful selection of interfaces and strong progress guarantees, Portals 4 is able to support multiple protocols without significant overhead. Recent developments with Portals 4, including development of MPI, SHMEM, and GASNet protocols are discussed.

Quantum Mechanical Simulations of Crystalline Helium Using High Performance Architectures

David D. Jenkins, Robert J. Hinde, Gregory D. Peterson (University of Tennessee, Knoxville)

With the rapid growth of emerging high performance architectures comes the ability for the acceleration of computational science applications. In this work, we present our approach to accelerating a Quantum Monte Carlo method called Variational Path Integral. Using many microprocessors and graphics processing units, this VPI implementation simulates the interactions of helium atoms in a crystalized structure at near zero temperature. This work uses an improved master-worker approach to increase scalability from tens to thousands of cores on the Kraken supercomputer. A single node of the Keeneland GPU cluster delivers performance equivalent to ten nodes of Kraken. High performance computing enables us to simulate larger crystals and many more simulations than were previously possible.

Multiple Pairwise Sequence Alignments with the Needleman-Wunsch Algorithm on GPU

Da Li, Michela Becchi (University of Missouri)

Pairwise sequence alignment is a method used in bioinformatics to determine the similarity between DNA, RNA and protein sequences. The Needleman-Wunsch algorithm is typically used to perform global alignment, and has been accelerated on Graphics Processing Units (GPUs) on single pairs of sequences. Many applications require multiple pairwise comparisons over sets of sequences. The large sizes of modern bioinformatics datasets leads to a need for efficient tools that allow a large number of pairwise comparisons. Because of their massive parallelism, GPUs are an appealing choice for accelerating these computations. In this paper, we propose an efficient GPU implementation of multiple pairwise sequence alignments based on the Needleman-Wunsch algorithm. Compared

to a well-known existing solution, our implementation improves the memory transfer time by a factor 2X, and achieves a ~3X speedup in kernel execution time.

GenASiS: An Object-Oriented Approach to High Performance Multiphysics Code with Fortran 2003

Reuben D. Budiardja (National Institute for Computational Sciences), Christian Y. Cardall, Eirik Endeve, Anthony Mezzacappa (Oak Ridge National Laboratory)

Many problems in astrophysics and cosmology are multiphysics and multiscale in nature. For problems with multiphysics components, the challenges facing the development of complicated simulation codes can be ameliorated by the principles of object-oriented design. GenASiS is a new code being developed to face these challenges from the ground up. Its object-oriented design and approach are accomplished with features of Fortran 2003 that support the object-oriented paradigm and can do so without sacrificing performance. Its initial primary target, although not exclusively, is the simulation of core-collapse supernovae on the world's leading capability supercomputers. We present an overview of GenASiS architecture, including its cell-by-cell refinement with multilevel mesh and object-oriented approach with Fortran 2003. We demonstrate its initial capabilities and solvers and show its scalability on the massively parallel supercomputer.

Exploring Design Space of a 3D Stacked Vector Cache

Ryusuke EGAWA, Yusuke Endo (Tohoku University), Jubee Tada (Yamagata University), Hiroyuki Takizawa, Hiroaki Kobayashi (Tohoku University)

This paper explores and presents a design method of a 3D integrated memory system using conventional EDA tools. In addition, to clarify the potential of TSVs, delay and power consumption of TSVs are quantitatively evaluated, and are compared with those of conventional 2D wires under various CMOS process technologies. The main contributions of this paper are: (1) clarifying the potential of TSVs based on the SPICE compatible simulations; (2) exploring the design methodology of a 3D integrated memory system using conventional EDA tools; and (3) quantitatively comparing 3D integrated cache memories with 2D ones.

A Disc-Based Decomposition Algorithm with Optimal Load Balancing for N-body Simulations

Akila Gothandaraman (University of Pittsburgh), Lee Warren (College of New Rochelle), Thomas Nason (University of Pittsburgh)

We propose a novel disc data decomposition algorithm for N-body simulations and compare its performance against a cyclic decomposition algorithm. We implement the data decomposition algorithms towards the calculation of three-body interactions in the Stillinger-Weber potential for a system of water molecules. The performance is studied in terms of load

balance and speedup from the MPI implementations of the two algorithms. We are also currently working on a performance study of the disc decomposition algorithm on graphics processing units (GPUs).

Remote Visualization for Large-Scale Simulation using Particle-Based Volume Rendering

Takuma Kawamura, Yasuhiro Idomura, Hiroko Miyamura, Hiroshi Takemiya (Japan Atomic Energy Agency)

With the recent development of supercomputers, it is required to efficiently visualize the result of super-large scale numerical simulations on a few hundreds to a few ten thousands of parallel processes. Conventional offline-processing of visualization give rise to difficult challenges such as transferring large-scale data and reassembly of extensive amount of computational result files, which is inevitable for sort-first or sort-last visualization methods. On the other hand, interactive use of a supercomputer is still limited. We propose a remote visualization system for solving these problems by applying Particle-based Volume Rendering which is sortless volume rendering technique and convert resulting data to rendering primitive particles. Proposed system can generate particles on same number of parallel processes of the numerical simulation and transfer the particles to a visualization server. Proposed system allows users can observe the resulting data with changing the camera position freely.

Tracking and Visualizing Evolution of the Universe: In Situ Parallel Dark Matter Halo Merger Trees

Jay Takle (Rutgers University), Katrin Heitmann, Tom Peterka (Argonne National Laboratory), Deborah Silver (Rutgers University), George Zagaris (Kitware, Inc.), Salman Habib (Argonne National Laboratory)

We present a framework to study the behavior and properties of cosmological structures called dark matter halos. As part of the framework, we build an evolution history, called halo merger trees, which follow the evolution of the halos over time. The entire process from tracking to building the merger tree is performed in parallel and in situ with the underlying cosmological N-body simulation. We are currently in the phase of parallelizing this process for merger tree analysis that will reside in situ in a cosmology application at the scale of millions of processes and is further integrated with a production visualization tool as part of an end-to-end computation/analysis/visualization pipeline. In the poster session we would like to explain the importance of implementing parallel in situ analyses along with effective visualization for cosmological studies and discovery.

Autonomic Modeling of Data-Driven Application Behavior

Steena D.S. Monteiro, Greg Bronevetsky, Marc Casas-Guix
(Lawrence Livermore National Laboratory)

Computational behavior of large-scale data driven applications are complex functions of their input, various configuration settings, and underlying system architecture. The resulting difficulty in predicting their behavior complicates optimization of their performance and scheduling them onto compute resources. Manually diagnosing performance problems and reconfiguring resource settings to improve performance is cumbersome and inefficient. We thus need autonomic optimization techniques that observe the application, learn from the observations, and subsequently successfully predict its behavior across different systems and load scenarios. This work presents a modular modeling approach for complex data-driven applications that uses statistical techniques to capture pertinent characteristics of input data, dynamic application behaviors and system properties to predict application behavior with minimum human intervention. The work demonstrates how to adaptively structure and configure the model based on the observed complexity of application behavior in different input and execution contexts.

Automated Mapping Streaming Applications onto GPUs

Andrei Hagiescu (National University of Singapore), Huynh Phung Huynh (Singapore Agency for Science, Technology and Research), Abhishek Ray (Nanyang Technological University), Weng-Fai Wong (National University of Singapore), Rick Goh Siow Mong (Singapore Agency for Science, Technology and Research)

Many parallel general purpose applications have been efficiently mapped to GPUs. Unfortunately, many stream processing applications exhibit unfavorable data movement patterns and low computation-to-communication ratio that may lead to poor performance. We describe an automated compilation flow that maps most stream processing applications onto GPUs by taking into consideration two important architectural features of nVidia GPUs, namely interleaved execution as well as the small amount of shared memory available in each streaming multiprocessors. Our scheme goes against the conventional wisdom of GPU programming which is to use a large number of homogeneous threads. Instead, it uses a mix of compute and memory access threads, together with a carefully crafted schedule that exploits parallelism in the streaming application, while maximizing the effectiveness of the memory hierarchy. We have implemented our scheme in the compiler of the StreamIt programming language, and our results show a significant speedup compared to the state-of-the-art solutions.

Planewave-Based First-Principles MD Calculation on 80,000-Node K-computer

Akiyoshi Kuroda, Kazuo Minami (RIKEN), Takahiro Yamasaki (Fujitsu Laboratories Ltd.), Jun Nara (National Institute for Materials Science), Junichiro Koga (ASMS), Tsuyoshi Uda (ASMS), Takahisa Ohno (National Institute for Materials Science)

We show the efficiency of a first-principles electronic structure calculation code, PHASE on the massively-parallel super computer, K, which has 80,000 nodes. This code is based on plane-wave basis set, thus FFT routines are included. We succeeded in parallelization of FFT routines needed in our code by localizing each FFT calculation into a small number of nodes, resulting in the decrease of communication time required for FFT calculation. We also introduced multi-axis parallelization for bands and plane waves and applied BLAS routines by transforming matrix-vector products into matrix-matrix products with the bundle of vectors. As a result, PHASE has very high parallel efficiency. By using this code, we have investigated the structural stability of screw dislocations in silicon carbide that has attracted much attention due to its semiconductor industry importance.

Bringing Task- and Data-Parallelism to Analysis of Climate Model Output

Robert Jacob, Jayesh Krishna, Xiabing Xu, Sheri Mickelson (Argonne National Laboratory), Kara Peterson (Sandia National Laboratories), Michael Wilde (Argonne National Laboratory)

Climate models are both outputting larger and larger amounts of data and are doing it on more sophisticated numerical grids. The tools climate scientists have used to analyze climate output, an essential component of climate modeling, are single threaded and assume rectangular structured grids in their analysis algorithms. We are bringing both task- and data-parallelism to the analysis of climate model output. We have created a new data-parallel library, the Parallel Gridded Analysis Library (ParGAL) which can read in data using parallel I/O, store the data on a compact representation of the structured or unstructured mesh and perform sophisticated analysis on the data in parallel. ParGAL has been used to create a parallel version of a script-based analysis and visualization package. Finally, we have also taken current workflows and employed task-based parallelism to decrease the total execution time.

Evaluating Asynchrony in Gibraltar RAID's GPU Reed-Solomon Coding Library

Xin Zhou, Anthony Skjellum (University of Alabama at Birmingham), Matthew L. Curry (Sandia National Laboratories)

This poster describes the Gibraltar Library, a Reed-Solomon coding library that is part of the Gibraltar software RAID implementation, and recent improvements to its applicability to

small coding tasks. GPUs have been well known for performing computations on large pieces of data with high performance, but some workloads are not able to offer enough data to keep the entire GPU occupied. In this work, we have updated the library to include a scheduler that takes advantage of new GPU features, including multi-kernel launch capability, to improve the performance of several small workloads that cannot individually take advantage of the entire GPU. This poster includes performance data that demonstrates significant improvement in throughput, which can translate directly to improvement in RAID performance for random reads and writes, as well as other non-RAID applications. These improvements will be released at <http://www.cis.uab.edu/hpcl/gibraltar>.

Matrix Decomposition Based Conjugate Gradient Solver for Poisson Equation

Hang Liu, howie Huang (George Washington University),
Jung-Hee Seo, Rajat Mittal (Johns Hopkins University)

Finding a fast solver for Poisson equation is important for many scientific applications. In this work, we develop a matrix decomposition based Conjugate Gradient (CG) solver, which leverages GPU clusters to accelerate the calculation of the Poisson equation. Our experiments show that the new CG solver is highly scalable and achieves significant speedups over a CPU based multi-grid solver.

Evaluating the Error Resilience of GPGPU Applications

Bo Fang, Jiesheng Wei, Karthik Pattabiraman, Matei Ripeanu (University of British Columbia)

GPUs have been originally designed for error-resilient workload. Today, GPUs are used in error-sensitive applications, e.g. General Purpose GPU (GPGPU) applications. The goal of this project is to investigate the error resilience of GPGPU applications and understand their reliability characteristics. To this end, we employ fault injection on real GPU hardware. We find that, compared to CPUs, GPU platforms lead to a higher rate of silent data corruption—a major concern since these errors are not flagged at runtime and often remain latent. We also find that out-of-bound memory accesses are the most critical reason of crashes on GPGPU applications.

Comparing GPU and Increment-Based Checkpoint Compression

Dewan Ibtesham (University of New Mexico), Dorian Arnold (University of New Mexico), Kurt B. Ferreira (Sandia National Laboratories), Ronald Brightwell (Sandia National Laboratories)

The increasing size and complexity of HPC systems have led to major concerns over fault frequencies and the mechanisms necessary to tolerate these faults. Previous studies have shown that state-of-the-field checkpoint/restart mecha-

nisms will not scale sufficiently for future generation systems. Therefore, Checkpoint/restart overheads must be improved to maintain feasibility for future HPC systems. Previously, we showed the effectiveness of checkpoint data compression for reducing checkpoint/restart latencies and storage overheads. In this work we (1) compare CPU-based and GPU-based checkpoint compression, (2) compare to increment-based checkpoint optimization, (3) evaluate the combination of checkpoint compression with incremental checkpointing and (4) motivate future GPU-based compression work by exploring various hypothetical scenarios.

The Magic Determination of the Magic Constants by ttgLib Autotuner

Michail Pritula (ttgLabs, LLC), Maxim Krivov (ttgLabs, LLC),
Sergey Grizan (ttgLabs, LLC), Pavel Ivanov (Lomonosov Moscow State University)

When the program is being optimized for execution on GPU, one has to introduce a lot of performance affected constants that define blocks parameters, data chunks size, parallelism granularity, etc. And the more software is optimized, the more magic constants it introduces. Furthermore, adding multi-GPU system support often requires usage of smart load balancing strategies that considers GPU-specific effects such as potential speed-up from ignoring some accelerators, time volatility of GPU-performance and others. As a result, performance of the target software can be significantly increased just by tuning to the hardware and processing data being used. The authors developed a means of determining the optimal values of these constants called ttgLib autotuner which is capable of monitoring the software at runtime and automatically tuning magic constants as well as performing dynamic load balancing between CPU and multiple GPUs. The performed tests showed the additional speedup upto 50-80% by tuning alone.

MemzNet: Memory-Mapped Zero-copy Network Channel for Moving Large Datasets over 100Gbps Networks

Mehmet Balman (Lawrence Berkeley National Laboratory)

High-bandwidth networks are poised to provide new opportunities in tackling large data challenges in today's scientific applications. However, increasing the bandwidth is not sufficient by itself; we need careful evaluation of future high-bandwidth networks from the applications' perspective. We have experimented with current state-of-the-art data movement tools, and realized that file-centric data transfer protocols do not perform well with managing the transfer of many small files in high-bandwidth networks, even when using parallel streams or concurrent transfers. We require enhancements in current middleware tools to take advantage of future networking frameworks. To improve performance and efficiency, we develop an experimental prototype, called MemzNet (Memory-mapped Zero-copy Network Channel), which uses a

block-based data movement method in moving large scientific datasets. We have implemented MemzNet that takes the approach of aggregating files into blocks and providing dynamic data channel management. In this work, we present our initial results in 100Gbps networks.

Evaluating Communication Performance in Supercomputers BlueGene/Q and Cray XE6

Huy Bui (University of Illinois at Chicago), Venkatram Vishwanath (Argonne National Laboratory), Jason Leigh (University of Illinois at Chicago), Michael E. Papka (Argonne National Laboratory)

Evaluation of the communication performance in supercomputers Blue Gene/Q and Cray XE6 using MPI and lower level libraries are presented. MPI is widely used for to its simplicity, portability and straightforwardness; however, it also introduces overhead that degrades communication performance. Some applications with performance constraints cannot tolerate that degradation. Recent supercomputers such as the Blue Gene/Q and CrayXE provide lower level communications libraries such as PAMI (Parallel Active Message Interface) and uGNI (User Generic Network Interface). Our experiments show that with different communications modes: one-sided, two-sided, inter-node, intra-node, we can achieve higher performance with certain message sizes. These results will enable us to develop a light-weight API for GLEAN which is a framework for I/O acceleration and in situ analysis, to get improved performance on these systems.

Statistical Power and Energy Modeling of Multi-GPU kernels

Sayan Ghosh (University of Houston), Sunita Chandrasekaran (University of Houston), Barbara Chapman (University of Houston)

Current high performance computing systems consume a lot of energy. Although there have been substantial increase in computation performance, the same is not reflected in case of energy efficiency. To have an exascale computer by end of this decade, tremendous improvements in energy efficiency is mandatory. It is not possible to have sophisticated instruments to measure energy or power at such a large scale, but estimation could be useful. In this work, we have developed a statistical model using limited performance counters providing an estimation of power/energy components. The data collected range from different types of application kernel such as FFT, DGEMM, Stencils and Pseudo-Random-Number-Generators; widely used in various disciplines of high performance computing. A power analyzer has been used to analyze/extract the electrical power usage information of the multi-GPU node under inspection. An API was also written to remotely interface with the analyzer and get the instantaneous power readings.

Virtual Machine Packing Algorithms for Lower Power Consumption

Satoshi Takahashi (University of Tsukuba)

VM (Virtual Machine)-based flexible capacity management is an effective scheme to reduce total power consumption in the data center. However, there have been the following issues, tradeoff of power-saving and user experience, decision of VM packing in feasible calculation time and collision avoidance of VM migration processes. In order to resolve these issues, we propose a matching-based and a greedy-type VM packing algorithm, which enables to decide a suitable VM packing plan in polynomial time. The experiments evaluate not only a basic performance, but also a feasibility of the algorithms by comparing with optimization solvers. The feasibility experiment uses a super computer trace data prepared by Center for Computational Sciences of University of Tsukuba. The basic performance experiment shows that the algorithms reduce total power consumption by between 18% and 50%.

PanDA: Next Generation Workload Management and Analysis System for Big Data

Mikhail Titov (University of Texas at Arlington)

In the real world any big science project implies to use a sophisticated Workload Management System (WMS) that deals with a huge amount of highly distributed data, which is often accessed by large collaborations. The Production and Distributed Analysis System (PanDA) is a high-performance WMS that is aimed to meet production and analysis requirements for a data-driven workload management system capable of operating at the Large Hadron Collider data processing scale. PanDA provides execution environments for a wide range of experimental applications, automates centralized data production and processing, enables analysis activity of physics groups, supports custom workflow of individual physicists, provides a unified view of distributed worldwide resources, presents status and history of workflow through an integrated monitoring system, archives and curates all workflow. PanDA is now being generalized and packaged, as a WMS already proven at extreme scales, for the wider use of the Big Data community.

Numerical Studies of the Klein-Gordon Equation in a Periodic Setting

Albert Liu (University of Michigan)

In contemporary physics research, there is much interest in modeling quantum interactions. The Klein-Gordon equation is a wave equation useful for such purposes. We investigate the equation by simulating different solutions of the equation using various initial conditions, with solutions that tend to zero at infinity being of special interest. The primary site used to perform these simulations is Trestles at SDSC, and we also studied the performance increase when running jobs on

supercomputing resources. This involved performing a scaling study involving the relationship between core/node number and performance increase. In addition to investigating the Klein-Gordon equation, another important goal of our project was to provide an undergraduate perspective on supercomputing. When considering undergraduate involvement in the field of high performance computing, the level of student engagement is very disappointing. From our experience with supercomputing resources, we look to provide new ways in enhancing student outreach and engagement.

ACM Student Research Competition Posters

On the Cost of a General GPU Framework - The Strange Case of CUDA 4.0 vs. CUDA 5.0

Matthew Wezowicz (University of Delaware)

CUDA reached its max performance with CUDA 4.0. Since its release, NVIDIA has started the re-design of the CUDA framework driven by the search for a framework whose compiler back-end is unified with OpenCL. However, our poster indicates that the new direction comes at a high performance cost. We use the MD code FENZI as our benchmark for our performance analysis. We consider two versions of FENZI: a first version that was implemented for CUDA 4.0 and an optimized version on which we performed additional code optimizations by strictly following NVIDIA's guidelines. For the first version we observed that CUDA 4.0 always outperforms CUDA 4.1, 4.2, and 5.0. We repeated the performance comparison for the optimized FENZI and the four CUDA variants. CUDA 5.0 provides the best performance; still its performance across GPUs and molecular systems is less than the performance of FENZI without optimizations for CUDA 4.0.

High Quality Real-Time Image-to-Mesh Conversion for Finite Element Simulations

Panagiotis Foteinos (College of William and Mary)

In this poster, we present a parallel Image-to-Mesh Conversion (I2M) algorithm with quality and fidelity guarantees achieved by dynamic point insertions and removals. Starting directly from an image, it is able to recover the surface and mesh the volume with tetrahedra of good shape. Our tightly-coupled shared-memory parallel speculative execution paradigm employs carefully designed memory and contention managers, load balancing, synchronization and optimizations schemes, while it maintains high single-threaded performance as compared to CGAL, the state of the art sequential I2M software we are aware of. Our meshes come also with theoretical guarantees: the radius-edge is less than 2 and the angles of the boundary triangles are more than 30 degrees.

The effectiveness of our method is shown on Blacklight, the NUMA machine of Pittsburgh-Supercomputing-Center. We observe a more than 90% strong scaling efficiency for up to 64 cores and a super-linear weak scaling efficiency for up to 128 cores.

Optimus: A Parallel Optimization Framework With Topology Aware PSO and Applications

Sarat Sreepathi (North Carolina State University)

This research presents a parallel metaheuristic optimization framework, Optimus (Optimization Methods for Universal Simulators) for integration of a desired population-based search method with a target scientific application. Optimus includes a parallel middleware component, PRIME (Parallel Reconfigurable Iterative Middleware Engine) for scalable deployment on emergent supercomputing architectures. Additionally, we designed TAPSO (Topology Aware Particle Swarm Optimization) for network based optimization problems and applied it to achieve better convergence for water distribution system (WDS) applications. The framework supports concurrent optimization instances, for instance multiple swarms in the case of PSO. PRIME provides a lightweight communication layer to facilitate periodic inter-optimizer data exchanges. We performed scalability analysis of Optimus on Cray XK6 (Jaguar) at Oak Ridge Leadership Computing Facility for the leak detection problem in WDS. For a weak scaling scenario, we achieved 84.82% of baseline at 200,000 cores relative to performance at 1000 cores and 72.84% relative to one core scenario.

An MPI Library Implementing Direct Communication for Many-Core Based Accelerators

Min Si (University of Tokyo)

DCFA-MPI is an MPI library implementation for manycore-based clusters, whose compute node consists of Intel MIC (Many Integrated Core) connected to the host via PCI Express with InfiniBand. DCFA-MPI enables direct data transfer between MIC units without the host assist. MPI_Init and MPI_Finalize functions are offloaded to the host side in order to initialize the InfiniBand HCA and inform its PCI-Express address to MIC. MPI communication primitives executed in MIC may transfer data directly to other MICs or hosts by issuing commands to HCA. The implementation is based on the Mellanox InfiniBand HCA and Intel's Knights Ferry, and compared with the Intel MPI + offload mode. Preliminary results show that DCFA-MPI outperforms the Intel MPI + offload mode by 1 ~ 4.2 times.

Massively Parallel Model of Evolutionary Game Dynamics

Amanda E. Peters Randles (Harvard University)

To study the emergence of cooperative behavior, we have developed a scalable parallel framework for evolutionary game dynamics. An important aspect is the amount of history that each agent can keep. We introduce a multi-level decomposition method that allows us to exploit both multi-node and thread-level parallel scaling while minimizing communication overhead. We present the results of a production run modeling up to six memory steps for populations consisting of up to 10^{18} agents, making this study one of the largest yet undertaken. The high rate of mutation within the population results in a non-trivial parallel implementation. The strong and weak scaling studies provide insight into parallel scalability and programmability trade-offs for large-scale simulations, while exhibiting near perfect weak and strong scaling on 16,384 tasks on Blue Gene/Q. We further show 99% weak scaling to 294,912 processors 82% strong scaling efficiency to 262,144 processors of Blue Gene/P.

Scalable Cooperative Caching with RDMA-Based Directory Management for Large-Scale Data Processing

Junya Arai (University of Tokyo)

Cooperative caching provides an extensive virtual file cache by combining file caches on all clients. We propose a novel cooperative caching system that addresses two problems of existing systems: lack of utilization of high-throughput, low-latency remote direct memory access (RDMA) and low scalability against concentrated request for a particular cache block. The proposed system uses only RDMA to look up a block location in cache directories distributed to servers and transfer a block from another node, reducing access times to a cache block. In addition, to provide more robust scalability, nodes are divided into groups and managed semi-independently, so that access concentration on a particular node is mitigated. We are going to run over 10,000 processes on the K computer for evaluation and present experimental results on the poster. To our knowledge, this will be the first study that investigates performance of cooperative caching with over 1,000 processes.

An Ultra-Fast Computing Pipeline for Metagenome Analysis with Next-Generation DNA Sequencers

Shuji Suzuki (Tokyo Institute of Technology)

Metagenome analysis is useful for not only understanding symbiotic systems but also watching environment pollutions. However, metagenome analysis requires sensitive sequence homology searches which require large computation time and it is thus a bottleneck in current metagenome analysis based on the data from the latest DNA sequencers generally called a next-generation sequencer (NGS). To solve the problem, we developed a new efficient GPU-based homology search program GHOSTM and a large-scale automated computing

pipeline for analyzing huge amount of metagenomic data obtained from a NGS. This pipeline enables us to analyze metagenomic data from a NGS in real time by utilizing huge computational resources on TSUBAME2. For homology search, the pipeline can utilize both GHOSTM and BLASTX, which has been generally used for previous metagenomic researches. By using the new program and pipeline, we achieved to process metagenome information obtained from a single run of a NGS in a few hours.

Reducing the Migration Times of Multiple VMs on WANs

Tae Seung Kang (University of Florida)

It is difficult to statically minimize the time required to transfer multiple VMs across a WAN. One approach is to migrate a large number of VMs concurrently, but this leads to long migration times of each individual VM. Long migration times are problematic in catastrophic circumstances where needed resources can fail within a short period of time. Thus, it is important to shorten both the total time required to migrate multiple VMs and the migration time of individual VMs. This work proposes a feedback-based controller that adapts the number of concurrent VM migrations in response to changes in a WAN. The controller implements an algorithm inspired by the TCP congestion avoidance algorithm in order to regulate the number of VMs in transit depending on network conditions. The experiments show that the controller shortens the individual migration time by up to 5.7 fold compared to that of the static VM migrations.

Performing Cloud Computation on a Parallel File System

Ellis H. Wilson (Pennsylvania State University)

The MapReduce (MR) framework is a programming environment that facilitates rapid parallel design of applications that process big data. While born in the Cloud arena, numerous other areas are now attempting to utilize it for their big data due to the speed of development. However, for HPC researchers and many others who already utilize centralized storage, MR marks a paradigm shift toward co-located storage and computation resources. In this work I attempt to reach the best of both worlds by exploring how to utilize MR on a network-attached parallel file system. This work is nearly complete and has unearthed key issues I've subsequently overcome to achieve desired high throughput. In my poster I describe many of these issues, demonstrate improvements possible with different architectural schemas, and provide reliability and fault-tolerance considerations for this novel combination of Cloud computation and HPC storage.

Crayons: An Azure Cloud-Based Parallel System for GIS Overlay Operations

Dinesh Agarwal (Georgia State University)

Geographic Information System (GIS) community has always perceived the processing of extremely large vector-based spatial datasets as a challenging research problem. This has not been for the lack of individual parallel algorithms, but as we discovered, it is because of the irregular and data intensive nature of the underlying computation. While effective systems for parallel processing of raster-based spatial data files are abundant in the literature, there is only meager amount of reported system work that deals with the complexities of vector (polygonal) data and none on cloud platform. We have created an open-architecture-based system named Crayons for Azure cloud platform using state-of-the-art techniques. The Crayons system scales well for sufficiently large datasets, achieving end-to-end absolute speedup of over 28-fold employing 100 Azure processors. For smaller, more irregular workload, it still yields over 9-fold speedup.

Pay as You Go in the Cloud: One Watt at a Time

Kayo Teramoto (Yale University)

Advancements in virtualization have led to the construction of large data centers that host thousands of servers and to the selling of virtual machines (VM) to consumers on a per-hour rate. This current pricing scheme employed by cloud computing providers ignores the disparities in consumer usage and in its related infrastructural costs of providing the service to different users. We thus propose a new pricing model based on the liable power consumption of the VM, which we correlate to the VM's proportion of CPU and disk I/O usage. In the poster, we evaluate the fairness and practicality of our accountable power consumption model on various machines and storage types. We then demonstrate the benefits of the proposed pricing model by looking at four consumer cases. Our work is undergoing further experimentation and we hope to expand our testing using cloud services.

Optimizing pF3D using Model-Based, Dynamic Parallelism

ChunYi Su (Virginia Tech)

Optimizing parallel applications for performance and power in current and future systems poses significant challenges. A single node today presents multiple levels of parallelism including multiple SMT-threads, cores, sockets, and memory domains. Determining the optimal concurrency and mapping of an application to the underlying processing units may be intractable for online optimization and challenging for efficient offline search. In this work, we present a framework to dynamically optimize the performance of parallel programs based on model predictions of the optimal configuration. We optimize the performance of kernels from pF3D, a real-world

multi-physics code used to simulate the interaction between a high-intensity laser and a plasma. Our results show that our approach predicts performance within 6% of the optimal in average and achieve performance improvements from 1.03x to 5.95x compared to the Linux default setting.

Norm-Coarsened Ordering for Parallel Incomplete Cholesky Preconditioning

Joshua D. Booth (Pennsylvania State University)

The preconditioned conjugate gradient method using incomplete Cholesky factors (PCG-IC) is a widely used iterative method for the scalable parallel solution of linear systems with a sparse symmetric positive definite coefficient matrix. Performance of the method depends on the ordering of the coefficient matrix which controls fill-in, exposes parallelism, and changes the convergence of conjugate gradient method. Furthermore, for a truly parallel solution, it is desirable that the ordering step itself can be parallelized. Earlier work indicates that orderings such as nested dissection and coloring that are suitable for parallel solution can often degrade the quality of the preconditioner. This work seeks to address this gap by developing a norm-coarsened ordering scheme that can be implemented in parallel while potentially improving convergence. Norm-coarsened ordering may improve the effective flops (iterations times nonzeros in the preconditioner) by as much 68% compared to nested dissection orderings and 34% compared to Reverse Cuthill-McKee.

Neural Circuit Simulation of Hodgkin-Huxley Type Neurons Toward Petascale Computers

Daisuke Miyamoto (University of Tokyo)

We ported and optimized simulation environment "NEURON" on K computer to simulate a insect brain as multi-compartment Hodgkin-Huxley type model. To use SIMD units of SPARC64VIIIfx (CPU of K computer), exchanged the order of the compartment loop and the ion channel loop and apply sector caches. These tunings improved single core performance 340 MFLOPS/core to 1080 MFLOPS/core (about 7% efficiency). Spike exchange method of "NEURON" (MPI_Allgather) demands large amount of time in case of 10,000 cores or more and simple asynchronous point-to-point method (MPI_Isend) is not effective too, because of a large number of function calls and long distance of interconnect pathway. To tackle these problems, we adopted MPI/OpenMP hybrid parallelization to reduce interconnect communications and we developed a program to optimize location of neurons on calculation nodes in the 3D torus network. As a these results, we preliminary obtained about 70 TFLOPS with 196,608 CPU cores.

Scientific Visualization Showcase

Tuesday, November 13

Reception & Exhibit: 5:15pm-7pm

Chair: Kelly Gaither (TACC)

Room: North Foyer

Wednesday, November 14-

Thursday, November 15

8:30am-5pm

Room: North Foyer

Computing The Universe From Big Bang to Stars

Bruno Thooris, Daniel Pomarède (CEA)

The movie "Computing The Universe" was performed in August 2011 at IRFU, the Research Institute on Fundamental Laws of the Universe, a CEA institute at Saclay, France. The movie, produced also in stereoscopic version, shows to the general public the results of the numerical simulations in astrophysics produced by the COAST project team; the team involves astrophysicists and computing science engineers, in particular for the visualization part. These results are produced after some weeks of calculation on the thousands of processors in parallel of supercomputers. Results are published in international astrophysics journals but the visualization of these data always creates a big enthusiasm in the general public.

Scenario: The movie follows a logic of decreasing sizes of astronomical phenomena and includes five parts: -Cosmology -Galaxies formation -Interstellar Medium -Magnetism of the Sun -Supernovae Remnants

Investigation of Turbulence in the Early Stages of a High Resolution Supernova Simulation

Robert Sisneros (National Center for Supercomputing Applications), Chris Malone, Stan Woosley (University of California, Santa Cruz), Andy Nonaka (Lawrence Berkeley National Laboratory)

Cosmologists have used the light curves of Type Ia supernovae (SN Ia) as tools for surveying vast distances. Previous simulations have used coarse resolution and artificial initial conditions that substantially influenced their outcome. Here, we have the unique advantage of being able to import the results from previous simulations of convection leading to ignition from our low Mach number code, MAESTRO, directly into our compressible code, CASTRO. These initial conditions include the location of ignition and the turbulence on the grid. In this video, we show the turbulence within the early "bubble" of a supernova via renderings of the magnitude of the vorticity within the simulation. We then focus on the highest values of the magnitude of vorticity to observe the formation of "vortex tubes." The video is available here: <http://www.ncsa.illinois.edu/~sisneros/sc video.html>

Two Fluids Level Set: High Performance Simulation and Post Processing

Herbert Owen, Guillaume Houzeaux, Cristobal Samaniego, Fernando Cucchietti, Guillermo Marin, Carlos Tripiana, Mariano Vazquez, Hadrien Calmet (Barcelona Supercomputing Center)

Flows with moving interfaces appear in a wide range of real world problems. This report, accompanying the video "Two fluids level set: High performance simulation and post processing" presents the implementation of a Level Set method for two fluid flows in the parallel finite element code Alya that can scale up to thousands of processors. To give an idea of the versatility of the implementation examples extending from the flushing of a toilet to the simulation of free surface flows around ship hulls are presented. The spatial discretization is based on unstructured linear finite elements, tetrahedras and prisms that allow a great degree of flexibility for complex geometries as will be shown in the examples. The time discretization uses a standard trapezoidal rule. The position of the moving interface is captured with the Level Set technique that is better suited for complex flows than interface tracking schemes. The jump in the fluid properties is smoothed in a region close to the interface. For ship hydrodynamics simulations the model has been coupled with the SST k-omega turbulence model.

SiO₂ Fissure in Molecular Dynamics

Aaron Knoll (Texas Advanced Computing Center)

This video demonstrates output of a large scale molecular dynamics computation of 4.7 million atoms, simulating glass fracture over the course of over a nanosecond. Postprocessed data consists of 500 frames at 3 GB per frame. We generate approximate charge density volume data to highlight the overall structure. Rendering is performed using Nanovol, a volume raycasting tool for computational chemistry.

Direct Numerical Simulations of Cosmological Reionization: Field Comparison: Density

Joseph A. Insley (Argonne National Laboratory), Rick Wagner, Robert Harkness, Michael L. Norman (San Diego Supercomputer Center), Daniel R. Reynolds (Southern Methodist University), Mark Hereld, Michael E. Papka (Argonne National Laboratory)

The light from early galaxies had a dramatic impact on the gasses filling the universe. This video highlights the spatial structure of the light's effect, by comparing two simulations: one with a self-consistent radiation field (radiative), and one without (non-radiative), each with a very high dynamic range. Looking at the simulations side-by-side it's hard to see any difference. However, because the simulations have the same initial conditions, we can directly compare them, by looking at the relative difference of the density. The coral-like blobs are regions where light has radiated out, heating the gas, and raising the pressure. The red regions show where the density is much higher in the radiative simulation, while the yellow

regions are where the non-radiative has more density, showing where gravity was able to pull the filaments into tighter cylinders, without having to work against pressure from stellar heating. This is the first known visualization of this process, known as Jeans smoothing.

Direct Numerical Simulations of Cosmological Reionization: Field Comparison: Ionization Fraction

Joseph A. Insley (Argonne National Laboratory), Rick Wagner, Robert Harkness, Michael L. Norman (San Diego Supercomputer Center), Daniel R. Reynolds (Southern Methodist University), Mark Hereld, Michael E. Papka (Argonne National Laboratory)

The light from early galaxies had a dramatic impact on the gasses filling the universe. This video highlights the spatial structure of the light's effect, by comparing two simulations: one with a self-consistent radiation field (radiative), and one without (non-radiative), each with a very high dynamic range. Ionization fraction is the amount of the gas that has been ionized. Looking at this quantity from the simulations side-by-side one can clearly see differences but it can be difficult to decipher how the regions of concentration in the two simulations relate to one another. However, because the simulations have the same initial conditions, we can directly compare them, by looking at the relative difference of the ionization fraction in a single view. The yellow and red regions show where the gas has been ionized in the radiative simulation, while at the center of these blobs are small blue regions where the ionized gas from the non-radiative simulation is concentrated. The purple illustrates the boundary at the advancing edge of the ionization from the radiative simulation, where the two simulations are the same.

Direct Numerical Simulation of Flow in Engine-Like Geometries

Martin Schmitt, Christos E. Frouzakis (ETH Zurich), Jean M. Favre (Swiss National Supercomputing Centre)

Internal combustion engine flows are turbulent, unsteady and exhibit high cycle-to-cycle variations. There are multiple turbulence generating mechanisms and their effects overlap in time and space, creating strong challenges for the turbulence models currently used, at least for the in-depth understanding of underlying mechanisms as well as for predictive purposes.

Using the highly scalable, parallel, spectral element flow solver nek5000, multiple cycles of the flow around an open valve induced by a moving piston are computed by solving the incompressible Navier-Stokes equations in the temporally-varying geometry. The visualization of the high resolution simulations results reveals cycle-to-cycle fluctuations due to differences in the jet breakup and position, which depend on the turbulence remaining in the cylinder at the top dead center (i.e. the piston position closest to the cylinder head). The fine flow

structures generated during the expansion stroke, and their suppression during the compression stroke are shown in the animations of the volume rendering of the velocity magnitude and the isocontours of the vorticity magnitude of the first two cycles.

Cosmology on the Blue Waters Early Science System

Dave Semeraro (University of Illinois at Urbana-Champaign)

The submitted visualization represents work performed by the Enzo PRAC team lead by Brian O'Shea on the Blue Waters Early Science system. A relatively small test calculation was performed followed by several much larger AMR cosmological runs. The overall scientific goal was to understand how galaxies in the early Universe (the first billion years or so after the Big Bang) grow and evolve, in several statistically-dissimilar environments. Specifically, we looked at a region that is statistically over dense (substantially more galaxies than the average per volume), under dense (the opposite), of average mean density, and, finally, a region that will make a Milky Way-like galaxy at the present day. For each calculation, we used two separate formulations for our sub grid models of star formation and feedback. The simulation visualized in these movies represents the "average" calculations, which are the most statistically comparable to galaxies that have been observed with the Hubble Space Telescope and large ground-based telescopes. The visualization was done using the VisIt volume renderer. The analysis was completely performed on the Blue Waters Early science system. Volume renderings of density and temperature are presented. Each of these simulations was substantially larger than any previous Enzo AMR calculation ever run (as well as larger than any other AMR cosmological simulation ever done). By the end of the run, the calculations have several billion unique resolution elements on six levels of adaptive mesh.

Explosive Charge Blowing a Hole in a Steel Plate Animation

Brad Carvey, Nathan Fabian, David Rogers (Sandia National Laboratories)

The animation shows a simulation of an explosive charge, blowing a hole in a steel plate. The simulation data was generated on Sandia National Lab's Red Sky Supercomputer. ParaView was used to export polygonal data, which was then textured and rendered using a commercial 3d rendering package.

Using ParaView's co-processing capability, data was captured directly from the memory of the running super computer simulation. We then created a set of seamless fragment surfaces extracted from the underlying cells' material volume fractions. ParaView outputs a sequence of models that are converted to obj polygonal objects, using NuGraf, a model format conversion program. The objects vary in size, with some

objects consisting of around a million polygons. One quarter of the animation was generated. The other 3 sections were instanced and textured. Instancing the imported geometry does not create more geometry, the original instance is used to generate the other sections. This saves memory and speeds up the rendering. A custom script controlled the loading and rendering of the sequential models. The final object sequence was rendered offline with Modo 6.0.

Computational Fluid Dynamics and Visualization

Michael A. Matheson (Oak Ridge National Laboratory)

The video showcases work in both Computational Fluid Dynamics (CFD) and Visualization with an example given of Smoothed Particle Hydrodynamics (SPH). This is a particle method that can run utilizing many types of parallelism such as OpenMP, MPI, and CUDA. The visualization shows many different types of techniques applied to the time dependent solution which also runs with many types of parallelism on the same hardware. Particle methods exhibit fairly high efficiency on hybrid supercomputers such as the OLCF's Titan and have desirable features for data analysis such as "free" path operations. The ability to process the particles in parallel since there is no explicit mesh topology may make these types of methods attractive at exascale where In Situ techniques will be required for efficient use of these systems.

Effect of Installation Geometry on Turbulent Mixing Noise from Jet Engine Exhaust

*Joseph A. Insley (Argonne National Laboratory),
Umesh Paliath, Sachin Premasathan (GE Global Research)*

Jet noise is one of the most dominant noise components from an aircraft engine that radiates over a wide frequency range. Empirical and semi-empirical models that rely on scaling laws and calibrations of constants based on the far field acoustic measurements are limited in their range of applicability, as universal calibration over variations in operating conditions and nozzle geometries is impossible. In recent years direct jet noise prediction using large eddy simulation (LES) and computational aero-acoustics methodology has attracted increasing attention in the jet noise community.

Over the past decades jet noise reduction has been achieved mainly through increase of engine bypass ratio, which lowers jet speed for a given thrust. As under-wing-mounted engine diameters increase, jet axes must move closer to the airframe, to maintain the same ground clearance. This close-coupling now means that installation noise plays a major part in community noise reduction matrix. It is essential to be able to predict and understand the altering of the noise source generation and propagation mechanisms in the presence of forward flight and installation geometries like pylon, wing and flap. Very few studies have been conducted to assess this aspect of

nozzle design. Such knowledge is vital to engine and airframe integration.

Virtual Rheoscopic Fluid for Large Dynamics Visualization

Paul A. Navrátil, William L. Barth (Texas Advanced Computing Center), Hank Childs (Lawrence Berkeley National Laboratory)

Visualizations of fluid hydraulics often use some combination of isosurfacing and streamlines to identify flow features. Such algorithms show surface features well, but nearby features and internal features can be difficult to see. However, virtual rheoscopic fluids (VRF) provide a physically-based visualization that can be analyzed like physical rheoscopic fluid and can show nearby and internal features. Whereas previous VRF implementations were custom serial implementations that could not scale to large dynamics problems, these visualizations demonstrate our scalable VRF algorithm operating on large fluid dynamics data. Our algorithm is implemented in VisIt and the movies were generated on TACC's Longhorn machine using 128 nodes (1024 cores).

Inside Views of a Rapidly Spinning Star

*Greg Foss, Greg Abram (Texas Advanced Computing Center),
Ben Brown (University of Wisconsin-Madison), Mark Miesch (University Corporation for Atmospheric Research)*

"Inside Views of a Rapidly Spinning Star" shows a sampler of visualized variables from a star simulation generated with ASH (Anelastic Spherical Harmonic, originally developed at the University of Colorado) on Ranger at the Texas Advanced Computing Center. This simulated star is similar to our Sun in mass and composition but spinning five times faster. The movie compares the variables radial velocity, enstrophy, and velocity magnitude.

A Dynamic Portrait of Global Aerosols

William Putman (NASA)

Through numerical experiments that simulate our current knowledge of the dynamical and physical processes that govern weather and climate variability of Earth's atmosphere, models create a dynamic portrait of our planet. The simulation visualized here captures how winds lift up aerosols from the Earth's surface and transport them around the globe. Such simulations allow scientists to identify the source and pathway of these tiny particulates that influence weather and climate. Each frame covers a 30-minute interval, from September 1, 2006 to January 31, 2007.

With a resolution of 10 kilometers per grid cell, among the highest resolutions for any global atmospheric model, the simulation represents a variety of features worldwide. Winds near the surface and aloft (white) lift up sea-salt (blue) from the

oceans, dust (red) from the earth's surface, disperse plumes of sulphates (ash brown) from volcanic eruptions and fossil fuel emissions, and carry organic and black carbon (green) within smoke from wildfires and human-initiated burning (red-yellow dots) as detected by NASA's MODIS satellite. These tiny particles can be transported large distances from their sources within the strong winds of the atmospheric circulation and have a significant impact on air quality, visibility and human health.

Probing the Effect of Conformational Constraints on Binding

Anne Bowen (Texas Advanced Computing Center), Yue Shi (University of Texas at Austin)

Increasing the strength of binding between a molecule and a receptor is an important technique in the design of effective drugs. One experimental technique to increase the strength of binding (called "binding affinity") is to synthesize molecules that are already in the shape that it will take when bound to a receptor. This technique works because it decreases the binding entropy which increases the overall binding affinity. A recent experimental study of a series of receptor-molecule complexes aimed to increase the binding affinity by introducing a bond constraint. However, the constrained molecules had less favorable binding entropies than their flexible counterparts. Yue Shi of the Ren lab at UT Austin aimed to probe the origin of this entropy paradox with molecular dynamics simulations which were run on Lonestar and Ranger at TACC. Their group used approximately 2 million CPU hours on Ranger and almost 1 million on Lonestar this past year. Their research addresses biological and medical challenges from single molecules to the genome with high performance computing and theory. In collaboration with other experimental groups, they utilize computer modeling and simulations to understand these complex biomolecular systems and to discover molecules for treating disease and improving human health. Effectively communicating the results of their computational studies to experimentalists is essential to the success of their collaborative efforts. Anne Bowen of the Texas Advanced Computing Center collaborated with Yue Shi to prepare animations and graphics to better explain the origins of the "entropy paradox" to experimentalists and the general public.

In-Situ Feature Tracking and Visualization of a Temporal Mixing Layer

Earl P.N. Duque, Daniel Heipler (Intelligent Light), Christopher P. Stone (Computational Sciences and Engineering LLC), Steve M. Legensky (Intelligent Light)

The flowfield for a temporal mixing layer was analyzed by solving the Navier-Stokes equations via a Large Eddy Simulation method, LESLIE3D, and then visualizing and post-processing the resulting flow features by utilizing the prototype visualization and CFD data analysis software system Intelligent In-Situ Feature Detection, Tracking and Visualization for Turbulent Flow Simulations (IFDT). The system utilizes volume rendering with an Intelligent Adaptive Transfer Function that allows the user to train the visualization system to highlight flow features such as turbulent vortices. A feature extractor based upon a Prediction-Correction method then tracks and extracts the flow features and determines the statistics of features over time. The method executes In-Situ with the flow solver via a Python Interface Framework to avoid the overhead of saving data to file. The movie submitted for this visualization showcase highlights the visualization of the flow such as the formation of vortex features, vortex breakdown, the onset of turbulence and then fully mixed conditions.



Tutorials

Tutorials offer attendees a variety of short courses on key topics and technologies relevant to high performance computing, networking, storage, analytics. Tutorials also provide the opportunity to interact with recognized leaders in the field and learn about the latest technology trends, theory, and practical techniques.

This year we offer 35 half- and full-day tutorials. These tutorials cover a spectrum of foundation skills, hot topics, and emerging technologies, with material appealing to beginning, intermediate, and advanced HPC professionals.

Tutorials

Tutorials

Sunday, November 11

How to Analyze the Performance of Parallel Codes 101 8:30am-12pm

Presenters: Martin Schulz (Lawrence Livermore National Laboratory), Jim Galarowicz (Krell Institute), Don Maghrak (Krell Institute), David Montoya (Los Alamos National Laboratory), Mahesh Rajan (Sandia National Laboratories), Matthew LeGrand (Lawrence Livermore National Laboratory)

Performance analysis is an essential step in the development of HPC codes. It will even gain in importance with the rising complexity of machines and applications that we are seeing today. Many tools exist to help with this analysis, but the user is too often left alone with interpreting the results. In this tutorial we will provide a practical road map for the performance analysis of HPC codes and will provide users step-by-step advice on how to detect and optimize common performance problems in HPC codes. We will cover both on-node performance and communication optimization and will also touch on accelerator-based architectures. Throughout this tutorial will show live demos using Open|SpeedShop, a comprehensive and easy-to-use performance analysis tool set, to demonstrate the individual analysis steps. All techniques will, however, apply broadly to any tool, and we will point out alternative tools where useful.

Hybrid MPI and OpenMP Parallel Programming 8:30am-12pm

Presenters: Rolf Rabenseifner (High Performance Computing Center Stuttgart), Georg Hager (Erlangen Regional Computing Center), Gabriele Jost (AMD)

Most HPC systems are clusters of shared memory nodes. Such systems can be PC clusters with single/multi-socket and multi-core SMP nodes, but also “constellation” type systems with large SMP nodes. Parallel programming may combine the distributed memory parallelization on the node interconnect with the shared memory parallelization inside of each node. This tutorial analyzes the strengths and weaknesses of several parallel programming models on clusters of SMP nodes. Multi-socket-multi-core systems in highly parallel environments are given special consideration. This includes a discussion on planned future OpenMP support for accelerators. Various hybrid MPI+OpenMP approaches are compared with pure MPI, and benchmark results on different platforms are presented. Numerous case studies demonstrate the performance-related aspects of hybrid MPI/OpenMP programming, and application categories that can take advantage of this model are identified. Tools for hybrid programming such as thread/process

placement support and performance analysis are presented in a “how-to” section.

Details: <https://fs.hlr.de/projects/rabenseifner/publ/SC2012-hybrid.html>

Large Scale Visualization with ParaView 8:30am-12pm

Presenters: Kenneth Moreland (Sandia National Laboratories), W. Alan Scott (Sandia National Laboratories), Nathan Fabian (Sandia National Laboratories), Utkarsh Ayachit (Kitware, Inc.), Robert Maynard (Kitware, Inc.)

ParaView is a powerful open-source turnkey application for analyzing and visualizing large data sets in parallel. 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 hands-on lessons. The tutorial features detailed guidance in visualizing the massive simulations run on today’s supercomputers and an introduction to scripting and extending ParaView. Attendees should bring laptops to install ParaView and follow along with the demonstrations.

Parallel Programming with Migratable Objects for Performance and Productivity 8:30am-12pm

Presenters: Laxmikant V. Kale, Eric J. Bohm (University of Illinois at Urbana-Champaign)

In this tutorial, we describe the migratable, message-driven objects (MMDO) execution model, which allows programmers to write high performance code productively. It empowers an adaptive runtime system (ARTS) to automate load balancing, tolerate faults, and support efficient composition of parallel modules. With MMDO, an algorithm is over-decomposed into objects encapsulating work and data. Objects are message-driven and communicate asynchronously, to automatically overlap communication with computation. MMDO also allows the ARTS to manage program execution. Attendees will gain practical experience with the MMDO paradigm through a number of examples written in the CHARM++ programming system. CHARM++ supports several scalable applications, and has been deployed effectively on multicore desktops and 300K core supercomputers alike. Therefore, it provides a mature and robust vehicle for the exposition of MMDO design principles.

Productive Programming in Chapel: A Language for General, Locality-aware Parallelism

8:30am-12pm

Presenters: Bradford L. Chamberlain, Sung-Eun Choi (Cray Inc.)

Chapel is an emerging parallel language being developed by Cray Inc. to improve the productivity of parallel programmers, from large-scale supercomputers to multicore laptops and workstations. Chapel aims to vastly improve programmability over current parallel programming models while supporting performance and portability that is comparable or better. Chapel supports far more general, dynamic, and data-driven models of parallel computation while also separating the expression of parallelism from that of locality/affinity control. Though being developed by Cray, Chapel is portable, open-source software that supports a wide spectrum of platforms including desktops (Mac, Linux, and Windows), UNIX-based commodity clusters, and systems sold by Cray and other vendors. This tutorial will provide an in-depth introduction to Chapel, from context and motivation to a detailed description of Chapel concepts via lecture, example computations, and live demos. We'll conclude by giving an overview of ongoing Chapel activities and collaborations and by soliciting participants for their feedback to improve Chapel's utility for their parallel computing needs.

A Hands-On Introduction to OpenMP

8:30am-5pm

Presenters: Tim Mattson (Intel Corporation), Mark Bull (Edinburgh Parallel Computing Centre)

OpenMP is the *de facto* standard for writing parallel applications for shared memory computers. With multi-core processors in everything from laptops to high-end servers, the need for multithreaded applications is growing and OpenMP is one of the most straightforward ways to write such programs. We will cover the full OpenMP 3.1 standard in this tutorial. This will be a hands-on tutorial. We expect students to use their own laptops (with Windows, Linux, or OS/X). We will have access to systems with OpenMP (a remote SMP server or virtual machines you can load onto your laptops), but the best option is for you to load an OpenMP compiler onto your laptop before the tutorial. Information about OpenMP compilers is available at www.openmp.org.

Debugging MPI and Hybrid-Heterogeneous Applications at Scale

8:30am-5pm

Presenters: Ganesh Gopalakrishnan (University of Utah), David Lecomber (Allinea Software), Matthias S. Mueller (Technische Universitaet Dresden), Bronis R. de Supinski (Lawrence Livermore National Laboratory), Tobias Hilbrich (Technische Universitaet Dresden)

MPI programming is error prone due to the complexity of MPI semantics and the difficulties of parallel programming. Difficulties are exacerbated by increasing heterogeneity (e.g., MPI plus OpenMP/CUDA), the scale of parallelism, non-determinism, and platform dependent bugs. This tutorial covers the detection/correction of errors in MPI programs as well as heterogeneous/hybrid programs. We will first introduce our main tools: MUST, that detects MPI usage errors at runtime with a high degree of automation; ISP/DAMPI, that detects interleaving-dependent MPI deadlocks through application replay; and DDT, a parallel debugger that can debug at large scale. We will illustrate advanced MPI debugging using an example modeling heat conduction. Attendees will be encouraged to explore our tools early during the tutorial to better appreciate their strengths/limitations. We will also present best practices and a cohesive workflow for thorough application debugging with all our tools. Leadership scale systems increasingly require hybrid/heterogeneous programming models—e.g., Titan (ORNL) and Sequoia (LLNL). To address this, we will present debugging approaches for MPI, OpenMP, and CUDA in a dedicated part of afternoon session. DDT's capabilities for CUDA/OpenMP debugging will be presented, in addition to touching on the highlights of GKLEE—a new symbolic verifier for CUDA applications.

Parallel I/O In Practice

8:30am-5pm

Presenters: Robert Latham, Robert Ross (Argonne National Laboratory), Brent Welch (Panasas), Katie Antypas (NERSC)

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 in detail (PFSs). We cover general concepts and examine four examples: GPFS, Lustre, PanFS, and PVFS. 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.

Productive, Portable Performance on Accelerators Using OpenACC Compilers and Tools

8:30am-5pm

Presenters: James Beyer, Luiz DeRose, Alistair Hart, Heidi Poxon (Cray Inc.)

The current trend in the supercomputing industry is towards hybrid systems with accelerators attached to multi-core processors. The current Top500 list has more than 50 systems with GPUs; ORNL and NCSA have plans to deploy large-scale hybrid systems by the end of 2012. The dominant programming models for accelerator-based systems (CUDA and OpenCL) offer the power to extract performance from accelerators, but with extreme costs in usability, maintenance, development, and portability. To be an effective HPC platform, these systems need a high-level software development environment to enable widespread porting and development of applications that run efficiently on either accelerators or CPUs. In this hands-on tutorial we present the high-level OpenACC parallel programming model for accelerator-based systems, demonstrating compilers, libraries, and tools that support this cross-vendor initiative. Using personal experience in porting large-scale HPC applications, we provide development guidance, practical tricks, and tips to enable effective and efficient use of these hybrid systems.

Scalable Heterogeneous Computing on GPU Clusters

8:30am-5pm

Presenters: Jeffrey Vetter (Oak Ridge National Laboratory), Allen Malony (University of Oregon), Philip Roth, Kyle Spafford, Jeremy Meredith (Oak Ridge National Laboratory)

This tutorial is suitable for attendees with an intermediate-level in parallel programming in MPI and with some background in GPU programming in CUDA or OpenCL. It will provide a comprehensive overview on the optimization techniques to port, analyze, and accelerate applications on scalable heterogeneous computing systems using MPI and OpenCL, CUDA, and directive-based compilers using OpenACC. First, we will review our methodology and software environment for successfully identifying and selecting portions of applications to accelerate with a GPU, motivated with several application case studies. Second, we will present an overview of several performance and correctness tools, which provide performance measurement, profiling, and tracing information about applications running on these systems. Third, we will present a set of best practices for optimizing these applications: GPU and NUMA optimization techniques, optimizing interactions between MPI and GPU programming models. A hands-on session will be conducted on the NSF Keeneland System, after each part to give participants the opportunity to investigate techniques and performance optimizations on such a system. Existing tutorial codes and benchmark suites will be provided to facilitate

individual discovery. Additionally, participants may bring and work on their own applications.

This Is Not Your Parents' Fortran: Object-Oriented Programming in Modern Fortran

8:30am-5pm

Presenters: Karla Morris, Damian Rouson (Sandia National Laboratories), Salvatore Filippone (University of Rome Tor Vergata)

Modern Fortran provides powerful constructs for multiple programming paradigms: Fortran 95, 2003, and 2008 explicitly support functional, object-oriented (OO), and parallel programming. User surveys across HPC centers in the U.S. and Europe consistently indicate the majority of users write Fortran but most write older Fortran dialects and almost all describe their programming language skills as "self-taught." Thus, while 2012 appears to be a watershed moment with burgeoning compiler support for the aforementioned constructs, most HPC users lack access to training in the associated programming paradigms. In this tutorial, three leaders of open-source, parallel OO Fortran libraries will give the students hands-on application programming experience at the level required to use these libraries to write parallel applications.

Using Application Proxies for Co-design of Future HPC Computer Systems and Applications

8:30am-5pm

Presenters: Michael A. Heroux (Sandia National Laboratories), Alice E. Koniges, David F. Richards (Lawrence Livermore National Laboratory), Richard F. Barrett (Sandia National Laboratories), Thomas Brunner (Lawrence Livermore National Laboratory)

The computing community is in the midst of disruptive architectural changes. The advent of manycore and heterogeneous computing nodes, increased use of vectorization, light-weight threads and thread concurrency, along with concerns about energy and resilience, force us to reconsider every aspect of the computer system, software and application stack, often simultaneously. Application proxies have emerged as an important collection of tools for exploring this complex design space. In this tutorial we first present a broad overview of available application proxies including traditional offerings (NAS Parallel Benchmarks, High Performance Linpack, etc.) in order to provide proper context. We then focus on a new collection of proxies called compact apps and miniapps. These two tools have proven especially effective in the past few years since they permit a broader collection of activities, including completely rewriting them. This tutorial is designed for anyone interested in the design of future computer systems, languages, libraries and applications. Hands-on activities will include the ability for attendees to download, compile and run

miniapps on their local machines. We will also provide access to NERSC resources and provide a web portal for modifying, compiling and running on a remote server.

An Overview of Fault-tolerant Techniques for HPC

1:30pm-5pm

Presenters: Thomas Hérault (University of Tennessee, Knoxville), Yves Robert (ENS Lyon)

Resilience is a critical issue for large-scale platforms. This tutorial provides a comprehensive survey on fault-tolerant techniques for high performance computing. It is organized along four main topics: (1) An overview of failure types (software/hardware, transient/fail-stop), and typical probability distributions (Exponential, Weibull, Log-Normal); (2) Application-specific techniques, such as ABFT for grid-based algorithms or fixed-point convergence for iterative applications; (3) General-purpose techniques, which include several checkpoint and rollback recovery protocols, possibly combined with replication; and (4) Relevant execution scenarios will be evaluated and compared through quantitative models (from Young's approximation to Daly's formulas and recent work). The tutorial is open to all SC12 attendees who are interested in the current status and expected promise of fault-tolerant approaches for scientific applications. There are no audience prerequisites: background will be provided for all protocols and probabilistic models. Only the last part of the tutorial devoted to assessing the future of the methods will involve more advanced analysis tools.

Basics of Supercomputing

1:30pm-5pm

Presenters: Thomas Sterling (Indiana University), Steven Brandt (Louisiana State University)

This is *the* crash course on supercomputers for everyone who knows almost nothing but wants to come up to speed fast. All the major topics are described and put into a meaningful framework. All of the terms, both technical and community-related, are presented and discussed in a practical, easy to grasp way. This tutorial requires no prior knowledge or prerequisites other than a need to gain a quantum leap in understanding. The technical foundations will be provided in all the basics, including supercomputer architecture and systems, parallel programming approaches and methods, tools for usage and debugging, and classes of applications. Also presented will be the basic HPC lexicon, the players in the community, the products leading the way, and what's likely to come next. Either first timers or those needing a timely review will benefit from this coverage of HPC.

C++ AMP: An Introduction to Heterogeneous Programming with C++

1:30pm-5pm

Presenters: Kelly Goss (Acceleware Ltd.)

Heterogeneous programming is a key solution to meeting performance goals for HPC algorithms. C++ AMP is a new open specification heterogeneous programming model, which builds on the established C++ language. This tutorial is designed for programmers who are looking to develop skills in writing and optimizing applications using C++ AMP. Participants will be provided with an introduction to the programming model, the tools, and the knowledge needed to accelerate data-parallel algorithms by taking advantage of hardware such as GPUs. A combination of lectures, programming demonstrations and group exercises will provide participants with: (1.) An introduction to the fundamentals of C++ AMP including basic functionality, syntax and data parallelism; (2.) An understanding of the Tiling API within C++ AMP for performance improvement; and (3.) An overview and instructions on how to use C++ AMP specialized features available in Visual Studio 2012, including the Parallel and GPU Debugger.

Developing Scalable Parallel Applications in X10

1:30pm-5pm

Presenters: David Grove, David Cunningham, Vijay Saraswat, Olivier Tardieu (IBM Research)

X10 is a modern object-oriented programming language specifically designed to support productive programming of large-scale, high-performance parallel applications. X10 is a realization of the Asynchronous PGAS (Partitioned Global Address Space) programming model in a Java-like language. The concepts and design patterns introduced in this tutorial via examples written in the X10 programming language will be applicable to other APGAS-based programming environments as well. The tutorial will include an introduction to the X10 language and its implementation and tooling, but will primarily focus on the effective use of X10 concepts to develop scalable, high-performance parallel applications. These concepts will be introduced and motivated by case studies of scalable X10 application programs, drawn from a number of domains including graph algorithms and machine learning. The usage of application frameworks such as the X10 array library, dynamic global load balancing framework, and distributed sparse matrix library will be highlighted in the examples. Participants will have the option of installing X10DT, a full-featured X10 IDE, and complete source code versions of all sample programs and applications presented in the tutorial to enable hands-on exploration of X10 concepts.

In-Situ Visualization with Catalyst

1:30pm-5pm

Presenters: Nathan D. Fabian (Sandia National Laboratories), Andrew C. Bauer (Kitware, Inc.), Norbert Podhorszki (Oak Ridge National Laboratory), Ron A. Oldfield (Sandia National Laboratories), Utkarsh Ayachit (Kitware, Inc.)

In-situ visualization is a term for running a solver in tandem with visualization. Catalyst is the new name for ParaView's co-processing library. ParaView is a powerful open-source turnkey application for analyzing and visualizing large data sets in parallel. By coupling these together, we can utilize HPC platforms for analysis while circumventing bottlenecks associated with storing and retrieving data in disk storage. We demonstrate two methods for in-situ visualization using Catalyst. The first is linking Catalyst directly with simulation codes. It simplifies integration with the codes by providing a programmatic interface to algorithms in ParaView. Attendees will learn how to build pipelines for Catalyst, how the API is structured, how to bind it to C, C++, Fortran, and Python and how to build Catalyst for HPC architectures. The second method uses a variety of techniques, known as data staging or in-transit visualization, that involve passing the data through the network to a second running job. Data analysis applications, written using Catalyst, can operate on this networked data from within this second job minimizing interference with the simulation but also avoiding disk I/O. Attendees will learn three methods of handling this procedure as well as the APIs for ADIOS and NESSIE.

Python in HPC

1:30pm-5pm

Presenters: Andy R. Terrel (Texas Advanced Computing Center), Travis Oliphant (Continuum Analytics), Aron J. Ahmadi (King Abdullah University of Science & Technology)

Python is a versatile language for the HPC community, with tools as diverse as visualizing large amounts of data, creating innovative user interfaces, and running large distributed jobs. Unfortunately, Python has a reputation for being slow and unfit for HPC computing. HPC Python experts and their sixty-five thousand cores disagree. As HPC increases its vision to big data and non-traditional applications, it must also use languages that are easier for the novice, more robust to general computing, and more productive for the expert. Using Python in a performance way moves HPC applications ever closer to these goals. This success has made Python a requirement for supporting users new to the HPC field and a good choice for practitioners to adopt.

In this tutorial, we give students practical experience using Python for scientific computing tasks from leaders in the field of Scientific Python. Coming from diverse academic backgrounds, we show common tasks that are applicable to all. Topics include linear algebra and array computing with NumPy, interactive and parallel software development with IPython, performance and painless low-level C linking with Cython, and the friendliest performance interfaces to MPI at SC this year.

Monday, November 12

A Tutorial Introduction to Big Data

8:30am-5pm

Presenters: Robert Grossman (University of Chicago), Alex Szalay (Johns Hopkins University), Collin Bennett (Open Data Group)

Datasets are growing larger and larger each year. The goals of this tutorial are to give an introduction to some of the tools and techniques that can be used for managing and analyzing large datasets. (1) We will give an introduction to managing datasets using databases, federated databases (Graywulf architectures), NoSQL databases, and distributed file systems, such as Hadoop. (2) We will give an introduction to parallel programming frameworks, such as MapReduce, Hadoop streams, pleasantly parallel computation using collections of virtual machines, and related techniques. (3) We will show different ways to explore and analyze large datasets managed by Hadoop using open source data analysis tools, such as R. We will illustrate these technologies and techniques using several case studies, including: the management and analysis of the large datasets produced by next generation sequencing devices, the analysis of astronomy data produced by the Sloan Digital Sky survey, the analysis of earth science data produced by NASA satellites, and the analysis of netflow data.

Advanced MPI

8:30am-5pm

Presenters: William Gropp (University of Illinois at Urbana-Champaign), Ewing Lusk, Robert Ross, Rajeev Thakur (Argonne National Laboratory)

MPI continues to be the dominant programming model for parallel scientific applications on all large-scale parallel machines, such as IBM Blue Gene and Cray XE/XK, as well as on clusters of all sizes. An important trend is the increasing number of cores per node of a parallel system, resulting in increasing interest in combining MPI with a threaded model within a node. The MPI standard is also evolving to meet the needs of future systems, and MPI 3.0 is expected to be released later

this year. This tutorial will cover several advanced features of MPI that can help users program such machines and architectures effectively. Topics to be covered include parallel I/O, multithreaded communication, one-sided communication, dynamic processes, and new features being added in MPI-3 for hybrid programming, one-sided communication, collective communication, fault tolerance, and tools. In all cases, we will introduce concepts by using code examples based on scenarios found in real applications. Attendees will leave the tutorial with an understanding of how to use these advanced features of MPI and guidelines on how they might perform on different platforms and architectures.

Advanced OpenMP Tutorial

8:30am-5pm

Presenters: Christian Terboven (RWTH Aachen University), Alejandro Duran, Michael Klemm (Intel Corporation), Ruud van der Pas (Oracle), Bronis R. de Supinski (Lawrence Livermore National Laboratory)

With the increasing prevalence of multicore processors, shared-memory programming models are essential. OpenMP is a popular, portable, widely supported and easy-to-use shared-memory model. Developers usually find OpenMP easy to learn. However, they are often disappointed with the performance and scalability of the resulting code. This disappointment stems not from shortcomings of OpenMP but rather with the lack of depth with which it is employed. Our “Advanced OpenMP Programming” tutorial addresses this critical need by exploring the implications of possible OpenMP parallelization strategies, both in terms of correctness and performance. While we quickly review the basics of OpenMP programming, we assume attendees understand basic parallelization concepts and will easily grasp those basics. We discuss how OpenMP features are implemented and then focus on performance aspects, such as data and thread locality on NUMA architectures, false sharing, and private versus shared data. We discuss language features in-depth, with emphasis on features recently added to OpenMP such as tasking. We close with debugging, compare various tools, and illustrate how to avoid correctness pitfalls.

Developing and Tuning Parallel Scientific Applications in Eclipse

8:30am-5pm

Presenters: Beth R. Tibbitts, Greg Watson (IBM), Jay Alameda, Jeff Overbey (National Center for Supercomputing Applications)

For many HPC developers, developing and tuning parallel scientific applications involves a hodgepodge of disparate command-line tools. Based on the successful open-source Eclipse

integrated development environment, the Eclipse Parallel Tools Platform (PTP) combines tools for coding, debugging, job scheduling, error detection, tuning, revision control and more into a single tool with a streamlined graphical user interface. PTP helps manage the complexity of HPC code development and optimization on diverse platforms. This tutorial will provide attendees with a hands-on introduction to Eclipse and PTP. (Compared to previous years, this year’s tutorial will contain substantially more material on performance tuning, and less introductory material.) Part 1 (Morning) will introduce code development in Eclipse: editing, building, launching and monitoring parallel applications in C and Fortran. It will also cover support for efficient development of code on remote machines, and developing and analyzing code with a variety of languages and libraries. Part 2 (Afternoon) focuses on parallel debugging and performance optimization tools. Participants will inspect and analyze a real application code, profiling its execution and performance. Access to parallel system(s) for the hands-on portions will be provided. NOTE: Bring a laptop and pre-install Eclipse and PTP. See <http://wiki.eclipse.org/PTP/tutorials/SC12> for installation instructions.

InfiniBand and High-speed Ethernet for Dummies

8:30am-12pm

Presenters: Dhabaleswar K. (DK) Panda (Ohio State University), Hari Subramoni (Ohio State University)

InfiniBand (IB) and High-speed Ethernet (HSE) technologies are generating a lot of excitement towards building next generation High-End Computing (HEC) systems including clusters, datacenters, file systems, storage, and cloud computing (Hadoop, HBase and Memcached) environments. RDMA over Converged Enhanced Ethernet (RoCE) technology is also emerging. This tutorial will provide an overview of these emerging technologies, their offered architectural features, their current market standing, and their suitability for designing HEC systems. It will start with a brief background behind IB and HSE. In-depth overview of the architectural features of IB and HSE (including iWARP and RoCE), their similarities and differences, and the associated protocols will be presented. Next, an overview of the emerging OpenFabrics stack which encapsulates IB, HSE and RoCE in a unified manner will be presented. Hardware/software solutions and the market trends behind IB, HSE and RoCE will be highlighted. Finally, sample performance numbers of these technologies and protocols for different environments will be presented.

Infrastructure Clouds and Elastic Services for Science**8:30am-5pm**

Presenters: John Bresnahan, Kate Keahey (Argonne National Laboratory), Patrick Armstrong, Pierre Riteau (University of Chicago)

Infrastructure-as-a-service cloud computing has recently emerged as a promising outsourcing paradigm: it has been widely embraced commercially and is also beginning to make inroads in scientific communities. Although popular, understanding how science can leverage it is still in its infancy. Specific and accurate information is needed for scientific communities to understand whether this new paradigm is worthwhile and how to use it. Our objective is to introduce infrastructure cloud computing and elastic tools to scientific communities. We will provide up-to-date information about features and services that benefit science and explain patterns of use that can best fit scientific applications. We will highlight opportunities, conquer myths, and equip the attendees with a better understanding of the relevance of cloud computing to their scientific domain. Our tutorial mixes the discussion of various aspects of cloud computing for science, such as performance, elasticity, privacy, with practical exercises using clouds and state-of-the-art tools.

Intro to PGAS---UPC and CAF--- and Hybrid for Multicore Programming**8:30am-5pm**

Presenters: Alice Koniges, Katherine Yelick (Lawrence Berkeley National Laboratory), Rolf Rabenseifner (High Performance Computing Center Stuttgart), Reinhold Bader (Leibniz Supercomputing Centre), David Eder (Lawrence Livermore National Laboratory)

PGAS (Partitioned Global Address Space) languages offer both an alternative to traditional parallelization approaches (MPI and OpenMP), and the possibility of improved performance on heterogeneous and modern architectures. In this tutorial we cover general PGAS concepts and give an in-depth presentation of two commonly used PGAS languages, Coarray Fortran (CAF) and Unified Parallel C (UPC). Hands-on exercises to illustrate important concepts are interspersed with the lectures. Basic PGAS features, syntax for data distribution, intrinsic functions and synchronization primitives are discussed. Advanced topics include optimization and correctness checking of PGAS codes with an emphasis on emerging and planned PGAS language extensions targeted at scalability and usability improvement. A section on migration of MPI codes using performance improvements from both CAF and UPC is given in a hybrid programming section. Longer examples, tools and performance data on the latest petascale systems round out the presentations. Further details and updates: <http://portal.nersc.gov/project/training/files/SC12/pgas> or <https://fs.hlr.de/projects/rabenseifner/publ/SC2012-PGAS.html>

Introduction to GPU Computing with OpenACC**8:30am-12pm**

Presenters: Michael Wolfe (Portland Group, Inc.)

GPUs allow advanced high performance computing with lower acquisition and power budgets, affecting scientific computing from very high-end supercomputing systems down to the departmental and personal level. Learn how to make effective use of the available performance on GPUs using OpenACC directives. We will present examples in both C and Fortran, showing problems and solutions appropriate for each language. We will use progressively more complex examples to demonstrate each of the features in the OpenACC API. Attendees can download exercises to do interactively or to take home.

Large Scale Visualization and Data Analysis with VisIt**8:30am-5pm**

Presenters: Cyrus Harrison (Lawrence Livermore National Laboratory), Jean M. Favre (Swiss National Supercomputing Centre), Hank Childs (Lawrence Berkeley National Laboratory), Dave Pugmire (Oak Ridge National Laboratory), Brad Whitlock, Harinarayan Krishnan (Lawrence Berkeley National Laboratory)

This tutorial will provide attendees with a practical introduction to VisIt, an open source scientific visualization and data analysis application. VisIt is used to visualize simulation results on wide range of platforms from laptops to many of the world's top supercomputers. This tutorial builds on the success of past VisIt tutorials with material updated to showcase the newest features and use cases of VisIt. We will show how VisIt supports five important scientific visualization scenarios: data exploration, quantitative analysis, comparative analysis, visual debugging, and communication of results. We begin with a foundation in basic principles and transition into several special topics and intermediate-level challenges. The last portion of the tutorial will discuss advanced VisIt usage and development, including writing new database readers, writing new operators, and how to couple VisIt with simulations executing on remote computers for in-situ visualization.

Linear Algebra Libraries for High-Performance Computing: Scientific Computing with Multicore and Accelerators

8:30am-5pm

Presenters: Jack Dongarra (University of Tennessee, Knoxville), James Demmel (University of California, Berkeley), Michael Heroux (Sandia National Laboratories), Jakub Kurzak (University of Tennessee, Knoxville)

Today, a desktop with a multicore processor and a GPU accelerator can already provide a TeraFlop/s of performance, while the performance of the high-end systems, based on multicores and accelerators, is already measured in PetaFlop/s. This tremendous computational power can only be fully utilized with the appropriate software infrastructure, both at the low end (desktop, server) and at the high end (supercomputer installation). Most often a major part of the computational effort in scientific and engineering computing goes in solving linear algebra subproblems. After providing a historical overview of legacy software packages, the tutorial surveys the current state-of-the-art numerical libraries for solving problems in linear algebra, both dense and sparse. PLASMA, MAGMA and Trilinos software packages are discussed in detail. The tutorial also highlights recent advances in algorithms that minimize communication, i.e. data motion, which is much more expensive than arithmetic.

Secure Coding Practices for Grid and Cloud Middleware and Services

8:30am-12pm

Presenters: Barton Miller (University of Wisconsin-Madison), Elisa Heymann (Universidad Autonoma de Barcelona)

Security is crucial to the software that we develop and use. With the growth of both Grid and Cloud services, security is becoming even more critical. This tutorial is relevant to anyone wanting to learn about minimizing security flaws in the software they develop. We share our experiences gained from performing vulnerability assessments of critical middleware. You will learn skills critical for software developers and analysts concerned with security. This tutorial presents coding practices subject to vulnerabilities, with examples of how they commonly arise, techniques to prevent them, and exercises to reinforce them. Most examples are in Java, C, C++, Perl and Python, and come from real code belonging to Cloud and Grid systems we have assessed. This tutorial is an outgrowth of our experiences in performing vulnerability assessment of critical middleware, including Google Chrome, Wireshark, Condor, SDSC Storage Resource Broker, NCSA MyProxy, INFN VOMS Admin and Core, and many others.

Supporting Performance Analysis and Optimization on Extreme-Scale Computer Systems

8:30am-12pm

Presenters: Martin Schulz (Lawrence Livermore National Laboratory), Bernd Mohr, Brian Wylie (Forschungszentrum Juelich GmbH)

The number of processor cores available in high-performance computing systems is steadily increasing. In the June 2012 list of the TOP500 supercomputers, only ten systems have less than 4,096 processor cores and the average is almost 27,000 cores, which is an increase of 9,000 in just one half year. Even the median system size is already over 13,000 cores. While these machines promise ever more compute power and memory capacity to tackle today's complex simulation problems, they force application developers to greatly enhance the scalability of their codes to be able to exploit it. To better support them in their porting and tuning process, many parallel tools research groups have already started to work on scaling their methods, techniques and tools to extreme processor counts. In this tutorial, we survey existing performance analysis and optimization tool covering both profiling and tracing techniques, demonstrate selected tools, report on our experience in using them in extreme scaling environments, review existing working and promising new methods and techniques, and discuss strategies for addressing unsolved issues and problems.

The Practitioner's Cookbook for Good Parallel Performance on Multi- and Manycore Systems

8:30am-5pm

Presenters: Georg Hager, Gerhard Wellein (Erlangen Regional Computing Center)

The advent of multi- and manycore chips has led to a further opening of the gap between peak and application performance for many scientific codes. This trend is accelerating as we move from petascale to exascale. Paradoxically, bad node-level performance helps to "efficiently" scale to massive parallelism, but at the price of increased overall time to solution. If the user cares about time to solution on any scale, optimal performance on the node level is often the key factor. Also, the potential of node-level improvements is widely underestimated, thus it is vital to understand the performance-limiting factors on modern hardware. We convey the architectural features of current processor chips, multiprocessor nodes, and accelerators, as well as the dominant MPI and OpenMP programming models, as far as they are relevant for the practitioner. Peculiarities like shared vs. separate caches, bandwidth bottlenecks, and ccNUMA characteristics are pointed out, and the influence of system topology and affinity on the performance of typical parallel programming constructs is demonstrated. Performance engineering is introduced as a powerful tool that helps the user assess the impact of possible code optimizations by establishing models for the interaction of the software with the hardware on which it runs.

Advanced GPU Computing with OpenACC**1:30pm-5pm***Presenters: Michael Wolfe (Portland Group, Inc.)*

For those who have experience with OpenACC directives, this tutorial will focus on performance optimization and advanced features. We will use examples in both C and Fortran, showing motivating problems and solutions for each language. Attendees can download exercises to do interactively or to take home. Topics will include performance measurement, performance tuning, programming multiple GPUs, asynchronous host and GPU computation, and mixed CUDA / OpenACC programs.

Asynchronous Hybrid and Heterogeneous Parallel Programming with MPI/OmpSs for Exascale Systems**1:30pm-5pm***Presenters: Jesus Labarta (Barcelona Supercomputing Center), Xavier Martorell (Technical University of Catalunya), Christoph Niethammer (High Performance Computing Center Stuttgart), Costas Bekas (IBM Zurich Research Laboratory)*

Due to its asynchronous nature and look-ahead capabilities, MPI/OmpSs is a promising programming model approach for future exascale systems, with the potential to exploit unprecedented amounts of parallelism, while coping with memory latency, network latency and load imbalance. Many large-scale applications are already seeing very positive results from their ports to MPI/OmpSs (see EU projects Montblanc, TEXT). We will first cover the basic concepts of the programming model. OmpSs can be seen as an extension of the OpenMP model. Unlike OpenMP, however, task dependencies are determined at runtime thanks to the directionality of data arguments. The OmpSs runtime supports asynchronous execution of tasks on heterogeneous systems such as SMPs, GPUs and clusters thereof. The integration of OmpSs with MPI facilitates the migration of current MPI applications and improves, automatically, the performance of these applications by overlapping computation with communication between tasks on remote nodes. The tutorial will also cover the constellation of development and performance tools available for the MPI/OmpSs programming model: the methodology to determine OmpSs tasks, the Ayudame/Temanejo debugging toolset, and the Paraver performance analysis tools. Experiences on the parallelization of real applications using MPI/OmpSs will also be presented. The tutorial will also include a demo.

Designing High-End Computing Systems with InfiniBand and High-Speed Ethernet**1:30pm-5pm***Presenters: Dhableswar K. (DK) Panda, Hari Subramoni (Ohio State University)*

As InfiniBand (IB) and High-Speed Ethernet (HSE) technologies mature, they are being used to design and deploy different kinds

of High-End Computing (HEC) systems: HPC clusters with accelerators (GPUs and MIC) supporting MPI and PGAS (UPC and OpenSHMEM), Storage and Parallel File Systems, Cloud Computing systems with Hadoop (HDFS, MapReduce and HBase), Multi-tier Datacenters with Web 2.0 (memcached) and virtualization, and Grid Computing systems. These systems are bringing new challenges in terms of performance, scalability, portability, reliability and network congestion. Many scientists, engineers, researchers, managers and system administrators are becoming interested in learning about these challenges, approaches being used to solve these challenges, and the associated impact on performance and scalability. This tutorial will start with an overview of these systems and a common set of challenges being faced while designing these systems. Advanced hardware and software features of IB and HSE and their capabilities to address these challenges will be emphasized. Next, case studies focusing on domain-specific challenges in designing these systems (including the associated software stacks), their solutions and sample performance numbers will be presented. The tutorial will conclude with a set of demos focusing on RDMA programming, network management infrastructure and tools to effectively use these systems.

The Global Arrays Toolkit - A Comprehensive, Production-Level, Application-Tested Parallel Programming Environment**1:30pm-5pm***Presenters: Bruce Palmer, Jeff Daily, Daniel G. Chavarria, Abhinav Vishnu (Pacific Northwest National Laboratory), Sriram Krishnamoorthy (Pacific Northwest National Laboratory)*

This tutorial provides an overview of the Global Arrays (GA) programming toolkit with an emphasis on the use of GA in applications, interoperability with MPI and new features and capabilities. New functionality will be highlighted, including a robust Python interface, user level control of data mapping to processors, and a new capability for creating global arrays containing arbitrary data objects called Global Pointers. The tutorial will begin with an overview of GA's array-oriented, global-view programming model and its one-sided communication basis. It will then describe the basic functionality and programming model provided by GA. Advanced features of GA with an emphasis on how these are used in actual applications will then be presented, followed by a discussion of a new GA-based implementation of the NumPy library (GAIN) that will illustrate how GA applications can be created using the popular & productive Python scripting language. The new Global Pointers functionality and user interface will be described and strategies for programming with Global Pointers to develop arbitrary global data structures (including sparse matrices) will be discussed. The tutorial will finish with a section on upcoming capabilities in GA to address challenges associated with programming on the next generation of extreme scale architectures.



Workshops

SC12 includes 17 full-day and 7 half-day workshops that complement the overall technical program events, expand the knowledge base of its subject area, and extend its impact by providing greater depth of focus. These workshops are geared toward providing interaction and in-depth discussion of stimulating topics of interest to the HPC community.

Workshops

Workshops

Sunday, November 11

Broader Engagement

Organizer: Tiki L. Suarez-Brown (Florida A&M University)
8:30am-5pm

The Broader Engagement (BE) Program workshop brings an eclectic mix of topics ranging from exascale computing to video-gaming. Some sessions are being organized in association with the HPC Educators Program. Two plenary sessions (8:30am to 10am on November 11 and 12) have been organized to introduce the audience to topics related to cloud computing and exascale computing. The current challenges and future directions in the areas of HPC and Big Data will be also discussed. Participants will receive hands-on experience on OpenMP and directives programming for the accelerators.

The main goal of the BE program is to broaden the participation of underrepresented groups in HPC in general and SC conferences in particular.

The workshop will provide multiple opportunities and sessions to interact with a diverse group of participants. The workshop participants also are invited to purchase tickets and attend the resource fair organized by the BE and HPC Educators Program.

3rd Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems

9am-5:30pm

Organizers: Vassil Alexandrov (Barcelona Supercomputing Center), Jack Dongarra (University of Tennessee, Knoxville), Al Geist, Christian Engelmann (Oak Ridge National Laboratory)

Novel scalable scientific algorithms are needed to enable key science applications to exploit the computational power of large-scale systems. This is especially true for the current tier of leading petascale machines and the road to exascale computing as HPC systems continue to scale up in compute node and processor core count. These extreme-scale systems require novel scientific algorithms to hide network and memory latency, have very high computation/communication overlap, have minimal communication, and have no synchronization points. Scientific algorithms for multi-petaflop and exaflop systems also need to be fault tolerant and fault resilient, since the probability of faults increases with scale. With the advent of heterogeneous compute nodes that employ standard processors and GPGPUs, scientific algorithms need to match these architectures to extract the most performance. Key science applications require novel mathematical models and system software that address the scalability and resilience challenges of current and future-generation extreme-scale HPC systems.

High Performance Computing, Networking and Analytics for the Power Grid

9am-5:30pm

Organizers: Daniel G. Chavarria, Bora Akyol, Zhenyu Huang (Pacific Northwest National Laboratory)

The workshop intends to promote the use of high performance computing and networking for power grid applications. Technological/policy changes make this an urgent priority. Sensor deployments on the grid are expected to increase geometrically in the immediate future, while the demand for clean energy generation is driving the use of non-dispatchable power sources such as solar and wind. New demands are being placed on the power infrastructure due to the introduction of plug-in vehicles. These trends reinforce the need for higher fidelity simulation of power grids and higher frequency measurement of their state. Traditional grid simulation and monitoring tools cannot handle the increased amounts of sensor data or computation imposed by these trends. The use of high performance computing and networking technologies is of paramount importance for the future power grid, particularly for its stable operation in the presence of intermittent generation and increased demands placed on its infrastructure.

High Performance Computing Meets Databases

9am-5:30pm

Organizers: Bill Howe, Jeff Gardner, Magdalena Balazinska (University of Washington), Kerstin Kleese-Van Dam, Terence Critchlow (Pacific Northwest National Laboratory)

Emerging requirements from HPC applications offer new opportunities for engagement between the database and HPC communities: higher level programming models, combined platforms for simulation, analysis, and visualization, ad hoc interactive query, and petascale data processing. Exascale HPC platforms will share characteristics with large-scale data processing platforms: relatively small main memory per node, relatively slow communication between nodes, and IO a limiting factor. Relevant database techniques in this setting include a rigorous data model, cost-based optimization, declarative query languages, logical and physical data, new applications (financial markets, image analysis, DNA sequence analysis, social networks) and new platforms (sensor networks, embedded systems, GPGPUs, shared-nothing commodity clusters, cloud platforms). But these techniques have only been minimally explored in the high-performance computing community.

Second Workshop on Irregular Applications - Architectures and Algorithms

9am-5:30pm

Organizers: John Feo, Antonino Tumeo, Oreste, Simone Secchi, Mahantesh Halappanavar (Pacific Northwest National Laboratory)

Many data-intensive scientific applications are by nature irregular. They may present irregular data structures, control flow or communication. Current supercomputing systems are organized around components optimized for data locality and regular computation. Developing irregular applications on them demands a substantial effort, and often leads to poor performance. Solving these applications efficiently, however, will be a key requirement for future systems. The solutions needed to address these challenges can only come by considering the problem from all perspectives: from micro- to system-architectures, from compilers to languages, from libraries to runtimes, from algorithm design to data characteristics. Only collaborative efforts among researchers with different expertise, including end users, domain experts, and computer scientists, could lead to significant breakthroughs. This workshop aims at bringing together scientists with all these different backgrounds to discuss, define and design methods and technologies for efficiently supporting irregular applications on current and future architectures.

The Second International Workshop on Network-aware Data Management

9am-5:30pm

Organizers: Mehmet Balman, Surendra Byna (Lawrence Berkeley National Laboratory)

Scientific applications and experimental facilities generate large amounts of data. In addition to increasing data volumes and computational requirements, today's major science requires cooperative work in globally distributed multidisciplinary teams. In the age of extraordinary advances in communication technologies, there is a need for efficient use of the network infrastructure to address increasing data and compute requirements of large-scale applications. Since the amount of data and the size of scientific projects are continuously growing, traditional data management techniques are unlikely to support future collaboration systems at the extreme scale. Network-aware data management services for dynamic resource provisioning, end-to-end processing of data, intelligent data-flow and resource coordination are highly desirable. This workshop will seek contribution from academia, government, and industry to discuss emerging trends in use of networking for data management, novel techniques for data representation, simplification of end-to-end data flow, resource coordination, and network-aware tools for the scientific applications. (URL: <http://sdm.lbl.gov/ndm/2012>)

The Third International Workshop on Data-Intensive Computing in the Clouds - DataCloud

9am-5:30pm

Organizers: Tevfik Kosar (University at Buffalo), Ioan Raicu (Illinois Institute of Technology), Roger Barga (Microsoft Corporation)

The third international workshop on Data-Intensive Computing in the Clouds (DataCloud 2012) will provide the scientific community a dedicated forum for discussing new research, development, and deployment efforts in running data-intensive computing workloads on cloud computing infrastructures. This workshop will focus on the use of cloud-based technologies to meet the new data intensive scientific challenges that are not well served by the current supercomputers, grids or compute-intensive clouds. We believe the workshop will be an excellent place to help the community define the current state, determine future goals, and present architectures and services for future clouds supporting data-intensive computing.

Third Annual Workshop on Energy Efficient High Performance Computing - Redefining System Architecture and Data Centers

9am-5:30pm

Organizers: Natalie Bates, Anna Maria Bailey (Lawrence Livermore National Laboratory), Josip Loncaric (Los Alamos National Laboratory), David Martinez (Sandia National Laboratories), Susan Coghlan (Argonne National Laboratory), James Rogers (Oak Ridge National Laboratory)

Building on last year's workshop "Towards and Beyond Energy Efficiency," this year we will dive deeper into the redefinition of system architecture that is required to get to exascale. The workshop starts by looking at the historical trends of power and supercomputing. Then, it will look forward to exascale challenges. After gaining this broad perspective, we will start to focus and drill down on potential solutions to particular aspects of the power challenges. The topics will cover both the data center infrastructure and system architecture. This annual workshop is organized by the Energy Efficient HPC Working Group (<http://eehpcwg.lbl.gov/>). Speakers include Peter Kogge (University of Notre Dame), John Shalf (LBNL), Satoshi Matsuo (Tokyo Institute of Technology), Herbert Huber (Leibniz Supercomputing Centre), Steve Hammond (NREL), Nicolas Dube (HP), Michael Patterson (Intel), and Bill Tschudi (LBNL). They are well known leaders in energy efficiency for supercomputing and promise a lively and informative session.

Monday, November 12

Broader Engagement Workshop

Chair: Tiki L. Suarez-Brown (Florida A&M University)

8:30am-5pm

The Broader Engagement (BE) Program workshop brings an eclectic mix of topics ranging from exascale computing to video-gaming. Some sessions are being organized in association with the HPC Educators Program. Two plenary sessions (8:30am to 10am on November 11 and 12) have been organized to introduce the audience to topics related to cloud computing and exascale computing. The current challenges and future directions in the areas of HPC and Big Data will be also discussed. Participants will receive hands-on experience on OpenMP and directives programming for the accelerators. The main goal of the BE program is to broaden the participation of underrepresented groups in HPC in general and SC conferences in particular. The workshop will provide multiple opportunities and sessions to interact with a diverse group of participants. The workshop participants also are invited to purchase tickets and attend the resource fair organized by the BE and HPC Educators Program.

3rd International Workshop on Performance Modeling, Benchmarking and Simulation of High Performance Computing Systems

9am-5:30pm

Organizers: Stephen Jarvis (University of Warwick), Simon Hammond (Sandia National Laboratories), Pavan Balaji (Argonne National Laboratory), Todd Gamblin (Lawrence Livermore National Laboratory), Darren Kerbyson (Pacific Northwest National Laboratory), Rolf Riesen (IBM Research), Arun Rodrigues (Sandia National Laboratories), Ash Vadgama (AWE Plc), Meghan Wingate McClelland (Los Alamos National Laboratory), Yunquan Zhang (Chinese Academy of Sciences)

This workshop deals with the comparison of HPC systems through performance modeling, benchmarking or through the use of tools such as simulators. We are particularly interested in the ability to measure and make tradeoffs in software/hardware co-design to improve sustained application performance. We are also concerned with the assessment of future systems to ensure continued application scalability through peta- and exascale systems. The aim of this workshop is to bring together researchers, from industry and academia, concerned with the qualitative and quantitative evaluation and modeling of HPC systems. Authors are invited to submit novel research in all areas of performance modeling, benchmarking and simulation, and we welcome research that brings together current theory and practice. We recognize that the coverage of the term ‘performance’ has broadened to include power consumption and reliability, and that performance modeling is practiced through analytical methods and approaches based on software tools and simulators.

3rd SC Workshop on Petascale Data Analytics: Challenges and Opportunities

9am-5:30pm

Organizers: Ranga Raju Vatsavai, Scott Klasky (Oak Ridge National Laboratory), Manish Parashar (Rutgers University)

The recent decade has witnessed a data explosion, and petabyte-sized data archives are not uncommon any more. It is estimated that organizations with high-end computing infrastructures and data centers are doubling the amount of data they are archiving every year. On the other hand, computing infrastructures are becoming more heterogeneous. The first two workshops held with SC10 and SC11 were a great success. Continuing on this success, in addition to the cloud focus, we are broadening the topic of this workshop with an emphasis on middleware infrastructure that facilitates efficient data analytics on big data. This workshop will bring together researchers, developers, and practitioners from academia, government, and industry to discuss new and emerging trends in high-end computing platforms, programming models, middleware and software services and outline the data mining and knowledge discovery approaches that can efficiently exploit this modern computing infrastructure.

5th Workshop on High Performance Computational Finance

9am-5:30pm

Organizers: Andrew Sheppard (Fountainhead), Mikhail Smelyanskiy (Intel Corporation), Matthew Dixon (Thomson Reuters), David Daly, Jose Moreira (IBM)

The purpose of this workshop is to bring together practitioners, researchers, vendors, and scholars from the complementary fields of computational finance and HPC, in order to promote an exchange of ideas, develop common benchmarks and methodologies, discuss future collaborations and develop new research directions. Financial companies increasingly rely on high performance computers to analyze high volumes of financial data, automatically execute trades, and manage risk. Recent years have seen the dramatic increase in compute capabilities across a variety of parallel systems. The systems have also become more complex with trends towards heterogeneous systems consisting of general-purpose cores and acceleration devices. The workshop will enable the dissemination of recent advances and learning in the application of high performance computing to computational finance among researchers, scholars, vendors and practitioners, and to encourage and highlight collaborations between these groups in addressing HPC research challenges.

7th Parallel Data Storage Workshop**9am-5:30pm**

Organizers: John Bent (EMC), Garth Gibson (Carnegie Mellon University)

Peta- and exascale computing infrastructures make unprecedented demands on storage capacity, performance, concurrency, reliability, availability, and manageability. This workshop focuses on the data storage problems and emerging solutions found in peta- and exascale scientific computing environments, with special attention to issues in which community collaboration can be crucial for problem identification, workload capture, solution interoperability, standards with community buy-in, and shared tools. This workshop seeks contributions on relevant topics, including but not limited to: performance and benchmarking, failure tolerance problems and solutions, APIs for high performance features, parallel file systems, high bandwidth storage architectures, wide area file systems, metadata intensive workloads, autonomies for HPC storage, virtualization for storage systems, archival storage advances, resource management innovations, and incorporation of emerging storage technologies.

Climate Knowledge Discovery Workshop**9am-5:30pm**

Organizers: Per Nyberg (Cray Inc.), Reinhard Budich (Max Planck Institute for Meteorology), John Feo (Pacific Northwest National Laboratory), Tobias Weigel (German Climate Computing Center), Karsten Steinhaeuser (University of Minnesota)

As we enter the age of data intensive science, knowledge discovery in simulation-based science rests upon analyzing massive amounts of data. In climate science, model-generated and observational data represent one of the largest repositories of scientific data. Geoscientists gather data faster than they can be interpreted. They possess powerful tools for stewardship and visualization, but not for data intensive analytics to understand causal relationships among simulated events. Such tools will provide insights into challenging features of the earth system, including anomalies, nonlinear dynamics and chaos, with the potential to play a significant role in future IPCC assessments. The breakthroughs needed to address these challenges will come from collaborative efforts involving several disciplines, including end-user scientists, computer and computational scientists, computing engineers, and mathematicians. This workshop brings together research scientists in these diverse disciplines to discuss the design and development of methods and tools for knowledge discovery in climate science.

The 5th Workshop on Many-Task Computing on Grids and Supercomputers**9am-5:30pm**

Organizers: Ioan Raicu (Illinois Institute of Technology), Ian Foster (Argonne National Laboratory), Yong Zhao (University of Electronic Science and Technology of China)

This workshop will provide the scientific community a dedicated forum for presenting new research, development, and deployment efforts of large-scale many-task computing (MTC) applications on large scale clusters, grids, Supercomputers, and cloud computing infrastructure. MTC, the theme of the workshop, encompasses loosely coupled applications which are generally composed of many tasks (both independent and dependent) to achieve some larger application goal. This workshop will cover challenges that can hamper efficiency and utilization in running applications on large-scale systems, such as local resource manager scalability and granularity, efficient utilization of raw hardware, parallel file-system contention and scalability, data management, I/O management, reliability at scale, and application scalability. We welcome paper submissions on all theoretical, simulations, and systems topics related to MTC, but we give special consideration to papers addressing petascale to exascale challenges.

The 7th Workshop on Ultrascale Visualization**9am-5:30pm**

Organizers: Kwan-Liu Ma (University of California, Davis), Venkatram Vishwanath (Argonne National Laboratory), Hongfeng Yu (University of Nebraska-Lincoln)

The output from leading-edge scientific simulations, experiments, and sensors is so voluminous and complex that advanced visualization techniques are necessary to make correct and timely interpretation of the results. Even though visualization technology has progressed significantly in recent years, we are barely capable of exploiting petascale data to its full extent, and exascale datasets are on the horizon. This workshop aims at addressing this pressing issue by fostering communication between visualization researchers and the users of visualization. Attendees will be introduced to the latest and greatest research innovations in large data visualization and analysis and also learn how these innovations impact scientific supercomputing and discovery process.

The Seventh Workshop on Workflows in Support of Large-Scale Science

9am-5:30pm

Organizers: Johan Montagnat (CNRS), Ian Taylor (Cardiff University)

Data Intensive Workflows (a.k.a. scientific workflows) are routinely used in most scientific disciplines today, especially in the context of parallel and distributed computing. Workflows provide a systematic way of describing the analysis and rely on workflow management systems to execute the complex analyses on a variety of distributed resources. This workshop focuses on the many facets of data-intensive workflow management systems, ranging from job execution to service management and the coordination of data, service and job dependencies. The workshop, therefore, covers a broad range of issues in the scientific workflow lifecycle that include: data intensive workflows representation and enactment; designing workflow composition interfaces; workflow mapping techniques that may optimize the execution of the workflow; workflow enactment engines that need to deal with failures in the application and execution environment; and a number of computer science problems related to scientific workflows such as semantic technologies, compiler methods, fault detection and tolerance.

Friday, November 16

Extreme-Scale Performance Tools

8:30am-12:30pm

Organizers: Felix Wolf (German Research School for Simulation Sciences)

As we approach exascale, the rising architectural complexity coupled with severe resource limitations with respect to power, memory and I/O, makes performance optimization more critical than ever before. All the challenges of scalability, heterogeneity, and resilience that application and system developers face will also affect the development of the tool environment needed to achieve performance objectives. This workshop will serve as a forum for application, system, and tool developers to discuss the requirements of future exascale-enabled performance tools and the roadblocks that need to be addressed on the way. The workshop is organized by the Virtual Institute - High Productivity Supercomputing, an international initiative of academic HPC programming-tool builders aimed at the enhancement, integration, and deployment of their products. The event will not only focus on technical issues but also on the community-building process necessary to create an integrated performance-tool suite ready for an international exascale software stack.

Multi-Core Computing Systems (MuCoCoS) Performance Portability and Tuning

8:30am-12:30pm

Organizers: Sabri Pllana (University of Vienna), Jacob Barhen (Oak Ridge National Laboratory)

The pervasiveness of homogeneous and heterogeneous multi-core and many-core processors in a large spectrum of systems, from embedded and general-purpose to high-end computing systems, poses major challenges to software industry. In general, there is no guarantee that software developed for a particular architecture will be executable (i.e., functional) on another architecture. Furthermore, ensuring that the software preserves some aspects of performance behavior (such as temporal or energy efficiency) across different such architectures is an open research issue. This workshop focuses on novel solutions for functional and performance portability as well as automatic tuning across different architectures. The topics of the workshop include but are not limited to: performance measurement, modeling, analysis and tuning; portable programming models, languages and compilation techniques; tunable algorithms and data structures; run-time systems and hardware support mechanisms for auto-tuning; case studies highlighting performance portability and tuning.

Preparing Applications for Exascale Through Co-design

8:30am-12:30pm

Organizers: Lorna Smith, Mark Parsons (Edinburgh Parallel Computing Centre), Achim Basermann (German Aerospace Center), Bastian Koller (High Performance Computing Center Stuttgart), Stefano Markidis (KTH Royal Institute of Technology), Frédéric Magoulès (Ecole Centrale Paris)

The need for exascale platforms is being driven by a set of important scientific drivers. These drivers are scientific challenges of global significance that cannot be solved on current petascale hardware, but require exascale systems. Example grand challenge problems originate from energy, climate, nanotechnology and medicine and have a strong societal focus. Meeting these challenges requires associated application codes to utilize developing exascale systems appropriately. Achieving this requires a close interaction between software and application developers. The concept of co-design dates from the late 18th century, and recognized the importance of a priori knowledge. In modern software terms, co-design recognizes the need to include all relevant perspectives and stakeholders in the design process. With application, software and hardware developers now engaged in co-design to guide exascale development, a workshop bringing these communities together is timely. Authors are invited to submit novel research and experience in all areas associated with co-design.

Python for High Performance and Scientific Computing 8:30am-12:30pm

Organizers: Andreas Schreiber (German Aerospace Center), William Scullin (Argonne National Laboratory)

Python is a high-level programming language with a growing community in academia and industry. It is a general-purpose language adopted by many scientific applications such as computational fluid dynamics, bio molecular simulation, artificial intelligence, statistics and data analysis, scientific visualization etc. More and more industrial domains are turning towards it as well, such as robotics, semiconductor manufacturing, automotive solutions, telecommunication, computer graphics, and games. In all fields, the use of Python for scientific, high performance parallel, and distributed computing, as well as general scripted automation is increasing. Moreover, Python is well-suited for education in scientific computing. The workshop will bring together researchers and practitioners from industry and academia using Python for all aspects of high performance and scientific computing. The goal is to present Python applications from mathematics, science, and engineering, to discuss general topics regarding the use of Python (such as language design and performance issues), and to share experience using Python in scientific computing education.

Sustainable HPC Cloud Computing 2012 8:30am-12:30pm

Organizers: Justin Y. Shi, Abdallah Khreishah (Temple University)

The proposed workshop focuses on HPC cloud computing with an emphasis on practice and experiences, programming methods/models that can tolerate volatile environments, virtualized GPU performance and reliability studies.

The First International Workshop on Data Intensive Scalable Computing Systems (DISCS) 8:30am-12:30pm

Organizers: Yong Chen (Texas Tech University), Xian-He Sun (Illinois Institute of Technology)

HPC is a major strategic tool for science, engineering, and industry. Existing HPC systems, however, are largely designed and developed for computation-intensive applications with a computing-centric paradigm. With the emerging and timely needs of supporting data intensive scientific discovery and innovations, there is a need of rethinking the system architectures, programming models, runtime systems, and tools available for data intensive HPC. This workshop provides a forum for researchers and developers in the high performance computing, data intensive computing, and parallel computing fields to take the Big Data challenges together and present innovative ideas, experiences, and latest developments that help address these challenges. (Visit <http://data.cs.ttu.edu/discs/> for more on this workshop.)

Workshop on Domain-Specific Languages and High-Level Frameworks for High Performance Computing 8:30am-12:30pm

Organizers: Sriram Krishnamoorthy (Pacific Northwest National Laboratory), J. Ramanujam (Louisiana State University), Ponnuswamy Sadayappan (Ohio State University)

Multi-level heterogeneous parallelism and deep memory hierarchies in current and emerging computer systems make their programming very difficult. Domain-specific languages (DSLs) and high-level frameworks (HLFs) provide convenient abstractions, shielding application developers from much of the complexity of explicit parallel programming in standard programming languages like C/C++/Fortran. However, achieving scalability and performance portability with DSLs and HLFs is a significant challenge. For example, very few high-level frameworks can make effective use of accelerators like GPUs and FPGAs. This workshop seeks to bring together developers and users of DSLs and HLFs to identify challenges and discuss solution approaches for their effective implementation and use on massively parallel systems.

Birds of a Feather

Don't just observe, ENGAGE! Birds of a Feather sessions (BOFs) are among the most interactive, popular, and well-attended sessions of the SC conference series. These sessions provide a non-commercial, dynamic venue for conference attendees to openly discuss current topics of focused mutual interest within the HPC community with a strong emphasis on audience-driven discussion, professional networking and grassroots participation. SC12 continues this tradition with a full schedule of exciting, informal, interactive sessions focused around a variety of special topics of mutual interest, including:

- Applications, Languages, and Programming Environments
- Computing, Storage, Networking, and Analysis
- Systems Administration and Data Center Operations
- Innovation in HPC and Emerging Technologies
- Government and Group Initiatives

BOF sessions are an excellent opportunity to connect and interact with other attendees with whom you share a mutual interest.

Birds of a Feather

Birds of a Feather

Tuesday, November 13

ACM SIGHPC First Annual Members Meeting

12:15pm-1:15pm

Room: 155-E

Primary Session Leader: Cherri Pancake (Oregon State University)

ACM SIGHPC (Special Interest Group on High Performance Computing) is the first international group devoted exclusively to the needs of students, faculty, and practitioners in high performance computing. Members and prospective members are encouraged to attend this first annual Members Meeting. SIGHPC officers and volunteers will share what has been accomplished to date, provide tips about resources available to members, and get audience input on priorities for the future. Join us for a lively discussion of what you think is important to advance your HPC activities.

Collaborative Opportunities with the Open Science Data Cloud

12:15pm-1:15pm

Room: 250-DE

Primary Session Leader: Robert Grossman (University of Chicago)
Secondary Session Leaders: Heidi Alvarez (Florida International University)

Scientists in a wide variety of disciplines are producing unprecedented volumes of data that is transforming science. Unfortunately, many scientists are struggling to manage, analyze, and share their medium to large size datasets. The Open Science Data Cloud (OSDC) was developed to fill this gap. It is a cloud-based infrastructure that allows scientists to manage, analyze, integrate and share medium to large size scientific datasets. It is operated and managed by the not-for-profit Open Cloud Consortium. Come to this session to learn more about the OSDC and how you can use the OSDC for your big data research projects.

Data and Software Preservation for Big-Data Science Collaborations

12:15pm-1:15pm

Room: 255-A

Primary Session Leader: Rob Roser (Fermi National Laboratory)
Secondary Session Leaders: Michael Hildreth (University of Notre Dame), Elisabeth M. Long (University of Chicago), Ruth Pordes (Fermi National Laboratory)

The objective of this BOF is to communicate information about and gain useful ideas and input towards data and software preservation for large-scale science communities. The BOF will present information across specific physics, astrophysics and digital library projects of their current goals, status and plans in this area. It will include a focus discussion to expose and explore common challenges, beneficial coordinated activities, and identify related research needs.

Exascale IO Initiative: Progress Status

12:15pm-1:15pm

Room: 155-F

Primary Session Leader: Toni Cortes (Barcelona Supercomputing Center)
Secondary Session Leaders: Peter Braam (Xyratex), André Brinkmann (Johannes Gutenberg University Mainz)

The ELOW intends to architect and ultimately implement an open source, upper level I/O middleware system suitable for exa-scale storage. It intends to be primarily motivated by the requirements of the applications, management and system architectures, and to a lesser extent by the constraints and traditions of the storage industry. The resulting middleware system is targeting adoption in the HPC community by creation of or integration into various HPC software components, such as libraries. We also target adoption by storage vendors to layer this on either existing or new products targeting scalable high performance storage.

Fifth Graph500 List

12:15pm-1:15pm

Room: 255-BC

Primary Session Leader: David A. Bader (Georgia Institute of Technology)
Secondary Session Leaders: Richard Murphy (Sandia National Laboratories), Marc Snir (Argonne National Laboratory)

Data intensive applications represent increasingly important workloads but are ill suited for most of today's machines. The Graph500 has demonstrated the challenges of even simple analytics. Backed by a steering committee of over 30 international HPC experts from academia, industry, and national laboratories, this effort serves to enhance data intensive workloads for the community. This BOF will unveil the fifth Graph500 list, and delve into the specification for the second kernel. We will further explore the new energy metrics for the Green Graph500, and unveil the first results.

Genomics Research Computing: The Engine that Drives Personalized Medicine Forward

12:15pm-1:15pm

Room: 251-A

Primary Session Leader: Christine Fronczak (Dell)

The purpose of this BOF is to organize the members in the HPC community interested in optimizing bioinformatics algorithms and HPC architectures for use in genomics analysis. We will discuss ongoing performance optimization efforts of these applications at the Virginia Bioinformatics Institute, Translational Genomics Institute and Dell. The goals of the BOF are to make researchers aware of these open source efforts, request application recommendations and development priorities, and invite them to join in development and testing. The current and future roles of coprocessors, e.g. MIC, GPUs and FPGAs, in HPC ecosystems for genomics will be discussed.

HDF5: State of the Union

12:15pm-1:15pm

Room: 250-C

Primary Session Leader: Quincey Koziol (HDF Group)

A forum for HDF5 developers and users to interact. HDF5 developers will describe the current status of HDF5 and discuss future plans, followed by an open discussion.

How the Government can Enable HPC and Emerging Technologies

12:15pm-1:15pm

Room: 355-D

Primary Session Leader: Ron Bewtra (NOAA)

Can the US Government help address the “missing middle”; enable emerging technologies; and promote innovation? Are there better ways to run acquisitions, incentivize vendors, and promote novel approaches to address the US’ computing needs. Come join leading experts, Government leaders, and industry advisors to discuss what can and should (and even should not!) be done.

Implementing Parallel Environments: Training and Education

12:15pm-1:15pm

Room: 251-D

Primary Session Leader: Charles Peck (Earlham College)
Secondary Session Leaders: Tom Murphy (Contra Costa College), Clay Breshears (Intel Corporation)

This BoF, co-hosted by Educational Alliance for a Parallel Future (EAPF), a community of industry, academics, research and professional organizations which have a stake in helping to ensure that parallel computing is integrated through the undergraduate and graduate curriculum, will be focused on

the challenges and opportunities created by the manycore, heterogeneous compute platforms that are now ubiquitous. We will discuss the kinds of training and retraining that will be necessary as parallel computing comes to terms with the rapid advancement of exascale computing hardware.

Interoperability in Scientific Cloud Federations

12:15pm-1:15pm

Room: 250-AB

Primary Session Leader: Christine Morin (INRIA)
Secondary Session Leaders: Kate Keahey (Argonne National Laboratory), Yvon Jegou, Roberto Cascellla (INRIA)

The uptake of cloud computing has as major obstacle in the heterogeneity of hardware and software, which make difficult the portability of applications and services. Interoperability among cloud providers is the only way to avoid vendor lock-in and open the way toward a more competitive market. Interoperability can be achieved either by using open standards and protocols or by a middleware service to adapt the application/service to a specific cloud provider. The audience will be guided through the major challenges for interoperability from the IaaS to PaaS model and discuss the potential approaches for the interoperability in scientific cloud federations.

MPICH: A High-Performance Open-Source MPI Implementation

12:15pm-1:15pm

Room: 155-B

Primary Session Leader: Darius Buntinas (Argonne National Laboratory)
Secondary Session Leaders: Pavan Balaji (Argonne National Laboratory), Rajeev Thakur (Argonne National Laboratory)

MPICH 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 MPICH as well as developers of MPI implementations derived from MPICH to discuss experiences and issues in using and porting MPICH. Future plans for MPI-3 support will be discussed. Representatives from MPICH-derived implementations will provide brief updates on the status of their efforts. MPICH developers will also be present for an open forum discussion.

Network Measurement

12:15pm-1:15pm

Room: 255-EF

Primary Session Leader: Jon Dugan (Energy Sciences Network)
Secondary Session Leaders: Aaron Brown (Internet2)

Networks are critical to high performance computing: they play a crucial role both within the data center and in providing

access to remote resources. It is imperative that these networks perform optimally. In order to understand the behavior and performance of these networks, they must be measured. This is a forum discussing network measurement, particularly as it relates to HPC. There will be presentations from experts in network performance measurement as well as time for questions, discussion and impromptu presentations. Potential topics include (but are not limited to) measurement tools, measurement frameworks, visualization, emerging standards and current research.

Obtaining Bitwise Reproducible Results - Perspectives and Latest Advances

12:15pm-1:15pm

Room: 251-E

Primary Session Leader: Kai Diethelm (Gesellschaft für numerische Simulation mbH)

Secondary Session Leader: Noah Clemons (Intel Corporation)

It is a widely known HPC reality that many optimizations and scheduling techniques require a change in the order of operations, creating results that are not bitwise reproducible (BWR). This BOF will bring together members of the HPC community who are affected by the (non-) reproducibility phenomenon in various ways. A number of leading experts from numerical software tools, academic/military, and commercial software development will present their points of view on the issue in short presentations and discuss the implications with the audience.

OpenACC API Status and Future

12:15pm-1:15pm

Room: 255-D

Primary Session Leader: Michael Wolfe (The Portland Group, Inc.)

Secondary Session Leader: Rob Farber (BlackDog Endeavors, LLC.)

The OpenACC API for programming host+accelerator systems was introduced at SC11. This session will present the status of current implementations by vendors, and give experiences by early users. It will close with updates on the specification and suggestions for future directions. If your current or future system includes accelerators such as NVIDIA GPUs, AMD GPUs or APUs, Intel MIC, or something else, come to discuss the leading high level programming standard for your system.

Parallel and Accelerated Computing Experiences for Successful Industry Careers in High-Performance Computing

12:15pm-1:15pm

Room: 251-F

Primary Session Leader: Eric Stahlberg (SAIC-Frederick / Frederick National Laboratory for Cancer Research)

Secondary Session Leaders: Melissa Smith (Clemson University), Steven Bogaerts (Wittenberg University)

Experience and knowledge of parallel and accelerated computing have become essential to successful careers involving high-performance computing. Yet, there remains an important gap between the educational experience of undergraduate students and the real needs of industry and academic research programs. Session organizers have been working to develop methods to bridge this gap across a spectrum of computer and computational science courses. Building upon recent classroom and internship experiences, this session will bring together industry and academia in a common forum to discuss and share experiences for students that will help close the gap between preparation and application of high-performance computing.

Python for High Performance and Scientific Computing

12:15pm-1:15pm

Room: 155-C

Primary Session Leader: Andreas Schreiber (German Aerospace Center)

Secondary Session Leaders: William R. Scullin (Argonne National Laboratory), Andy R. Terrel (Texas Advanced Computing Center)

This BoF is a forum for presenting projects, ideas and problems. Anyone can present short lightning talks (5 minutes each). The goal is to get in contact with other colleagues for further discussion and joint activities. All presentations should be related to Python in some way, for example: introduction of existing software using Python for HPC applications, experience reports with advantages or drawbacks of Python for HPC, announcements of events related to Python and HPC, proposals for projects where Python plays a role, or request for collaboration and search for partners.

Scalable Adaptive Graphics Environment (SAGE) for Global Collaboration

12:15pm-1:15pm

Room: 251-C

Primary Session Leader: Jason Leigh (University of Illinois at Chicago)

Secondary Session Leader: Maxine Brown (University of Illinois at Chicago)

SAGE, the Scalable Adaptive Graphics Environment, receives funding from the National Science Foundation to provide persistent visualization and collaboration services for global cyberinfrastructure. It is a widely used open-source platform and

the scientific community's defacto software operating environment, or framework, for managing content on scalable-resolution tiled display walls. The SC BOF provides an unparalleled opportunity for the SAGE global user community, and potential users, to meet and discuss current and future development efforts, and to share examples of community-developed use cases and applications.

System-wide Programming Models for Exascale

12:15pm-1:15pm

Room: 355-BC

Primary Session Leader: Kathryn O'Brien (IBM Research)

Secondary Session Leader: Bronis de Supinsky (Lawrence Livermore National Laboratory)

The challenge of programmability is acknowledged as a fundamental issue in the ongoing design of future exascale systems. The perceived notion that heterogeneity will be a key ingredient for node level architectures has intensified the debate on the most appropriate approach for portable and productive programming paradigms. However, the continued focus on node level programming models means less attention is being paid to the need for a more radical rethinking of higher level, system wide programming approaches. This BoF will feature presentations on application requirements and future, system-wide programming models that address them for exascale.

The 2012 HPC Challenge Awards

12:15pm-1:15pm

Room: 355-A

Primary Session Leader: Piotr Luszczek (University of Tennessee, Knoxville)

Secondary Session Leaders: Jeremy Kepner (MIT Lincoln Laboratory)

The 2012 HPC Challenge Awards BOF is the 8th edition of an award ceremony that awards high performance results in broad categories taken from the HPC Challenge benchmark as well as elegance and efficiency of parallel programming and execution environments. The performance results are taken from the HPCC public database of submitted results that are unveiled at the time of BOF. The competition for the most elegant and efficient code takes place during the BOF and is judged on the spot with winners revealed at the very end of the BOF. Judging and competing activities are interleaved to save time.

Computing Research Testbeds as a Service: Supporting Large-scale Experiments and Testing

5:30pm-7pm

Room: 251-E

Primary Session Leader: Geoffrey Fox (Indiana University)

Secondary Session Leader: José A.B. Fortes (University of Florida)

This BOF discusses the concept of a Computing Testbed as a Service supporting application, computer science, education and technology evaluation usages that have different requirements from production jobs. We look at lessons from projects like Grid5000, FutureGrid, OpenCirrus, PlanetLab and GENI. We discuss 1) the requirements that Computing Testbeds as a Service need to address 2) The software needed to support TestbedaaS and a possible open source activity and 3) interest in federating resources to produce large scale testbeds and what commitments participants may need to make in such a federation.

Critically Missing Pieces in Heterogeneous Accelerator Computing

5:30pm-7pm

Room: 155-A

Primary Session Leader: Pavan Balaji (Argonne National Laboratory)

Secondary Session Leader: Satoshi Matsuoka (Tokyo Institute of Technology)

Heterogeneous architectures play a massive role in architecting the largest systems in the world. However, much of the interest in these architectures is an artifact of the hype associated with them. For such architectures to truly be successful, it is important that we look beyond this hype and learn what these architectures provide and what is critically missing. This continuing BoF series brings together researchers working on aspects of accelerator architectures---including data management, resilience, programming, tools, benchmarking, and auto-tuning---to identify critical gaps in the accelerator ecosystem.

Cyber Security's Big Data, Graphs, and Signatures

5:30pm-7pm

Room: 250-AB

Primary Session Leader: Daniel M. Best (Pacific Northwest National Laboratory)

Cyber security increases in complexity and network connectivity every day. Today's problems are no longer limited to malware using hash functions. Interesting problems, such as coordinated cyber events, involve hundreds of millions to billions of nodes and similar or more edges. Nodes and edges go beyond single attribute objects to become multivariate entities depicting complex relationships with varying degree of impor-

tance. To unravel cyber security's big data, novel and efficient algorithms are needed to investigate graphs and signatures. We bring together domain experts from various research communities to talk about current techniques and grand challenges being researched to foster discussion.

Energy Efficient High Performance Computing

5:30pm-7pm

Room: 155-C

Primary Session Leader: Simon McIntosh-Smith (University of Bristol)

Secondary Session Leader: Kurt Keville (Massachusetts Institute of Technology)

At SC'11 we launched the Energy Efficient HPC community at the first EEHPC BOF. This brought together researchers evaluating the use of technologies from the mobile and embedded spaces for use in HPC. One year on much progress has been made, with the launch of several server systems based on low power consumer processors. Several EEHPC systems are in development and advances have been made in HPC software stacks for mobile processors. This BOF will look at the EEHPC report card one year on, review what progress has been made, and identify where there are still challenges to be met.

Exascale Research – The European Approach

5:30pm-7pm

Room: 255-A

Primary Session Leader: Alex Ramirez (Barcelona Supercomputing Center)

Secondary Session Leader: Hans-Christian Hoppe, Marie-Christine Sawley (Intel Corporation)

To deliver Exascale performance before 2020 poses serious challenges and requires combined action to be taken now. Europe is undertaking a set of leading initiatives to accelerate this progress. On one hand, the European Commission is funding its own Exascale initiative, currently represented by three complementary research projects focusing on hardware-software co-design, scalable system software, and novel system architectures. On the other hand, Intel is conducting a number of collaborative research activities through its European Exascale Labs. The session will present the advances and results from these initiatives and open the floor to discuss the results and the approach taken.

High Performance Computing Programming Techniques For Big Data Hadoop

5:30pm-7pm

Room: 251-F

Primary Session Leader: Gilad Shainer (HPC Advisory Council)

Hadoop MapReduce and High Performance Computing share many characteristics such as, large data volumes, variety of data types, distributed system architecture, required linear performance growth with scalable deployment and high CPU utilization. RDMA capable programming models enable efficient data transfers between computation nodes. In this session we will discuss a collaborative work done among several industry and academic partners on porting Hadoop MapReduce framework to RDMA, the challenges, the techniques used, the benchmarking and testing.

High Productivity Languages for High Performance Computing

5:30pm-7pm

Room: 251-B

Primary Session Leader: Michael McCool (Intel Corporation)

Two kinds of languages promise high productivity in scientific applications: scripting languages and domain specific languages (DSLs). Scripting languages typically resolve details such as typing automatically to simplify software development, while domain specific languages include features targeting specific application classes and audiences. It is possible for both kinds of languages to be used for building high-performance applications. The implementation technology for scripting languages has improved tremendously, and DSLs are often built on top of a scripting language infrastructure. This BOF will bring together developers and users of scripting languages so that recent developments can be presented and discussed.

High-level Programming Models for Computing Using Accelerators

5:30pm-7pm

Room: 250-DE

Primary Session Leader: Duncan Poole (NVIDIA)

Secondary Session Leader: Douglas Miles (The Portland Group, Inc.)

There is consensus in the community that higher-level programming models based on directives or language extensions have significant value for enabling accelerator programming by domain experts. The past year has brought the introduction of several directive-based GPU and co-processor compilers, more progress within the OpenMP accelerator sub-committee, and continued development of GPU language extensions by large industry players and academics. This BoF will provide an overview of the status of various implementations, and explore and debate the merits of the current options and approaches for high-level heterogeneous manycore programming.

HPC Cloud: Can Infrastructure Clouds Provide a Viable Platform for HPC?

5:30pm-7pm

Room: 355-BC

Primary Session Leader: Kate Keahey (Argonne National Laboratory)

Secondary Session Leaders: Franck Cappello (University of Illinois at Urbana-Champaign), Peter Dinda (Northwestern University), Dhabaleswar Panda (Ohio State University), Lavanya Ramakrishnan (Lawrence Berkeley National Laboratory)

Clouds are becoming an increasingly popular infrastructure option for science but despite achieving significant milestones their viability for HPC applications is still in question. This BOF is designed to bring together the following communities: (1) the HPC community (2) scientists evaluating HPC workloads on infrastructure clouds, and (3) scientists creating technology that would make HPC clouds possible/efficient. The objective of the BOF is to assess the state of the art in the area as well as to formulate crisp challenges for HPC clouds (as in: “Clouds will become a viable platform for HPC when...”).

HPC Runtime System Software

5:30pm-7pm

Room: 255-EF

Primary Session Leader: Rishi Khan (E.T. International, Inc.)

Secondary Session Leader: Thomas Sterling (Indiana University)

Future extreme-scale systems may require aggressive runtime strategies to achieve both high efficiency and high scalability for continued performance advantage through Moore’s Law. Exploitation of runtime information will support dynamic adaptive techniques for superior resource management and task scheduling through introspection for better system utilization and scaling. This BoF brings together runtime system software developers, application programmers, and members of HPC community to discuss the opportunities and challenges of using new runtime system software. Presentations will discuss how runtime systems (OCR, ParalleX-HPX, and SWARM) can handle heterogeneity of architectures, memory/network subsystems, energy consumption, and continued execution when faults occur.

Hybrid Programming with Task-based Models

5:30pm-7pm

Room: 251-C

Primary Session Leader: Bettina Krammer (Université de Versailles St-Quentin-en-Yvelines / Exascale Computing Research)

Secondary Session Leaders: Rosa M. Badia (Barcelona Supercomputing Center), Christian Terboven (RWTH Aachen University)

OpenMP, cilk, OmpSs) from the point of view of application developers and runtime developers, alternating with discussion. Most programmers are aware of the fact that, with the advent of manycore processors and hardware accelerators, hybrid programming models may exploit underlying (heterogeneous) hardware better and deliver higher performance than pure MPI codes. In this BoF we give a forum to application and runtime developers presenting and discussing different approaches combining MPI with a task-based programming model such as OpenMP, cilk, or OmpSs.

Large-Scale Reconfigurable Supercomputing

5:30pm-7pm

Room: 255-D

Primary Session Leader: Martin Herbordt (Boston University)

Secondary Session Leaders: Alan George, Herman Lam (University of Florida)

Reconfigurable computing is characterized by hardware that adapts to match the needs of each application, offering unique advantages in speed per unit energy. With a proven capability of 2 PetaOPS at 12KW, large-scale reconfigurable supercomputing has an important role to play in the future of high-end computing. The Novo-G in the NSF CHREC Center at the University of Florida is rapidly moving towards production utilization in various scientific and engineering domains. This BOF introduces the architecture of such systems, describes applications and tools being developed, and provides a forum for discussing emerging opportunities and issues for performance, productivity, and sustainability.

Managing Big Data: Best Practices from Industry and Academia

5:30pm-7pm

Room: 255-BC

Primary Session Leader: Steve Tuecke (University of Chicago)

Secondary Session Leaders: Vas Vasiladis (University of Chicago), William Mannel (SGI), Rachana Ananthakrishnan, Ravi Madduri, Raj Kettimuthu (Argonne National Laboratory)

We will explore the challenges faced by companies and scientific researchers managing “big data” in increasingly diverse IT environments including local clusters, grids, clouds, and supercomputers, and present architectural and technology options for solving these challenges, informed by use cases

from commercial organizations and academic research institutions. Participants will discuss key questions related to data management, and identify innovative technologies and best practices that researchers should consider adopting. Organizations including Ebay/Paypal, NASA, SGI and IDC will present commercial use cases; the University of Chicago and Argonne National Laboratory will discuss examples from bioinformatics, imaging based science, and climate change research.

OCI-led Activities at NSF

5:30pm-7pm

Room: 155-E

Primary Session Leader: Daniel S. Katz (National Science Foundation)

Secondary Session Leader: Barry Schneider (National Science Foundation)

This BOF will inform the SC community about NSF activities that are being led by OCI. OCI's areas include Advanced Computing Infrastructure (ACI), Computational and Data-Enabled Science (CDSE), Data, Learning & Workforce Development, Networking, and Software. A number of these areas have developed vision documents, which will be discussed, in addition to discussion of open programs and solicitations, and open challenges.

OpenMP: Next Release and Beyond

5:30pm-7pm

Room: 355-A

Primary Session Leader: Barbara M. Chapman (University of Houston)

Secondary Session Leaders: Bronis R. de Supinski (Lawrence Livermore National Laboratory), Matthijs van Waveren (Fujitsu Laboratories Ltd.)

Now celebrating its 15th birthday, OpenMP has proven to be a simple, yet powerful model for developing multi-threaded applications. OpenMP continues to evolve, adapt to new requirements, and push at the frontiers of parallelization. A comment draft of the next specification version, which includes several significant enhancements, will be released at this BoF. We will present this draft as well as plans for the continued evolution of OpenMP. A lively panel discussion will critique the draft. We will solicit audience opinions on what features are most needed for future specifications. Ample time will be allowed for attendee participation and questions.

Policies and Practices to Promote a Diverse Workforce

5:30pm-7pm

Room: 253

Primary Session Leader: Fernanda Foertter (Oak Ridge National Laboratory)

Secondary Session Leaders: Rebecca Hartman-Baker (Interactive Virtual Environments Centre), Hai Ah Nam, Judy Hill (Oak Ridge National Laboratory)

Organizations compete to recruit and retain the best talent to achieve its mission focus, whether in academia, industry or federal agencies. Demographics show an increasingly diverse workforce, but often workplace policies lag behind and can be exclusionary. Not surprisingly, successful policies and practices that encourage diversity are beneficial to all employees, as they convey the organization's recognition that people are its most valuable asset. Clear policies that are equally applied encourage diversity in the workplace. In this BOF, participants will share management strategies and experiences that have been successful at creating a diverse environment and are productive for all employees.

Scientific Application Performance in Heterogeneous Supercomputing Clusters

5:30pm-7pm

Room: 155-F

Primary Session Leader: Wen-mei Hwu (University of Illinois at Urbana-Champaign)

Secondary Session Leaders: Jeffrey Vetter (Oak Ridge National Laboratory), Nacho Navarro (Polytechnic University of Catalonia)

Many current and upcoming supercomputers are heterogeneous CPU-GPU computing clusters. Accordingly, applications groups are porting scientific applications and libraries to these heterogeneous supercomputers. Industry vendors have also been actively collaborating with system teams as well as application groups. Although these teams have been working on diverse applications and targeting different systems, countless shared lessons and challenges exist. With this BOF, we aim to bring together system teams and application groups to discuss their experiences, results, lessons, and challenges to date. We hope to form a collaborative community moving forward.

SPEC HPG Benchmarks For Next Generation Systems**5:30pm-7pm****Room: 155-B***Primary Session Leader: Kalyan Kumaran (Argonne National Laboratory)**Secondary Session Leader: Matthias Mueller (Technical University Dresden)*

The High Performance Group (HPG) at the Standard Performance Evaluation Corporation (SPEC) is a standard consortium comprising of vendors, research labs and universities. It has a long history of producing standard benchmarks with metrics for comparing the latest in HPC systems and software. The group has successfully released, or is working on, benchmarks, and supporting run rules to measure performance of OpenCL, OpenACC, OpenMP, and MPI programming models and their implementation in various science applications running on hybrid (CPU+accelerator) architectures. This BOF will include presentations on the existing benchmarks and results and a discussion on future benchmarks and metrics.

The Apache Software Foundation, Cyberinfrastructure, and Scientific Software: Beyond Open Source**5:30pm-7pm****Room: 251-A***Primary Session Leader: Marlon Pierce (Indiana University)**Secondary Session Leader: Suresh Marru (Indiana University), Chris Mattmann (NASA Jet Propulsion Laboratory)*

Many cyberinfrastructure and scientific software projects are adopting free and open source software practices and licensing, but establishing an open community around the software requires additional thought. The key to building and running these communities is a well-chosen governance model. In this Birds of a Feather session, we examine the Apache Software Foundation as a governance model for open source scientific software communities. The BOF will discuss both general and specific governance requirements for research software communities through interactive discussions motivated by case study presentations. The outcome will be a summary white paper co-authored by BOF volunteers.

The Eclipse Parallel Tools Platform**5:30pm-7pm****Room: 250-C***Primary Session Leader: Beth R. Tibbitts (IBM)**Secondary Session Leaders: Greg Watson (IBM)*

The Eclipse Parallel Tools Platform (PTP, <http://eclipse.org/ptp>) is an open-source project providing a robust, extensible workbench for the development of parallel and scientific codes. PTP makes it easier to develop, build, run, optimize, and debug parallel codes on a variety of remote clusters using a single unified interface. PTP includes support for MPI, OpenMP, UPC, Fortran, and other libraries as well.

The BOF will consist of brief demos and discussions about PTP, and an overview of upcoming features. Information from contributors and vendors is welcome concerning integrations with PTP. How to get involved in the PTP project will be covered.

TOP500 Supercomputers**5:30pm-7pm****Room: Ballroom-EFGH***Primary Session Leader: Erich Strohmaier (Lawrence Berkeley National Laboratory)*

The TOP500 list of supercomputers serves as a “Who’s Who” in the field of HPC. It started as a list of the most powerful supercomputers in the world and has evolved to a major source of information about trends in HPC. The 40th TOP500 list will be published in November 2012. This BoF will present detailed analyses of the TOP500 and discuss the changes in the HPC marketplace during the past years. The BoF is meant as an open forum for discussion and feedback between the TOP500 authors and the user community.

TORQUE, RPMs, Cray and MIC**5:30pm-7pm****Room: 251-D***Primary Session Leader: Kenneth Nielson (Adaptive Computing)**Secondary Session Leaders: David Beer (Adaptive Computing), Michael Jennings (Lawrence Berkeley National Laboratory)*

This BoF will update the user community concerning changes in the TORQUE tarball to better support RPM installs, the new simplified support for Cray, and plans to support the Intel MIC architecture.

XSEDE User Meeting**5:30pm-7pm****Room: 355-D***Primary Session Leader: John Towns (National Center for Supercomputing Applications)**Secondary Session Leader: Glenn Brook (National Institute for Computational Sciences)*

The Extreme Science and Engineering Discovery Environment (XSEDE) is the most advanced, collection of integrated advanced digital resources and services in the world. It is a single virtual system that scientists can use to interactively share computing resources, data, and expertise. This BOF brings together users, potential users, and developers of XSEDE services for an open, candid discussion about the present state and future plans of XSEDE. This BOF explores user opinions of XSEDE activities and solicits feedback on the XSEDE user experience. The session concludes with an open discussion of topics raised by participants in a town-hall format.

Wednesday, November 14

Architecture and Systems Simulators

12:15pm-1:15pm

Room: 155-A

Primary Session Leader: Stephen Poole (Oak Ridge National Laboratory)

Secondary Session Leader: Bruce Childers (University of Pittsburgh)

The goal of this BOF is to engage academia, government and industry to make sure that there is an open framework allowing interoperability of simulation tools for Computer Architecture. It is important to bring together and build a community of researchers and implementers of architecture simulation, emulation, and modeling tools. This BOF will inform Supercomputing attendees and solicit their involvement and feedback in an effort to build an interoperable, robust, community-supported simulation and emulation infrastructure.

Building an Open Community Runtime (OCR) framework for Exascale Systems

12:15pm-1:15pm

Room: 255-EF

Primary Session Leader: Vivek Sarkar (Rice University)

Secondary Session Leaders: Barbara Chapman (University of Houston), William Gropp (University of Illinois at Urbana-Champaign)

Exascale systems will impose a fresh set of requirements on runtime systems that include targeting nodes with hundreds of homogeneous and heterogeneous cores, as well as energy, data movement and resiliency constraints within and across nodes. The goal of this proposed BOF session is to start a new community around the development of open runtime components for exascale systems that we call the Open Community Runtime (OCR). Our hope is that OCR components will help enable community-wide innovation in programming models above the OCR level, in hardware choices below the OCR level, and in runtime systems at the OCR level.

Chapel Lightning Talks 2012

12:15pm-1:15pm

Room: 255-A

Primary Session Leader: Sung-Eun Choi (Cray Inc.)

Secondary Session Leader: Bradford L. Chamberlain (Cray Inc.)

Are you a scientist considering a modern high-level language for your research? Are you a language enthusiast who wants to stay on top of new developments? Are you wondering what the future of Chapel looks like after the end of the DARPA HPCS program? Then this is the BOF for you! In this BOF, we will hear "lightning talks" on community activities involving

Chapel. We will begin with a talk on the state of the Chapel project, followed by a series of talks from the broad Chapel community, wrapping up with Q&A and discussion.

Early Experiences Debugging on the Blue Gene/Q

12:15pm-1:15pm

Room: 155-E

Primary Session Leader: Chris Gottbrath (Rogue Wave Software)

Secondary Session Leaders: Dong Ahn Ahn (Lawrence Livermore National Laboratory)

This BOF will highlight user experiences debugging on the new Blue Gene/Q architecture. Existing programs may need to be changed to take full advantage of the new architecture or the new architecture may allow bugs that haven't previously shown up to manifest. The speakers will share experiences and insights from the process of porting applications over to the new Blue Gene/Q. Blue Gene/Q users Dong Ahn (Lawrence Livermore National Labs), Ray Loy (Argonne National Labs), and Bernd Mohr (Jülich Supercomputer Center) have expressed interest in speaking. Other community members interested in reserving time in this session should contact Chris. Gottbrath@roguewave.com.

International Collaboration on System Software Development for Post-petascale Computing

12:15pm-1:15pm

Room: 355-BC

Primary Session Leader: William Harrod (DOE Office of Advanced Scientific Computing Research)

Secondary Session Leaders: Takahiro Hayashi (Japanese Ministry of Education, Culture, Sports, Science and Technology), Yutaka Ishikawa (University of Tokyo)

The US DOE (Department of Energy) and MEXT (Ministry of Education, Culture, Sports, Science and Technology), Japan have agreed to pursue cooperation between the U.S. and Japan on system software for post-petascale computing, including collaborative R&D and international standardization of system software. Standardization is a double-edged sword: it can facilitate interactions but might hinder innovation. Discussions are needed to determine how such international collaborations should be managed to ensure that synergies are achieved without slowing down research. The purpose of this BOF is to discuss the benefits and issues of such international collaboration, and obtain feedback from the community.

Open MPI State of the Union**12:15pm-1:15pm****Room: 155-B**

Primary Session Leader: Jeffrey M. Squyres (Cisco Systems)
Secondary Session Leaders: George Bosilca (University of Tennessee, Knoxville)

MPI-3.0 is upon us. The Open MPI community members have been heavily involved in the MPI Forum. Come hear what we have done both in terms of standardization and implementation of the new MPI-3.0 specification. Open MPI community's unique blend of academics, researchers, system administrators, and vendors provide many different viewpoints for what makes an MPI implementation successful. Work is ongoing in many areas that are directly applicable to real-world HPC applications and users. Join us at the BOF to hear a state of the Union for Open MPI. New contributors are welcome!

PGAS: The Partitioned Global Address Space Programming Model**12:15pm-1:15pm****Room: 355-EF**

Primary Session Leader: Tarek El-Ghazawi (George Washington University)
Secondary Session Leader: Lauren Smith (US Government)

PGAS, the Partitioned Global Address Space, programming model can provide ease-of-use through a global address space while emphasizing performance through locality awareness. For this, the PGAS model has been gaining rising attention. A number of PGAS languages such as UPC, CAF, Chapel and XO are either becoming available on high-performance computers or active research areas. In addition, modern multicore chips are exhibiting NUMA effects, as core count increases, which requires locality aware parallel programming even at the chip level. This BoF will bring together developers, researchers as well as current and potential users for exchange of ideas and information.

PRObE: A 1000-Node Facility for Systems Infrastructure Researchers**12:15pm-1:15pm****Room: 255-BC**

Primary Session Leader: Garth Gibson (Carnegie Mellon University)

The NSF-funded Parallel Reconfigurable Observational Environment (PRObE) (www.newmexicoconsortium.org/probe) facility is making thousands of computers available to systems researchers for dedicated use in experiments that are not compelling at a smaller scale. Using retired equipment donated by DOE and Los Alamos National Laboratory, two staging clusters are available now (marmot.nmc-probe.org and denali.nmc-probe.org) and a 1024 node cluster (Kodiak)

will be available by SC12. Using Emulab (www.emulab.net) researchers will have complete control of all software and hardware while running experiments for days. PRObE encourages systems researchers to attend this BoF and communicate your needs and interests.

Science-as-a-Service: Exploring Clouds for Computational and Data-Enabled Science and Engineering**12:15pm-1:15pm****Room: 155-C**

Primary Session Leader: Manish Parashar (Rutgers University)
Secondary Session Leaders: Geoffrey Fox (Indiana University), Kate Keahey (Argonne National Laboratory), David Lifka (Cornell University)

Clouds are rapidly joining high-performance computing system, clusters and grids as viable platforms for scientific exploration and discovery. As a result, understanding application formulations and usage modes that are meaningful in such a hybrid infrastructure, and how application workflows can effectively utilize it, is critical. This BoF will explore how clouds can be effectively used to support real-world science and engineering applications, and will discuss key research challenges (from both, a computer science as well as an applications perspective) as well as a community research agenda.

Setting Trends for Energy Efficiency**12:15pm-1:15pm****Room: 250-AB**

Primary Session Leader: Natalie Bates (Energy Efficient HPC Working Group)
Secondary Session Leader: Wu Feng (Virginia Tech), Erich Strohmaier (Lawrence Berkeley National Laboratory)

You can only improve what you can measure, but that can be easier said than done.

The Green500, Top500, Green Grid and Energy Efficient HPC Working Group are collaborating on specifying workloads, methodologies and metrics for measuring the energy efficiency of supercomputing systems for architecture design and procurement decision-making processes. A current focus on this collaboration is to improve the methodology for measuring energy in order to get an 'apples to apples' comparison between system architectures. An improved methodology is under development and beta testing. This BoF will report on and review the results of the beta tests.

The Way Forward: Addressing the Data Challenges for Exascale Computing**12:15pm-1:15pm****Room: 355-A***Primary Session Leader: Scott Klasky (Oak Ridge National Laboratory)**Secondary Session Leaders: Hasan Abbasi (Oak Ridge National Laboratory)*

This BOF will bring together application scientists, middleware researchers, and analysis and visualization experts to discuss the major challenges in the coordinated development of new techniques to meet the demands of exascale computing. In situ and in transit computations in a data pipeline has been proposed as an important abstraction in the next generation of I/O systems. We intend to explore the impact of this paradigm on algorithms and applications, as well as the evolution of middleware to achieve these goals.

Unistack: Interoperable Community Runtime Environment for Exascale Systems**12:15pm-1:15pm****Room: 355-D***Primary Session Leader: Pavan Balaji (Argonne National Laboratory)**Secondary Session Leaders: Laxmikant Kale (University of Illinois at Urbana-Champaign)*

This BoF session will discuss a new community initiative to develop a unified runtime infrastructure that is capable of supporting multiple programming models in an interoperable and composable manner. The session will spend some time on presentations by MPI, Charm++, Global Arrays, UPC/CAF runtime developers, describing the current state of practice in this area. The rest of the session will be kept open for discussions and feedback from the larger community.

XSEDE Metrics on Demand (XDMoD) Technology Auditing Framework**12:15pm-1:15pm****Room: 250-C***Primary Session Leader: Thomas R. Furlani (SUNY Buffalo)**Secondary Session Leaders: Matthew D. Jones (SUNY Buffalo), Steven M. Gallo (SUNY Buffalo)*

XSEDE Metrics on Demand (XDMoD) is an open-source tool designed to audit and facilitate the utilization of XSEDE cyber-infrastructure, providing a wide range of metrics on XSEDE resources and services. Currently supported metrics include allocations and computing utilization, allowing a comprehensive view of both current and historical utilization, and scientific/engineering application profiling (via application kernels) for quality of service. XDMoD (<https://xdmod.ccr.buffalo.edu>) uses a role-based scheme to tailor the presentation of

information to the public, individual users, principal investigators, service providers, campus champions, and program managers. At this BOF the current state of XDMoD will be demonstrated and discussed.

Application Grand Challenges in the Heterogeneous Accelerator Era**5:30pm-7pm****Room: 355-BC***Primary Session Leader: Satoshi Matsuoka (Tokyo Institute of Technology)**Secondary Session Leaders: Pavan Balaji (Argonne National Laboratory)*

Accelerators have gained prominence as the next disruptive technology with a potential to provide a non-incremental jump in performance. However, the number of applications that have actually moved to accelerators is still limited because of many reasons, arguably the biggest of which is the gap in understanding between accelerator and application developers. This BoF is an application oriented session that aims to bring the two camps of application developers and accelerator developers head-to-head.

Co-design Architecture and Co-design Efforts for Exascale: Status and Next Steps**5:30pm-7pm****Room: 355-A***Primary Session Leader: Sudip Dosanjh (Lawrence Berkeley National Laboratory)**Secondary Session Leaders: Marie-Christine Sawley (Intel Corporation), Gilad Shainer (HPC Advisory Council)*

Pathfinding for exascale recently started in many nations: a DOE program in U.S.A., exascale projects funded by FP7 in Europe, partnering initiatives between Intel and European government research institutions, national efforts in Japan/China are good examples. The co-design concept developed by the embedded community is a key HPC strategy for reaching exascale. This BoF will focus on the most relevant experiences and commonalities in applying co-design to HPC, and identify gaps in open research that fuel further developments. A particular example of co-design discussed is the co-development of application communication libraries and the underlying hardware interconnect to overcome scalability issues.

Common Practices for Managing Small HPC Clusters**5:30pm-7pm****Room: 355-D***Primary Session Leader: David Stack (University of Wisconsin-Milwaukee)**Secondary Session Leaders: Roger Bielefeld (Case Western Reserve University)*

This is an opportunity for those responsible for deploying and managing campus-scale, capacity clusters to share their techniques, successes and horror stories. Attendees will discuss the results of a pre-conference survey that ascertained which compute, storage, submit/compile and scheduler environments are common in systems of this size. Attendees will also share how they provide end-user support and system administration services for their clusters.

Cool Supercomputing: Achieving Energy Efficiency at the Extreme Scales**5:30pm-7pm****Room: 155-A***Primary Session Leader: Darren J. Kerbyson (Pacific Northwest National Laboratory)**Secondary Session Leaders: Abhinav Vishnu (Pacific Northwest National Laboratory), Kevin J. Barker (Pacific Northwest National Laboratory)*

Power consumption is a major concern for future generation supercomputers. Current systems consume around a Megawatt per Petaflop. Exascale levels of computation will be significantly constrained if power requirements scale linearly with performance. The optimization of power and energy at all levels, from application to system software and to hardware, is required. This BOF will discuss state-of-the-art tools and techniques for observing and optimizing energy consumption. The challenges ahead are many-fold. Increasing parallelism, memory systems, interconnection networks, storage and uncertainties in programming models all add to the complexities. The interplay between performance, power, and reliability also leads to complex tradeoffs.

Cyberinfrastructure Services for Long Tail Research**5:30pm-7pm****Room: 253***Primary Session Leader: Ian Foster (Argonne National Laboratory)**Secondary Session Leaders: Bill Howe (University of Washington), Carl Kesselman (University of Southern California)*

Much research occurs in small and medium laboratories (SMLs) that may comprise a PI and a few students/postdocs.

For these small teams, the growing importance of cyberinfrastructure for discovery and innovation is as much problem as

opportunity. With limited resources and expertise, even simple data discovery, collection, analysis, management, and sharing tasks are difficult. Thus in this "long tail" of science, modern computational methods often are not exploited, valuable data goes unshared, and too much time is consumed by routine tasks. In this BOF, we aim to spur a discussion around needs and designs for cyberinfrastructure services targeted at SMLs.

DARPA's High Productivity Computing Systems Program: A Final Report**5:30pm-7pm****Room: 255-D***Primary Session Leader: Lauren L. Smith (National Security Agency)**Secondary Session Leaders Dolores A. Shaffer (Science and Technology Associates, Inc.)*

The DARPA High Productivity Computing Systems (HPCS) program has been focused on providing a new generation of economically viable high productivity computing systems for national security, scientific, industrial and commercial applications. This program was unique because it focused on system productivity that was defined to include enhancing performance, programmability, portability, usability, manageability and robustness of systems as opposed to just being focused on one execution time performance metric. The BOF is for anyone interested in learning about the two HPCS systems and how productivity in High Performance Computing has been enhanced.

Exploiting Domain Semantics and High-Level Abstractions in Computational Science**5:30pm-7pm****Room: 155-B***Primary Session Leader: Milind Kulkarni (Purdue University)**Secondary Session Leaders: Arun Prakash, Samuel Midkiff (Purdue University)*

In the last thirty years, much discovery and progress in science and engineering disciplines has been possible because of the development of simulations and modeling software, a development that has made software productivity a major factor in scientific progress. To enable high productivity programming and high performance applications across a variety of application domains, it is essential to leverage the high-level abstractions provided to libraries to provide compilers with semantic information. This BOF will provide a venue for application, compiler and runtime system developers to meet and discuss the design of systems to provide these benefits across multiple disciplines.

HPC Centers**5:30pm-7pm****Room: 155-E**

Primary Session Leader: David E. Martin (Argonne National Laboratory)

Secondary Session Leader: Robert M. Whitten (Oak Ridge National Laboratory)

This BoF brings together support and outreach staff from HPC centers around the world to discuss common concerns, explore ideas and share best practices. Special focus will be given to industrial use of HPC centers, with users giving examples of successful interactions. This BoF is a meeting of the HPC Centers Working Group, but is open to all.

Intel MIC Processors and the Stampede Petascale Computing System**5:30pm-7pm****Room: 255-A**

Primary Session Leader: John (Jay) R. Boisseau (Texas Advanced Computing Center)

Secondary Session Leaders: Dan C. Stanzione, Karl W. Schulz (Texas Advanced Computing Center)

The Intel Xeon Phi will offer tremendous performance and efficiency for highly data-parallel applications when it becomes available in late 2012. The Xeon Phi will debut for HPC in the NSF-funded Stampede system, which will offer ~10PF peak performance in January 2013. Programming and porting for Xeon Phi coprocessors are accomplished with standard, widely used programming tools—Fortran, C/C++, OpenMP, and MPI. In addition, a variety of usage modes are available which offer tradeoffs between code porting speed and raw application performance. This session will introduce attendees to the Phi processor, various programming models, and the Stampede system at TACC.

OpenCL: Supporting Mainstream Heterogeneous Computing**5:30pm-7pm****Room: Ballroom-A**

Primary Session Leader: Timothy G. Mattson (Intel Corporation)

Secondary Session Leaders: Simon McIntosh-Smith (University of Bristol), Ben Gaster (AMD)

OpenCL is an industry standard for programming heterogeneous computers (e.g. CPUs + GPUs). If you do heterogeneous computing and you don't want to be locked into a single vendor's products, you need to learn about OpenCL. At this BOF, we will share the latest developments in OpenCL. More importantly, however, we will launch the OpenCL user's group. This group will be an independent community of users who use OpenCL, build OpenCL tools, and want to influence the evolution of OpenCL. Attend this BOF so you can get in on the ground floor of this exciting new development in OpenCL.

Operating Systems and Runtime Technical Council**5:30pm-7pm****Room: 355-EF**

Primary Session Leader: Ron Brightwell (Sandia National Laboratories)

Secondary Session Leaders: Pete Beckman (Argonne National Laboratory)

The US DOE recently convened an Operating Systems and Runtime (OS/R) Technical Council to develop a research agenda and a plan to address OS/R software challenges associated with extreme-scale systems. The Council's charter is to summarize the challenges, assess the impact on requirements of facilities, applications, programming models, and hardware architectures, describe a model to interact with vendors, and identify promising approaches. This BOF will detail the findings of the Council, which will also include a summary of a workshop held in early October.

Power and Energy Measurement and Modeling on the Path to Exascale**5:30pm-7pm****Room: 255-EF**

Primary Session Leader: Dan Terpstra (University of Tennessee, Knoxville)

Secondary Session Leaders: Laura Carrington (San Diego Supercomputer Center), Rob Fowler (Renaissance Computing Institute), Rong Ge (Marquette University), Andres Marquez (Pacific Northwest National Laboratory), Kazutomo Yoshii (Argonne National Laboratory)

Power and energy consumption have been identified as key "speed bumps" on the path to exascale computing. Members of the research community and industry will present current state-of-the-art and limitations in measuring and modeling power and energy consumption and their effect on HPC application performance. An open discussion about future directions for such work will follow, with the intention of creating a "wish list" of feature requests to HPC vendors. Open questions for discussion include: vendor support for power and energy measurement; power measurement infrastructures and granularity; Dynamic Voltage and Frequency Scaling issues; and software control of other power-related features.

PRACE Future Technologies Evaluation Results**5:30pm-7pm****Room: 250-AB**

Primary Session Leader: Sean Delaney (Irish Centre For High-End Computing)

Secondary Session Leader: Torsten Wilde (Leibniz Supercomputing Centre)

The Partnership for Advanced Computing in Europe (PRACE) explores a set of prototypes to test and evaluate promising new technologies for future multi-Petaflop/s systems. These

include GPUs, ARM processors, DSPs and FPGAs as well as novel I/O solutions and hot water cooling. A common goal of all prototypes is to evaluate energy-consumption in terms of “energy-to-solution” to estimate the suitability of those components for future high-end systems. For this purpose, the “Future Technologies” work package developed an energy-to-solution benchmark suite. A synopsis of the assessments and selected results will be presented in a short series of presentations and discussions.

The Green500 List

5:30pm-7pm

Room: 255-BC

Primary Session Leader: Wu Feng (Virginia Tech)

Secondary Session Leader: Kirk Cameron (Virginia Tech)

The Green500, now entering its sixth year, seeks to encourage sustainable supercomputing by raising awareness in the energy efficiency of such systems. This BoF will present (1) new metrics, methodologies, and workloads for measuring the energy efficiency of a HPC system, (2) highlights from the latest Green500 List, and (3) trends across the history of the Green500, including the trajectory towards exascale. In addition, the BoF will solicit feedback from the HPC community to enhance the impact of the Green500 on energy-efficient HPC design. The BoF will close with an awards presentation, recognizing the most energy-efficient supercomputers in the world.

The Ground is Moving Again in Paradise: Supporting Legacy Codes in the New Heterogeneous Age

5:30pm-7pm

Room: 155-F

Primary Session Leader: Ben Bergen (Los Alamos National Laboratory)

Secondary Session Leaders: Guillaume Colin de Verdiere (CEA), Simon McIntosh-Smith (University of Bristol)

We are now at least five years into the heterogeneous age of computing, and it is still very much a moving target, with no clear path forward for the evolution of legacy codes. This is due to the many challenges that legacy developers face in this endeavor: Limited Support for Fortran; Application and Data Structure Design Issues; User Base (support vs. refactoring); Developer Base (may not be colocated). This BOF will focus on these challenges and on potential strategies for supporting legacy codes in current and future HPC environments.

Using Application Proxies for Exascale Preparation

5:30pm-7pm

Room: 250-C

Primary Session Leader: Richard Barrett (Sandia National Laboratories)

Secondary Session Leaders: Allen McPherson (Los Alamos National Laboratory), Bert Still (Lawrence Livermore National Laboratory)

Application proxies (mini-, skeleton-, and compact applications, etc) provide a means of enabling rapid exploration of the parameter space that spans the performance of complex scientific application programs designed for use on current, emerging, and future architectures. This meeting will include presentations describing work encompassing a broad set of issues impacting a broad set of application programs critical to the mission of the Department of Energy’s ASC campaign, particularly with regard to exascale preparations. This meeting is also designed to encourage participation by an expanding set of collaborators, including those from universities, vendors, and other research institutions.

What Next for On-Node Parallelism? Is OpenMP the Best We Can Hope For?

5:30pm-7pm

Room: 155-C

Primary Session Leader: Jim Cownie (Intel Corporation)

OpenMP is the de-facto standard for on-node parallelism, but it is big, prescriptive and composes poorly. Newer standards like Cilk™ Plus or TBB propose more dynamic exploitation of parallelism while being much less prescriptive. In this BOF we’ll invite the OpenMP experts Michael Wolfe (PGI) and Tim Mattson (Intel) and Cilk experts Robert Geva (Intel) and Bradley Kuszmaul (MIT) to present their view of the future requirements, how we can meet them, and whether OpenMP is sufficient.

The aim is a lively debate with a lot of audience participation moderated by MPI expert Rusty Lusk (ANL).

Thursday, November 15

Charm++: Adaptive Runtime-Assisted Parallel Programming

12:15pm-1:15pm

Room: 255-A

Primary Session Leader: Laxmikant Kale (University of Illinois at Urbana-Champaign)

Secondary Session Leaders: Ramprasad Venkataraman, Eric Bohm (University of Illinois at Urbana-Champaign)

A BoF for the community interested in parallel programming using Charm++, Adaptive MPI, and the associated ecosystem (mini-languages, tools, etc.), along with parallel applications developed using them. Intended to engage a broader audience and drive adoption. Charm++ is a parallel programming system with increasing usage. Next to MPI (and now, possibly OpenMP) it is one of the most used systems deployed on parallel supercomputers, using a significant fraction of CPU cycles. A unified programming model with multicore and accelerator support, its abilities include: dynamic load balancing, fault tolerance, latency hiding, interoperability with MPI, and overall support for adaptivity and modularity.

Data Analysis through Computation and 3D Stereo Visualization

12:15pm-1:15pm

Room: 355-EF

Primary Session Leader: Jason T. Haraldsen (Los Alamos National Laboratory)

Secondary Session Leader: Alexander V. Balatsky (Los Alamos National Laboratory)

We present and discuss the advancement of data analysis through computation and 3D active stereo visualization. Technological innovations have begun to produce larger and more complex data than can be analyzed through traditional methods. Therefore, we demonstrate the combination of computation and 3D stereo visualization for the analysis of large complex data sets. We will present specific examples of theoretical molecular dynamics, density functional, and inelastic neutron scattering simulations as well as experimental data of scanning tunneling microscopy and atom probe tomography. We will also present an open discussion of visualization and the new frontier of data analysis.

Discussing Biomedical Data Management as a Service

12:15pm-1:15pm

Room: 250-C

Primary Session Leader: Ian M. Foster (Argonne National Laboratory)

Secondary Session Leaders: Raimond L. Winslow (Johns Hopkins University), Ravi K. Madduri (Argonne National Laboratory)

The biomedical community needs software tools for managing diverse types of biomedical data. The vast majority of biomedical studies lack IT support, and therefore study sites are not able to install and operate complex software applications on their own. The Cardiovascular Research Grid (CVRG) has therefore moved towards the Software-as-a-Service approach, delivering powerful data management and analysis tools to our users that are accessed through the web browser. Through this Bird-of-a-Feather session, the CVRG team would like to discuss topics of security, data sharing and data integration. We would like to share what we have learned and hear community ideas.

Graph Analytics in Big Data

12:15pm-1:15pm

Room: 255-EF

Primary Session Leader: Amar Shan (YarcData, Inc.)

Secondary Session Leader: Shoaib Mufti (Cray Inc.)

Data intensive computing, popularly known as Big Data, has grown enormously in importance over the past 5 years. However, most data intensive computing is focused on conventional analytics: searching, aggregating and summarizing the data set. Graph analytics goes beyond conventional analytics to search for patterns of relationships, a capability that has important application in many HPC areas ranging from climate science to healthcare and life sciences to intelligence. The purpose of this BOF is to bring together practitioners of graph analytics. Presentations and discussions will include system architectures and software designed specifically for graph analytics; applications; and benchmarking.

HPC Advisory Council University Award Ceremony

12:15pm-1:15pm

Room: 155-B

Primary Session Leader: Gilad Shainer (HPC Advisory Council)

Secondary Session Leader: Brian Sparks (HPC Advisory Council)

The HPC Advisory Council is a leading worldwide organization for high-performance computing research, development, outreach and education activities. One of the HPC Advisory Council's main activities is community and education outreach, in particular to enhance students' computing knowledge-base as early as possible. As such, the HPC Advisory Council has established a university award program in which universities are

encouraged to submit proposals for advanced research around high-performance computing, as well as the Student Cluster Competition at the International Supercomputing conference. During the session we will announce the winning proposals for both programs.

In-silico Bioscience: Advances in the Complex, Dynamic Range of Life Sciences Applications

12:15pm-1:15pm

Room: 155-F

Primary Session Leader: Jill Matzke (SGI)

Secondary Session Leader: Simon Appleby (SGI)

Few disciplines are facing exponential growth in both algorithm and dataset size and complexity as is the case in life sciences. From genome assembly to high content screening and integrative systems modeling, the demands on both the software and hardware side require a dynamic range in computing capability and therefore present a major HPC challenge. This session details some of the ground-breaking results in cancer research and other areas, gained from solutions such as massive, in-memory computing and presented by prominent research institutions in collaboration with leading solution vendors working to advance the science.

New Developments in the Global Arrays Programming Model

12:15pm-1:15pm

Room: 155-E

Primary Session Leader: Bruce J. Palmer (Pacific Northwest National Laboratory)

Secondary Session Leader: Abhinav Vishnu (Pacific Northwest National Laboratory)

The session will be an informal discussion describing current developments in GA and the underlying ARMCI runtime, including a complete restructuring of the ARMCI runtime, the implementation of a new Global Pointers capability, and a new compatibility port to ARMCI using the standard MPI two-sided libraries. We will also discuss plans to publish a standard for the ARMCI interface that will allow other development teams to write ports for ARMCI. The session will then open up to comments and discussion by session participants, including feedback and user experiences from the GA programming community.

OpenSHMEM: A standardized SHMEM for the PGAS community

12:15pm-1:15pm

Room: 155-C

Primary Session Leader: Steve Poole (Oak Ridge National Laboratory)

Secondary Session Leader: Tony Curtis (University of Houston)

The purpose of this meeting is to engage collaboration and input from users and developers of systems, libraries, and applications to further expand an open organization and specification for OpenSHMEM. The initial specification is based on the existing SGI API, but we are now discussing concrete ideas for extensions and expect more to come as this new API is ported to a large variety of platforms. We will also talk about other PGAS frameworks and their relationship with OpenSHMEM, plus the OpenSHMEM “ecosystem” of implementations, applications and tool support.

Petascale Systems Management

12:15pm-1:15pm

Room: 355-BC

Primary Session Leader: William R. Scullin (Argonne National Laboratory)

Secondary Session Leader: Adam D. Yates (Louisiana State University), Adam J. Hough (Petroleum Geo-Services)

Petascale systems often present their administrators yottascale problems. This BOF is a forum for the administrators, systems programmers, and support staff behind some of the largest machines in the world to share solutions and approaches to some of their most vexing issues and meet other members of the community. This year we are focusing on the social, ethical, and policy issues that arise in HPC systems administration discussing such topics as: the environmental and social impact of a Top500 system, devising a fair allocation process, user management, and communication in a competition-laden field.

Resilience for Extreme-scale High-performance Computing**12:15pm-1:15pm****Room: 255-BC**

Primary Session Leader: Christian Engelmann (Oak Ridge National Laboratory)

Secondary Session Leader: Nathan DeBardeleben (Los Alamos National Laboratory)

This session will include a small number of presentations followed by a short discussion at the end. It is targeted toward the Resilience, i.e., providing efficiency and correctness in the presence of faults, is one of the most important exascale computer challenges as systems are expected to scale up in component count and component reliability is expected to decrease. The goal of this BoF is to provide the research community with a coherent view of the HPC resilience challenge, ongoing HPC resilience efforts, and upcoming funding opportunities.

SLURM User Group Meeting**12:15pm-1:15pm****Room: 155-A**

Primary Session Leader: Morris Jette (SchedMD)

Secondary Session Leaders: Danny Auble (SchedMD), Eric Monchalin (Bull)

The SLURM is an open source job scheduler used many on TOP500 systems and provides a rich set of features including topology aware optimized resource allocation, the ability to expand and shrink jobs on demand, the ability to power down idle nodes and restart them as needed, hierarchical bank accounts with fair-share job prioritization and many resource limits. The meeting will consist of three parts: The SLURM development team will present details about changes in the new version 2.5, describe the SLURM roadmap, and solicit user feedback. Everyone interested in SLURM use and/or development is encouraged to attend.

The MPI 3.0 Standard**12:15pm-1:15pm****Room: 355-A**

Primary Session Leader: Richard Graham (Mellanox Technologies)

A new version of the MPI standard, MPI 3.0, has recently been released, and is the culmination of several years of work by the forum. This version introduces large enhancements to this standard, including nonblocking collective operations, new Fortran bindings, neighborhood collectives, enhanced RMA support, a new Tools information interface, as well as many smaller changes to the standard. In this BoF an overview of the new functionality will be provided.

The UDT Forum: A Community for UDT Developers and Users**12:15pm-1:15pm****Room: 255-D**

Primary Session Leader: Robert Grossman (University of Chicago)

Secondary Session Leader: Allison Heath (University of Chicago)

UDT is an open source library supporting high performance data transport. It is used by a growing number of cyberinfrastructure projects and has been commercialized by over 12 companies. UDStar is an application that integrates UDT with common utilities, such as rsync and scp. In this session, we will provide a UDT and UDStar roadmap and have a discussion with the SC12 attendees from the UDT community of ways that the UDT core developers can support the community of UDT developers and users.



Awards

Each year, SC showcases not only the best and brightest stars of HPC, but also its rising stars and those who have made a lasting impression. SC Awards is one way these people are recognized at SC.

Awards

Awards

Ken Kennedy Award

The Ken Kennedy Award recognizes substantial contributions to programmability and productivity in computing and substantial community service or mentoring contributions. The award honors the remarkable research, service, and mentoring contributions of the late Ken Kennedy and includes a \$5,000 honorarium. This award is co-sponsored by ACM and IEEE Computer Society.

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 certificate and \$10,000 honorarium. This award is sponsored by the IEEE Computer Society.

Sidney Fernbach Memorial Award

The Sidney Fernbach Memorial Award honors innovative uses of high performance computing in problem solving. A certificate and \$2,000 honorarium are given to the winner. This award is sponsored by the IEEE Computer Society.

ACM Gordon Bell Prize

The Gordon Bell Prize is awarded each year to recognize outstanding achievement in HPC. Administered by the Association of Computing Machinery (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 HPC to applications in science. Gordon Bell prizes have 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.

IEEE Reynolds B. Johnson Storage Systems Award

The IEEE Reynold B. Johnson Information Storage Systems Award, sponsored by Hitachi Data Systems, was established in 1992 for outstanding contributions to information storage systems, with emphasis on computer storage systems. The award was named in honor of Reynold B. Johnson, renowned as a pioneer of magnetic disk technology, and founding manager of the IBM San Jose, California, Research and Engineering Laboratory in 1952, where IBM research and development in the field was centered.

For leadership in the development of innovative storage systems for heterogeneous open and mainframe servers, business continuity solutions, and virtualization of heterogeneous storage systems, *the winner of the 2012 Reynold B. Johnson Information Storage Systems Award is Dr. Naoya Takahashi.*

Student Awards

George Michael Memorial HPC Fellowship Program

The ACM, IEEE Computer Society and the SC Conference series established the HPC Ph.D. Fellowship Program to honor exceptional Ph.D. students throughout the world with the focus areas of high performance computing, networking, storage and analysis. Fellowship recipients are selected based on: overall potential for research excellence; the degree to which their technical interests align with those of the HPC community; their academic progress to date, as evidenced by publications and endorsements from their faculty advisor and department head as well as a plan of study to enhance HPC-related skills; and the demonstration of their anticipated use of HPC resources.

ACM Student Research Competition

The Association for Computing Machinery Student Research Competition (ACM SRC) provides an opportunity for undergraduate and graduate students to present original research at several ACM-sponsored or ACM co-sponsored conferences throughout the year. The first round of the competition is held during the Tuesday evening Poster session. Semi-finalists selected based on the poster session then give a brief presentation about their research in a session on Wednesday afternoon. Students selected as SC12 SRC award winners are given an ACM medal, a monetary prize, and an opportunity to compete in the ACM SRC Grand Finals.

ACM Gordon Bell Finalists

Tuesday, November 13

ACM Gordon Bell Prize I

1:30pm-3pm

Room: 155-E

Billion-Particle SIMD-Friendly Two-Point Correlation on Large-Scale HPC Cluster Systems

Authors: Jatin Chhugani, Changkyu Kim (Intel Corporation), Hemant Shukla (Lawrence Berkeley National Laboratory), Jongsoo Park, Pradeep Dubey (Intel Corporation), John Shalf, Horst D. Simon (Lawrence Berkeley National Laboratory)

Correlation analysis is a widely used tool in a range of scientific fields, ranging from geology to genetics and astronomy. In astronomy, Two-point Correlation Function (TPCF) is commonly used to characterize the distribution of matter/energy in the universe. Due to the large amount of computation with massive data, TPCF is a compelling benchmark for future exascale architectures.

We propose a novel algorithm that significantly reduces the computation and communication requirement of TPCF. We exploit the locality of histogram values and thus achieve near-linear scaling with respect to number of cores and SIMD-width.

On a 1600-node Zin supercomputer at Lawrence Livermore National Laboratory (1.06 Petaflops), we achieve 90% parallel efficiency and 96% SIMD efficiency and perform computation on a 1.7 billion particle dataset in 5.3 hours (35–37X faster than previous approaches). Consequently, we now have line-of-sight to achieving the processing power for correlation computation to process billion+ particles telescopic data.

Toward Real-Time Modeling of Human Heart Ventricles at Cellular Resolution: Simulation of Drug-Induced Arrhythmias

Authors: Arthur A. Mirin, David F. Richards, James N. Glosli, Erik W. Draeger, Bor Chan, Jean-luc Fattebert, William D. Krauss, Tomas Oppelstrup (Lawrence Livermore National Laboratory), John Jeremy Rice, John A. Gunnels, Viatcheslav Gurev, Changhoan Kim, John Magerlein (IBM T.J. Watson Research Center), Matthias Reumann (IBM Research Collaboratory for Life Sciences), Hui-Fang Wen (IBM T.J. Watson Research Center)

We have developed a highly efficient and scalable cardiac electrophysiology simulation capability that supports groundbreaking resolution and detail to elucidate the mechanisms of sudden cardiac death from arrhythmia. We can simulate thousands of heartbeats at a resolution of 0.1 mm, comparable to the size of cardiac cells, thereby enabling scientific inquiry not previously possible. Based on scaling results from the partially deployed Sequoia IBM Blue Gene/Q machine at Lawrence Livermore National Laboratory and planned optimizations,

we estimate that by SC12 we will simulate 8-10 heartbeats per minute - a time-to-solution 400-500 times faster than the state-of-the-art. Performance between 8 and 11 PFlop/s on the full 1,572,864 cores is anticipated, representing 40–55 percent of peak. The power of the model is demonstrated by illuminating the subtle arrhythmogenic mechanisms of anti-arrhythmic drugs that paradoxically increase arrhythmias in some patient populations.

Extreme-Scale UQ for Bayesian Inverse Problems Governed by PDEs

Authors: Tan Bui-Thanh (University of Texas at Austin), Carsten Burstedde (University of Bonn), Omar Ghattas, James Martin, Georg Stadler, Lucas Wilcox (University of Texas at Austin)

Quantifying uncertainties in large-scale simulations has emerged as the central challenge facing CS&E. When the simulations require supercomputers, and uncertain parameter dimensions are large, conventional UQ methods fail. Here we address uncertainty quantification for large-scale inverse problems in a Bayesian inference framework: given data and model uncertainties, find the pdf describing parameter uncertainties. To overcome the curse-of-dimensionality of conventional methods, we exploit the fact that the data are typically informative about low-dimensional manifolds of parameter space to construct low rank approximations of the covariance matrix of the posterior pdf via a matrix-free randomized method. This results in a method that scales independently of the forward problem dimension, the uncertain parameter dimension, the data dimension, and the number of processors. We apply the method to the Bayesian solution of an inverse problem in 3D global seismic wave propagation with a million parameters, for which we observe three orders of magnitude speedups.

Wednesday, November 14

Ken Kennedy / Sidney Fernbach / Seymour Cray Award Talks

10:30am-12pm

Room: 155-E

Ken Kennedy Award Recipient:

Mary Lou Soffa (University of Virginia)

Seymour Cray Award Recipient:

Peter Kogge (University of Notre Dame)

Sidney Fernbach Award Recipients:

Laxmikant Kale and Klaus Schulten (University of Illinois)

ACM Gordon Bell Prize II**1:30pm-3pm****Room: 155-E****The Universe at Extreme Scale - Multi-Petaflop Sky Simulation on the BG/Q**

Authors: Salman Habib, Vitali Morozov, Hal Finkel, Adrian Pope, Katrin Heitmann, Kalyan Kumar, Tom Peterka, Joseph Insley (Argonne National Laboratory), David Daniel, Patricia Fasel, Nicholas Frontiere (Los Alamos National Laboratory), Zarija Lukic (Lawrence Berkeley National Laboratory)

Remarkable observational advances have established a compelling cross-validated model of the universe. Yet, two key pillars of this model—dark matter and dark energy—remain mysterious. Next-generation sky surveys will map billions of galaxies to explore the physics of the ‘Dark Universe’. Science requirements for these surveys demand simulations at extreme scales; these will be delivered by the HACC (Hybrid/Hardware Accelerated Cosmology Code) framework. HACC’s novel algorithmic structure allows tuning across diverse architectures, including accelerated and multi-core systems.

Here we describe our efforts on the IBM BG/Q, attaining unprecedented performance and efficiency (2.52 PFlops, more than 50% of peak on a prototype system, 4X expected for the final submission) at extreme problem sizes, larger than any cosmological simulation yet performed—more than a trillion particles. HACC simulations at these scales will for the first time enable tracking individual galaxies over the entire volume of a cosmological survey.

4.45 Pflops Astrophysical N-Body Simulation on K Computer - The Gravitational Trillion-Body Problem

Authors: Tomoaki Ishiyama, Keigo Nitadori (Tsukuba University), Junichiro Makino (Tokyo Institute of Technology)

As an entry for the 2012 Gordon Bell performance prize, we report performance results of astrophysical N-body simulations of one trillion particles performed on the full system of K computer. This is the first gravitational trillion-body simulation in the world. We describe the scientific motivation, the numerical algorithm, the parallelization strategy, and the performance analysis. Unlike many previous Gordon-Bell prize winners that used the tree algorithm for astrophysical N-body simulations, we used the hybrid TreePM method, for similar level of accuracy in which the short-range force is calculated by the tree algorithm, and the long-range force is solved by the particle-mesh algorithm. We developed a highly-tuned gravity kernel for short-range forces, and a novel communication algorithm for long-range forces. The average performance on 24576 and 82944 nodes of K computer are 1.53 and 4.45 Pflops, which correspond to 49% and 42% of the peak speed.

ACM Student Research Competition Semi-Finals**Chair: Torsten Hoefler (ETH Zurich)****1:30pm-3pm; 3:30pm-5pm****Room: 250-C**

The ACM Student Research Competition poster presentations are taking place during the poster session. A jury will select posters for the semi-finals where each selected student presents a 10 minute talk about his poster. Those talks are then again evaluated by a jury and the winners for the ACM SRC graduate and undergraduate medals are chosen and presented at the SC12 award ceremony. This is the first of two sessions that will be filled depending on the number of finalists.

Thursday, November 15**George Michael Memorial HPC Ph.D.Fellowship Presentation****Chair: Bruce Loftis (University of Tennessee, Knoxville)****10:30am-11am****Room: 155-E****Efficient and Scalable Runtime for GAS Programming Models on Petascale Systems**

Xinyu Que (Auburn University)

Global Address Space (GAS) programming models enable a convenient, shared memory-style addressing model. Typically, this is realized through one-sided operations that can enable asynchronous communication and data movement. On the petascale systems, the underlying runtime systems face critical challenges in (1) scalably managing resources (such as memory for communication buffers), and (2) gracefully handling unpredictable communication patterns and any associated contention. This talk will first present a Hierarchical Cooperation architecture for Scalable communication in GAS programming models. Then, it will cover the techniques used for the implementation on a popular GAS runtime library, Aggregate Remote Memory Copy Interface (ARMCI). Finally, experimental results will be discussed to show that our architecture is able to realize scalable resource management and achieve resilience to network contention, while at the same time maintaining and/or enhancing the performance of scientific applications.

SC12 Conference Award Presentations**Chair: Bernd Mohr (University of Tennessee, Knoxville)****12:30pm-1:30pm****Room: Ballroom-EFGH**

The awards managed by the SC12 Conference will be presented. These include: The ACM Gordon Bell Prize; Best Paper, Best Student Paper and Best Poster Awards; George Michael Memorial HPC Ph.D. Fellowship; ACM Student Research Competition; and Student Cluster Competition.



Exhibitor Forum

The Exhibitor Forum showcases the latest advances in the industry, such as new products and upgrades, recent research and development initiatives, and future plans and roadmaps. Industry leaders will give insight into the technology trends driving strategies, the potential of emerging technologies in their product lines, or the impact of adopting academic research into their development cycle. Topics will cover a wide range of areas, including heterogeneous computing, interconnects, data centers, networking, applications and innovations. Come visit us in rooms 155 B and C, close to the exhibit area.

Exhibitor Forum

Tuesday, November 13

Interconnect and Advanced Architectures I

10:30am-12pm

Room: 155-B

PCI Express as a Data Center Fabric

Presenter: Ajoy Aswadhati (PLX Technology)

The presentation highlights how PCI Express (PCIe) is evolving by extending its dominance within the box—to be an external connectivity of choice within the rack to create clusters of server, switch and storage appliances. A PCI Express Fabric based on PCIe Gen3 and Gen4 inside the rack is complementary to InfiniBand and Ethernet in next-generation cloud-driven data centers. PCIe does not replace the existing network itself, but instead extends the benefits of PCIe outside the box by moving network interface cards (NICs) to the top of the rack—or edge of the cluster—thereby reducing cost and power while maintaining features offered by other network fabrics like InfiniBand and Ethernet. PCIe is the lowest power, lowest cost solution, and it negates the cumbersome need to translate multiple interconnects, thus resulting in lower latency and higher performance.

Mellanox Technologies—Paving the Road to Exascale Computing

Presenters: Todd Wilde, Michael Kagan (Mellanox Technologies)

Today's large-scale supercomputer systems span tens of thousands of nodes, requiring a high level of performance and scalability from the interconnect. Cluster sizes continue to increase, and with the advent of accelerators, bandwidth and latency requirements continue to increase at a rapid pace. Moreover, the ability for the network to offload communication elements and to provide advanced features to increase efficiencies and scalability continues to become a critical aspect in delivering desired performance of the supercomputer. The Mellanox ScalableHPC solution provides the necessary technology to accelerate MPI and PGAS environments with communication offloading, scalable transport techniques, and communication accelerations such as GPUDirect RDMA, which allows direct communication between the GPU and the network, bypassing CPU involvement in the communication. This presentation will cover the latest technology advancements from Mellanox, including details on the new 100Gb/s Connect-IB HCA architecture built to drive the most power supercomputers in the world.

Affordable Shared Memory for Big Data

Presenter: Einar Rustad (Numascale AS)

Numascale's NumaConnect technology enables building scalable servers with the functionality of enterprise mainframes at the cost level of clusters and is ideally suited for handling large scale "Big-Data" applications. The NumaConnect hardware implements a directory based global cache coherence protocol and allows large shared memory systems to be controlled by standard operating systems like Linux. Systems based on NumaConnect will efficiently support all classes of applications. Maximum system size is 4k multicore nodes. The cache coherent shared memory size is limited by the 48-bit physical address range provided by the Opteron processors resulting in a total system main memory of 256 TBytes. At the heart of NumaConnect is NumaChip; a single chip that combines the cache coherent shared memory control logic with an on-chip 7-way switch. This eliminates the need for a separate, central switch and enables linear capacity and cost scaling with well-proven 2-D or 3-D Torus topologies.

HPC in the Cloud I

10:30am-12pm

Room: 155-C

Taking HPC to the Cloud—Overcoming Complexity and Accelerating Time-to-Results with Unlimited Compute

Presenter: Adam Jacob (Opscode)

Research projects, financial analysis, and other types of quantitative exercises are increasing in complexity and often require a massive degree of computational power to successfully complete. These types of projects are rarely straightforward and frequently contain transient workloads and dynamic problem sets requiring constant revision. With widespread access and increasing adoption of cloud computing, organizations of all types and sizes now have access to theoretically limitless compute resources. However, unprecedented compute power also creates exponentially more complex management challenges. This session will detail best practices for solving the management problems presented by moving HPC to the cloud. Leveraging real-world examples involving tens of thousands of cloud compute cores, Jacob will present the route to not only overcoming complexity, but to maximizing the potential of cloud computing in dramatically accelerating time-to-results for HPC projects.

HPC Cloud ROI and Opportunities Cloud Brings to HPC

Presenter: Brady Kimball (Adaptive Computing)

Advancing cloud technology presents opportunities for HPC data centers to maximize ROI. HPC system managers can leverage the benefits of cloud for their traditional HPC environments. Incorporating HPC Cloud into an HPC system enables you to meet workload demand, increase system utilization, and expand the HPC system to a wider community. Extended

ROI and opportunities include: (1) Scale to support application and job needs with automated workload-optimized node OS provisioning; (2) Provide simplified self-service access for broader set of users reducing management and training costs; (3) Accelerate collaboration or funding by extending HPC resources to community partners without their own HPC system; (4) Enable showback/chargeback reporting for actual resource usage by user, group, project, or account; (5) Support using commercial HPC service providers for surge and peak load requirements to accelerate results; (6) Enable higher cloud-based efficiency without the cost and disruption of ripping and replacing existing technology

The Technical Cloud: When Remote 3D Visualization Meets HPC

Presenter: Andrea Rodolico (NICE)

Data access speed, rapid obsolescence, heat, noise and application availability are just some of the issues current workstation-based technical users (designers, engineers, scientists, etc.) face in their pre-post processing activities. In other IT domains, virtualization and remote desktop solutions today address most of these issues, but it is not broadly applied to technical computing because in traditional VDIs, the GPU cannot be virtualized and shared among users. NICE Desktop Cloud Visualization (DCV) addresses these limitations and allows technical users to run fully accelerated, off-the-shelf OpenGL applications on Windows or Linux in a right-sized “virtual workstation” on or near the HPC system. In this “technical cloud” pixels are transferred instead of data, boosting application performance, security and manageability. We will analyze multiple usage scenarios, including physical and virtual deployments with dedicated GPU, shared GPU as well as acceleration by an external “GPU appliance.”

Heterogeneous Computing I

1:30pm-3pm

Room: 155-B

Transforming HPC Yet Again with NVIDIA Kepler GPUs

Presenter: Roy Kim (NVIDIA)

GPUs revolutionized the HPC industry when NVIDIA introduced the Fermi GPU computing architecture in 2009. With NVIDIA GPUs, scientists and engineers can develop more promising cancer treatments, higher fidelity automotive designs, and a deeper understanding of fundamental science. The next generation NVIDIA GPU, codenamed “Kepler,” is designed to revolutionize the HPC industry yet again. Attend this talk to learn about the soul of the Kepler design, its innovative features, and the world-leading performance it delivers.

Addressing Big Data Challenges with Hybrid-Core Computing

Presenter: Kirby Collins (Convey Computer)

The spread of digital technology into every facet of modern life has led to a corresponding explosion in the amount of data that is stored and processed. Understanding the relationships between elements of data has driven HPC beyond numerically-intensive computing into data-intensive algorithms used in fraud detection, national security, and bioinformatics. In this session Convey Computer will present the latest innovations in Hybrid-Core computing and describe how high bandwidth, highly parallel reconfigurable architectures can address these and other applications with higher performance, lower energy consumption, and lower overall cost of ownership compared to conventional architectures.

Flash Memory and GPGPU Supercomputing: A Winning Combination

Presenter: David A. Flynn (Fusion-io)

Implementing GPUs in supercomputers is a great way to further accelerate raw processing power. Another great way is implementing flash as a persistent, high-capacity memory tier that acts like RAM. Connecting flash as a memory tier, directly to the system BUS provides the ability for supercomputers to run applications natively on flash memory, which greatly accelerates performance. Companies that have used this type of technology have been able to make more potent supercomputers that rival those in the Top500. This presentation will share the details behind using flash as a new memory tier, how it works with GPGPU, what this means for companies, and what environments can benefit from this new architecture.

Software Development Tools I

1:30pm-3pm

Room: 155-C

Faster, Better, Easier Tools: The Shortcut to Results

Presenter: David Lecomber (Allinea Software)

HPC systems continue to grow in size and complexity and—with many architectures to choose from—good software is needed more than ever! Tools, such as the scalable debugger, Allinea DDT, that reach across both hybrid and homogeneous platforms, are leading the charge. In this talk we will reveal exciting new developments to be released at SC12.

Scalable Debugging with TotalView for Xeon Phi, BlueGene/Q, and more

Presenter: Chris Gottbrath (Rogue Wave Software)

Writing scalable software is hard, and substantial projects are almost always developed over many years and ported to a wide range of different processor and cluster architectures. These challenges are the key motivations for Rogue Wave Software’s commitment to providing high-quality tools and libraries across the range of important HPC architectures. The

Xeon Phi co-processor and BlueGene/Q supercomputer are two innovative architectures to which we've recently ported our TotalView scalable parallel debugger. This talk will highlight some of the challenges that developers may have when working with these new architectures and show how TotalView can be used to overcome those issues. It will also provide an update on scalability, usability, CUDA/OpenACC support, the port of the ReplayEngine deterministic reverse debugging feature to the Cray XE platform, and the ThreadSpotter product.

Advanced Programming of Many-Core Systems Using CAPS OpenACC Compiler

Presenter: Stephane Bihan (CAPS)

The announcement last year of the new OpenACC directive-based programming model supported by CAPS, CRAY and PGI compilers has opened up the door to more scientific applications that can be ported on many-core systems. Following a porting methodology, this talk will first present the three key principles of programming with OpenACC and then the advanced features available in the CAPS HMPP compiler to further optimize OpenACC applications. As a source-to-source compiler, HMPP uses hardware vendors' backends, such as NVIDIA CUDA and OpenCL making CAPS products the only OpenACC compilers supporting various many-core architectures.

Storage and File Systems I

3:30pm-5pm

Room: 155-B

Integrating ZFS RAID with Lustre Today

Presenter: Josh Judd (WARP Mechanics Ltd.)

The long-term future of Lustre is integration with ZFS. It may take years, however, to fully integrate the code. And even then, there are scalability, performance, and fault-isolation benefits to keeping the RAID layer physically separate from the OSS. Learn how to architect a production-grade Lustre network today using ZFS RAID arrays, which will move you towards the long-term architecture while maintaining the benefits of a layered design.

The End of Latency: A New Storage Architecture

Presenter: Michael Kazar (Avere Systems)

In the last decade, a number of developments in storage and infrastructure technology have changed the IT landscape forever. The adoption of solid-state storage media, the virtualization of servers and storage, and the introduction of Cloud have impacted all aspects of how organizations will build out the storage architecture for the future. Mike Kazar will look at not just how these major developments have combined to create unexpected leaps in performance and scalability, but he will also identify the biggest technical roadblock to successfully deploying each of these and why traditional storage architectures will always at best be a compromise and at worst a dead end.

The Expanding Role of Solid State Technology in HPC Storage Applications

Presenter: Brent Welch (Panasas)

As enterprises and research organizations strive to increase the velocity with which they can acquire, access and analyze ever-expanding data sets, the role of solid state technology is expanding in HPC storage applications. Brent Welch, chief technology officer at Panasas, will discuss the ramifications of SSD technology as it relates to high performance parallel storage environments.

Memory Systems

3:30pm-5pm

Room: 155-C

How Memory and SSDs can Optimize Data Center Operations

Presenters: Sylvie Kadivar, Ryan Smith (Samsung Semiconductor, Inc.)

Memory and solid-state drives (SSD) are critical components for the server market. This presentation will highlight the latest advancements in DRAM (main system memory) and solid-state drives, from a power savings and a performance perspective. We will take a close look at the different types of Green Memory, while pinpointing their performance and power-saving advantages in cloud infrastructures, rack-designed systems and virtualized environments. The presentation will spotlight Samsung's most advanced DDR3 and upcoming DDR4 memory, as well as its new generation of SSDs. Samsung will discuss how choosing the right memory and storage can have a major impact on the efficiency of client/server operations, by increasing performance, lowering energy costs and reducing overall CO2 emissions. We will also help you to understand the best ways to use SSDs as a primary way of addressing IT performance bottlenecks. (Note: Large-file-size supportive PDF collateral available upon request.)

Beyond von Neumann With a 1 Million Element Massively Parallel Cognitive Memory

Presenter: Bruce McCormick (Cognimem Technologies, Inc.)

Presentation details a demonstrable and scalable pattern recognition system solution ("Beyond"). It is based on 1000 chips connected in parallel. Each chip integrates 1024 fully parallel and general-purpose pattern recognition and machine-learning memory processing elements. These 1 million memory processing elements are non-linear classifiers (kNN and Radial Basis Functions) implemented as a three-layer network. The system provides a constant 10usec latency of fuzzy or exact vector comparison of 1 versus up to 1 million 256 byte vectors offering equivalent of .13 petaflops of performance at a miserly 250 watts. It eliminates the bottleneck in traditional von Neumann architectures between processing and memory. Applications range from DNA, iris, fingerprint, hash matching to scientific modeling, video analytics & data mining including techniques such as anomaly detection, clustering, sorting,

general-purpose pattern recognition and more. Green Graph 500 benchmark will be verified and presented.

Hybrid Memory Cube (HMC): A New Paradigm for System Architecture Design

Presenter: Todd Farrell (Micron Technology)

DRAM technology has been utilized as main memory in microprocessor-based systems for decades. From the early days of frequency scaling, the gap has been growing between the DRAM performance improvement rate versus the processor data consumption rate. This presentation will show several advantages of using Micron's Hybrid Memory Cube (HMC) technology. HMC is a three-dimensional structure with a logic device at its base and a plurality of DRAMs vertically stacked above it using through-silicon via (TSV) connections. The HMC concept is completely re-architected, redistributing the normal DRAM functions while delivering: Scalable System Architectures: Flexible topologies (expandability) Abstraction - Future memory process scaling and challenges; Performance: Higher effective DRAM bandwidth Lower DRAM system latency Increased DRAM random request rate; Energy (Power-Efficient Architectures): Lower DRAM energy per useful unit of work done Reduced data movement; Dependability (RAS): In-field repair capability Internal DRAM ECC.

Wednesday, November 14

Novel and Future Architectures I

10:30am-12pm

Room: 155-B

Hybrid Solutions with a Vector-Architecture for Efficiency

Presenter: Shigeyuki Aino (NEC Corporation)

The important topic of HPC these days is sustained performance on real applications per total cost of ownership. NEC's target is to optimize this by providing hybrid systems, addressing the diversity of user requirements in the optimal way. NEC's scalar product line, the LX-Series, is based on industry-standard components and very successful because of NEC's ability to maximize application performance on x86-based and GPU-accelerated compute clusters. Vector architectures feature a high efficiency by design. Therefore NEC is developing the next generation NEC SX vector system focusing on TCO-enhancements through superior computational and power efficiency. The real challenge lies in the seamless integration of a variety of such components into a hybrid system. NEC has installed such systems at big customer sites, getting valuable feedback, and is working to enhance the seamless integration to improve the personal efficiency of the scientific user.

Appro's Next Generation Xtreme-X Supercomputer

Presenter: John Lee (Appro)

This presentation will focus on the value of the next generation, Appro Xtreme-X Supercomputer, based on the Appro GreenBlade2 platform. This new server platform is the foundation of Appro's energy efficient and scalable supercomputing systems which combines high performance capacity computing with superior fault-tolerant capability computing. It will also cover the benefits of Appro's latest system hardware architecture and how integration of enhancements in power and cooling system that supports 480V input as well as micro-power management are important for the overall system to achieve TCO reductions as well as achieve system efficiency.

Innovation and HPC Transformation

Presenter: Scott Misage (Hewlett-Packard)

HP delivers break-through innovation and scale, built on technology and affordability that has transformed HPC. HP technologies and services enable new levels of performance, efficiency and agility, and new game-changing architectures re-invent the traditional server paradigm, and provide the infrastructure for Exascale. We'll outline recent developments in systems, infrastructure and software that improve performance and performance-density and efficiency, and describe technologies that will revolutionize HPC. We'll discuss the next generation of HP ProLiant servers purpose-built for HPC, and integration and support for accelerators and Intel Xeon Phi. We'll describe deployments at major sites featuring these technologies. We will also provide more information on Project Moonshot, designed to unlock the promise of extreme low-energy server technology by pooling resources in a highly-federated environment. The initial platforms, Redstone and Gemini incorporate more than 2,800 servers in a single rack and are the foundation for a new generation of hyperscale computing solutions.

Software Platforms

10:30am-12pm

Room: 155-C

Create Flexible Systems As Your Workload Requires

Presenter: Shai Fultheim (ScaleMP)

ScaleMP is a leader in virtualization for high-end computing, providing higher performance and lower Total Cost of Ownership (TCO). The innovative Versatile SMP (vSMP) architecture aggregates multiple x86 systems into a single virtual x86 system, delivering an industry-standard, high-end symmetric multiprocessor (SMP) computer. Using software to replace custom hardware and components, ScaleMP offers a new, revolutionary computing paradigm. vSMP Foundation is a software-only solution that eliminates the need for extensive R&D or proprietary hardware components in developing high-end x86 systems, and reduces overall end-user system cost and operational expenditures. vSMP Foundation can aggregate

up to 128 x86 systems to create a single system with 4 to over 16,000 processors and up to 256 TB of shared memory. ScaleMP will discuss how creating systems based on the application requirements can benefit users. In addition, recent work with the Intel Xeon Phi coprocessor will be described.

The OpenOnload User-level Network Stack

Presenter: Dave Parry, Steve Pope (Solarflare)

The architecture of conventional networked systems has remained largely constant for many years. Some specialized application domains have, however, adopted alternative architectures. For example, the HPC community uses message passing libraries which perform network processing in user space in conjunction with the features of user-accessible network interfaces. Such user-level networking reduces networking overheads considerably without sacrificing the security and resource management functionality that the operating system normally provides. Supporting user-level TCP/UDP/IP networking for a more general set of applications poses considerable challenges, including: intercepting system calls, binary compatibility with existing applications, maintaining security, supporting fork() and exec(), passing sockets through Unix domain sockets and advancing the protocol when the application is not scheduled. This talk presents the OpenOnload architecture for user-level networking, which is rapidly becoming the de-facto standard for user-space protocol processing of TCP and UDP, particularly in latency sensitive applications. Performance measurements and real world deployment-cases will be discussed.

Runtimes and Applications for Extreme-Scale Computing

Presenter: Rishi Khan (E.T. International, Inc.)

Future generation HPC systems comprising many-core sockets and GPU accelerators will impose increasing challenges in programming, efficiency, heterogeneity and scalability for extreme-scale computing. Emerging execution models using event-driven, task-based parallelism; dynamic dependency and constraint analysis; locality-aware computation; and resource-aware scheduling show promise in addressing these challenges. Applications utilizing these innovative runtime systems have shown significant gains in performance and utilization of computational resources over conventional methodologies in a wide variety of application domains. For example, ETI's SWARM (SWift Adaptive Runtime Machine) has shown 3X improvement over OpenMP on N-Body Problems, 2-10x improvements over MPI on Graph500, and 50% improvement over Intel's MKL ScaLapack. SWARM has also been selected as a technology for a large US Department of Energy program to build an exascale software ecosystem. This presentation will articulate the challenges extreme scale systems pose to applications and runtime developers and highlight some of the solutions to these problems.

Novel and Future Architectures II

1:30pm-3pm

Room: 155-B

Cray's Adaptive Supercomputing Vision

Presenter: William Blake (Cray Inc.)

Cray's Adaptive Supercomputing vision is focused on meeting the market demand for realized performance and helping customers surpass current technological limitations by delivering innovative, next-generation products that integrate diverse processing technologies into a unified architecture. Adaptive Supercomputing is both our vision and a very public statement of purpose. With each product launch since its introduction, the company has introduced innovations in every aspect of HPC from scalability and performance to software and storage, moving it closer to realizing truly adaptive supercomputing. But where is it going from here? This presentation will look at where Adaptive Supercomputing has been and the strategy going forward, including how HPC and Big Data can be brought together into an integrated whole.

Findings from Real Petascale Computer Systems and Fujitsu's Challenges in Moving Towards Exascale Computing

Presenter: Toshiyuki Shimizu (Fujitsu)

Fujitsu offers a range of HPC solutions with the high-end PRIMEHPC FX10 supercomputer, the PRIMERGY x86 clusters and associated ecosystems. The presentation will highlight performance evaluations for the K computer and PRIMEHPC FX10, including details of the technologies implemented. The technologies include a SIMD extension for HPC applications (HPC-ACE), automatic parallelization support for multi-threaded execution (VISIMPACT), a direct network based on a 6D mesh/torus interconnect (Tofu), and hardware support for collective communication (via the Tofu & its dedicated algorithms). The evaluations confirm these features contribute to scalability, efficiency, and usability of application programs in a massively parallel execution environment. In addition, initial indications show that future platforms, which will have more processor cores with wider SIMD capability, will exhibit similar characteristics. Fujitsu's recent activities and collaboration framework for technology development toward Exascale computing will also be introduced. This includes co-design of architecture and system software with selected applications.

On Solution-Oriented HPC

Presenter: Stephen Wheat (Intel Corporation)

HPC vibrancy continues as the technology advances to address new challenges spanning the full spectrum of science and engineering to sociology. The astounding breadth of HPC benefit to everyday people around the world is a powerful motivation for the HPC ecosystem. The underlying complexity of an HPC-based solution, if anything, is also increasing. Regardless of system size, confidence in a given solution requires a well-choreographed interplay between the ingredients, from silicon

to software, each of which is innovating at its own pace. When we look to emerging opportunities for HPC, this confidence becomes even more imperative. Intel invests broadly in hardware platforms and software to enable HPC developers and administrators to deliver solutions to the user. We will review Intel's vision for scalable HPC solutions and introduce Intel's latest technology, products, and programs designed to simplify the development of HPC solutions and their application to the user's needs.

Storage and File Systems II

1:30pm-3pm

Room: 155-C

A Plague of Petabytes

Presenter: Matthew Drahzal (IBM)

In HPC, it is stunningly easy for users to create/store data, but these same users are completely unaware of challenges and costs of all this spinning disk. They treat all created data as equally important and critical to retain, whether or not this is true. Since the rate of growth of data stored is higher than the areal-density growth rate of spinning disks, organizations are purchasing more disk and spending more IT budget on managing data. While cost for computation is decreasing, cost to store, move, and manage the resultant information is ever expanding. IBM Research and Development are working on new technologies to shift data cost curves fundamentally lower, use automation to manage data expansion, and leverage diverse storage technologies to manage efficiencies---all "behind the scenes"---nearly invisible to end users. This presentation will describe new data technologies being developed and perfected, and how these changes may fundamentally reset data costs lower.

Big Data, Big Opportunity: Maximizing the Value of Data in HPC Environments

Presenter: Vinod Muralidhar (EMC Isilon)

According to a recent research report by the International Data Corporation (IDC), global data will grow to 2.7 zettabytes in 2012---up 48% from 2011. IDC also predicts this figure to balloon to eight zettabytes worth of data by 2015, while file-based data will grow 75X in the next decade. With such staggering growth rates, it is clear there has never been more data available - or a greater imperative to access, analyze and distribute it more efficiently. Especially in HPC environments, data stores can grow extremely rapidly and though compute server technology has kept pace, storage has not, creating a barrier between researchers and their data. This session will examine how implementing Isilon scale-out storage can eliminate the storage bottleneck in HPC and put data immediately into the hands of those who need it most.

FhGFS - Parallel Filesystem Performance at the Maximum

Presenter: Christian Mohrbacher (Fraunhofer ITWM)

FhGFS is the parallel file system from the Fraunhofer Competence Center for High Performance Computing. It has been designed to deliver highest performance and to provide the scalability that is required to run today's most demanding HPC applications. Large systems and metadata-intensive applications can greatly profit from the support for distributed metadata and from the avoidance of architectural bottlenecks. Furthermore, Fraunhofer sets a high value on ease of use and flexibility, which makes it possible to run the filesystem in a lot of different scenarios. The software has proven to be reliable in installations of all kinds and sizes, ranging from a handful of nodes to current Top500 systems. This talk will give an overview on FhGFS and demonstrate the advantages of its modern design by showing latest benchmarking results.

HPC in the Datacenter I

3:30pm-5pm

Room: 155-B

The HPC Advisory Council Outreach and Research Activities

Presenter: Gilad Shainer, Brian Sparks (HPC Advisory Council)

The HPC Advisory Council (www.hpcadvisorycouncil.com) is a distinguished body representing the high performance computing ecosystem that includes more than 320 worldwide organizations as members from OEMs, strategic technology suppliers, ISVs and end-users across the entire range of the HPC segments. The HPC Advisory Council's mission is to bridge the gap between HPC use and its potential, bring the beneficial capabilities of HPC to new users for better research, education, innovation and product manufacturing, bring users the expertise needed to operate HPC systems, provide application designers with the tools needed to enable parallel computing, and to strengthen the qualification and integration of HPC system products. The HPC Advisory Council operates a centralized support center providing end users with easy access to leading edge HPC systems for development and benchmarking, and a support/advisory group for consultations and technical support.

Difference on Cold and Hot Water Cooling on CPU and Hybrid Supercomputers

Presenter: Giovanbattista Mattiussi (Eurotech)

Eurotech is a publicly listed global embedded electronics and supercomputer company, which manufactures HPCs, from boards to systems, and delivers HPC solutions. Having engineered the first liquid cooled HPC in 2005 and shipped its first hot liquid cooled supercomputer to a customer in 2009, Eurotech is a real pioneer in hot liquid cooling. After many years of HPC deliveries, Eurotech has gained experience with many aspects of the liquid cooling technology and with a variety of

customer situations. Eurotech would like to share their experience highlighting a parallelism between hot and cold liquid cooling, discussing about pros and cons of the 2 approaches, different types of liquid cooling, climate zones, implementation, facilities, TCO and liquid cooling for accelerators like GPUs and MIC. The presentation will benefit who wants to approach liquid cooling, giving an overview of the state of art of the available alternatives and describing some real application cases.

100% Server Heat Recapture in Data Centers is Now a Reality *Presenter: Christiaan Best (Green Revolution Cooling)*

Green Revolution Cooling, an Austin-based manufacturer of performance submersion cooling systems for OEM servers, now has two European installations that can capture 100% of server heat for reuse. Christiaan Best, Founder & CEO of Green Revolution Cooling, will discuss the company's CarnotJet System product offering and the company's new advanced in heat recapture. The CarnotJet System is a total submersion cooling solution for data center servers that requires less investment than air cooling (even free cooling) while reducing data center energy use by 50% and providing cooling overhead of up to 100 kW of true (fanless) computer per rack. The submersion cooling technology is server agnostic, accepting servers from any OEM, including Dell, HP, IBM, and Supermicro. With a quickly growing list of distinguished customers, including five of the Top 100 supercomputing sites, Green Revolution Cooling is driving the next phase of data center evolution.

Interconnect and Advanced Architectures II

3:30pm-5pm

Room: 155-C

Something New in HPC? EXTOLL: A Scalable Interconnect for Accelerators

Presenter: Ulrich Bruening (University of Heidelberg)

EXTOLL is a new, scalable and high-performance interconnection network originally developed at the University of Heidelberg. It implements a 3-D torus topology designed for low latency and reliable transmission of messages. Cut-through switching allows for low latency beyond nearest neighbor communication. The feature rich network EXTOLL was selected by the EU-Project DEEP to implement a separate network allowing for the virtualization and aggregation of accelerator resources into a so called Booster. EXTOLL's unique feature of a PCIe root-complex allows for autonomous boot and operation of accelerators within the Booster sub-system.

QFabric Technology: Revolutionizing the HPC Interconnect Architecture

Presenter: Masum Mir (Juniper Networks)

QFabric, a key part of this year's SCinet network, is a revolutionary new architecture for HPC interconnect based on extremely low latency, highly scalable any-to-any fabric. QFabric, which scales to thousands of ports, is a single fabric with 1GbE, 10GbE, 40GbE, and Fiber Channel connectivity for Layer-2 and Layer-3 switching, optimized for distributed compute and storage clusters that allows MPI, storage access and management networks to be consolidated into a single fabric. This session discusses this unique, robust fabric architecture that builds upon the principles of Layer-2/Layer-3 Ethernet switch fabric design and will touch on how QFabric is being used in SCinet and in various demonstrations throughout the SC12 exhibit floor. Learn about the metamorphosis of 10 & 40 gigabit Ethernet into a true fabric for next generation supercomputing and how to increase bandwidth and scale in cluster environments to enable the next generation computing clusters.

Deploying 40 and 100GbE: Optics and Fiber Media

Presenter: Anshul Sadana (Arista Networks)

As servers move to 10G and 40G, there is a need for faster uplinks and interconnecting clusters with 40G/100G. While moving from 10G to 40G seems straight forward over MMF, 100G largely relies on SMF. In this session, we will cover the fundamentals shaping future technology and what needs to be done to be ready for 40G and 100G in data centers and supercomputer clusters worldwide.

Thursday, November 15

Heterogeneous Computing II

10:30am-12pm

Room: 155-B

An OpenCL Application for FPGAs

Presenter: Desh Singh (Altera)

OpenCL is a framework that enables programmers to produce massively parallel software in C. OpenCL has been adopted by CPU and GPU developers as a way to accelerate their hardware by exploiting parallelism on their chip. FPGAs by their nature are fine-grained, massively parallel arrays that process information in a significantly different manner from traditional CPU- or GPU-based systems and are a natural hardware platform to target an OpenCL program. OpenCL and the parallelism of FPGAs enable a new level of hardware acceleration and faster time-to-market for heterogeneous systems. During this presentation, Altera will show how OpenCL is being used by customers to map data parallel algorithms to FPGA-based devices and achieve high-performance FPGA applications in a fraction of the time. We will also show how to transform initial code that is functionally correct into a highly optimized implementation that maximizes the throughput on the FPGA.

Acceleration of ncRNA Discovery via Reconfigurable Computing Platforms

Presenter: Nathaniel McVicar (University of Washington)

Non-coding RNAs (ncRNAs) are biologically important RNAs transcribed from portions of DNA that do not code for proteins. Instead, ncRNAs directly interact with the body's metabolic processes. Finding and understanding ncRNAs may reveal important answers for human biology, species evolution, and other fields. Unlike coding regions, many of an ncRNA's bases have purely structural rolls where the requirement for a specific structure is that certain sequence positions be complementary. This makes the identification of ncRNAs more computationally intensive than finding proteins. A team at the University of Washington, in collaboration with Pico Computing, is developing an FPGA-based system for detecting ncRNAs. By leveraging the massive fine-grained parallelism inside FPGAs, along with system design innovations, our system will improve performance by up to two orders of magnitude. In this talk we will present our current status, demonstrate advances made so far, and discuss how reconfigurable computing helps enable genomics applications.

HPC in the Datacenter II

10:30am-12pm

Room: 155-C

Addressing HPC Compute Center Challenges with Innovative Solutions

Presenter: Alan Powers, Donna Klecka (Computer Sciences Corporation)

For more than 50 years, CSC has developed innovative solutions to solve our clients' toughest challenges, demonstrating a commitment to excellence and a passion for exceeding expectations. We have proven experience in HPC delivering HPC solutions and services to DOD, Civil (NASA, NOAA) and commercial customers. CSC currently manages over 2 Pet-aFlops HPC systems as well as managing over 120 Petabytes of archive data for their customers. These include the top three SGI/DMF sites at NOAA and NASA. For over 20 years CSC's dedicated HPC focused team has architected solutions to customer challenges and continues to produce dozens of innovative, cost and time saving solutions annually. Topics: Big Data, Upgrading one of the largest data archive (40+ PB) in the world; Big Compute, New Facility to Production; Windows HPC Server in the Private Cloud; Expanding the Largest Infiniband (70 miles: FDR, QDR, and DDR) Compute Cluster.

Achieving Cost-Effective HPC with Process Virtualization

Presenter: Amy Wyron, Dori Exterman (IncrediBuild)

Discover how to achieve cost-effective supercomputing performance with existing hardware. Most organizations already have all of the processing power they need sitting idle on existing PCs and servers in the network. By harnessing idle CPU cycles in the local network and/or public cloud, and distributing an application's computational-intensive processes

to those resources for parallel processing, IncrediBuild's solutions leverage existing hardware to create what can best be described as a virtual supercomputer. IncrediBuild uses unique process virtualization technology to virtualize processes on-demand, eliminating the need to maintain virtual environments on remote machines, and resulting in an out-of-the-box solution that requires no integration effort and no maintenance. Discover how process virtualization works, and how we're using it to easily leverage existing IT infrastructure and achieve cost-effective supercomputing performance.

Reducing First Costs and Improving Future Flexibility in the Construction of HPC Facilities

Presenter: Rich Hering, Brian Renner, Colin Scott (M+W Group)

Economic realities place limits on the funding available to construct new HPC space. Changes in cooling technologies brought about by the ever increasing densities have pushed the industry to search for more flexible designs. In this paper, we present the various methods for reducing HPC initial construction cost, while providing flexibility, adaptability, and energy efficiency with a focus on total cost of ownership. We will present current and future trends in HPC requirements for space environment and infrastructure needs. Using this base criteria we will evaluate several best in class approaches to constructing new HPC space, including 1) Site selection 2) air and water based cooling technologies 3) right sizing and phased construction 4), modular and containerized spaces 5) Expandable and flexible infrastructure. Among other interesting results, we demonstrate the cost savings available in these scenarios, while allowing for flexibility and expandability.

HPC in the Cloud II

1:30pm-3pm

Room: 155-B

HPC Cloud and Big Data Analytics - Transforming High Performance Technical Computing

Presenter: Chris Porter, Scott Campbell (IBM Platform Computing)

Big Data and Cloud Computing have transitioned from being buzz words to transforming how we think about high-performance technical computing. Traditional technical computing implementations are deployed through purpose-built cluster and grid resources, resulting in monolithic silos, which are often either not fully utilized or overloaded. However, with the rise of Cloud Computing and new techniques for managing Big Data Analytics workloads, this scenario is changing. This presentation explores how Cloud and Workload Management solutions provide a mechanism to transform isolated technical computing resources into a shared resource pool for both compute- and data-intensive applications. On-demand access to resources that can be rapidly provisioned based on workload requirements provides research flexibility, easy access to resources, reduced management overhead, and optimal infrastructure utilization.

Compute Software and Remote Visualization for a Globalized Market

Presenter: Devarajan Subramanian (Gompute)

The Globalized Market brings a lot of challenges to High Performance Computing, with data centers and engineers spread across the globe. Enormous amounts of generated data and requirement for collaboration and visualization, combined with facts such as high latency and long distances have a huge impact on productivity and utilization of resources. Compute software addresses these issues and provides a comprehensive solution to these challenges.

Windows HPC in the Cloud

Presenter: Alex Sutton (Microsoft Corporation)

Windows Azure is a powerful public cloud-computing platform that offers on-demand, pay-as-you-go, access to massively scalable compute and storage resources. Microsoft HPC Pack 2012 and Windows Azure provide a comprehensive and cost-effective solution that delivers high performance while providing a unified solution for running compute and data intensive HPC applications on Windows Azure and on premises; you have the ability to rapidly deploy clusters comprising of thousands of nodes in Windows Azure and on premises.. Learn about Microsoft's latest innovations around big compute and big data solutions on Windows Azure . Visit us at Booth #801!

Software Development Tools II

1:30pm-3pm

Room: 155-C

SET - Supercomputing Engine Technology

Presenter: Dean Dauger (Advanced Cluster Systems, Inc.)

Avoid multithreading headaches! Our patented Supercomputing Engine Technology (SET) is an MPI-based library that simplifies parallel programming and attracts a much wider range of software writers, allowing their codes to scale more efficiently to hundreds or thousands of cores. Many scientific algorithms are data parallel, and in most cases, these are amenable to parallel formulation. Conventional parallel programming, however, suffers numerous pitfalls. When writing MPI or multithreading code, too many programmers are stuck with code that either performs badly or simply does not perform at all. Programmability is where SET makes a difference. SET raises the abstraction level, making parallel programming easier, and steers its users to good parallel code, all while executing efficiently on current hardware. SET has successfully parallelized Mathematica, video encoding, Scilab, and more. Take control of scaling your own code and come see us. We present SET's structure, API, and examples of its use.

The Future of OpenMP

Presenter: Michael Wong (OpenMP ARB)

Now celebrating its 15th birthday, OpenMP has proven to be a simple, yet powerful model for developing multi-threaded applications. OpenMP continues to evolve, adapt to new requirements, and push at the frontiers of parallelization. It is developed by the OpenMP Architecture Review Board, a group of 23 vendors and research organizations. A comment draft of the next specification version will be released at or close to SC12. It will include several significant enhancements, including support for accelerators, error handling, thread affinity, tasking extensions and support for Fortran 2003. We will give an overview of the new specifications, after having described the process to get to this new specification.

Let The Chips Fall Where They May - PGI Compilers & Tools for Heterogeneous HPC Systems

Presenter: Michael Wolfe (Portland Group, Inc.)

Diversity of processor architectures and heterogeneity of HPC systems is on the rise. CPU and Accelerator processors from AMD, ARM, IBM, Intel, NVIDIA and potentially other suppliers are in the mix. Some HPC systems feature "big" CPU cores running at the fastest possible clock rates. Others feature larger numbers of "little" CPU cores running at moderate clock rates. Some HPC systems feature CPUs and Accelerators with separate memories. Others feature CPUs and Accelerators on the same chip die with shared virtual or physical memory. Nearly every possible combination of these processor types is being evaluated or proposed for future HPC systems. In this talk, we will discuss the challenges and opportunities in designing and implementing languages to maximize productivity, portability and performance on current and future heterogeneous HPC systems.

Special Event-Orange FS Drop-in

3:30pm-5pm

Room: 155-B

OrangeFS Drop-In

Presenters: Boyd Wilson, Randy Martin, Walt Ligon (Omnibond, Clemson University)

Visit the OrangeFS Drop-In, a casual gathering where you can meet with members of the OrangeFS community and development team. Experts and leaders from the various areas of development and support will be on hand to discuss the current release, as well as directions and designs for the future. This is a great opportunity for us to get together on a more personal level. And it's a drop-in, so you can come any time during the session.



Communities

The Communities Program seeks to attract, train, and encourage tomorrow's high performance computing professionals. SC12 encourages students, researchers and faculty from among underrepresented communities, educators, early-career students, professionals, and international groups to participate in the SC conference through one or more of the following programs: Broader Engagement, HPC Educators Program, Student Cluster Competition, Doctoral Showcase, International Ambassadors and Student Volunteers, as well as other special programs being planned, such as the Student Job Fair, George Michael Memorial Ph.D. Fellowship, and the Mentoring Program.

HPC Educators

Sunday, November 11

Broader Engagement and Education in the Exascale Era

Chair: Almadena Chtchelkanova (National Science Foundation)

8:30am-10am

Room: Ballroom-D

Presenter: Thomas Sterling (Indiana University)

The once rarefied field of HPC now encompasses the challenges increasingly responsible for the mainstream. With the advent of multicore technology permeating not just the one million core+ apex of but the entire spectrum of computing platforms down to the pervasive cell phone, parallelism now challenges all computing applications and systems. Now education in HPC is education in all scalable computing. Without this, Moore's Law is irrelevant. The mastery and application of parallel computing demands the broadest engagement and education even as HPC confronts the challenges of exascale computing for real world application. This presentation will describe the elements of change essential to the effective exploitation of multicore technologies for scientific, commercial, security, and societal needs and discuss the tenets that must permeate the shared future of education in HPC and the broader computing domain.

Supercomputing in Plain English

10:30am-12pm

Room: 255-A

Presenter: Henry J. Neeman (University of Oklahoma)

This session provides a broad overview of HPC. Topics include: what is supercomputing; the fundamental issues of supercomputing (storage hierarchy, parallelism); hardware primer; introduction to the storage hierarchy; introduction to parallelism via an analogy (multiple people working on a jigsaw puzzle); Moore's Law; the motivation for using HPC. Prerequisite: basic computer literacy.

A Nifty Way to Introduce Parallelism into the Introductory Programming Sequence

1:30pm-5pm

Room: 255-A

Presenters: David Valentine (Slippery Rock University), David Mackay (Intel)

Introductory programming classes feed a broad cross section of STEM disciplines, especially those engaged in HPC. This half day session will be a hands-on experience in adding parallelism

to several of the ACM SIGCSE Nifty programming assignments. These assignments have already been designated as exceedingly clever and engaging by the SIGCSE membership, and so they are a great way to expose our introductory students to a parallel programming paradigm. We will begin with completed versions of the Nifty Programs and then use Intel's Parallel Studio to identify the hot spots that will benefit from parallelism. Finally, we will show how OpenMP can be added easily to the serial program. Thus we teach our introductory students how to grab the "low hanging fruit" and boost the productivity of their (already working) project. The session will be of particular use to educators wanting to introduce parallelism into introductory programming classes.

Introducing Computational Science in the Curriculum

1:30pm-5pm

Room: 255-D

Presenter: Steven I. Gordon (Ohio Supercomputer Center)

Computer modeling is an essential tool for discovery and innovation in science and engineering. It can also be used as a valuable educational tool, involving students in inquiry-based problems which simultaneously illustrate scientific concepts, their mathematical representation, and the computational techniques used to solve problems. A number of institutions have begun undergraduate computational science minor and certificate programs. One major challenge in such programs is to introduce students from a wide variety of backgrounds to the principles of modeling, the underlying mathematics, and the programming and analytical skills necessary to provide the foundation for more advanced modeling applications in each student's major area. This session will review the organization and course materials from such a course. Participants will use a variety of models that are used to illustrate modeling, mathematical, and scientific principles critical to beginning work in computational science.

SC12 Communities - Conference Orientation

5:15pm-6:15pm

Room: Ballroom-D

This session will provide an overview of the SC12 conference and describe the different areas, activities, and events that Communities participants can take advantage of during the conference. It is recommended for first-time and returning Communities program attendees.

Broader Engagement/HPC Educators Joint Resource Reception

7pm-10pm

Room: Ballroom-ABCE

This event gives participants an opportunity to network and to share examples of mentoring, engagement, and educational materials from their classrooms and programs.

Monday, November 12

Broader Engagement/HPC Educators Plenary Session

Chair: Valerie Taylor (Texas A&M University)
8:30am-10am
Room: Ballroom-D

The Fourth Paradigm - Data-Intensive Scientific Discovery

Presenter: Tony Hey (Microsoft Research)

There is broad recognition within the scientific community that the emerging data deluge will fundamentally alter disciplines in areas throughout academic research. A wide variety of scientists—biologists, chemists, physicists, astronomers, engineers – will require tools, technologies, and platforms that seamlessly integrate into standard scientific methodologies and processes. “The Fourth Paradigm” refers to the data management techniques and the computational systems needed to manipulate, visualize, and manage large amounts of scientific data. This talk will illustrate the challenges researchers will face, the opportunities these changes will afford, and the resulting implications for data-intensive researchers.

Python for Parallelism in Introductory Computer Science Education

10:30am-12pm
Room: 255-D

Presenter: Steven A. Bogaerts (Wittenberg University), Joshua V. Stough (Washington and Lee University)

Python is a lightweight high-level language that supports both functional and object-oriented programming. The language has seen rapid growth in popularity both in academe and industry, due to the ease with which programmers can implement their ideas. Python’s easy expressiveness extends to programming using parallelism and concurrency, allowing the early introduction of these increasingly critical concepts in the computer science core curriculum. In this two-hour session we describe and demonstrate an educational module on parallelism and its implementation using Python’s standard multiprocessing module. We cover such key concepts as speedup, divide and conquer, communication, and concurrency. We consider how these concepts may be taught in the context of CS1 and CS2, and we provide extensive hands-on demonstrations of parallelized integration, Monte Carlo simulations, recursive sorting schemes, and distributed computing.

LittleFe Buildout Workshop

10:30am-5pm
Room: 255-A

Presenter: Charlie Peck (Earlham College)

The LittleFe buildout will consist of participants assembling their LittleFe unit from a kit; installing the Bootable Cluster CD (BCCD) Linux distribution on it; learning about the curriculum modules available for teaching parallel programming, HPC and CDESE that are built-in to the BCCD; and learning how to develop new curriculum modules for the LittleFe/BCCD platform.

Going Parallel with C++11

1:30pm-5pm
Room: 255-D

Presenter: Joe Hummel (University of California, Irvine)

As hardware designers turn to multi-core CPUs and GPUs, software developers must embrace parallel programming to increase performance. No single approach has yet established itself as the “right way” to develop parallel software. However, C++ has long been used for performance-oriented work, and it’s a safe bet that any viable approach involves C++. This position has been strengthened by ratification of the new C++0x standard, officially referred to as “C++11”. This interactive session will introduce the new features of C++11 related to parallel programming, including type inference, lambda expressions, closures, and multithreading. The workshop will close with a brief discussion of other technologies, in particular higher-level abstractions such as Intel Cilk Plus and Microsoft PPL.

Tuesday, November 13

Invited Talk

10:30am-12pm
Room: 255-D

TCCP: Parallel and Distributed Curriculum Initiative

Speaker: Sushil K. Prasad (Georgia State University)

The goal of core curricular guidelines on parallel and distributed computing (PDC) effort is to ensure that all students graduating with a bachelor’s degree in computer science/computer engineering receive an education that prepares them in the area of parallel and distributed computing, increasingly important in the light of emerging technology, in which their coverage might find an appropriate context. Roughly six dozen early-adopter institutions worldwide are currently trying out this curriculum. They will receive periodically updated versions of the guidelines. The early adopters have been awarded

stipends through four rounds of competitions (Spring and Fall 2011, and Spring and Fall 2012) with support from NSF, Intel, and NVIDIA. Additional competitions are planned for Fall 2013 and Fall 2014. A Center for Parallel and Distributed Computing Curriculum Development and Educational Resources (CDER) is being established to carry the work forward, possibly due to a new NSF grant.

Unveiling Parallelization Strategies at Undergraduate Level

10:30am-12pm

Room: 255-A

Presenters: Eduard Ayguadé (Polytechnic University of Catalonia), Rosa Maria Badia (Barcelona Supercomputing Center)

We are living the “real” parallel computing revolution. Something that was the concern of a “few” forefront scientists has become mainstream and of concern to every single programmer. This HPC Educator Session proposes a set of tools to be used at undergraduate level to discover parallelization strategies and their potential benefit. Tareador provides a very intuitive approach to visualize different parallelization strategies and understand their implications. The programmer needs to use simple code annotations to identify tasks and Tareador will dynamically build the computation task graph, identifying all data-dependencies among the annotated tasks. Tareador also feeds Dimemas, a simulator to predict the potential of the proposed strategy and visualize an execution timeline (Paraver). Using the environment, we show a top-down approach that leads to appropriate parallelization strategies (task decomposition and granularity) and helps to identify tasks interactions that need to be guaranteed when coding the application in parallel.

GPU Computing as a Pathway to System-conscious Programmers

1:30pm-5pm

Room: 255-A

Presenter: Daniel J. Ernst (Cray Inc.)

This session will explore GPU computing as one pathway for creating undergraduate students with a broad sense of performance-awareness, and who are ready to tackle the architectural changes that developers will face now and in the near future. The presentation will specifically address what important concepts GPU computing exposes to students, why GPUs provide a more motivating educational tool than traditional CPUs, where to approach inserting these topics into a curriculum, and how to effectively present these concepts in a classroom. The session will include hands-on exercises and classroom-usable demonstrations, as well as time to discuss the issues that arise in integrating these kinds of materials into a diverse set of curricular circumstances. The presenter will use Nvidia’s CUDA for the session, but the topics translate well to other throughput computing platforms.

Test-Driven Development for HPC Computational Science & Engineering

1:30pm-5pm

Room: 255-D

Presenter: Jeffrey Carver (University of Alabama)

The primary goal of this half-day session is to teach HPC Educators, especially those that interact directly with students in Computational Science & Engineering (CSE) domains, about Test-Driven Development (TDD) and provide them with the resources necessary to educate their students in this key software engineering practice. The session will cover three related topics: Unit Testing, Automated Testing Frameworks, and Test-Driven Development. For each session topic, there will be a lecture followed by a hands-on exercise. Both the lecture material and the hands-on exercise materials will be provided to attendees for subsequent use in their own classrooms. Due to the difficulties in testing many CSE applications, developers are often not able to adequately test their software. This session will provide attendees with an approach that will enable them to teach students how to develop well-tested software.

Wednesday, November 14

Computational Examples for Physics (and Other) Classes Featuring Python, Mathematica and eTextBook, and More

10:30am-12pm

Room: 255-A

Presenters: Rubin H. Landau (Oregon State University), Richard G. Gass (University of Cincinnati)

This tutorial provides examples of research-level, high performance computing that can be used in courses throughout the undergraduate physics curriculum. At present, such examples may be found in specialty courses in Computational Physics, although those courses too often focus on programming and numerical methods. In contrast, physics classes tend to use computing as just pedagogic tools to teach physics without attempting to provide understanding of the computation. The examples presented will contain a balance of modern computational methods, programming, and interesting physics and science. The Python examples derive from an eTextBook available from the National Science Digital Library that includes video-based lectures, Python programs, applets, visualizations and animations. The Mathematica examples will focus on non-linear dynamic, quantum mechanics and visualizations. Whether using Mathematica or Python, the session looks inside the computation black box to understand the algorithms and to see how to scale them to research-level HPC.

An Educator's Toolbox for CUDA**10:30am-5pm****Room: 255-D**

Presenters: Karen L. Karavanic (Portland State University), David Bunde (Knox College), Barry Wilkinson (University of North Carolina at Charlotte), Jens Mache (Lewis and Clark College)

GPUs (graphical processing units) with large numbers of cores are radically altering how high performance computing is conducted. With the introduction of CUDA for general-purpose GPU programming, we can now program GPUs for computational tasks and achieve orders of magnitude improvement in performance over using the CPU alone. The importance of this approach, combined with the easy and inexpensive availability of hardware, combine to make this an excellent classroom topic. How to get started? The purpose of this workshop is to provide CS educators with the fundamental knowledge and hands-on skills to teach CUDA materials. Four session leaders with a combined total of over four decades of teaching experience will present short lectures, exercises, course materials, and panel discussions.

Cyber-Physical Systems**1:30pm-5pm****Room: 255-A**

Presenters: Xiaolin Andy Li, Pramod Khargonekar (University of Florida)

Cyber-physical systems (CPS) have permeated into our daily lives before we realize them—from smartphones, mobile services, transportation systems, to smart grids and smart buildings. This tutorial introduces the basic notion of CPS, its key features, its design space, and its life cycle— from sensing, processing, decision-making, to control and actuation. Emerging research on clouds of CPS will also be discussed. Although CPS research is still in infancy, CPS applications are abundant. Through selected case studies, we attempt to distill key understanding of emerging CPS models and methods. The case studies include CPS applications in smart grids, health-care, assistive robots, and mobile social networks. Large-scale CPS systems are typically data-intensive and involve complex decision-making. We will also introduce big data processing and programming paradigms to support CPS systems.

HPC: Suddenly Relevant to Mainstream CS Education?**1:30pm-5pm****Room: 355-A**

Presenter: Matthew Wolf (Georgia Institute of Technology)

Significant computer science curriculum initiatives are underway, with parallel and distributed computing and the impacts of multi-/many-core infrastructures and ubiquitous cloud computing playing a pivotal role. The developing guidelines will impact millions of students worldwide, and many emerg-

ing geographies are looking to use them to boost competitive advantage. Does this mainstream focus on ubiquitous parallelism draw HPC into the core of computer science, or does it make HPC's particular interests more remote from the cloud/gaming/multi-core emphasis? Following the successful model used at SC10 and SC11, the session will be highly interactive. An initial panel will lay out some of the core issues, with experts from multiple areas in education and industry. Following this will be a lively, moderated discussion to gather ideas from participants about industry and research needs as well as the role of academia in HPC.

Thursday, November 15**Computational Examples for Physics (and Other) Classes Featuring Python, Mathematica an eTextBook, and More****10:30am-12pm****Room: 255-A**

Presenters: Rubin H. Landau (Oregon State University), Richard G. Gass (University of Cincinnati)

This tutorial provides examples of research-level, high performance computing that can be used in courses throughout the undergraduate physics curriculum. At present, such examples may be found in specialty courses in Computational Physics, although those courses too often focus on programming and numerical methods. In contrast, physics classes tend to use computing as just pedagogic tools to teach physics without attempting to provide understanding of the computation. The examples presented will contain a balance of modern computational methods, programming, and interesting physics and science. The Python examples derive from an eTextBook available from the National Science Digital Library that includes video-based lectures, Python programs, applets, visualizations and animations. The Mathematica examples will focus on non-linear dynamic, quantum mechanics and visualizations. Whether using Mathematica or Python, the session looks inside the computation black box to understand the algorithms and to see how to scale them to research-level HPC.

Teaching Parallel Computing through Parallel Prefix**10:30am-12pm****Room: 255-D**

Presenter: Srinivas Aluru (Iowa State University)

Some problems exhibit inherent parallelism which is so obvious that learning parallel programming is all that is necessary to develop parallel solutions for them. A vast majority of problems do not fall under this category. Much of the teaching in parallel computing courses is centered on ingenious solutions developed for such problems, and this is often difficult for students to grasp. This session will present a novel way of teaching parallel computing through the prefix sum problem. The

session will first introduce prefix sums and present a simple and practically efficient parallel algorithm for solving it. Then a series of interesting and seemingly unrelated problems are solved by clever applications of parallel prefix. The applications range from generating random numbers, to computing edit distance between two strings using dynamic programming, to the classic N-body simulations that have long been a staple of research in the SC community.

CSinParallel: An Incremental Approach to Adding PDC throughout the CS Curriculum

1:30pm-5pm

Room: 255-A

Presenters: Richard A. Brown (St. Olaf College), Elizabeth Shoop (Macalester College), Joel Adams (Calvin College)

CSinParallel.org offers an incremental approach for feasibly adding PDC (parallel and distributed computing) to existing undergraduate courses and curricula, using small (1-3 class days) teaching modules. Designed for flexible use in multiple courses throughout a CS curriculum, and typically offering a choice of programming languages, CSinParallel modules have minimal syllabus cost, yet contribute significantly to student understanding of PDC principles and practices. This HPC Educator Session explores three modules that focus on introducing notions in concurrency, multi-threaded programming, and map-reduce “cloud” computing, and will briefly survey other modules in the series. Featured modules are suitable for use in courses ranging from the beginning to advanced levels, and each module’s presentation includes hands-on experience for participants and reports from experience teaching those modules. We welcome participants with any level of exposure to PDC. Publicly available and affordable PDC computational platforms will be provided.

High-level Parallel Programming using Chapel

1:30pm-5pm

Room: 255-D

Presenters: David P. Bunde (Knox College), Kyle Burke (Wittenberg University)

Chapel is a parallel programming language that provides a wide variety of tools to exploit different kinds of parallelism. It is flexible, supporting both OO programming and a low-overhead style similar to scripting languages. Of particular note is its expressiveness; a single keyword launches an asynchronous task or performs a parallel reduction. Data parallelism is easily expressed using domains, index sets that can be grown, translated, or intersected. The availability of high-level parallel operations such as these makes Chapel well-suited for students since concise examples help them focus on the main point and students can quickly try different parallel algorithms. This session features a demonstration of the basics of Chapel, including hands-on exercises, followed by a discussion of ways it can benefit a wide variety of courses.

Broader Engagement

Sunday, November 11

Broader Engagement and Education in the Exascale Era

Chair: Almadena Chtchelkanova (National Science Foundation)

8:30am-10am

Room: Ballroom-D

Presenter: Thomas Sterling (Indiana University)

The once rarefied field of high performance computing (HPC) now encompasses the challenges increasingly responsible for the mainstream. With the advent of multicore technology permeating not just the 1 million core+ apex of but the entire spectrum of computing platforms down to the pervasive cell phone, parallelism now challenges all computing applications and systems. Now education in HPC is education in all scalable computing. Without this, Moore’s Law is irrelevant. The mastery and application of parallel computing demands the broadest engagement and education even as HPC confronts the challenges of exascale computing for real world application. This presentation will describe the elements of change essential to the effective exploitation of multicore technologies for scientific, commercial, security, and societal needs and discuss the tenets that must permeate the shared future of education in HPC and the broader computing domain.

Broadening Participation in HPC and Supercomputing R&D

Chair: Dorian C. Arnold (University of New Mexico)

10:30am-12pm

Room: 355-A

The Importance of Broader Engagement for HPC

Author: Valerie Taylor (Texas A&M University)

It is recognized that broader engagement is important to the field of high performance computing as different perspectives often result in major breakthroughs in a field. This talk will focus on the need to leverage from different cultural perspectives within HPC. With respect to cultural perspective, the presentation will focus on underrepresented minorities—African Americans, Hispanics, and Native Americans. The talk will build upon the experiences of the presenter.

Programming Exascale Supercomputers

Presenter: Mary Hall (University of Utah)

Predictions for exascale architectures include a number of changes from current supercomputers that will dramatically impact programmability and further increase the challenges faced by high-end developers. With heterogeneous processors, dynamic billion-way parallelism, reduced storage per flop, deeper and configurable memory hierarchies, new memory technologies, and reliability and power concerns, the costs of software development will become unsustainable using current approaches. In this talk, we will explore the limitations of current approaches to high-end software development and how exascale architecture features will exacerbate these limitations. We will argue that the time is right for a shift to new software technology that aids application developers in managing the almost unbounded complexity of mapping software to exascale architectures. As we rethink how to program exascale architectures, we can develop an approach that addresses all of the productivity, performance, power and reliability.

Gaming and Supercomputing

Chair: Sadaf R. Alam (Swiss National Supercomputing Centre)

1:30pm-3pm

Room: 355-A

An Unlikely Symbiosis: How the Gaming and Supercomputing Industries are Learning from and Influencing Each Other

Presenter: Sarah Tariq (NVIDIA)

Over the last couple of decades video games have evolved from simple 2D sprite-based animations to nearly realistic cinematic experiences. Today's games are able to do cloth, rigid body and fluid simulations, compute realistic shading and lighting, and apply complex post processing effects including motion blur and depth of field, all in less than a 60th of a second. The hardware powering these effects, the Graphics Processing Unit, has evolved over the last 15 years from a simple fixed-function triangle rasterizer to a highly programmable general purpose massively-parallel processor with high-memory bandwidth and high performance per watt. These characteristics make the GPU ideally suited for typical supercomputing tasks. GPUs have become widely adopted accelerators for high performance computing. The game industry has continued to push the increase in visual fidelity; many algorithms that were once only useful in the HPC world are becoming interesting for real-time use.

L33t HPC: How Teh Titan's GPUs Pwned Science

Presenter: Fernanda Foertter (Oak Ridge National Laboratory)

For a very long time, scientific computing has been limited to the economics of commodity hardware. While special purpose processors have always existed, squeezing performance

meant burdensome code re-writes and very high costs. Thus, the use of commodity processors for scientific computing was a compromise between cheaper hardware and portable code and consistent performance. Gaming changed the economics of these special purpose graphic processors, but the advent of frameworks and APIs turned them into general purpose GPUs, whose architecture is ideal for massively parallel scientific computing. Oak Ridge National Lab will accelerate time-to-results by using Kepler GPUs to a peak theoretical performance of over 20PF in a new system named Titan. This presentation will show preliminary results of GPU accelerated applications on Titan and other details of the system.

Big Data and Visualization for Scientific Discoveries

Chair: Linda Akli (Southeastern Universities Research Association)

3:30pm-5pm

Room: 355-A

Visual Computing - Making Sense of a Complex World

Author: Chris Johnson (University of Utah)

Computers are now extensively used throughout science, engineering, and medicine. Advances in computational geometric modeling, imaging, and simulation allow researchers to build and test models of increasingly complex phenomena and to generate unprecedented amounts of data. These advances have created the need to make corresponding progress in our ability to understand large amounts of data and information arising from multiple sources. To effectively understand and make use of the vast amounts of information being produced is one of the greatest scientific challenges of the 21st Century. Visual computing—which relies on, and takes advantage of, the interplay among techniques of visualization, large-scale computing, data management and imaging—is fundamental to understanding models of complex phenomena, which are often multi-disciplinary in nature. This talk provides examples of interdisciplinary visual computing and imaging research at the Scientific Computing and Imaging (SCI) Institute as applied to important problems in science, engineering, and medicine.

XSEDE (Extreme Science and Engineering Discovery Environment)

Presenter: John Towns (National Center for Supercomputing Applications)

The XSEDE (Extreme Science and Engineering Discovery Environment) is one of the most advanced, powerful, and robust collections of integrated advanced digital resources and services in the world. It is a single virtual system that scientists can use to interactively share computing resources, data and expertise. Scientists, engineers, social scientists, and humanists around the world use advanced digital resources and services every day. Supercomputers, collections of data, and

software tools are critical to the success of those researchers, who use them to make our lives healthier, safer, and better. XSEDE integrates these resources and services, makes them easier to use, and helps more people use them. XSEDE supports 16 supercomputers and high-end visualization and data analysis resources across the country. Researchers use this infrastructure to handle the huge volumes of digital information that are now a part of their work.

Monday, November 12

Energy Efficient HPC Technologies

Chair: Sadaf R. Alam (Swiss National Supercomputing Centre)
10:30am-12pm
Room: 355-A

The Sequoia System and Facilities Integration Story

Presenter: Kim Cupps (Lawrence Livermore National Laboratory)

Sequoia, a 20PF/s Blue Gene/Q system, will serve National Nuclear Security Administration's Advanced Simulation and Computing (ASC) program to fulfill stockpile stewardship requirements through simulation science. Problems at the highest end of this computational spectrum are a principal ASC driver as highly predictive codes are developed. Sequoia is an Uncertainty Quantification focused system at Lawrence Livermore National Laboratory (LLNL). Sequoia will simultaneously run integrated design code and science materials calculations enabling sustained performance of 24 times ASC's Purple calculations and 20 times ASC's Blue Gene/L calculations. LLNL prepared for Sequoia's delivery for over three years. During the past year we have been consumed with the integration challenges of siting the system and its facilities and infrastructure. Sequoia integration continues, acceptance testing begins in September, and production level computing is expected in March 2013. This talk gives an overview of Sequoia and its facilities and system integration victories and challenges.

Using Power Efficient ARM-Based Servers for Data Intensive HPC

Presenter: Karl Freund (Calxeda)

After two years of anticipation and rumors, the 1st ARM-based servers are now available and being tested for various workloads across the industry. This talk will cover some of the early experiences with Calxeda-based servers in various labs and institutions, as well as internal benchmarking conducted by Calxeda. A brief look ahead will provide planning assumptions for Calxeda's roadmap.

Accelerator Programming

Chair: Sadaf R. Alam (Swiss National Supercomputing Centre)
1:30pm-3pm
Room: 355-A

OpenMP: The "Easy" Path to Shared Memory Computing

Presenter: Tim Mattson (Intel Corporation)

OpenMP is an industry standard application programming interface (API) for shared-memory computers. OpenMP is an attempt to make parallel application programming "easy" and embraces the oft-quoted principle: "Everything should be as simple as possible, but not simpler." OpenMP was first released in 1997 with a focus on parallelizing the loop-oriented programs common in scientific programming. Under continuous development since then, it now addresses a much wider range of parallel algorithms including divide-and-conquer and producer-consumer algorithms. This talk is an introduction to OpenMP for programmers. In addition to covering the core elements of this API, we will explore some of the key enhancements planned for future versions of OpenMP.

OpenACC, An Effective Standard for Developing Performance Portable Applications for Future Hybrid Systems

Presenter: John Levesque (Cray Inc.)

For the past 35 years, comment-line directives have been used effectively to give the compiler additional information for optimizing the target architecture. Directives have been used to address vector, super-scalar, shared memory parallelization and now Hybrid architectures. This talk will show how the new proposed OpenACC directives have been effectively utilized for a diverse set of applications. The approach for using OpenACC is to add the OpenACC directives to an efficient OpenMP version of the application. The OpenACC directives are a superset containing the features of OpenMP for handling shared and private data plus additional directives for handling data management between the host and accelerator memory. The resultant code can then be run on MPP systems containing multi-core nodes, with or without an accelerator. Future systems will undoubtedly have a large slow memory and a smaller faster memory. OpenACC can effectively be utilized to handle such hybrid memory systems

HPC Challenges and Directions**Chair: Dorian C. Arnold (University of New Mexico)****3:30pm-5pm****Room: 355-A****The Growing Power Struggle in HPC***Presenter: Kirk Cameron (Virginia Tech)*

The power consumption of supercomputers ultimately limits their performance. The current challenge is not whether we will achieve an exaflop by 2018, but whether we can do it in less than 20 megawatts. The SCAPE Laboratory at Virginia Tech has been studying the tradeoffs between performance and power for over a decade. We've developed an extensive tool chain for monitoring and managing power and performance in supercomputers. We will discuss the implications of our findings for exascale systems and some research directions ripe for innovation.

Heading Towards Exascale—Techniques to Improve Application Performance and Energy Consumption Using Application-Level Tools*Presenter: Charles Lively (Oak Ridge National Laboratory)*

Power consumption is an important constraint in achieving efficient execution on HPC multicore systems. As the number of cores available on a chip continues to increase, the importance of power consumption will continue to grow. In order to achieve improved performance on multicore systems, scientific applications must make use of efficient methods for reducing power consumption and must further be refined to achieve reduced execution time. Currently, more tools are becoming available at the application-level for power and energy consumption measurements. The available tools allow for the performance measurements obtained to be used for modeling and optimizing the energy consumption of scientific applications. This talk will describe efforts in the Multiple Metrics Modeling Infrastructure (MuMMI) project to build an integrated infrastructure for measurement, modeling, and prediction of performance and power consumption, including E-AMOM, Energy-Aware Modeling and Optimization Methodology.

Mentor-Protégé Mixer**Chair: Raquell Holmes (improvsience)****5pm-7pm****Room: Ballroom-A**

The Mentor-Protégé program was initiated by the Broader Engagement Committee to support the development of new members and leaders in HPC-related fields. During pre-conference registration attendees of the SC conference elect to mentor students and professionals who participate in the SC Communities programs (HPC Educators, Student Volunteers,

Student Cluster Challenge, and Broader Engagement). The Mentor-Protégé mixer is the dedicated time for Mentor-Protégé pairs to meet, discuss ways to take advantage of the conference, identify potential areas of overlapping interest and suggest paths for continued involvement in the SC technical fields. This session is explicitly for participants of the Mentor-Protégé program.

Tuesday, November 13**Mentoring: Building Functional Professional Relationships****Chair: Raquell Holmes (improvsience)****10:30am-12pm****Room: 355-A**

As our technical and scientific fields increase in diversity, we have the challenge and opportunity to build professional relationships across apparent socio-cultural differences (gender, race, class, sexuality and physical ability). Successful mentor-protégé, advisor-advisee, or manager-staff relationships are dynamic and each relationship, despite common characteristics, is unique. Seeing the improvisational nature of such relationships can help us transform awkward, tentative alliances into functional units. In this session, we explore the challenges and opportunities that come with mentoring relationships and highlight the improvisational skills that help build functional, developmental relationships throughout our professional careers.

Panel: The Impact of the Broader Engagement Program—Lessons Learned of Broad Applicability**Chair: Roscoe C. Giles (Boston University)****1:30pm-3pm****Room: 355-A**

Panelists will provide information about programs used to broaden participation in HPC. Lessons learned and new ideas used will be presented. Discussion among panelists and audience is encouraged.

Wednesday, November 15

Broader Engagement Wrap-up Session: Program Evaluation and Lessons Learned

Chair: Tiki L. Suarez-Brown (Florida A&M University)

3:30pm-5pm

Room: 355-A

During this session the organizers request feedback from this year's participants. Participants should provide information about their experiences, both good and bad. Future committee members will use this information to continue to improve the program to enhance its benefit to the community.

Doctoral Showcase

Thursday, November 14

Doctoral Showcase I – Dissertation Research Showcase

Chair: Yong Chen (Texas Tech University)

10:30am-12pm

Room: 155-F

Analyzing and Reducing Silent Data Corruptions Caused By Soft-Errors

Presenter: Siva Kumar Sastry Hari (University of Illinois at Urbana-Champaign)

Hardware reliability becomes a challenge with technology scaling. Silent Data Corruptions (SDCs) from soft-errors pose a major threat in commodity systems space. Hence significantly reducing the user-visible SDC rate is crucial for low-cost in-field reliability solutions. This thesis proposes a program-centric approach to identify application locations that cause SDCs and convert them to detections using low-cost program-level error detectors. We developed Relyzer to obtain a detailed application resiliency profile by systematically analyzing all application fault-sites without performing time-consuming fault injections on all of them. It employs novel fault pruning techniques to lower the evaluation time by 99.78% for our workloads. Using Relyzer, we obtained and analyzed the comprehensive list of SDC-causing instructions. We then developed program-level error detectors that on average provide a much lower-cost alternative to a state-of-the-art solution for all SDC rate targets. Overall, we provide practical and flexible choice points on the performance vs. reliability trade-off curves.

Fast Multipole Methods on Heterogeneous Architectures

Presenter: Qi Hu (University of Maryland)

The N-body problem, in which the sum of N kernel functions centered at N source locations with strengths are evaluated at M receiver locations, arises in a number of contexts, such as stellar dynamics, molecular dynamics, etc. Particularly, in our project, the high fidelity dynamic simulation of brownout dust clouds by using the free vortex wake method requires millions of particles and vortex elements. Direct evaluations have quadratic cost, which is not practical to solve such dynamic N-body problems in the order of millions or larger. My dissertation is mainly on combining the algorithmic and hardware acceleration approaches to speed-up N-body applications: develop effective fast multipole methods (FMM) algorithms on the heterogeneous architectures. Our major contributions are the novel FMM parallel data structures on GPU, the fully distributed heterogeneous FMM algorithms with the state-of-art implementations, and their adaptations with novel reformulations to vortex methods as well as other applications.

Algorithmic Approaches to Building Robust Applications for HPC Systems of the Future

Presenter: Joseph Sloan (University of Illinois at Urbana-Champaign)

The increasing size and complexity of High Performance Computing systems is making it increasingly likely that individual circuits will produce erroneous results, especially when operated in a low energy mode. We describe one general approach for converting applications into more error tolerant forms by recasting these applications as numerical optimization problems. We also show how both intrinsically robust algorithms as well as fragile applications can exploit this framework and in some cases provide significant energy reduction. We also propose a set of algorithmic techniques for detecting faults in sparse linear algebra. These techniques are shown to yield up to 2x reductions in performance overhead over traditional ABFT checks. We also propose algorithmic error localization and partial recomputation as an approach for efficiently making forward progress. This approach improves the performance of the CG solver in high error scenarios by 3x- 4x and increases the probability of successful completion by 60%.

Total Energy Optimization for High Performance Computing Data Centers

Presenter: Osman Sarood (University of Illinois at Urbana-Champaign)

Meeting energy and power requirements for huge exascale machines is a major challenge. Energy costs for data centers can be divided into two main categories: machine energy and cooling energy consumptions. This thesis investigates reduction in energy consumption for HPC data centers in both these

categories. Our recent work on reducing cooling energy consumption shows that we can reduce cooling energy consumption by up to 63% using our temperature aware load balancer. In this work, we also demonstrate that data centers can reduce machine energy consumption by up to 28% by running different parts of the applications at different frequencies. The focus of our work is to gauge the potential for energy saving by reducing both machine and cooling energy consumption separately (and their associated execution time penalty) and then come up with a scheme that combines them optimally to reduce total energy consumption for large HPC data centers.

Parallel Algorithms for Bayesian Networks Structure Learning with Applications to Gene Networks

Presenter: Olga Nikolova (Iowa State University)

Bayesian networks (BNs) are probabilistic graphical models which have been used to model complex regulatory interactions in the cell (gene networks). However, BN structure learning is an NP-hard problem and both exact and heuristic methods are computationally intensive with limited ability to produce large networks. To address these issues, we developed a set of parallel algorithms. First, we present a communication efficient parallel algorithm for exact BN structure learning, which is work- and space-optimal, and exhibits near perfect scaling. We further investigate the case of bounded node in-degree, where a limit d on the number of parents per variable is imposed. We characterize the algorithm's runtime behavior as a function of d . Finally, we present a parallel heuristic approach for large-scale BN learning, which aims to combine the precision of exact learning. We evaluate the quality of the learned networks using synthetic and real gene expression data.

Exploring Multiple Levels of Heterogeneous Performance Modeling

Presenter: Vivek V. Pallipuram Krishnamani (Clemson University)

Heterogeneous performance prediction models are valuable tools to accurately predict application runtime, allowing for efficient design space exploration and application mapping. Existing performance prediction models require intricate computing architecture knowledge and do not address multiple levels of design space abstraction. Our research aims to develop an easy-to-use and accurate multi-level performance prediction suite that addresses two levels of design space abstraction: low-level with partial implementation details and system specifications; and high-level with minimum implementation details and high-level system specifications. The proposed multi-level performance prediction suite targets synchronous iterative algorithms (SIAs) using our synchronous iterative execution models for GPGPU and FPGA clusters. The current work focuses on low-level abstraction modeling for GPGPU clusters using regression analysis and achieves over

90% prediction accuracy for the chosen SIA case studies. Our continued research aims at complete construction of the heterogeneous performance modeling suite and validation using additional SIA case studies.

Doctoral Showcase II

Dissertation Research Showcase

Chair: Yong Chen (Texas Tech University)

1:30pm-3pm

Room: 155-F

Automatic Selection of Compiler Optimizations Using Program Characterization and Machine Learning

Presenter: Eunjung Park (University of Delaware)

Selecting suitable optimizations for a particular class of applications is difficult because of the complex interactions between the optimizations themselves and the involved hardware. It has been shown that machine-learning based driven optimizations often outperform bundled optimizations or human-constructed heuristics. In this dissertation, we propose to use different modeling techniques and characterizations to solve the current issues in machine-learning based selection of compiler optimizations. In the first part, we evaluate two different state-of-the-art predictive modeling techniques against a new modeling technique we invented, named the tournament predictor. We show that this novel technique can outperform the other two state-of-the-art techniques. In the second, we evaluate three different program characterization techniques including performance counters, reactions, and source code features. We also propose a novel technique using control flow graphs (CFG), which we named graph-based characterization. In the last part, we explored different graph-based IRs other than CFGs to characterize programs.

High Performance Non-Blocking and Power-Aware Collective Communication for Next Generation InfiniBand Clusters

Presenter: Krishna Kandalla (Ohio State University)

The design and development of current generation supercomputing systems is fueled by the increasing use of multi-core processors, accelerators and high-speed interconnects. However, scientific applications are unable to fully harness the computing power offered by current generation systems. Two of the most significant challenges are communication/synchronization latency and power consumption. Emerging parallel programming models offer asynchronous communication primitives that can, in theory, allow applications to achieve latency hiding. However, delivering near perfect communication/computation overlap is non-trivial. Modern hardware components are designed to aggressively conserve power during periods of inactivity. However, supercomputing systems are rarely idle and software libraries need to be designed in a power-aware manner. In our work, we address both of critical

problems. We propose hardware-based non-blocking MPI collective operations and re-design parallel applications to achieve better performance through latency hiding. We also propose power-aware MPI collective operations to deliver fine-grained power-savings to applications with minimal performance overheads.

Virtualization of Accelerators in High Performance Clusters

Presenter: Antonio J. Peña (Polytechnic University of Valencia)

In this proposal, GPU-accelerated applications are enabled to seamlessly interact with any GPU of the cluster independently of its exact physical location. This provides the possibility of sharing accelerators among different nodes, as GPU-accelerated applications do not fully exploit accelerator capabilities all the time, thus reducing power requirements. Furthermore, decoupling GPUs from nodes, creating pools of accelerators, brings additional flexibility to cluster deployments and allows accessing a virtually unlimited amount of GPUs from a single node, enabling, for example, a GPU-per-core configuration. Depending on the particular cluster needs, GPUs may be either distributed among computing nodes or consolidated into dedicated GPGPU servers, analogously to disk servers. In both cases, this proposal leads to energy, acquisition, maintenance, and space savings. Our performance evaluations employing the rCUDA Framework, developed as a result of the research conducted during the PhD period, demonstrate the feasibility of this proposal within the HPC arena.

Heterogeneous Scheduling for Performance and Programmability

Presenter: Thomas R. W. Scogland (Virginia Tech)

Heterogeneity is increasing at all levels of computing, with the rise of accelerators such as GPUs, FPGAs, and other co-processors into everything from desktops to supercomputers. More quietly it is increasing with the rise of NUMA systems, hierarchical caching, OS noise, and a myriad of other factors. As heterogeneity becomes a fact of life at every level of computing, efficiently managing heterogeneous compute resources is becoming a critical task; correspondingly however it increases complexity. Our work seeks to improve the programmability of heterogeneous systems by providing runtime systems, and proposed programming model extensions, which increase performance portability and performance consistency while retaining a familiar programming model for the user. The results so far are an extension to MPICH2 which automatically increases the performance consistency of MPI applications in unbalanced systems and a runtime scheduler which automatically distributes the iterations of Accelerated OpenMP parallel regions across CPU and GPU resources.

Integrated Parallelization of Computation and Visualization for Large-Scale Weather Applications

Presenter: Preeti Malakar (Indian Institute of Science)

Critical applications like cyclone tracking require simultaneous high-performance simulations and online visualization for timely analysis. These simulations involve large-scale computations and generate large amount of data. Faster simulations and simultaneous visualization enable scientists provide real-time guidance to decision makers. However, resource constraints like limited storage and slow networks can limit the effectiveness of on-the-fly visualization. We have developed an integrated user-driven and automated steering framework InSt that simultaneously performs simulations and efficient online remote visualization of critical weather applications in resource-constrained environments. InSt considers application dynamics like the criticality of the application and resource dynamics like the storage space and network bandwidth to adapt various application and resource parameters like simulation resolution and frequency of visualization. We propose adaptive algorithms to reduce the lag between the simulation and visualization times. We also improve the performance of multiple high-resolution nested simulations like simulations of multiple weather phenomena, which are computationally expensive.

Programming High Performance Heterogeneous Computing Systems: Paradigms, Models and Metrics

Presenter: Ashwin M. Aji (Virginia Tech)

While GPUs are computational powerhouses, GPU clusters are largely inefficient due to multiple data transfer costs across the PCIe bus. I have developed MPI-ACC, a high performance communication library for end-to-end data movement in CPU-GPU systems, where MPI-ACC is an extension to the popular MPI parallel programming paradigm. I provide a wide range of optimizations for point-to-point communication within MPI-ACC, which can be seamlessly leveraged by the application developers. I also show how MPI-ACC can further enable new application-specific optimizations, like efficient CPU-GPU co-scheduling and simultaneous CPU-GPU computation and network-GPU communication for improved system efficiency. I have also developed performance models to predict realistic performance bounds for GPU kernels, and this knowledge is used for optimal task distribution between the CPUs and GPUs for better efficiency. Lastly, I define a general efficiency metric for heterogeneous computing systems and show how MPI-ACC improves the overall efficiency of CPU-GPU based heterogeneous systems.

Doctoral Showcase III - Early Research Showcase

Chair: Wojtek James Goscinski (Monash University)

3:30pm-5pm

Room: 155-F

A Cloud Architecture for Reducing Costs in Local Parallel and Distributed Virtualized Environments

Presenter: Jeffrey M. Galloway (University of Alabama)

Deploying local cloud architectures can be beneficial to organizations that wish to maximize their computational and storage resources. Also, this architecture can be beneficial to organizations that do not wish to meet their needs using public vendors. The problem with hosting a private cloud environment includes overall cost, scalability, maintainability, and customer interfacing. Computational resources can be more utilized by using virtualization technology. Even then, there is room for improvement by using aggressive power saving strategies. Current open resource cloud implementations do not employ aggressive strategies for power reduction. My research in this area focuses on reducing power while maintaining high availability to compute resources. Clusters and clouds rely on the storage of persistent data. Deploying a power aware strategy for hosting persistent storage can improve performance per watt for the organization when combined with a power savings strategy for computational resources

Towards the Support for Many-Task Computing on Many-Core Computing Platforms

Presenter: Scott Krieder (Illinois Institute of Technology)

Current software and hardware limitations prevent Many-Task Computing (MTC) from leveraging hardware accelerators (NVIDIA GPUs, Intel MIC) boasting Many-Core Computing architectures. Some broad application classes that fit the MTC paradigm are workflows, MapReduce, high-throughput computing, and a subset of high-performance computing. MTC emphasizes using many computing resources over short periods of time to accomplish many computational tasks (i.e. including both dependent and independent tasks), where the primary metrics are measured in seconds. MTC has already proven successful in Grid Computing and Supercomputing on MIMD architectures, but the SIMD architectures of today's accelerators pose many challenges in the efficient support of MTC workloads on accelerators. This work aims to address the programmability gap between MTC and accelerators, through an innovative middleware that will enable MIMD workloads to run on SIMD architectures. This work will enable a broader class of applications to leverage the growing number of accelerated high-end computing systems.

Software Support for Regular and Irregular Applications in Parallel Computing

Presenter: Da Li (University of Missouri)

Today's applications can be divided into two categories: regular and irregular. Within regular applications, data are represented as arrays and stored in continuous memory. The arithmetic operations and memory access patterns are also very regular. In terms of irregular applications, data are represented as pointer-based tree and graphs. The memory access patterns are hard to predict so that utilizing locality may be infeasible. The parallelism is hard to extract at compile time because the data dependences are determined at runtime and more perplexing than regular ones. My research focus on following aspects: 1. Parallelization of regular & irregular algorithm on parallel architectures 2. Programming model and runtime system design for regular algorithm & irregular on parallel architectures

Towards Scalable and Efficient Scientific Cloud Computing

Presenter: Iman Sadooghi (Illinois Institute of Technology)

Commercial clouds bring a great opportunity to the scientific computing area. Scientific applications usually need huge resources to run on. However not all of the scientists have access to significant high-end computing systems, such as those found in the Top500. A good solution to this problem is using customized cloud services that are optimized for scientific application workloads. This work is investigating the ability of clouds to support the characteristics of scientific applications. These applications have grown accustomed to a particular software stack, namely one that supports batch scheduling, parallel and distributed POSIX-compliant file systems, and fast and low latency networks such as 10Gb/s Ethernet or Infini-band. This work will explore low overhead virtualization techniques (e.g. Palacios VMM), investigate network performance and how it might affect network bound applications, and explore a wide range of parallel and distributed file systems for their suitability of running in a cloud infrastructure.

On Bandwidth Reservation for Optimal Resource Utilization in High-Performance Networks

Presenter: Poonam Dharam (University of Memphis)

The bulk data transfer needs in network-intensive scientific applications necessitate the development of high-performance networks that are capable of provisioning dedicated channels with reserved bandwidths. Depending on the desired level of Quality of Service (QoS), some applications may request an advance reservation (AR) of bandwidth in a future time slot to ensure uninterrupted transport service, while others may request an immediate reservation (IR) of bandwidth upon availability to meet on-demand needs. The main issue in these networks is the increased possibility of preempting ongoing

IRs at the time of the activation of an AR due to the lack of available bandwidth. Such preemptions cause service degradation or discontinuity of IRs, hence impairing QoS promised to be provided for IRs at the time of their reservation confirmation. We propose a comprehensive bandwidth reservation solution to optimize network resource utilization by exploring the interactions between advance and immediate reservations.

Distributed File Systems for Exascale Computing

Presenter: Dongfang Zhao (Illinois Institute of Technology)

It is predicted that exascale supercomputers will be emerging by 2019. The current storage architecture of High Performance Computing (HPC) would unlikely work well due to the high degree of concurrency at exascales. By dispersing the storage throughout the compute infrastructure, storage resources (also processing capabilities and network bandwidth) will increase linearly with larger scales. We believe the co-location of storage and compute resources is critical to the impending exascale systems. This work aims to develop both theoretical and practical aspects of building an efficient and scalable distributed file system for HPC systems that will scale to millions of nodes and billions of concurrent I/O requests. A prototype has been developed and deployed on a 64-node cluster, and tested with micro benchmarks. We plan to deploy it on various supercomputers, such as the IBM Bluegene/P and Cray XT6, and run benchmarks and real scientific applications at petascales and beyond.

Dynamic Load-Balancing for Multicores

Presenter: Jesmin Jahan Tithi (Stony Brook University)

A very important issue for most parallel applications is efficient load-balancing. My current focus in PhD research is load-balancing for multicores and clusters of multicores. We have shown that sometimes an optimistic parallelization approach can be used to avoid the use of locks and atomic instructions during dynamic load balancing on multicores. We have used this approach to implement two high-performance parallel BFS algorithms based on centralized job queues and distributed randomized work-stealing, respectively. Our implementations are highly scalable and faster than most state-of-the-art multi-core parallel BFS algorithms. In my work on load-balancing on clusters of multicores, I have implemented distributed-memory and distributed-shared-memory parallel octree based algorithms for approximating polarization energy of molecules by extending existing work on shared-memory architectures. For large enough molecules our implementations show a speedup factor of about 10k w.r.t. the naïve algorithm with less than 1% error using as few as 144 cores (=12nodesX12cores/node).

Numerical Experimentations in the Dynamics of Particle-Laden Supersonic Impinging Jet Flow

Presenter: Paul C. Stegeman (Monash University)

The proposed thesis will be a numerical investigation of the physical mechanisms of under-expanded supersonic impinging (USI) jet flow and particle transport within it. A numerical solver has been developed implementing the LES methodology on a cylindrical curvilinear grid using both compact and explicit finite differencing. A hybrid methodology is used to accurately resolve discontinuities in which a shock selection function determines if the solver should apply a shock-capturing algorithm or a differencing algorithm. Particle transport within the USI jet will be studied by modeling the individual particles in a Lagrangian phase. Biased interpolation schemes with similar properties to the shock-capturing schemes are being developed. The use of these shock-optimized interpolation schemes will also be dependent on the shock-selection function that is used in the fluid phase. This aims to improve the accuracy of the particles' dynamics as they travel through shock waves.

An Efficient Runtime Technology for Many-Core Device Virtualization in Clusters

Presenter: Kittisak Sajjapongse (University of Missouri)

Many-core devices, such as GPUs, are widely adopted as part of high-performance, distributed computing. In every cluster setting, efficient resource management is essential to maximize the cluster utilization and the delivered performance, while minimizing the failure rate. Currently software components for many-core devices, such as the CUDA, provide limited support to concurrency and expose low-level interfaces which do not scale well and are therefore not suitable to cluster and cloud environments. In our research, we aim to develop runtime technologies that allow managing tasks in large-scale heterogeneous clusters, so to maximize the cluster utilization and minimize the failures exposed to end users. As manufacturers are marketing a variety of many-core devices with different hardware characteristics and software APIs, we will propose a unified management component hiding the peculiarities of the underlying many-core devices to the end users. Our study will focus on algorithms and mechanisms for scheduling and fault recovery.

Simulating Forced Evaporative Cooling Utilising a Parallel Direct Simulation Monte Carlo Algorithm

Presenter: Christopher Jon Watkins (Monash University)

Atomic spins moving in position-dependent magnetic fields are at the heart of many ultracold atomic physics experiments. A mechanism known as a "Majorana spin flip," causes loss from these magnetic traps. Though often avoided by various means, they play a larger role in a new form of hybrid trap - comprising a magnetic quadrupole superposed on an optical dipole

potential. Previously, the spin-flip mechanism was modelled with a finite sized “hole” from which atoms are expelled from the trap. Instead, we numerically model the coupled spin and motional dynamics of an ensemble of atoms in a magnetic quadrupole field. Directly tracking the spin dynamics provides insight into the spin-flip mechanism and its effect on the velocity distribution of atoms remaining in the trap. The high computational demand of this simulation has prompted the parallelisation on Graphics Processing Units using NVIDIA’s Compute Unified Device Architecture.

Autonomic Modeling of Data-Driven Application Behavior

Presenter: Steena D.S. Monteiro (Lawrence Livermore National Laboratory)

Computational behavior of large-scale data driven applications are complex functions of their input, various configuration settings, and underlying system architecture. The resulting difficulty in predicting their behavior complicates optimization of their performance and scheduling them onto compute resources. Manually diagnosing performance problems and reconfiguring resource settings to improve performance is cumbersome and inefficient. We thus need autonomic optimization techniques that observe the application, learn from the observations, and subsequently successfully predict its behavior across different systems and load scenarios. This work presents a modular modeling approach for complex data-driven applications that uses statistical techniques to capture pertinent characteristics of input data, dynamic application behaviors and system properties to predict application behavior with minimum human intervention. The work demonstrates how to adaptively structure and configure the model based on the observed complexity of application behavior in different input and execution contexts.

Metrics, Workloads and Methodologies for Energy Efficient Parallel Computing

Presenter: Balaji Subramaniam (Virginia Tech)

Power and energy efficiency have emerged as first-order design constraints in parallel computing due to the tremendous demand for power and energy efficiency in high-performance, enterprise and mobile computing. Despite the interest surrounding energy efficiency, there is no clear consensus on the metric(s) and workload(s) (i.e., benchmark(s)) for comparing efficiency of systems. The metric and workload must be capable of providing a system-wide view of energy efficiency in light of the predictions that the energy consumed by the non-computational elements will dominate the power envelope of the system. There is also a commensurate need to maximize the usage of a system under a given power budget in order to amortize energy related expenses. Effective methodologies are required to maximize usage under a power budget.

This dissertation work revolves around the above mentioned three dimensions (i.e., metrics, workloads and methodologies) for enabling energy efficient parallel computing.

Adaptive, Resilient Cloud Platform for Dynamic, Mission-Critical Dataflow Applications

Presenter: Alok Gautam Kumbhare (University of Southern California)

With the explosion in both real time streams and volume of accumulated data, we see proliferation of applications that need to perform long running, continuous data processing, some with mission critical requirements. These applications exhibit dynamism in changing data rates and even their composition in response to domain triggers. Clouds offer a viable platform for running these applications but are inhibited by the inherent limitations of Clouds, especially, non-uniform performance and fallible commodity hardware. This “infrastructure dynamism” coupled with “application dynamism” present a number of challenges for deploying resilient mission critical applications. We propose a class of applications, termed Dynamic Mission-Critical Dataflow (DMD) applications and propose to make three major contributions: programming abstractions and declarative goals to describe and model DMD applications, time-variant performance models to predict the variability and failure of Cloud resources, and a dataflow execution platform that optimally orchestrates DMD applications on Clouds to meet their goals.

Using Computational Fluid Dynamics and High Performance Computing to Model a Micro-Helicopter Operating Near a Flat Vertical Wall

Presenter: David C. Robinson (Monash University)

This research aims to use Computational Fluid Dynamics (CFD) and High Performance Computing (HPC) to model the aerodynamic forces acting on a micro-helicopter that is operating near a flat vertical wall. Preliminary CFD simulation results are presented which show that operating a micro-helicopter near a flat vertical wall will create an imbalance in the lift generated by each rotor blade which will adversely affect the stability of the helicopter. This CFD simulation will be repeated many times to develop a parametric model that characterises how helicopter stability varies with distance from the wall and rotor attitude. This modelling will be computationally intensive due to the large number of simulations required. Computation time will be reduced significantly by running the simulations on the Monash Campus Grid and by using the Nimrod parametric modelling toolkit (developed at Monash University) to optimally manage the setup, scheduling and result collation for each individual simulation.

Paving the Road to Exascale with Many-Task Computing

Presenter: Ke Wang (Illinois Institute of Technology)

Exascale systems will bring significant challenges: concurrency, resilience, I/O and memory, heterogeneity, and energy, which are unlikely being addressed using current techniques. This work addresses the first four, through the Many-Task Computing (MTC) paradigm, by delivering data-aware resource management and fully asynchronous distributed architectures. MTC applications are structured as graphs of discrete tasks, with explicit dependencies forming the edges. Tasks may be uniprocessor or multiprocessor, the set of tasks and volume of data maybe extremely large. The asynchronous nature of MTC makes it more resilient than traditional HPC approaches as the system MTTF decreases. Future highly parallelized hardware is well suited for achieving high throughput with large-scale MTC applications. This work proposes a distributed MTC execution fabric for exascale, MATRIX, which adopts work stealing to achieve load balance. Work stealing was studied through Sim-Matrix, a scalable simulator, which supports exascale (millions of nodes, billions of cores, and trillions of tasks).

High Performance Computing in Simulating Carbon Dioxide Geologic Sequestration

Presenter: Eduardo Sanchez (San Diego State University)

Traditionally, in studying CCUS, numerical codes that simulate water-rock interaction and reactive transport sequentially solve an elemental mass balance equation for each control volume that represents a discretized lithology containing charged aqueous solute species. However, this formulation is not well suited for execution on many-core distributed clusters. Here, we present the theory and implementation of a numerical scheme whereby all solute concentrations in all control volumes are solved simultaneously by constructing a large block-banded sparse matrix of dimension N_a times N_x , where N_x is the number of control volumes and N_a is the number species for which their diffusion, advection, and reaction processes are of interest. These are very large matrices that fulfill the requirements needed in order to be factored with SuperLU_DIST from Lawrence Berkeley National Laboratory (LBNL). Performance metrics are considered on the ~10K core XSEDE cluster trestles.sdsc.edu which is located at the San Diego Supercomputer Center (SDSC).

Uncovering New Parallel Program Features with Parallel Block Vectors and Harmony

Presenter: Melanie Kambadur (Columbia University)

The efficient execution of well-parallelized applications is central to performance in the multicore era. Program analysis tools support the hardware and software sides of this effort by exposing relevant features of multithreaded applications. This

talk describes Parallel Block Vectors, which uncover previously unseen characteristics of parallel programs. Parallel Block Vectors, or PBVs, provide block execution profiles per concurrency phase (e.g., the block execution profile of all serial regions of a program). This information provides a direct and fine-grained mapping between an application's runtime parallel phases and the static code that makes up those phases. PBVs can be collected with minimal application perturbation using Harmony, an instrumentation pass for the LLVM compiler. We have already applied PBVs to uncover novel insights about parallel applications that are relevant to architectural design, and we also have ideas for using PBVs in fields such as software engineering, compilers, and operating systems.

New Insights into the Colonization of Australia Through the Analysis of the Mitochondrial Genome

Presenter: Nano Nagle (La Trobe University)

Colonization of Australia by the ancestors of contemporary Aboriginal peoples and their subsequent isolation is a crucial event in understanding the human migration(s) from Africa ~70,000 - 50,000 years ago. However, our knowledge of genetic structure in contemporary Australian Aboriginal peoples is extremely limited and our understanding of the route taken by the migrants and their genetic composition is poorly understood. As part of the Genographic Project, we have investigated mitochondrial (mt) DNA sequence variation in ~ 300 unrelated samples of Aboriginal Australians from previously unsampled geographic locations. This is by far the largest and most geographically widespread DNA sample of indigenous Australians. Using Next Generation Sequencing to acquire whole mitochondrial genomes we are gaining new insights into the colonization of Australia. The next step is to utilise high performance computing to infer times of arrival of the unique haplogroups in Australia, ages of the haplogroups and possible coalescence.

A Meshfree Particle Based Model for Microscale Shrinkage Mechanisms of Food Materials in Drying Conditions

Presenter: Chaminda Prasad Karunasena Helambage (Queensland University of Technology)

Cells are the fundamental building block of plant based food materials and many of the food processing born structural changes can fundamentally be derived as a function of the deformations of the cellular structure. In food dehydration the bulk level changes in porosity, density and shrinkage can be better explained using cellular level deformations initiated by the moisture removal from the cellular fluid. A novel approach is used in this research to model the cell fluid with Smoothed Particle Hydrodynamics (SPH) and cell walls with Discrete Element Methods (DEM), that are fundamentally known to be robust in treating complex fluid and solid mechanics. High

Performance Computing (HPC) is used for the computations due to its computing advantages. Comparing with the deficiencies of the state of the art drying models, the current model is found to be robust in replicating drying mechanics of plant based food materials in microscale.

Reproducibility and Scalability in Experimentation through Cloud Computing Technologies

Presenter: Jonathan Klinginsmith (Indiana University)

In application and algorithm research areas, reproducing controlled experimental conditions is a challenging task for researchers. Reproducing an experiment requires the investigator to recreate the original conditions including operating system, software installations and configurations, and the original data set. The challenge is compounded when the experiment must be run in a large scale environment. We present our early research on creating a reproducible framework for the construction of large-scale computing environments on top of cloud computing infrastructures. The goal of our research is to reduce the challenges of experimental reproducibility for large-scale high performance computing areas.

Programming and Runtime Support for Enabling In-Situ/In-Transit Scientific Data Processing

Presenter: Fan Zhang (Rutgers University)

Scientific simulations running at extreme scale on leadership class systems are generating unprecedented amount of data, which has to be analyzed and understood to enable scientific discovery. However, the increasing gap between computation and disk IO speeds makes traditional data analytics pipelines based on post-processing cost prohibitive and often infeasible. This research explores hybrid in-situ/in-transit data staging and online execution of scientific data processing as coupled workflow. Specifically, we investigate the programming support for composing the coupled workflow from heterogeneous computation components, and the runtime framework for distributed data sharing and task execution. The framework employs data-centric task placement to map workflow computations onto processor cores to reduce network data movement and increase intra-node data reuse. In addition, the framework implements the shared space programming abstraction with specialized one-sided asynchronous data access operators and can be used to express coordination and data exchanges between the coupled components.

Ensemble-Based Virtual Screening to Expand the Chemical Diversity of LSD1 Inhibitors

Presenter: James C. Robertson (University of Utah)

The vast number of experiments that are needed to test new inhibitors that target relevant biomolecules generally hampers drug discovery and design. Novel computational screening approaches offer extraordinary help using docking algorithms that evaluate favorable binding poses between libraries of small-molecules and a biological receptor. However, including receptor flexibility is still challenging and new biophysical approaches are being developed in our group to improve the predictive power of computational screening. In this work, we present new methods to target lysine specific demethylase (LSD1), which associates with the co-repressor protein (CoREST), and plays an epigenetic-based role in a number of solid-tumor cancers. Molecular dynamics simulations of LSD1/CoREST offer new routes for including receptor flexibility in virtual screening performed on representative ensembles of LSD1/CoREST conformations at reduced costs, including conformational selection and induced-fit effects. Promising inhibitor candidates selected from virtual screening results are being experimentally tested (Mattevi lab; Pavia).

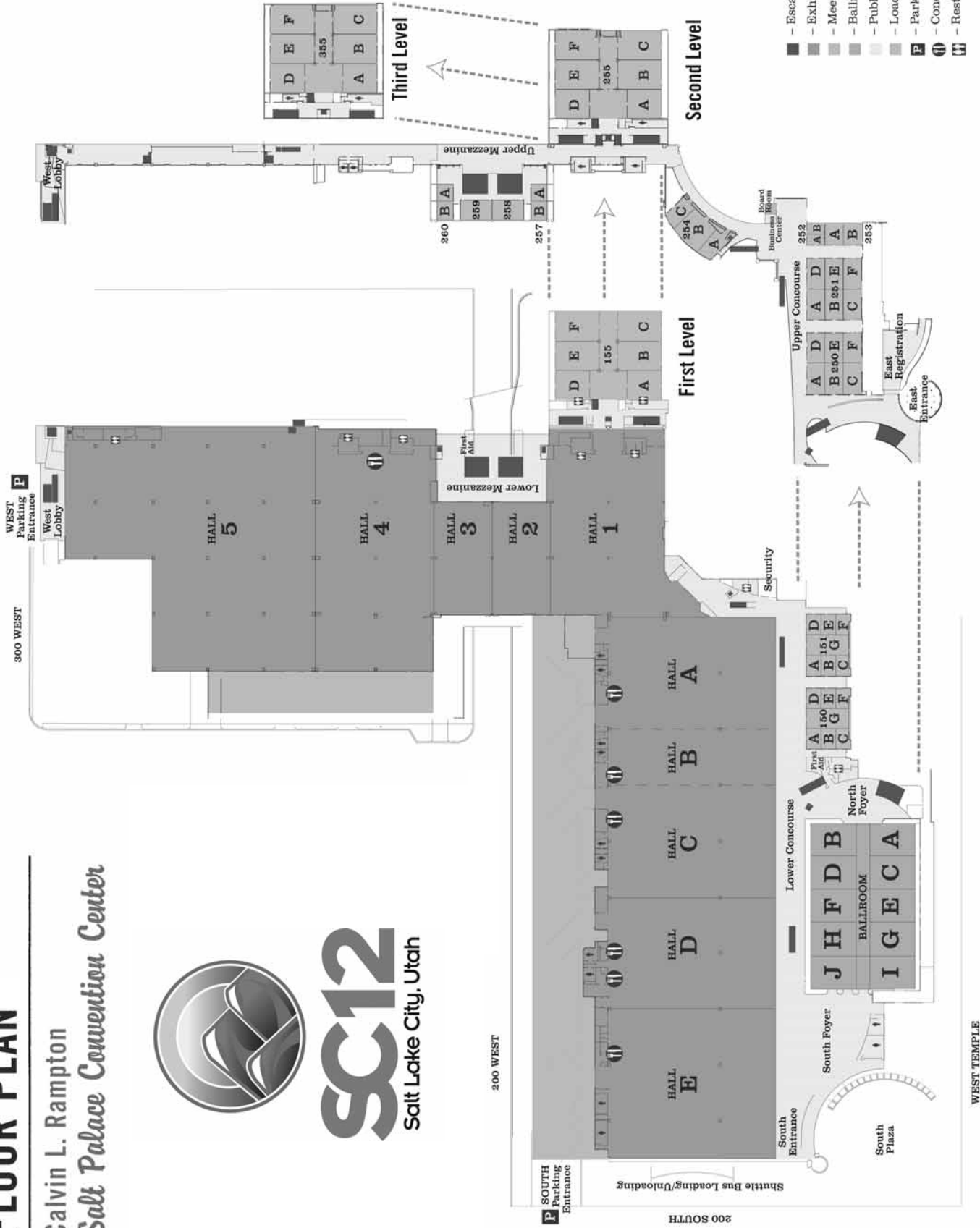
This image shows a full page of white paper with horizontal green ruling lines. The lines are evenly spaced and run across the width of the page. In the top-left corner, the word "Notes" is written in a bold, black, sans-serif font.

FLOOR PLAN

Calvin L. Rampton
Salt Palace Convention Center



SC12
Salt Lake City, Utah



WEST TEMPLE



Call for Participation

*The International Conference for High Performance Computing,
Networking, Storage and Analysis*

Why participate in SC13?

Not only is the exhibit space larger, the technical program innovative, and the city "Mile High," you'll have the opportunity to participate in new technologies, see next year's breakthroughs, and be "dared to see Denver in a whole new way."

SC's continuing goal is to provide an informative, high-quality technical program that meets the highest academic standards, with every aspect of the SC13 technical program being rigorously peer reviewed. Moreover, we are streamlining the opportunities for participants to immerse

themselves in the complex world of HPC, such as earlier acceptance deadlines for workshop program improvement and new programs to highlight innovative technologies.

It's not the Denver you remember – this great city has countless restaurants, hotels, and more, all within walking distance of the convention center, making it all the more convenient to collaborate and connect.

Visit sc13.supercomputing.org for more information on programs and deadlines.

**With HPC everywhere,
why not participate?**

Sponsors:



*Follow us on
Facebook
and Twitter*





SC12

Salt Lake City, Utah

Sponsors:



Association for
Computing Machinery

