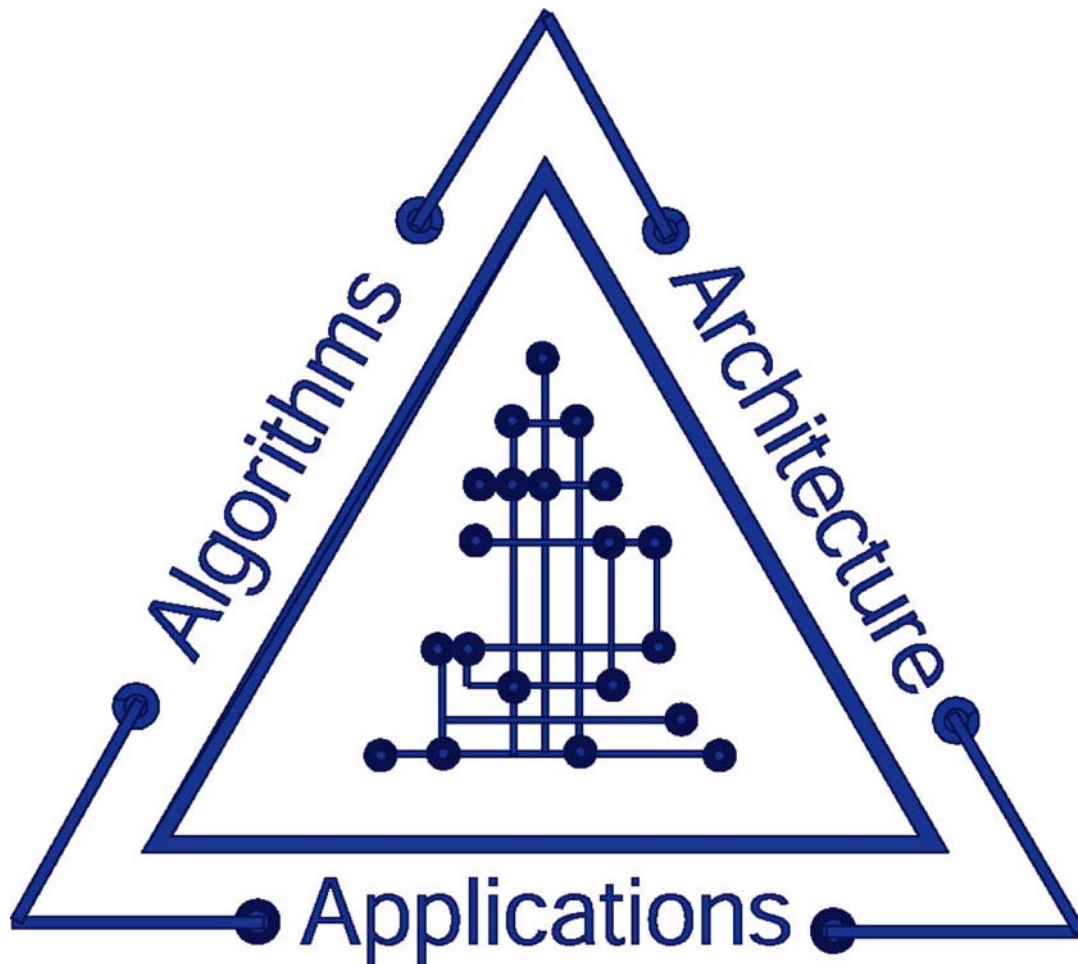


The Salishan Conference on
HIGH-SPEED COMPUTING

LANL / LLNL / SNL



April 21 – 24, 2008

*Salishan Lodge
Gleneden Beach, Oregon*

The Salishan Conference on High-Speed Computing *at a glance*

	Monday	Tuesday	Wednesday	Thursday
8:00 AM		Registration Opens Breakfast	Introduction to Sessions Breakfast	Introduction to Sessions Breakfast
8:30 AM		Session 1: <i>Chair – Jim Ang</i> <i>Exascale - The Next Great Challenge</i> ----- <i>Paving the Road from Petascale to Exascale with Many-Core Processors and Fast Interconnect Fabrics</i>	Session 3: <i>Chair – Alice Koniges</i> <i>SciDAC and the Path Toward Exascale</i> ----- <i>Kinetic Plasma Modeling with VPIC: Status and Future Plans on Hybrid Architectures</i>	Session 4: <i>Chair – Richard Murphy</i> <i>Programming Models and Languages for High Performance Computing</i> ----- <i>Toward an Open and Unified Model for Heterogeneous and Accelerated Multicore Computing</i>
9:50 AM		Break	Break	Break
10:10 AM		<i>The Role of Accelerated Computing in the Multi-Core Era</i> ----- <i>Why CPU's Have to Evolve: From Homogeneous to Heterogeneous Chips, A Brief Overview</i>	<i>Coping with Petascale Architectures</i> ----- <i>Auto-tuned Optimization of Scientific Kernels on Leading Multicore Systems</i>	<i>Transactional Memory and Threads – Sun</i> ----- <i>Software Invasion from Outer Space</i>
11:30 AM		Panel Discussion	Panel Discussion	Panel Discussion
NOON		Lunch: Council House	Lunch on Your Own	Lunch: Council House
1:30 PM		Session 2: <i>Chair – Adolfo Hoisie</i> <i>Systems Software Challenges and Strategies for the Petascale/Exascale Era</i> ----- <i>The Role of Compilers and Programming Languages for Client-Side Multicore Systems</i>	No Scheduled Session	Session 5: <i>Chair – Manuel Vigil</i> <i>The Institute for Advanced Architectures and Algorithms</i> ----- <i>Sequoia Architectural Requirements</i>
2:50 PM		Break		Break
3:10 PM	Registration 3:30-7:00 PM (Salal Room)	<i>Quad-core Catamount and R&D in Multi-core LWKs</i> ----- <i>Petascale Communication is not Business as Usual</i>		<i>The Cray Roadmap to Cascade</i> ----- <i>Moore, More Cores, and More Application Performance</i>
4:30 PM		Panel Discussion		Panel Discussion
6:00 PM	Welcome/Keynote Address (Long House) <i>Multicore Meets Exascale: The Catalyst for a Software Revolution</i>	Working Dinner/Speaker (Council House) <i>Multicore: Hey, Wait a Minute?</i>	6:30 PM <i>Random Access (Sign up to speak for 10 minutes)</i> (Long House)	Informal Discussions Council House
8:00 PM	Informal Discussions Council House	Informal Discussions Cedar Tree Room	Informal Discussions Council House	

THIS PAGE LEFT BLANK
INTENTIONALLY

Welcome

Welcome to the Salishan Conference on High-Speed Computing. This conference was founded in 1981 as a means of getting experts in computer architecture, languages, and algorithms together to improve communication, develop collaborations, solve problems of mutual interest, and provide effective leadership in the field of high-speed computing. Attendance at the conference is by invitation; we limit attendance to about 150 of the world's brightest people. Attendees are from national laboratories, academia, government, and private industry. We keep the conference small to preserve the level of interaction and discussion among the attendees.

The conference agenda and selection of participants has been designed to focus discussion on technical issues of relevance to our conference theme, "HPC in the Era of Ubiquitous parallelism: Multicore and Hybrid Architectures." The talks have been selected to give attendees information about the latest technologies and issues facing high-speed computing. The evening sessions are structured to encourage informal discussions and networking among all of the participants.

If you have any comments or suggestions for future topics and/or speakers, we encourage you to speak to any of the conference committee members.

We hope you find this conference stimulating, challenging, and also relaxing – enjoy!

Conference Committee

Jim Ang & Richard Murphy, *SNL* Manuel Vigil & Adolfo Hoisie, *LANL* Alice Koniges, *LLNL*

Logistics

Conference sessions and the Random Access session will be held in the Long House. Lunches and the working dinner will be held in the Council House.

For administrative support, please speak to any of the individuals located in the registration area (Salal room). If you have specific questions regarding audiovisual equipment or network connectivity, please seek out Tom Pratt or Bob Brothers.

Next conference dates:

April 27-30, 2009
April 26-29, 2010

Sponsorship

The Salishan Conference on High-Speed Computing is organized and hosted by Lawrence Livermore, Los Alamos, and Sandia National Laboratories. Additional sponsorship for the evening portions of our program is provided by the corporations listed here.

One of the highlights of the conference is the informal discussions held each evening. These sessions help us to go beyond the formal presentations to exchange ideas, solve problems, and develop friendships.

This year the following companies are helping to sponsor the evening sessions:

Advanced Micro Devices, Inc.

Cray, Inc.

Hewlett-Packard Company

IBM Corporation

Intel Corporation

Microsoft

NVIDIA Corporation

The Portland Group, Inc.

Silicon Graphics, Inc.

Sun Microsystems Inc.

We would like to express our thanks to these companies for their generous support.

Table of Contents

The Salishan Conference on High-Speed Computing

at a Glance	Inside Cover
Welcome and Logistics	1
Sponsorship	2
Conference Theme	5
Conference Program	
Monday Keynote	7
Tuesday Session 1: Processor Architecture Roadmap	8
Tuesday Session 2: System Software	9
Wednesday Session 3: Applications.....	11
Thursday Session 4: Programming Models/Environment	13
Thursday Session 5: System Architectures.....	14
Abstracts	15
Attendees	31
Conference Notes	46

THIS PAGE LEFT BLANK
INTENTIONALLY

Conference Theme

HPC in the Era of Ubiquitous Parallelism: Multicore and Hybrid Architectures

A new era in computer architecture has begun with the advent of multicore processor designs and hybrid architectures. For the last couple of decades, Moore's Law tracked the advances in microprocessor architecture, triggered by the exponential increase in the number of transistors on a chip, and by the constant increases in clock frequency. However, heat dissipation severely limits significant new gains from clock rates. Many cores on silicon emerged as the new architectural solution allowing us to maintain a Moore's Law pace of progress. It is now widely believed that we are embarking on a new trend that will double the number of cores with every silicon generation. In addition, many system tasks such as graphics or network interfaces that were previously accomplished outside of the processing elements, are frequently embedded on the multicore chips, leading to hybrid designs.

The emergence of increased on-chip parallelism poses significant opportunities and challenges. As we learned from the many decades of parallel computing in the scientific arena, which this conference is very much linked with and has chronicled with accuracy, parallelism is not "gain with no pain". Multicore and hybrid designs have the potential to modify the ways in which we think of parallelism, the way in which we program, develop system and application software, and integrate systems. A new dynamics will be created, from the interaction of the deep understanding of parallel computing in our community and its specific needs with the emergence of the grassroots, widespread availability of parallelism that the new architectural trend will enable. We plan to address this new landscape of architectures, software, and the infusion of new ideas and solutions in presentations and discussions at our conference.

Multicore and hybrid designs are destined to dominate the architecture landscape for some years to come, and it is essential that in high-performance computing we consider its effects on our future. In particular, we will address the following questions in five half-day sessions:

- What are the implications of multicore architectures on the ways in which we think of parallelism? Is parallelism at multiple scales a prerequisite for achieving efficiency on the new architectures?
- What are the implications of multicore and hybrid designs on the system software? Do we need to drastically re-think operating systems? What about compilers, runtime libraries, communication libraries? How much of that re-thinking is due to the sheer increase in parallelism, and how much of it to higher complexity of multicore and hybrid designs?

-
-
- What are the best ways to integrate new systems in the petascale regime from these new kinds of building blocks? How much of the resources enabled by the availability of many cores should be dedicated to system tasks, such as interfacing with the network or running the operating system?
 - What are the implications on the programming environment? Do we need new languages? Are communication libraries in their current design able to cope with the new realities? What are the tradeoffs needed, and what are the steps towards a programming environment that allows us to harness the complexity and the scale we are dealing with?
 - What are the impacts on application software design? Is it business as usual as far as applications go? Are current best practices on software engineering still applicable within the new architectural paradigm? What is the staying power of the new trend under consideration, and with that in mind, should we pay now the cost of re-factoring large applications? How do we leverage the investments in software for future hybrid platforms as well as other multicore systems?

These questions are addressed in five half-day sessions that are organized into the following areas:

1. Processor Architecture Roadmap
2. System Software
3. Applications
4. Programming Models/Environment
5. System Architectures

Conference Program

HPC in the Era of Ubiquitous Parallelism: Multicore and Hybrid Architectures

Monday, April 21, 2008

3:30 -7:00 PM Registration (Salal Room)

6:00 PM Welcome/Keynote Address

Title: Multicore Meets Exascale: The Catalyst for a Software Revolution

Speaker: Kathy Yelick, *Lawrence Berkeley National Laboratory*

8:00 PM Informal Discussions (Council House)

Tuesday, April 22, 2008

8:00 AM Registration Opens (Salal Room)

Breakfast available (Terrace)

8:30 AM Session 1: Processor Architecture Roadmap

Title: Exascale – The Next Great Challenge

Speaker: Peter Kogge, *University of Norte Dame* &
William Harrod, *DARPA*

**Title: Paving the Road from Petascale to Exascale
with Many-Core Processors and Fast
Interconnect Fabrics**

Speaker: William J. Camp, *Intel Corporation*

9:50 AM Break

Refreshments available (Terrace)

10:10 AM Session 1: Processor Architecture Roadmap

**Title: The Role of Accelerated Computing in the
Multi-Core Era**

Speaker: Charles Moore, *AMD*

**Title: Why CPU's Have to Evolve: From
Homogeneous to Heterogeneous Chips, A Brief
Overview**

Speaker: Michael Paolini, *IBM Corporation*

11:30 AM Panel Discussion

Tuesday, April 22, 2008 (cont.)

- Noon** **Lunch (Council House)**
- 1:30 PM** **Session 2: System Software**
- Title:** **Systems Software Challenges and Strategies for the Petascale/Exascale Era**
- Speaker: Fred Johnson, *DOE Office of Advanced Scientific Computing Research*
- Title:** **The Role of Compilers and Programming Languages for Client-Side Multicore Systems**
- Speaker: Vikram Adve, *University of Illinois, Urbana-Champaign*
- 2:50 PM** **Break**
- Refreshments available (Terrace)
- 3:10 PM** **Session 2: System Software**
- Title:** **Quad-core Catamount and R&D in Multi-core Lightweight Kernels**
- Speaker: Kevin Pedretti, *Sandia National Laboratories*
- Title:** **Petascale Communication is not Business as Usual**
- Speaker: Al Geist, *Oak Ridge National Laboratory*
- 4:30 PM** **Panel Discussion**

Tuesday, April 22, 2008 (cont.)

6:00 PM **Working Dinner/Speaker (Council House)**

Title: Multicore: Hey, Wait a Minute?

Speaker: Dan Reed, *Microsoft*

8:00 PM **Informal Discussions (Cedar Tree Room)**

Student Poster Session

Wednesday, April 23, 2008

8:00 AM **Introduction to Sessions**

Breakfast available (Terrace)

8:30 AM **Session 3: Applications**

Title: **SciDAC and the Path Toward Exascale**

Speaker: Walter Polansky, *Office of Advanced Scientific Computing Research, Office of Science*

Title: **Kinetic Plasma Modeling with VPIC: Status and Future Plans on Hybrid Architectures**

Speaker: Brian Albright, *Los Alamos National Laboratory*

9:50 AM **Break**

Refreshments available (Terrace)

10:10 AM **Session 3: Applications**

Title: **Coping with Petascale Architectures**

Speaker: Bronis R. de Supinski, *Lawrence Livermore National Laboratory*

Title: **Auto-tuned Optimization of Scientific Kernels on Leading Multicore Systems**

Speaker: Leonid Oliker, *Lawrence Berkeley National Laboratory*

11:30 AM **Panel Discussion**

Wednesday, April 23, 2008 (cont.)

Noon

Lunch on Your Own

1:30 PM

No Scheduled Session

6:30 PM

Random Access (Long House)

The Random Access session consists of timely communications from participants on areas of interest to the Conference. Presentations are strictly limited to 10 minutes. A sign-up board is provided in the registration area.

8:00 PM

Informal Discussions (Council House)

Thursday, April 24, 2008

- 8:00 AM** **Introduction to Sessions**
Breakfast available (Terrace)
- 8:30 AM** **Session 4: Programming Models/Environment**
- Title:** **Programming Models and Languages for High Performance Computing**
Speaker: Marc Snir, *University of Illinois, Urbana-Champaign*
- Title:** **Toward an Open and Unified Model for Heterogeneous and Accelerated Multicore Computing**
Speaker: Catherine Crawford, *IBM Corporation*
- 9:50 AM** **Break**
Refreshments available (Terrace)
- 10:10 AM** **Session 4: Programming Models/Environment**
- Title:** **Transactional Memory for a Modern Microprocessor**
Speaker: Marc Tremblay, *Sun Microsystems, Inc.*
- Title:** **Software Invasion from Outer Space**
Speaker: David Callahan, *Microsoft*
- 11:30 AM** **Panel Discussion**

Thursday, April 24, 2008 (cont.)

- Noon** **Lunch (Council House)**
- 1:30 PM** **Session 5: System Architectures**
- Title:** **The Institute for Advanced Architectures and Algorithms**
Speakers: Sudip Dosanjh, *Sandia National Laboratories*
 Jeff Nichols, *Oak Ridge National Laboratory*
- Title:** **Sequoia Architectural Requirements**
Speaker: Matt Leininger, *Lawrence Livermore National Laboratory*
- 2:50 PM** **Break**
Refreshments available (Terrace)
- 3:10 PM** **Session 5: System Architectures**
- Title:** **The Cray Roadmap to Cascade**
Speaker: John Levesque, *Cray, Inc.*
- Title:** **Moore, More Cores, and More Application Performance**
Speaker: Darren Kerbyson, *Los Alamos National Laboratory*
- 4:30 PM** **Panel Discussion**
- 6:00 PM** **Wrap-Up and Informal Discussions (Council House)**

Abstracts

Keynote Address

Multicore Meets Exascale: The Catalyst for a Software Revolution

Kathy Yelick, Lawrence Berkeley National Laboratory

Petascale systems will soon be available to the computational science community at multiple sites. These systems will represent a variety of architectural models, but with one common component, which is an increasing reliance on multicore technology as the building block for these machines. At the same time, the entire field of computing is shifting towards some form of multicore technology, either chip multiprocessors or heterogeneous processors that rely on data parallelism. The “View from Berkeley” paper lays out some of the research challenges for the general computing community, but many of these problems are also evident in high end computing. In this talk I will look at implications of the hardware trends on the kinds of algorithms, programming models, and applications that we can expect to scale across future machine generations. I will describe some programming approaches targeted at different programming communities, from performance and parallelism specialists to application developers and domain specialists. This will include shared address space models for efficiency, and domain-specific languages that hide parallelism for the productivity. These techniques must simultaneously address the problems of correctness, performance and ease of use.

Session 1: Processor Architecture Roadmap

Exascale – The Next Great Challenge

Peter Kogge, *University of Norte Dame*
William Harrod, *DARPA*

With petascale machines nearing production, the next great barrier for computing is exascale – a thousand times more computational capability. Given that it will have taken over 14 years from the first petaflops workshop in 1994 to real hardware, an obvious set of questions to ask is whether or not there is another three orders of magnitude left in silicon, whether or not architectures can utilize such technologies in an efficient manner, and what are the challenges if we were to try to halve the time from peta to exa over the prior tera to peta. This talk will investigate what headroom is left in silicon, and extrapolate several different architectures to exascale, including a “clean sheet.” From these extrapolations will arise several major challenges that must be addressed in a coordinated fashion over the next few years.

Paving the Road from Petascale to Exascale with Many-Core processors and Fast Interconnect Fabrics

William J. Camp, *Intel Corporation*

Any Exascale computer will involve many millions of processing elements and hundreds of millions of processing threads. This seems inevitable given that we are reaching a frequency asymptote for CMOS devices. Many-core processors without sufficiently fast memory hierarchies will not achieve acceptable single-socket efficiencies. In addition, efficient many-core processors without sufficiently fast interconnect fabrics and I/O systems will not achieve acceptable parallel efficiencies. Finally fast hardware without fast and programmable software will not achieve acceptable delivered applications performance. Determining sufficiency is a task that will vary depending in part on: the application characteristics, the size of the system, the size of the application on that system, and the degree of clumpiness of the computational/communication fabric. We will look at how the foreseeable advances in underlying technologies and architectures could take us down the road to Exascale. We will also discuss the interplay of market forces with the HPC community plans to reach Exascale applications performance in the middle of the next decade.

The Role of Accelerated Computing in the Multi-Core Era

Charles Moore, *AMD*

The computer industry is driven by a virtuous cycle of adding value to entice new purchases, which then fuel the technology development process that ultimately offers new value. In recent years, we have seen a decline in the rate of improvement on several traditional drivers of value in computer systems, namely transistor performance, wire delays, the return on deep pipelining, and techniques for extracting high numbers of instructions per cycle. As new techniques for adding value are explored, there are some important questions about the hardware/software contract, complexity management, and overall system-level maturity that come into play. In this talk, I will highlight the implications of some of these shifts and make some observations about the emergence of a new framework for future innovation.

Why CPU's have to evolve: From homogeneous to heterogeneous chips, a brief overview

Michael Paolini, *IBM*

Today's CPU's have to live within the confines of power and thermal envelopes while approaching the fundamental limits of our technology and physics while simultaneously delivering enough increase in compute performance to meet the demands of an increasingly analytic world. This raises the question. "Is an array of massive homogeneous 'Jack of All Trades' cores better than using the transistor area to mix and match specialized cores for different tasks and gaining greater compute speed-ups while simultaneously lowering power consumption?" Will CPU's follow the biological model of and evolve from collections of single cell entities to multicell entities, where some cells are specialized?

Session 2: System Software

Systems Software Challenges and Strategies for the Petascale/Exascale Era

Fred Johnson, *DOE, Office of Advanced Scientific Computing Research*

Leadership class computing is having a profound impact on the state of computational science in the Office of Science. Contemporary applications face challenges of scaling to tens or hundreds of thousands of cores, and efforts have begun to understand the opportunities and requirements of next generation etascale codes. At the system software level we face challenges both of new applications and of architectures that are rapidly evolving in both size and complexity, and there is wide recognition that something beyond "business as usual" is necessary to enable applications to harness the potential of next generation systems. This talk will give a snapshot of our current thoughts and plans and encourage a dialog on an evolving systems software research agenda for the petascale/exascale era.

The Role of Compilers and Programming Languages for Client-Side Multicore Systems

Vikram Adve, *University of Illinois, Urbana-Champaign*

An important strategy for simplifying parallel programming is to make it (nearly) like sequential programming: eliminate non-determinism and expose a guaranteed sequential semantics in which the application programmer need not be concerned with complexities like atomicity, data races, deadlock, or strong or weak memory models. At Illinois, we are developing a programming strategy that provides such guarantees, building on a combination of language and compiler technologies. The language guarantees determinism not only in cases like pure data-parallelism but also for modern object-oriented (O-O) programming styles with inheritance, aliasing, and concurrent updates to shared data. With a careful language design, the compiler can identify the sources of parallelism and guarantee that the program is deterministic using only simple, local reasoning and no complex interprocedural analysis (even in the presence of such complex O-O constructs). Nevertheless, sophisticated compiler technology can play two important roles in this context. First, it can be valuable in optimizing parallel program performance in the "back end" by enhancing locality and guiding run-time partitioning and load-balancing. Second, sophisticated concurrency discovery algorithms can be incorporated into interactive porting tools to assist programmers in porting existing sequential or parallel programs to the new language. Although such algorithms are inherently fragile (small changes in the code can affect whether they discover parallelism or not), this is not

a problem in an interactive setting: the programmer can get immediate feedback and rewrite the code or add more information to help the compiler discover the parallelism. In this talk, we will focus on the language design and briefly discuss the role of compiler technology for supporting deterministic parallel programming.

Quad-core Catamount and R&D in Multi-core Lightweight Kernels

Kevin Pedretti, *Sandia National Laboratories*

ASC capability supercomputers are massively complex, both in software and hardware. General-purpose operating systems have grown so complicated that they significantly impede the innovation that will be necessary to take full advantage of future multi-core architectures, which are likely to incorporate heterogeneous and hierarchical computing elements. This talk focuses on the compute node operating system and the work Sandia is doing to keep it simple, efficient, and functional. The case will be made that general-purpose operating systems, even slimmed down ones, add unnecessary complexity to the system and are detrimental to performance.

Two of our parallel efforts will be presented. The first will be an overview of the development project to add support for quad-core processors to the Catamount lightweight kernel (LWK) operating system that runs on Cray XT systems. Catamount is the latest in a series of specialized HPC operating systems that are descendant from SUNMOS, a LWK developed by Sandia and the University of New Mexico in 1990 for the 1024 processor nCube-2 system. Quad-core Catamount results from application testing on a Cray XT4 system will be presented.

The second portion of the talk will discuss our effort to create a new open source LWK that addresses short-comings of previous implementations and is well-suited for use in multi-core systems. This LWK is heavily based on Linux, but rewinds it to a much earlier design point. Unnecessary complexity such as demand paging has been replaced by simpler mechanisms. Enough of the Linux Application Binary Interface (ABI) is implemented to support HPC applications that are built with standard toolchains. Additionally, work is underway to support more full-featured guest operating systems through a simple hypervisor.

Petascale Communication is not Business as Usual

Al Geist , Oak Ridge National Laboratory

Multicore and hybrid architecture designs dominate the landscape for systems that are 1 to 20 petaflops peak performance. As such the systems software must be adapted to effectively use these types of architectures. This talk will address some of the new developments and research directions in the area of communication libraries. While applications may continue to use MPI, it is not business as usual in how communication libraries are being changed to effectively exploit the new petascale systems.

The talk will cover a number of areas being explored, including hierarchical algorithm designs, hybrid algorithm designs, and hardware support in memory management and NIC chips to improve communication performance. Hierarchical algorithm designs seek to consolidate information at different levels of the architecture to reduce the number of messages and contention on the interconnect. Natural places for such consolidation include the socket level, the node level, the cabinet level, and multiple-cabinet level.

Hybrid algorithm designs use different algorithms at different levels of the architecture, for example, an ALL_GATHER may use a shared memory algorithm across the node and a message passing algorithm between nodes, in order to better exploit the different data movement capabilities. A more complex type of communication library is to use adaptive algorithms. An adaptive communication library may dynamically select from a set of collective communication algorithms based on the number of nodes being sent to, where they are located in the system, the size of the message being sent, and the physical topology of the computer.

This talk will also describe things that ORNL's Leadership computing facility (LCF) has put in place so that science teams can better exploit the communication and IO capabilities of the Cray XT4 systems there. This includes assigning computational science liaisons to each science team. The liaison has knowledge of both the systems and the science, providing a bridge to improved communication patterns. The LCF also has a Cray Center of Excellence and a SUN Lustre Center of Excellence on site. These centers provide Cray and SUN engineers who work directly with the science teams to improve the performance of their applications. Finally this talk will look at the possibilities of future architectures incorporating advanced communication features such as atomic memory ops and collective communication into hardware.

Dinner Speaker

Multicore: Hey, Wait a Minute!

Dan Reed, *Microsoft*

Let's step back from our current analysis of GPUs and multicore processors and their deployment and think about the longer term future. Where is the technology going and what are the HPC implications? What did we do right or wrong to get here and what can we do about it? What architectures are appropriate for 100-way or 1000-way multicore designs? Is multicore itself a community failure of architectural vision or an inevitable and logical outcome? How do we develop and support software? This dinner talk will muse on some of the technical, economic and political forces that are pushing us down the multicore path and what we might or might not do about it.

Session 3: Applications

SciDAC and the Path Toward Exascale

Walter M. Polansky,
Office of Advanced Scientific Computing Research
Office of Science

Beyond the scientific computing research embedded throughout the Office of Science (SC) core research programs is Scientific Discovery through Advanced Computing (SciDAC); a portfolio of coordinated research efforts directed at exploiting the capabilities of terascale and emerging petascale computing resources. SciDAC research projects involve teams of physical scientists, mathematicians, computer scientists, and computational scientists working on major software and algorithm development for solving problems in high-energy physics, nuclear physics, climate, groundwater, fusion, life sciences, materials, chemistry and accelerator design. The SciDAC program was inaugurated in 2001 and recompleted in 2006. SciDAC is producing significant results across its entire domain-applied mathematics, computer science, software tools and computational science and is emerging as a model for future endeavors. However, that model, which will be described in this presentation, is about to be tested.

Fueled by continuing, rapid advances in technology, the mere possibility of enabling scientific advances through computing at the exascale has transitioned from a dream to a challenge in less than a year. Thoughtful formulations of the scientific challenges to be addressed at the exascale will determine success. Further, advances in basic research, coupled with lessons learned from existing simulation programs, including SciDAC, will underpin the breadth and the depth of successful research collaborations, and partnerships at the exascale.

Kinetic Plasma Modeling with VPIC: Status and Future Plans on Hybrid Architectures

Brian J. Albright, *Los Alamos National Laboratory*

VPIC is a first-principles three-dimensional kinetic plasma modeling code that has been designed at the Los Alamos National Laboratory and modified recently to run efficiently on the Roadrunner heterogeneous multi-core supercomputer. Roadrunner, scheduled to arrive at LANL this year, will be the first supercomputer capable of sustaining a petaflop/second, that is, a million billion operations per second and will enable “Science at Scale” simulations at unprecedented size and fidelity.

In work this past year several design changes were made to VPIC to enable use of existing and future hybrid/multicore platforms. In this talk, the VPIC physics algorithm will be discussed, including the physics modeled and associated computational science assumptions that we can make based on the physics. (For example, the finite speed of light automatically guarantees a degree of data locality). VPIC has been designed to operate efficiently in memory-bandwidth-starved environments, which has natural advantages for its deployment on hybrid architectures. Modifications to VPIC to enable platform flexibility and use of future hybrid systems will be described, as well as plans for the future.

Finally, science applications of VPIC in the next year and beyond will be summarized, including science runs on Roadrunner. These include weapons science studies relevant to thermonuclear burn and boost, application to inertial confinement fusion experiments on the National Ignition Facility, and magnetic reconnection, a basic physics problem of importance to magnetic fusion and space and astrophysics. Many of these applications pose challenges, e.g., I/O requirements for diagnostics and checkpointing, of concern for future high performance computing systems.

Work performed under the auspices of the U.S. Dept. of Energy by the Los Alamos National Security LLC Los Alamos National Laboratory under contract No. DE-AC52-06NA2536 and was supported in part by the ASC Program, the Science Campaigns, and the Laboratory Directed Research and Development (LDRD) Program.

Coping with Petascale Architectures

Bronis R. de Supinski, *Lawrence Livermore National Laboratory*

Although sustained petaflop performance for real applications is still some years away, many architecture trends are emerging that will shape how we will achieve that goal. We expect these systems to have millions of processor cores spread across nodes composed of chips with multiple, possibly heterogeneous, cores with novel mechanisms to assist in achieving the on-chip parallelism required for good single node performance. Further, compared to terascale systems, petascale systems are likely to have much less off-chip and off-node bandwidth per core as well as significantly smaller main memories per core. These trends will necessitate significant changes in applications and the development environment that supports them. We will require new mechanisms to target applications to these architectures, to identify and to solve software defects that arise in those applications and to understand and to improve their performance. In this talk, I will detail the overall NNSA ASC development environment strategy for petascale systems and several novel directions that we are pursuing as part of that strategy.

Auto-tuned Optimization of Scientific Kernels on Leading Multicore Systems

Leonid Oliker, *Lawrence Berkeley National Laboratory*

The computing industry is moving rapidly away from exponential scaling of clock frequency toward chip multiprocessors in order to better manage trade-offs among performance, energy efficiency, and reliability. Understanding the most effective hardware design choices and code optimizations strategies to enable efficient utilization of these systems is one of the key open questions facing the computational community today. Our work presents an auto-tuning approach for optimizing application performance on emerging multicore architectures. The methodology extends the idea of search-based performance optimizations, popular in linear algebra and FFT libraries, to application-specific computational kernels. We apply this strategy to both a lattice Boltzmann application (LBMHD), as well as the sparse matrix-vector multiplication (SpMV) kernel. Historically, these kernels have made poor use of scalar microprocessors due to their complex data structures and memory access patterns. Our work explores performance via auto-tuning optimizations on a broad set of multicore architectures including the Intel Xeon (Clovertown), AMD Opteron (X2), Sun Victoria Falls (Maramba), and the IBM Cell Broadband Engine. Overall results show that this approach results in substantial performance improvements, while amortizing tuning efforts across the machines. Additionally, we present detailed analysis of each optimization, which reveal surprising hardware bottlenecks and software challenges for future multicore systems and applications.

Session 4: Programming Models/Environment

Programming Models and Languages for High Performance Computing

Marc Snir, *University of Illinois, Urbana-Champaign*

For more than two decades, high performance computing systems have been built by assembling hardware and software components developed for mass markets, and adding relatively few HPC-specific technologies to the mix. Economic realities are likely to ensure this stays so in the foreseeable future. Parallelism is becoming now pervasive in the mass client and game markets. As a result, parallelism will be an essential ingredient of the hardware and software bricks used in building future HPC systems. Up to now, the hardware and software support for parallelism outside HPC was mostly driven by the server market; in the future it will be driven by the needs of a client-oriented mass market. The forms of parallelism that are most useful for client applications are quite different from the forms of parallelism that evolved for server applications and, quite possibly, closer to the needs of the HPC community. This is likely to have a significant impact on the evolution of programming languages and tools in support of High Performance Computing.

Our talk will discuss the above thesis in more detail; we shall discuss plausible directions on the evolution of HPC programming models and languages and how those will be impacted by multi-core technology.

Toward an Open and Unified Model for Heterogeneous and Accelerated Multicore Computing

Catherine Crawford, *IBM Corporation*

In recent years, more and more systems are being proposed which combine judicious exploitation of multi-core and multi-process technology in conjunction with the implementation of libraries and computational kernels on accelerators which offer a more efficient use of silicon in terms of area and power consumption. In this talk, we will describe one software enablement approach to utilizing the compute power of the both a system on a chip version of an accelerated system, the Cell Broadband Engine processor, as well as a cluster composed of x86_64 and PowerXCell8i processors integrated within a single hybrid “compute node”, a.k.a. the Roadrunner architecture. We begin with a review of historical approaches to concurrent multicore computing which includes a summary of many tools within the IBM Software Development Kit for Multicore Acceleration. The review is used to provide motivation for our development of the Data Communication and

Synchronization (DaCS) Library and Accelerated Library Framework (ALF) which are designed to allow developers to create new applications and adapt existing applications to exploit hybrid computing platforms. We present examples of usage of both ALF and DaCS on the Cell Broadband engine processor as well as the integrated hybrid nodes to demonstrate both the ubiquity and the limitations that these frameworks have in their current form. Finally, the applicability of DaCS and ALF to other multicores, e.g. x86_64 based symmetric memory processors, and accelerator frameworks, e.g. GPGPUs, is discussed.

Transactional Memory for a Modern Microprocessor

Marc Tremblay, *Sun Microsystems Inc.*

Transactional Memory has emerged as a leading technique that enables applications to better take advantage of multi-threaded, multi-core microprocessors. Setting goals for the scope of an implementation of Transactional Memory is a key milestone that has a pervasive impact upon the overall architecture of a modern microprocessor (codenamed Rock). In this talk, a description of what we believe is the first hardware implementation of Transactional Memory will be given. The synergy between a modern pipeline capable of handling today's memory latency as well as supporting sophisticated multithreading, is the key enabler of our approach to Transactional Memory.

Software Invasion from Outer Space

David Callahan, *Microsoft*

When major qualitative shifts such as the emergence of the graphical user interface (GUI), the Internet, mobile devices, and software services transformed the computing industry, Microsoft has successfully adapted the company, products, and business models to enable the next generation of computing experiences. Each previous shift has made computing more personal, social, and mobile. The recent advances in microelectronic technology and the advent of multi-core and manycore processors are a signal that another large industry change is on the horizon. The computational power of manycore processors, new programming models and platform, and advanced research in usability promises to change the way people interact with computers. This talk describes Microsoft's Parallel Computing Initiative and near term evolution of Windows and Visual Studio to support task-oriented parallel programming in a general-purpose environment. These are the first steps to take advantage of the "manycore shift" by enabling a new generation of responsive and scalable applications.

Session 5: System Architectures

The Institute for Advanced Architectures and Algorithms

Sudip Dosanjh, *Sandia National Laboratories*

Jeff Nichols, *Oak Ridge National Laboratory*

In the next few years, tremendous increases in computing speeds will revolutionize the way supercomputers are used. Predictive computer simulations will play a critical role in assuring a safe and reliable 21st century nuclear stockpile, revolutionize scientific discovery, and significantly impact national competitiveness, homeland security and quality of life issues. This dramatic increase in computing power will be driven by a rapid escalation in the parallelism incorporated in microprocessors. The transition from massively parallel architectures to hierarchical systems (hundreds of processor cores per CPU chip) will be as profound and challenging as the change from vector architectures to massively parallel computers that occurred in the early 1990's. Quickly overcoming this hurdle will provide game changing opportunities in the national security, scientific, and commercial sectors. Without DOE leadership, the chasm between peak speed and sustained performance will grow exponentially, and the societal benefits of advances in component technologies will be delayed and greatly diminished. With DOE leadership of a collaborative effort between the Laboratories and key university and industrial partners, the architectural bottlenecks that limit supercomputer scalability and performance can be overcome. The nation needs an enduring, focused activity that enables supercomputing technology transitions to occur efficiently, assuring that the United States achieves the maximum benefit from technical advances in computing.

To meet these challenges Sandia and Oak Ridge are establishing an Institute for Advanced Architectures and Algorithms (IAA). IAA will be a physically distributed center with sites in Albuquerque, NM and Knoxville, TN. Initial IAA focus areas will include:

- Interconnection Network Technologies
- Memory Systems
- Processor Microarchitecture
- RAS/Resilience
- System Software
- Architecture/Algorithm Co-Design

Sequoia Architectural Requirements

Matt Leininger, *Lawrence Livermore National Laboratory*

With several petascale sized systems nearing deployment the R&D focus has shifted to exascale, yet significant challenges remain in fielding and utilizing these petascale platforms to deliver predictive scientific simulations for national benefit. For example, although the list of potential petascale applications is large, very few applications today can take advantage of order one to three million processor cores/threads. Other challenges include improving the basic scientific models, mathematical descriptions of those models (e.g. turbulence), numerical techniques for solving those mathematical descriptions (e.g. scalable iterative methods for solving large sparse linear systems), and the verification and validation of the resulting petascale multi-physics/engineering and multi-scale applications. Another example is the daunting challenge of IO subsystems. Today's IO subsystems are straining under the load of terascale platforms. Significant changes in IO subsystems will be necessary to achieve balanced petascale simulation environments. In this talk we propose workable strategies to deal with petascale system deployments for productive programmatic usage and discuss how these experiences will contribute to future lessons on the road to exascale.

The Cray Roadmap to Cascade

John Levesque, *Cray, Inc.*

Over the next several years Cray will roll out a series of massively parallel systems that will culminate in the DARPA HPCS Cascade system. From the current Cray XT4, the system will transition to a more heterogeneous system in the XT5, which includes multiple choices for nodes, from the XT4 to the X2 system. As the system evolves innovative cooling will allow for packaging to become denser and field upgradable to new node and interconnects as they become available. The Cascade system itself will be comprised of a Granite node, which will be a custom node and a Marble node which will be the then fastest node from the XT line of MPPs. The custom interconnect will support global shared memory across the different node types, making hybrid parallel programming easier with the use of PGAS languages.

In addition, to the hardware, a matured Cray Linux Environment, compilers, libraries, programming tools and debuggers will be delivered that allows users to effectively employ all types of nodes on a single application.

Moore, More Cores, and More Application Performance

Darren J. Kerbyson, *Los Alamos National Laboratory*

Multi-core, heterogeneity, as well as memory and network hierarchies are already here. As a famous 20th thinker once said: “The future will be like the present only more so” [1]. In this talk we will examine a number of issues that we have observed in current multi-core systems, from single nodes, up to some of the largest systems available. Current multi-core processors have their own strengths and weaknesses, which we analyze. System topologies can impact performance such as causing contention for particular application communication patterns. We illustrate this for meshes and Infiniband, and we propose solutions with other rich topologies such as optical circuit switching or multi-hop direct-connect networks. Achieving performance is the key – it can be impeded by the capabilities of a socket, configuration of a node, or system connectivity. As the depth of system hierarchies and complexity increase, the challenges of achieving high application performance will increase many-fold also. But with challenges come opportunities, and we use performance modeling to bring it all together.

[1] Groucho Marx