The Alliance for Computing at the Extreme Scale

James Ang[2], Doug Doerfler[2], Sudip Dosanjh[2], Scott Hemmert[2] Ken Koch[1], John Morrison[1] and Manuel Vigil[1]

[1] *Los Alamos National Laboratory, Los Alamos, NM*
[2] *Sandia National Laboratories, Albuquerque, NM*

ABSTRACT: *Los Alamos and Sandia National Laboratories have formed a new high performance computing center, the Alliance for Computing at the Extreme Scale (ACES). The two labs will jointly architect, develop, procure and operate capability systems for DOE's Advanced Simulation and Computing Program. This presentation will discuss a petascale production capability system, Cielo, that will be deployed in late 2010, and a new partnership with Cray on advanced interconnect technologies.*

KEYWORDS: Cielo, Petascale

1. Introduction

Los Alamos National Laboratory and Sandia National Laboratories[1] have collaborated to create a New Mexico center for high performance computing, the Alliance for Computing at the Extreme Scale (ACES). ACES is funded by the U.S. Department of Energy's Advanced Simulation and Computing (ASC) program and was formed to enable the solution of critical national security problems through the development and deployment of high performance computing technologies. Current ACES efforts include (1) developing and deploying a 2010 production petascale supercomputer, Cielo, (2) an advanced interconnect development project and (3) architecting a 2015 system, Trinity. Other industrial partnerships are being pursued both within the context of reaching exascale and providing production capability computing.

Cielo will replace the Purple supercomputer [1] at Lawrence Livermore National Laboratory. Many targeted national security problems are extremely large and will require most of the nodes on Cielo for a single simulation. Consequently, hardware and software scalability are critical concerns. Reliability is also a key because the mean time between interrupts for an application executing on Cielo decreases with the number of nodes it uses [4]. Another design consideration was effectively supporting existing ASC computer codes with little or no modification. That is, applications that ran on Purple must execute efficiently on Cielo. The overall goal is to provide an order of magnitude increase in capability over Purple. After a competitive procurement process, Cray has been awarded the contract. Cielo will be an instantiation of Cray's Baker supercomputer architecture.

ACES is also partnering with Cray on an advanced interconnect project. An important goal is to insure that future interconnects continue to effectively support ASC codes. Of particular interest is enabling efficient implementations of MPI. In the next decade these codes are expected to evolve. Reaching exascale will require applications to deal with billion-way parallelism and manage locality [2]. Whether a new unified programming model emerges or the programming paradigm becomes MPI plus a node level model remains an open research question. In this project, modelling and system simulation

---

[3] will be used to analyze the performance of ASC codes and microapplications on new architectures and characterize the impact of design alternatives.

Many DOE national security, energy and science applications require significant increases in supercomputer performance over today's systems. Consequently, there is extensive planning for reaching exascale in the 2018-2021 timeframe. It is unlikely that this thousand-fold increase in performance over today's systems can be achieved in a single step. Consequently, 2015-2016 pre-exascale systems are being targeted as an intermediate step to enable hardware, system software and application scaling. Since not all DOE applications will be ready to run on a pre-exascale advanced architecture in 2015, a production capability Cielo follow-on will also be needed.

2. Application Drivers and Requirements

The ASC Program needs an improved production capability system for the Stockpile Stewardship Program to fulfil its mission. Studies of the ASC workload requirements established a need for greater than 2 petaFLOPS of computational resources, starting in 2010 with increasing demands through 2015. The ACES Cielo system will be the signature platform for Los Alamos, Livermore, and Sandia national laboratories' during its planned operation during this timeframe. Cielo will provide more than 1 petaFLOPS and an order of magnitude increase in capability over the ASC Purple system, soon to be retired at LLNL. As a capability class system Cielo will be used principally for the most challenging problems and largest parallel applications. ASC will use its Capability Computing Campaign work package approach for allocating time on Cielo; the proposed work packages are reviewed and prioritized based on relevance, importance, and technical merit. Simulation efforts using Cielo will be required to have one or more of their major planned simulations running on a significant portion of the entire machine.

Cielo is targeted to directly support the ASC roadmap goal of "*Establishing a validated predictive capability for key physical phenomena.*" The challenge to ASC and Stockpile Stewardship is to accurately assess the behavior of weapons and to ensure robust and reliable performance while maintaining the nuclear testing moratorium using science-based predictive simulations. This drives the need for fine-scale numerical resolution and advanced models for physics and material behavior for both system simulations and for the sciences supporting the models in those simulations. The workload anticipated for Cielo

will bring significant computer, algorithmic, and physics and engineering challenges.

A major planned use of Cielo will be for 3D full system weapon calculations with high spatial-temporal resolution and/or higher fidelity and more computationally complex physics models. A vast array of other simulation needs which Cielo will be used for include:

- New 3D baseline models, simulations and validation suites comparing high-resolution calculations to past underground nuclear and non-nuclear experiments and supporting stockpile annual re-certifications.
- Better understanding and the targeted elimination of four key physics modeling "knobs" related to weapons performance as called out in the Predicative Capability Framework.
- Improved 3D details and improved physics to allow Significant Finding Investigations (SFIs) to achieve an average closure time of just 2 years.
- Understanding the boost process and its initial conditions, which is a principle issue for weapons performance.
- Theory-based computation of material properties and nuclear physics, especially regimes and conditions unobtainable or infeasible by experiments.
- Large 3D simulations for assured safety and surety, performance, and survivability of non-nuclear weapons system components.
- Uncertainty Quantification efforts which have as a goal the prediction of important parameters at a 95% confidence level and the sensitivity of such parameters to various input data and model assumptions.
- Improved NIF target designs, implosion stability studies, and beam propagation and laser-plasma interactions.

The system architecture for Cielo was specifically targeted to provide ease of running the existing suite of ASC simulation and science codes with minimal required changes. As such, it emphasized the need for a robust MPI-everywhere programming approach on a petascale machine comprised of thousands of conventional 8-way or larger multi-core processors, which is the programming model used in most of the current ASC codes. Although the physical node count, size, and power of Cielo does not stress the technological envelope, the MPI rank count of 100,000 to 150,000 for the full Cielo system will almost certainly stress simulation scalability due to coding issues, parallelism strategies, and possibly some of the algorithms themselves. The Cray Baker architecture and

system software is well suited for the MPI-everywhere programming model, while also supporting quite a rich set of additional parallel programming strategies which codes may begin to explore or move to (e.g. MPI plus OpenMP or Pthreads, Shmem, CoArray Fortran, and PGAS languages like UPC & Chapel).

The number one hardware requirement drawn from the ASC applications developers and users was to have 2 GB of memory per core. Another design requirement was to maximize the aggregate system-wide memory bandwidth since studies have shown that many of the ASC applications performance benefit the most from increased memory bandwidth and not FLOPS rates. The Cielo design using AMD Magny-Cours 8-core rather than 12-core processors better matched these two design requirements with good cost effectiveness. Another Cielo design goal was to maximize MPI message injection rate, and the Gemini-based interconnect of the Cray Baker architecture excel in this metric. A highly robust and predicable user and operations experience was also a goal that the Baker design provides. Overall the Cielo architecture and its 1+ petaFLOPS size is an excellent match to the requirements and goals and will provide an excellent production capability system for ASC.

3. Cielo Architecture

The ACES design team was focused on a few key attributes that drove the development of the requirements and the selection process: reliability, power, hardware scaling, system software scaling, and application scaling. In this section, an overview of the high-level Cielo requirements will be provided and the corresponding Cray Baker architecture specification or feature.

Table 1: Cielo High-Level Aggregate Performance Metrics

Performance Metric	RFP Specification	Cielo
Peak FP	1.0 PF	1.37 PF
Total Memory Capacity	> 200 TB	298 TB
Memory per core	> 2 GB	2 GB[a]
Peak Memory BW	> 400 TB/s	763 TB/s
Sustained Bisection BW	> 20 TB/s	15.3 TB/s
Sustained Msg Injection Rate	> 50 GMsgs/s	71.5 GMsgs/s
Sustained Off Platform I/O BW	> 200 GB/s (160 GB/s PFS)	271 GB/s[b] (160 GB/s DVS)
System Power (max)	< 8 MW	4.4 MW (est)
Full System Job MTBI	> 25 hours	25 hours
System MTBI	> 200 hours	200 hours

System Hardware Requirements
The Cielo acquisition architecture requirements were numerous, but the following high-level metrics were used as the primary drivers, summarized in Table 1. The corresponding Cielo specification is also provided.

The application teams required at least 2 GB of memory per processor core, or more appropriately per MPI rank. Many of the current codes store large amounts of state for every MPI rank and in order to support the MPI-everywhere programming model, 2 GB is the minimum threshold that the application teams felt they could tolerate without significant code modifications.

In addition, the design team had the philosophy that memory bandwidth translates into application performance, and this metric was used to size the platform. This is in contrast to many acquisitions that use peak floating-point performance as the sizing metric. Total memory capacity is an artifact of the 2 GB/core requirement and less dependent on any particular application need.

The high-speed network performance is essential for good scaling characteristics. The design team was not too concerned with any particular topology, as the ASC applications have been shown to scale well on a fat-tree, a 3D mesh, a 3D torus, and hypercubes. It's more important to have enough bandwidth in the network to support a large amount of traffic and avoid congestion. And although raw bandwidth is highly desirable, other features such as routing algorithms and the ability to handle a failure in the network become more important. The sustained bisection bandwidth metric was a difficult parameter to specify, as it is extremely topology dependent. The 3D mesh and torus are topologies that inherently have a low minimum bisection when compared to a fat-tree for example. In the case of the Cielo requirements, it was a way to specify an adequate level of bandwidth in the network, but not drive any particular topology.

The message injection rate metric is a measure of how fast data can be injected into the high-speed network. The higher the message injection rate, the better the utilization of the network bandwidth. Injection rate is a small message metric. For all practical networks, it is not possible to sustain the full network bandwidth until message sizes become 10's to 100's of KB. As the message injection rate increases, full network bandwidth can be achieved with smaller and smaller message sizes. Many of the ASC applications use the bulk synchronous parallel communications model and have bundled multiple data structures into a single message in order to

increase the message size and hence make better use of the network. However, this is at the cost of additional memory bandwidth usage as many copies are required to build the "large" message. As the messaging rate gets better, it will be possible to avoid these copies and send each data structure individually and reduce total time to solution by eliminating the additional CPU and memory cycles necessary for large message formation. The ASC community has been asking for improvements in messaging rate for many years and it was included in the Cielo acquisition in order to reinforce the message. The latest generation of high-speed networks has excellent raw message rate capabilities, including Cray's Gemini and the quad data rate InfiniBand solutions. We would like to believe our message is being heard, but in fact it is probably due to the need for efficient networks for global address space programming models.

Off platform I/O bandwidth is primarily for the parallel file system. All of the ASC codes use application driven checkpoint restart. The frequency of a checkpoint is a function of the reliability of the platform and the rate at which checkpoint data can be written [4]. Based on the capability of Cielo and its expected MTBI a sustained I/O rate of 160 GB/s was chosen for the parallel file system. There are many file system performance metrics in the Cielo requirements, but in short the platform needed to sustain 160 GB/s for many application tasks to many files (N to N) and many applications tasks to a single file (N to 1), reading and writing.

For the Cielo acquisition, the parallel file system was procured separate from the machine. As such, Cielo also has a sustained network I/O bandwidth requirement of 200 GB/s. The design team felt that if the platform could sustain 200 GB/s of TCP/IP traffic, that 160 GB/s of file system performance was achievable.

The Full System Job (application) MTBI (mean time between interrupt) was specified to be at least 25 hours. The design team chose 25 hours as the minimum acceptable time that an application can run without interrupt and still be productive. This is a productivity metric and not something chosen based on machine size, technology or architecture. The System MTBI of greater than 200 hours specifies the minimum time between interrupts for the whole platform. That is, the platform becomes unavailable, unusable, or significantly degraded in resources and requires significant administrative intervention, such as a reboot. Again, System MTBI is a practical number for a capability platform and was chosen as a minimum time period that must be met in order for the platform to be productive.

System Software and Tools Requirements
In defining the system software requirements for Cielo, two operating system functionalities were specified, a fully-featured operating system (FFOS) and a light-weight operating system (LWOS). The FFOS and the LWOS do not have to be distinctly different, i.e. the LWOS did not have to be a light-weight kernel such as the Catamount OS used on the XT systems and Red Storm [5]. It was perfectly acceptable for the FFOS and the LWOS to be derivatives of a single base operating system with different features configured, if necessary, to meet the requirements.

The FFOS supports service applications and functions, for example login, external I/O, batch scheduling, etc. The FFOS was required to be configurable to support as many features as required for a given service or hardware configuration in Cielo.

The intent of the LWOS is to facilitate application scalability. For example, LWOS features that promote application scalability are: low overhead (noise) in the kernel, compact size, efficient execution, and minimal but sufficient kernel capabilities to execute current ASC capability workloads. Some applications will require more features than others, e.g. support for dynamic linking, dlopen(), OpenMP, POSIX threads, and Python. The LWOS was required to be configurable to support one or a combination of these features. Some of these OS features do not support scalable execution or scalable implementations may not exist. In this case, if the feature was configured into the OS, it was required to not impede the performance of an application that did not utilize it.

Cielo Hardware Architecture: Cray Baker Architecture
The Cielo platform will be an instantiation of Cray's Baker supercomputer architecture. It is outside the scope of this paper to describe the details of the Cray Baker architecture, but the Cielo configuration and options will be expanded upon in the following sections.

Overview
Cielo will be deployed in two phases due to the platform funding profile. The 1st phase will be deployed in the 4th quarter of 2010, while the 2nd phase will increase the size of the platform by 33% and will be integrated into Phase 1 in the 2nd quarter of 2011.

Processor Configuration
The primary Baker architecture option is the choice of the AMD 6100 Series (Magny-Cours) processor [6]. AMD offers 8-core and 12-core models of the Magny-Cours. The ACES design team chose to use the 8-core model

6136 processor.[2] The 12-core model would have required using the more expensive 8 GB DIMM in order to meet

AMD Magny-Cours Performance Predictions (relative to 2.6 Ghz, 6-core AMD Istanbul)
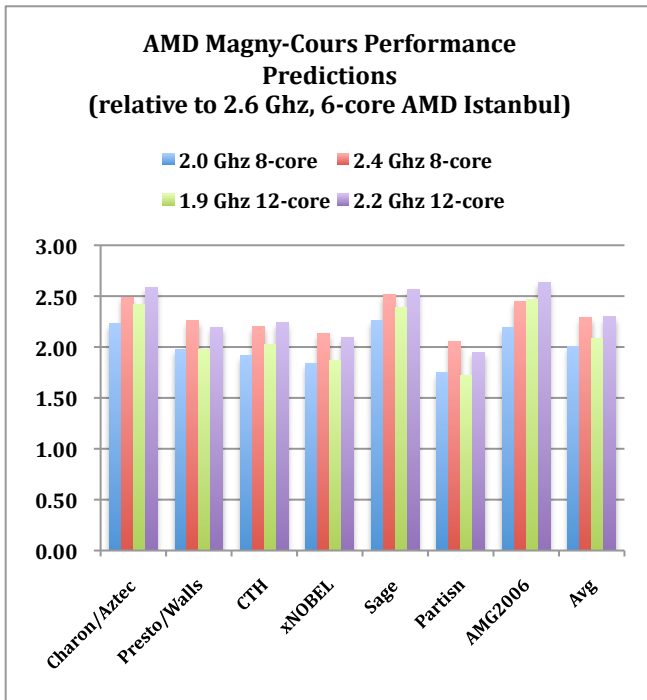
the 2 GB/core memory requirement. The ACES design team was also focused on maximizing the memory bandwidth of the platform. The 8-core models use the same DDR3-1333 Mhz memory as the 12-core models. In addition, the 12-core part is priced significantly higher than the 8-core part when comparing similar frequency bins.

The performance modelling results show an advantage to using the 2.4 Ghz 8-core processor as opposed to the price equivalent 1.9 Ghz 12-core processor [7]. In fact, on average there was no advantage for the 2.2 Ghz 12-core processor, which cost nearly 60% more [6].

Based on the 2 GB/core memory requirement and the performance modeling analysis, the model 6136 8-core processor was chosen for Cielo. Had the ACES design team better understood the pricing structure at the time of negotiations, strong consideration may have been given to the lower frequency model 6128 8-core processor.

Visualization and Data Analysis Partition
Four cabinets of the compute section will have double the memory of the rest of the compute partition. This is

---

[2] The AMD 6136 specifications are: 2.4 Ghz, 12 MB L3 cache, 6.4 GT/s HT3 link, DDR3-1333 Mhz DRAM.

primarily to support visualization and data analysis applications, but it may also be used for those applications requiring more memory per core. This sub-partition will be configured with 4 GB of memory per core, as opposed to the 2 GB per core for the rest of the compute partition. The four large memory cabinets will be configured to be a 4x2(4)x24 sub-mesh in the Cielo topology.

Parallel File System and Integration into the LANL PaScalBB
Cielo will be integrated into the LANL Parallel Scalable Backbone Global Parallel File System (PaScalBB). Cielo has been configured to provide greater than 200 GB/s of sustained TCP/IP bandwidth to the PaScalBB. The PaScalBB will be expanded to include an additional 10 PB of user available storage capacity and an additional 160 GB/s of sustained file system performance.

In Cielo's final configuration, 272 service nodes, each with two 10 GigE connections, will be connected to the PaScalBB. Each service node will provide more than 1.2 GB/s of sustained network bandwidth.

4. Cielo Schedule and Facilities

Cielo will be delivered in two phases. Phase 1, consisting of over a petaFLOPS capability, will be delivered in the third quarter CY2010. Phase 2, consisting of an additional 0.33 petaFLOPS, is scheduled for delivery in the second quarter of CY2011. The two phases will be integrated as one system in the third quarter of CY2011. The following chart provides a high-level overview of the Cielo schedule.

CIELO High Level Schedule
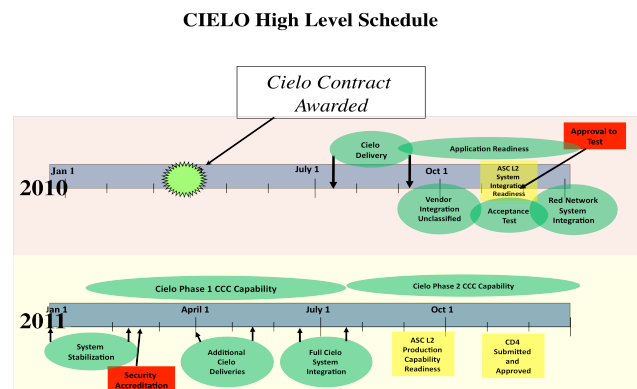
Once delivery is completed the Phase 1 system will be integrated into the unclassified network at Los Alamos. Acceptance testing is scheduled for the November 2010 timeframe. Once acceptance testing is completed Cielo will be "moved" to the classified environment for integration into LANL's classified network for initial

stabilization work. Allocations for use of the machine, starting in the first quarter of CY2011, will be based on the NNSA ASC Capability Computing Campaigns model. Cielo will be operated as the NNSA's National User Facility for capability computing.

Additional smaller system and applications testbeds are part of the acquisition to provide initial systems for getting selected applications ready to use the system in preparation for the full Cielo system deliveries. Phase 2 will be procured and tested in CY2011 before being integrated with the Phase 1 system to provide the final Cielo system.

Cielo was acquired under the NNSA Office of the Chief Information Office (OCIO) Project Execution Model (PEM) for IT Investments. This model provides for several Critical Decision (CD) milestones to be approved for acquisition and operation. Cielo must also meet several ASC programmatic milestones before it is formally approved for production capability. The full platform will consist of 96 cabinets using less than 1500 sq. ft. and is targeted to use less than 4MW of power for operation. The Cielo platform will be air-cooled with bottom to top airflow through the cabinet. The site preparation for Cielo has been completed and is ready for system installation.

The Cielo platform will be housed at the Strategic Computing Complex facility (also known as the Nicholas C. Metropolis Center) at Los Alamos National Laboratory (shown below).

5. Interconnect Development and Engineeering Collaboration

In the fall of 2008, NNSA/ASC asked ACES to consider definition of, and technical oversight for a technology development and engineering project with Cray. By early 2009, after a systematic review of a number of technology development opportunities, ACES settled on a project in advanced interconnection network technology development. ACES worked in collaboration with Cray to define a statement of work and deliverables that matched the ASC funding profile. This statement of work was finalized in the fall of 2009. The final agreement on contractual terms and conditions and issues regarding treatment of Intellectual Property were resolved in late spring, 2010.

The Cray-ACES collaborative Interconnection Network Project (INP) focuses on a potential interconnection network, which Cray refers to as the *Pisces* Interconnect. While the Pisces Interconnect is not currently on Cray's roadmap, the intent of this project is to analyze potential capabilities for Cray to include in Pisces that will result in significant performance impact on a suite of ASC applications and integrated codes. Assuming our collaborative Pisces interconnect research and development effort culminates successfully, Cray would plan to make the Pisces Interconnect available in its future commercially available computer systems. Based on current projected timetables this would not occur before CY 2015.

Cray already has several generations of network interconnect in its products or on the drawing board:

- Generation 1 = Seastar (in production) – collaboratively designed and developed with Sandia National Labs
- Generation 2 = Gemini (in prototype debug, production 3Q CY2010)
- Generation 3 = Aries (in design, production late 4Q CY2012)

Assuming the Pisces Interconnect effort comes to fruition, the Pisces Interconnect will become Cray's 4th Generation of network interconnect and would leverage all knowledge and experience from the first 3 generations. It is anticipated that Pisces will incorporate a derivative of the network interface controller first used in Gemini and the network topology first used in Aries. The project will involve three (3) stages of effort.

Stage One (1) NIC Studies/Analysis:
Cray and ACES will analyze the performance characteristics of the *Gemini* interconnect (2nd generation) and look for areas of improvement that can be leveraged to enhance the capabilities of the Pisces interconnect. This effort will focus on the Gemini network interface controller (NIC), with a particular emphasis on occupancy, latency, and MPI message throughput and

independent progress. All of these characteristics have an important impact on application scalability. In this effort, Cray and ACES will leverage experience garnered from Cielo. Baseline analysis will be accomplished by running on real Gemini hardware, while architectural exploration will be accomplished through simulation.

Stage Two (2) Router and Network Studies/Analysis:
Cray will analyse the performance characteristics of the *Aries* interconnect (3rd generation) and look for areas of improvement that can be fed into and used to enhance the Pisces Interconnect development. Specifically, Cray will analyse the "network routing" portion of Aries. This work will encompass:

- Cray and ACES will jointly define the application trace format that will be used throughout this INP effort.
- Aries "network routing" simulations using ASC application traces received from and important to ACES.
- In order to do realistic and adequate simulation of ASC application traces on a complex Aries network, a dedicated and computationally capable system will be used by Cray. Individual simulations are expected to run for days to weeks on such a dedicated platform and many variants/trials of each simulation will need to be completed, analyzed and retried. ACES simulators will also be used to complement Cray's simulation efforts.
- Aries "network routing" evaluations on Cray's commercially available computer systems (this is post the conclusion of the HPCS program) using real Aries hardware and the Aries SW stack (Nile SW Stack).

Stage Three (3) Pisces:
Cray will perform a comparative study between state of the art InfiniBand interconnects (available in CY2015) and a potential Pisces interconnect (targeted for production in CY2015). Cray will validate and quantify the value of the Pisces interconnect relative to commodity IB interconnects. In addition, there is the potential to compare the potential Pisces interconnect to 1 to 2 other CY2015 interconnects. The decision to invest in this comparative analysis will be made jointly by Cray and ACES.

The Initial Pisces Architectural Specification will be crafted in collaboration with ACES and will be based on the joint architectural explorations performed in Stages 1 and 2. It will include the specifications for the NIC and router portions of the Pisces chip architecture and will fully describe all functions, features, performance targets,

error handling mechanisms, reset mechanisms, user and kernel accessible registers/programmability for the NIC and router. This initial specification will be the working blueprint for the Pisces architecture.

6. Conclusion

ACES is partnering with Cray to deploy a production petascale capability platform, Cielo, in 2010. Cielo will be used to solve critical DOE/NNSA national security problems. Another partnership with Cray is focused on developing advanced interconnects for pre-Exascale systems in the 2015-2016 timeframe.

Acknowledgments

About the Authors

James Ang is the Technical Manager of the Scalable Computer Architectures Department at Sandia National Laboratories. James is the Leader of the ACES Architecture Office. James is interested in applying Open Innovation principles to develop collaborative teams spanning national labs, industry and academia to codesign exascale architectures, algorithms and applications.

Douglas Doerfler is a Principle Member of Technical Staff at Sandia National Laboratories. Doug is the ACES Cielo Architect and his research interests include high-performance computer architectures and performance analysis.

Sudip Dosanjh is the head of computer and software systems at Sandia National Laboratories. He is co-director of ACES and the Institute for Advanced Architectures and Algorithms. He has served on DOE's Exascale Initiative Steering Committee.

Scott Hemmert is a Senior Member of Technical Staff at Sandia National Laboratories, where he leads the advanced interconnects research in the scalable architectures department. Recent research has focused on network interface architectures for enabling independent progress and high message rate MPI.

Kenneth Koch is an R&D Scientist-5 at Los Alamos National Laboratory working on current and future high-performance computer architectures, including programming approaches and performance. Ken is the Technical Manager of the Roadrunner Cell-accelerated supercomputer at LANL and the Deputy Leader of the

Architecture Office of ACES.

John Morrison leads the High Performance Computing Division at Los Alamos National Laboratory. He has worked in the high performance computing field from the early days of the Cray-1. John is co-director of ACES and has served on DOE's Exascale Initiative Steering Committee.

Manuel Vigil is a Program-Project Director at Los Alamos National Laboratory working on project management, planning, acquisition, and integration of current and future high performance computing systems. Manuel is the Project Manager for the Cielo system as part of the ACES Collaboration.

References

[1] ASC,Purple, http://en.wikipedia.org/wiki/ASC_Purple, 2010.

[2] Alvin, K. et al., "On the Path to Exascale," to appear in the *International Journal of Distributed Systems and Technologies*, 2010.

[3] Underwood, K.D., Levenhagen, M., Rodrigues, A.: "Simulating Red Storm: Challenges and successes in building a system simulation," *21st International Parallel and Distributed Processing Symposium*, March 2007.

[4] Daly, J., "A higher order estimate of the optimum checkpoint interval for restart dumps," *Future Generation Computer Systems*, 2006, pp. 303-312.

[5] Tomkins, J. et al., "The red storm architecture and early experiences with multi-core processors," to appear in the *International Journal of Distributed Systems and Technologies*, 2010.

[6] "The AMD Opteron™ 6000 Series Platform," http://www.amd.com/us/products/server/processors/6000-series-platform/Pages/6000-series-platform.aspx.

[7] Douglas Doerfler, Courtenay Vaughan, Mahesh Rajan, Paul Lin, "Predicting Application Performance of AMD's Next Generation Opteron Processor, Magny-Cours", internal Sandia white paper.