SANDIA REPORT

SAND2009-6496
Unlimited Release
Printed October 2009

Evaluation of the Impact Chip Multiprocessors have on SNL Application Performance

Doug Doerfler

Electronic version available at:
http://www.sandia.gov/L2_milestone_report_2009/SAND2009-6496.pdf

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865)576-8401
Facsimile: (865)576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800)553-6847
Facsimile: (703)605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2009-6496
Unlimited Release
Printed October 2009

Evaluation of the Impact Chip Multiprocessors have on SNL Application Performance

Douglas W. Doerfler
Sandia National Laboratories
PO Box 5800
Albuquerque, NM 87185-1319

Abstract

This report describes trans-organizational efforts to investigate the impact of chip multiprocessors (CMPs) on the performance of important Sandia application codes. The impact of CMPs on the performance and applicability of Sandia's system software was also investigated. The goal of the investigation was to make algorithmic and architectural recommendations for next generation platform acquisitions.

[*]Electronic version with navigational hyperlinks available at:
http://www.sandia.gov/L2_milestone_report_2009/SAND2009-6496.pdf

Acknowledgments

Funding for these efforts was provided by ASC/CSSE.

The author acknowledges the assistance of NNNSA/ASC HQ, CSSE program managers, platform design team members, IC program managers, and application developers.

Performance Modeling and Analysis Team members were instrumental in data collection and analysis.

Table of Contents

Evaluation of the Impact Chip Multiprocessors have on SNL Application Performance

Executive Summary

Sandia has met the requirements of the ASC Level 2 Chip Multi-Processors Milestone (Milestone #3158). Investigations have had a major impact on 1) Red Storm upgrade to dual-core processors, 2) Red Storm upgrade to quad-core processors, 3) improved performance and scalability of the TLCC platform, 4) the acquisition and deployment of Red Sky, and 5) the development of the technical requirements for the NNSA Zia Capability Computing Platform.

This effort was a result of a multi-disciplinary effort with contributions from CSSE System Software and Tools, CSSE Advanced Systems, CSRF and LDRD programs. Staff included members of three Centers, 1400, 9200 and 1500.

Customers of this Milestone are NNSA/ASC HQ, CSSE program managers and platform design team members, IC program managers, and application developers.

Milestone Objective Success Criteria

The focal points in the Milestone were to investigate the impact of chip multi-processors (multi-cores) on the performance of important SNL application codes and the impact on the performance and applicability of SNL's system software. In addition, algorithmic and architectural recommendations were to be made for next generation platform acquisitions. Results were to be formally presented in a program review and documented in a report. All of these goals have been met.

In addition to an impact on platforms, this milestone made contributions to and had collaborations with the ASC algorithm and application groups. These collaborations led to development of the Mantevo mini-applications, multi-core capabilities in the Trilinos solver libraries, and performance improvements to the ALEGRA code.

CMP research and investigations associated with this milestone led to numerous publications, journal articles, invited talks/presentations and seminars. The Catamount N-way with SMARTMAP technology, part of this milestone effort, won a prestigious R&D 100 award in 2009.

To meet the success criteria, SNL worked closely with the vendor community to understand their processor roadmaps and technologies that are likely be deployed in the ACES Zia capability initiative. SNL worked closely with vendors to acquire early examples of technologies for evaluation. Also, computer architects have worked closely with application and algorithm developers, including Intel and the Portland Group compiler vendors, to research and develop techniques and methodologies to extract optimal performance from multi-core processors.

As well, criteria were developed and compared with benchmarking results where applicable to provide metrics for successful completion of this milestone. For Sandia's TLCC systems, these efforts had significant impact, a few of which are listed below:

- Performance analysis led to recommendation to use MVAPICH instead of OpenMPI
- Processor and memory affinity analysis at scale led to deployment of a job-launch wrapper that uses numactl to do process and memory placement
- Identification of the impact of cache coherency overhead on multi-socket node performance

Within CSSE, the impact of multi-core spanned both hardware and software tasks. The system software team ensured that the Red Storm system software kept pace with the Red Storm dual-core and quad-core upgrades. In addition, interaction between the system software team and the performance analysis team was essential to understanding and quantifying the effect of processor and memory affinity on application runtime.

CSRF and LDRD projects led by Mike Heroux provided many of the algorithm contributions and were instrumental to the architecture and performance analysis teams. These projects provided mini-applications and consultations in understanding the impact of architectural features of multi-core at the application and algorithm level.

Future impact of this milestone will be evident in the next-generation Zia platform. For example, performance analyses of candidate processors have provided quantitative data for technology decisions. Also, the 6X application performance benchmarks for Zia are driving code teams to address higher levels of scalability and parallelism. Our research also identified mesh generation for large problems as a major application readiness issue to use Zia at scale. Resolutions to all of these issues will prove crucial for Zia and beyond.

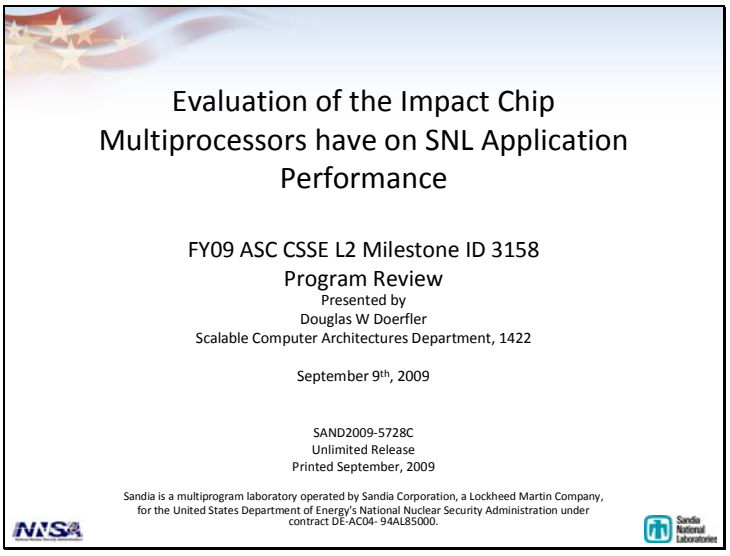
Multi-core Algorithmic and Architectural Recommendations for Next Generation Platforms

- Quantitative analysis has validated the importance of "effective" memory bandwidth for the Tri-Labs applications, and continuing architecture evaluations are driving processor choices for Zia. Multi-core is increasing the gap between computation and memory performance (i.e. the memory wall).

- Processor and memory affinity control is essential for extracting maximum performance from multi-core architectures. As we make the transition from multi-core to many-core (10's to 100's of cores interconnected at the silicon level by high-speed interconnects similar to today's HPC systems), this will become even more critical.

- The deployment of multi-core, and eventually many-core, in capability class platforms will drive much higher levels of parallelism for applications. The 6x application suite is forcing code teams to work these issues early and identify deficiencies, e.g., mesh generation for problem sizes necessary to fully utilize Zia at scale. Future machines will require at least two orders of magnitude more parallelism than is being deployed on current platforms.

Slide 1: Executive summary



Executive Summary

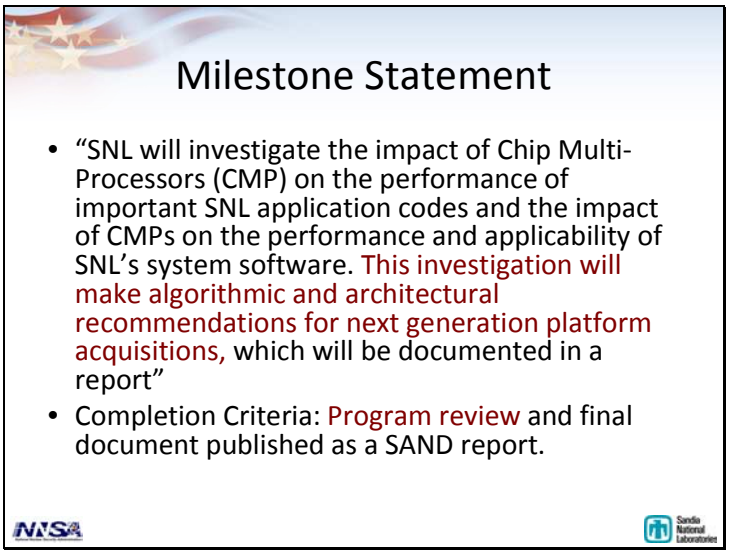
Sandia has met the requirements of the ASC Level 2 Chip Multi-Processors Milestone (Milestone #3158). Investigations have had a major impact on 1) Red Storm upgrade to dual-core processors, 2) Red Storm upgrade to quad-core processors, 3) improved performance and scalability of the TLCC platform, 4) the acquisition and deployment of Red Sky, and 5) the development of the technical requirements for the NNSA Zia Capability Computing Platform.

In addition to an impact on platforms, this milestone made contributions to and had collaborations with the ASC algorithm and application groups. These collaborations led to development of the Mantevo mini-applications, multi-core capabilities in the Trilinos solver libraries, and performance improvements to the ALEGRA code.

CMP research and investigations associated with this milestone led to numerous publications, journal articles, invited talks/presentations and seminars. The Catamount N-way with SMARTMAP technology won a prestigious R&D 100 award in 2009.

This effort was a result of a multi-disciplinary effort with contributions from CSSE System Software and Tools, CSSE Advanced Systems, CSRF and LDRD programs. Staff included members of three Centers, 1400, 9200 and 1500.

Slide 2: Milestone statement



This milestone is an FY09 ASC Level 2 for the CSSE subprogram.

Completion Date: Sept-09

Customer: NNSA/ASC HQ, CSSE program managers and platform design team members, IC subprogram managers, and application developers.

Milestone Certification Method: 1) a program review is conducted and results are documented, and 2) Professional documentation is prepared as a record of milestone completion.

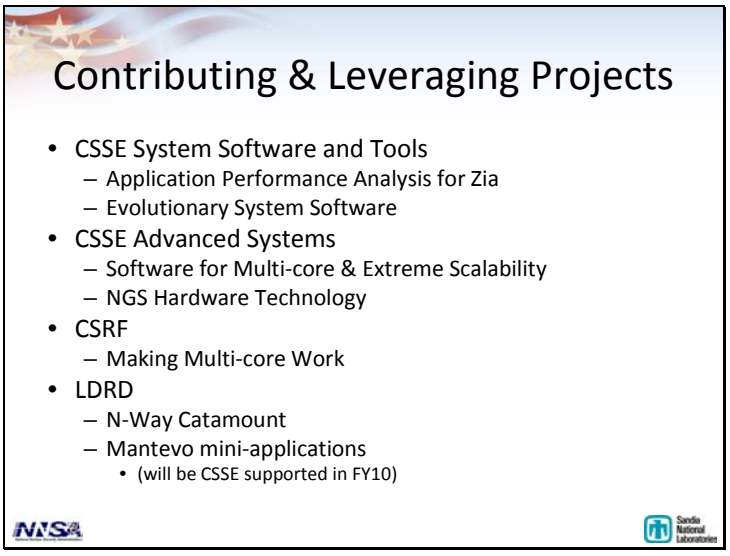
Supporting Resources: ASC capacity and capability platforms, test beds and early deliverables acquired commercially and via CSSE/Advanced Systems projects.

Codes/Simulation Tools Employed: This effort leverages the performance analysis and modeling tools developed and/or supported by ASC/CSSE, including but not limited to O|SS and SST.

Contribution to the ASC Program: Platform acquisitions will be better equipped to deliver architectures and systems to meet the mission needs of NNSA's computing campaigns.

Contribution to Stockpile Stewardship: Platforms

Slide 3: Contributing & leveraging projects

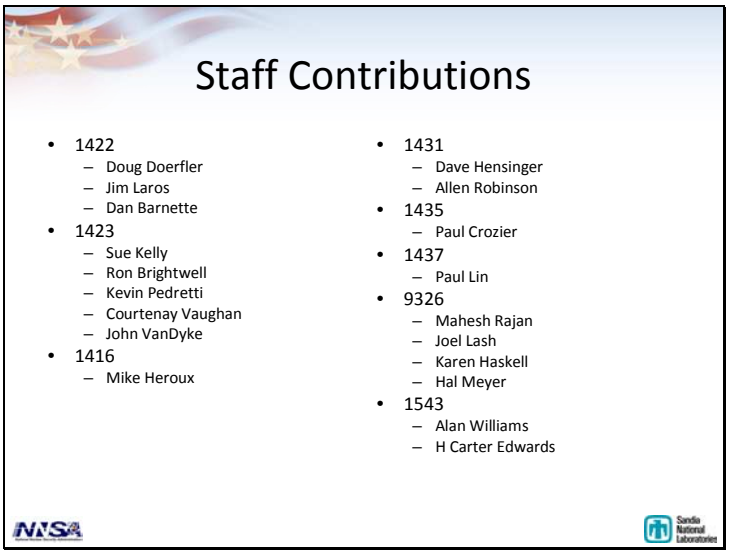


The multi-core work was encompassed by multiple projects and efforts within and external to CSSE.

Within CSSE, the impact of multi-core spanned both hardware and software tasks. The system software team ensured that the Red Storm system software kept pace with the Red Storm dual-core and quad-core upgrades. In addition, interaction between the system software team and the performance analysis team was essential in order to understand and quantify the effect of processor and memory affinity on application runtime.

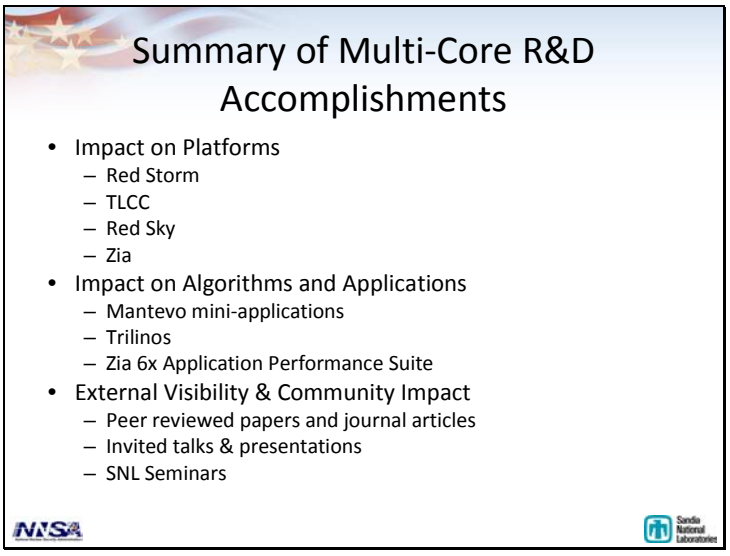
Mike Heroux's CSRF and LDRD projects provided many of the algorithm contributions and were instrumental to the architecture and performance analysis teams in providing mini-applications and consultation in understanding the impact of architectural features of multi-core at the application and algorithm level.

Slide 4: Staff contributions



Contributors spanned multiple 1400 organizations in addition to centers 9300 & 1500.

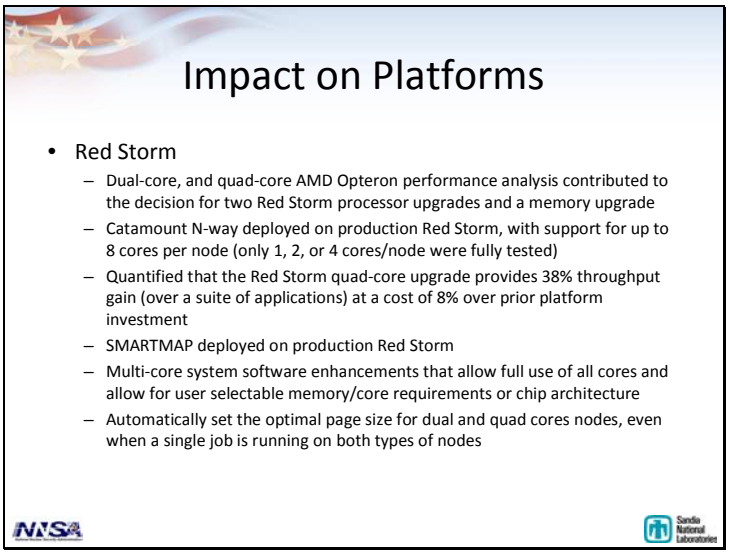
Slide 5: Summary of multi-Core R&D accomplishments



The format of the review is to summarize the impact on platforms, applications, and external visibility as a result of the multi-core initiatives within CSSE. The amount of material that has resulted from this effort is much greater than can be communicated in the timeframe of this review, but highlights of some of the major contributions will be presented.

To demonstrate external visibility and community impact, a bibliography of publications, presentations and seminars is supplied as a separate appendix.

Slide 6: Impact on platforms



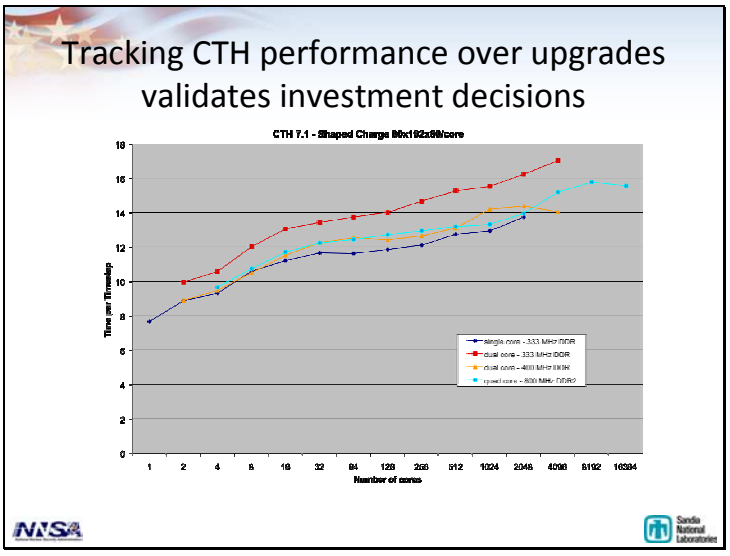
Red Storm has gone through two major upgrades over its lifetime. This includes two separate multi-core upgrades and numerous updates and revisions to the Catamount light weight kernel and associated runtime.

Many of the system software upgrades allowed applications to take advantage of the multiple cores without having to make any application changes. E.g. in going from a dual-core to a quad-core processor AMD changed the number of small and large page TLB entries. By automatically setting the optimal page size based on processor type, the application developer and/or analyst did not have to know what page size to specify for best application performance for each run.

The suite of applications for the throughput improvement measurement included CTH, SAGE, LAMMPS, POP, and ALEGRA.

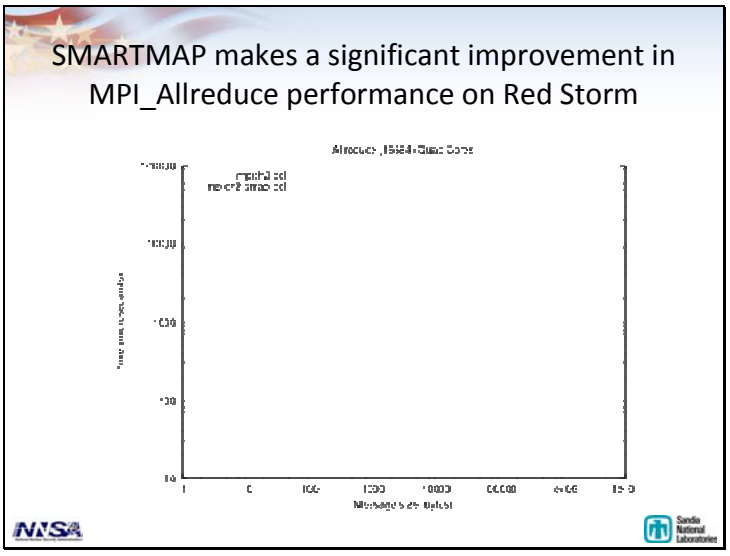
Catamount N-way with SMARTMAP technology won a prestigious R&D 100 award in 2009.

Slide 7: Tracking CTH performance over upgrades



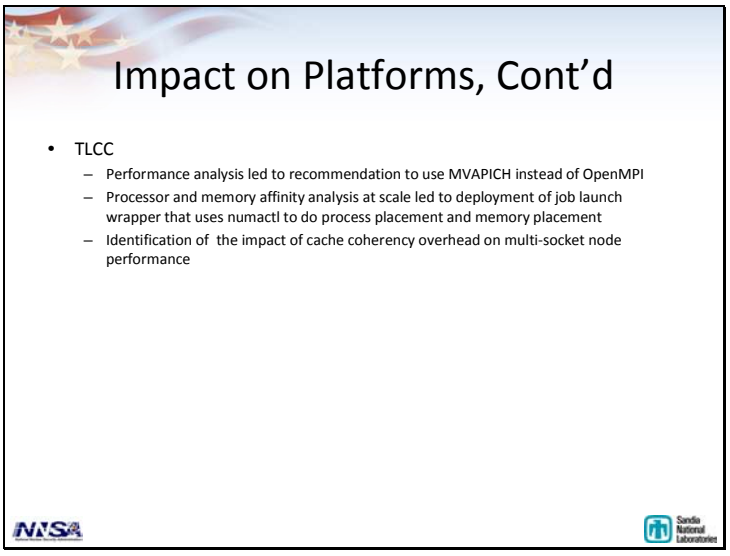
CTH performance was tracked to validate performance after each major Red Storm upgrade. The first phase of Red Storm used single-core AMD Opterons and 333 MHz DDR1 memory. During the first upgrade, dual-core processors replaced the single cores, with memory speeds remaining at 333 MHz. For a fixed size MPI job, performance degraded as the two cores contended for the memory subsystem. A fifth row was added to Red Storm, and those nodes received 400 MHz DDR1 memory. This bump in memory performance helped the dual-core processors to be bridge the gap to the single core performance, while utilizing one-half as many nodes. The upgrade to quad-core processors, and their associated 800 MHz memory subsystem, demonstrated essentially equivalent performance to the single-core and dual-core configurations. Another 2x reduction in the number nodes for the quad-core run is one-quarter the number of nodes of the single-core results. This achievement is due to not only hardware improvements, but also the necessary and associated software improvements.

Slide 8: SMARTMAP make significant improvement



A key benefit to applications provided by SMARTMAP technology is the improvement seen in MPI_allreduce operations. Allreduce performance is critical in many of NNSA's science and engineering codes, and is particularly important in many sciences codes such as weather and ocean model applications.

Slide 9: Impact on platforms, cont'd

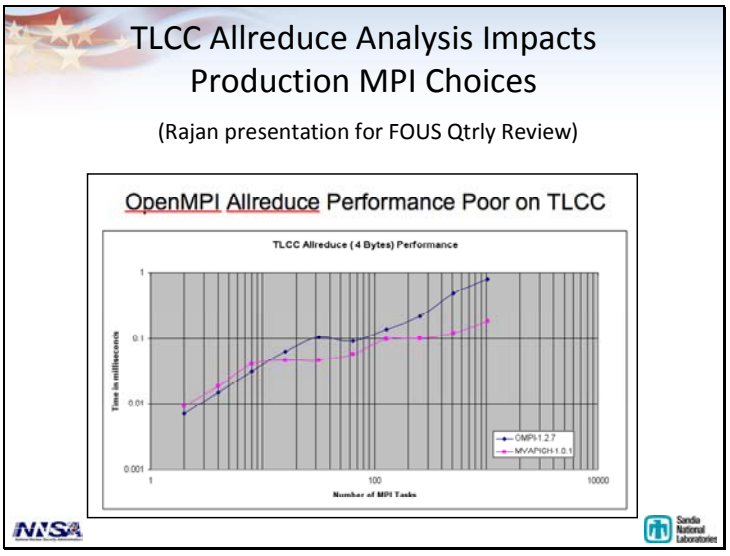


TLCC is the NNSA's Tri-Lab capacity cluster acquisition which was used to deploy multiple clusters at all three Labs. The TLCC platforms use an AMD quad-core, quad-socket node and DDR Infiniband for the high-speed interconnect. Early performance analysis identified several issues that were preventing applications from running well at scale. One was the poor All_reduce performance of the OpenMPI implementation. Studies quantified the issue and led to feedback to the OpenMPI community, with subsequent algorithmic fixes. MVAPICH demonstrated much better collective performance, and as a result, MVAPICH is the recommended MPI implementation on TLCC platforms.

Early multi-core studies on CSSE test beds identified the need to ensure proper memory affinity for best performance with NUMA architectures. The need for proper affinity was made even more evident when analyzing application performance at scale on TLCC. The deployment of job launch wrapper scripts to force proper affinity on the TLCC platforms led to substantial performance gains and more predictable runtimes. The following slides demonstrate the effects of not using memory and processor affinity controls.

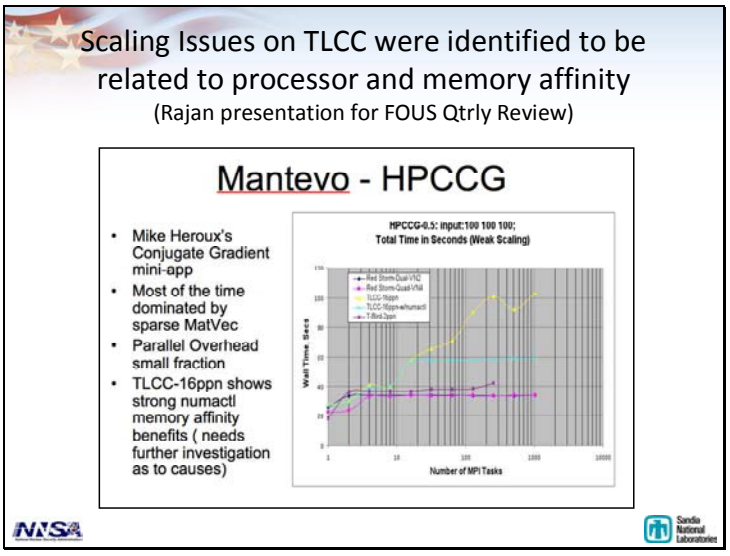
The AMD Barcelona processor demonstrates reduced memory subsystem performance in quad-socket configurations. This is due to a poor cache coherency protocol design that has since been fixed in the AMD Istanbul processor. The affects of the reduced performance were quantified by the CSSE application performance analysis team.

Slide 10: TLCC allreduce analysis



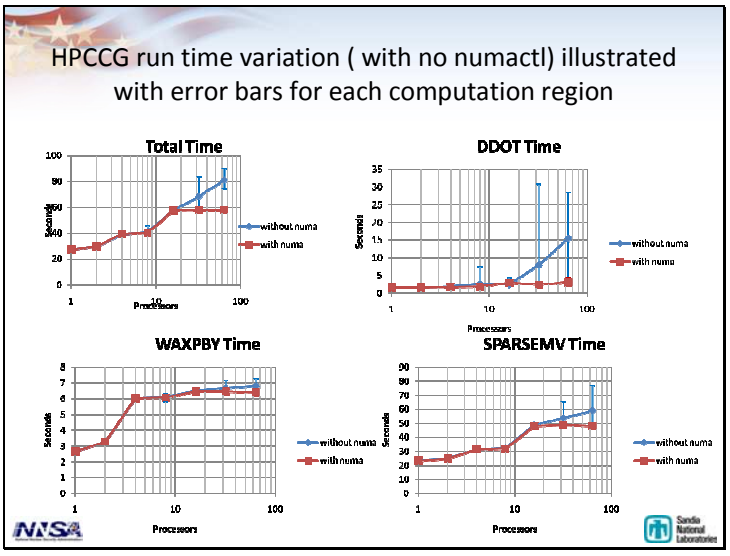
This chart demonstrates the reduced Allreduce performance demonstrated by OpenMPI. Note that the vertical axis is in Log scale, and at 1024 MPI ranks performance is approaching an order of magnitude performance degradation as compared to the MVAPICH results. These early tests by the CSSE performance analysis team led to the recommendation to use the MVAPICH MPI for production applications.

Slide 11: Scaling issues on TLCC



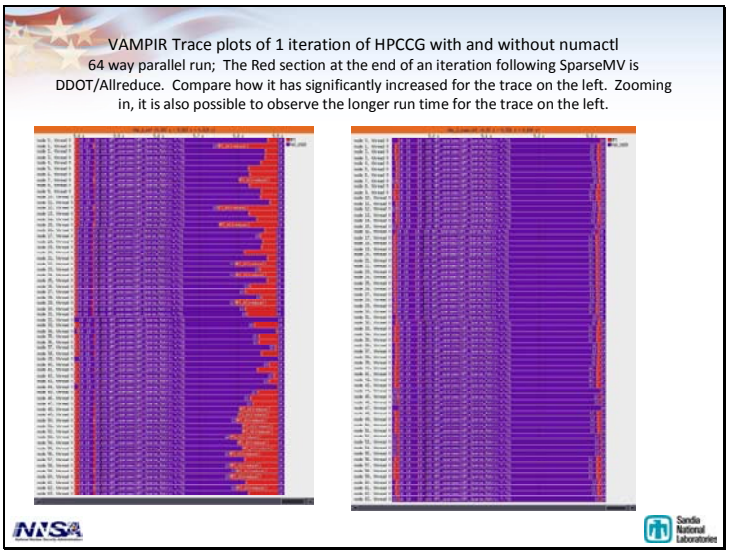
The Mantevo HPCCG mini-application demonstrates the effects of improper memory and processor affinity control. Although this chart does not contain error bars, it has also been shown that the variation is dramatically reduced when using proper affinity. In this instance, at 1024 cores, by using memory affinity controls the runtime is reduced ~40% and follows a similar trend to the expected behavior as demonstrated on the other platforms.

Slide 12: HPCCG run time variation



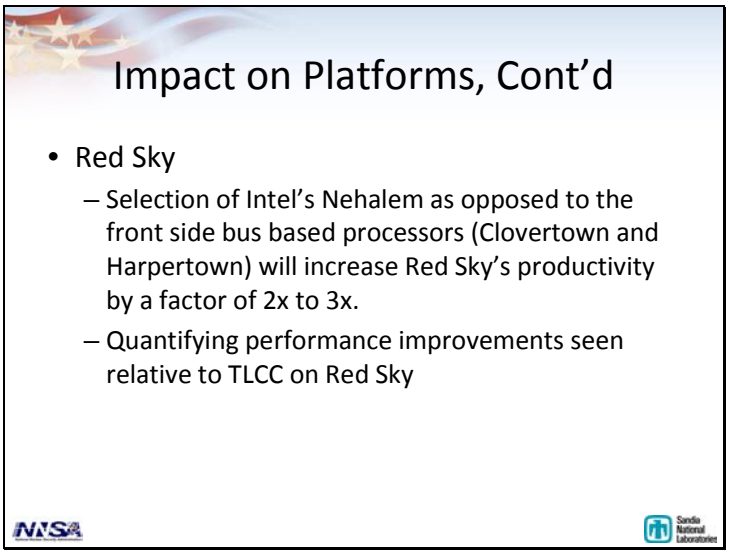
A more detailed analysis of the HPCCG mini-application and the effects of affinity control. Although runtime variations can be seen with as little as 8 processors (cores), at 16 processors (cores) the effects become dramatically apparent. This is the point at which multiple TLCC nodes are required, which indicates that the effect on communication performance is a factor in addition to computational performance.

Slide 13: VAMPIR trace plots



Using performance analysis tools, such as VAMPIR Trace, it is possible to get a visual representation of the effect of affinity control on runtime behavior. Note that this is only a 64 rank job! The effect is similar to load imbalance due to domain decomposition in parallel algorithms. Another analogy is an increase in system noise, which impacts some processors more than others, and hence reduces the per iteration performance of all the participating processors to that of the slowest node.

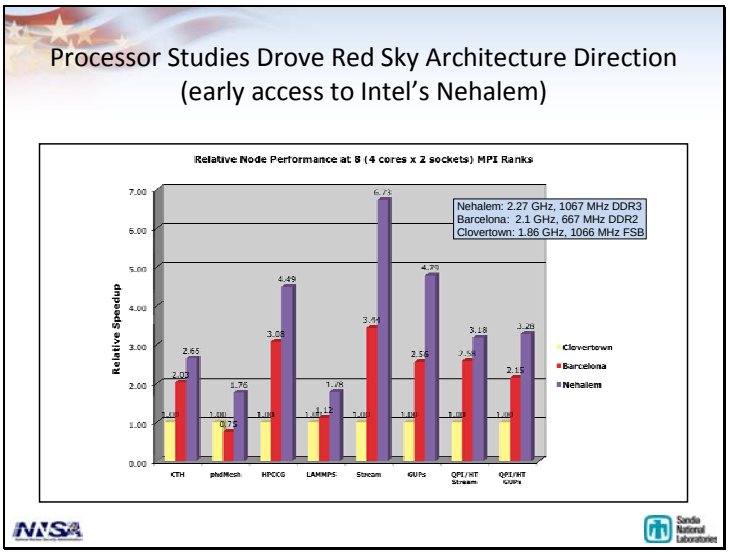
Slide 14: Impact on platforms, cont'd



Red Sky is Sandia's newest institutional computing platform. Early performance analysis of Intel's front side bus (FSB) based processors quantified the inability of Intel's powerful cores to scale beyond 2 cores per socket. This is primarily a limitation of the FSB architecture and poor FSB chipset implementations. This led to the recommendation to not select a proposal that used this architecture. Subsequent performance analysis of Intel's Nehalem processor, with integrated memory controllers, showed excellent performance improvements to not only Intel's FSB based solutions, but also to AMD's Opteron processor. This led to selection of a Nehalem-based architecture for Red Sky.

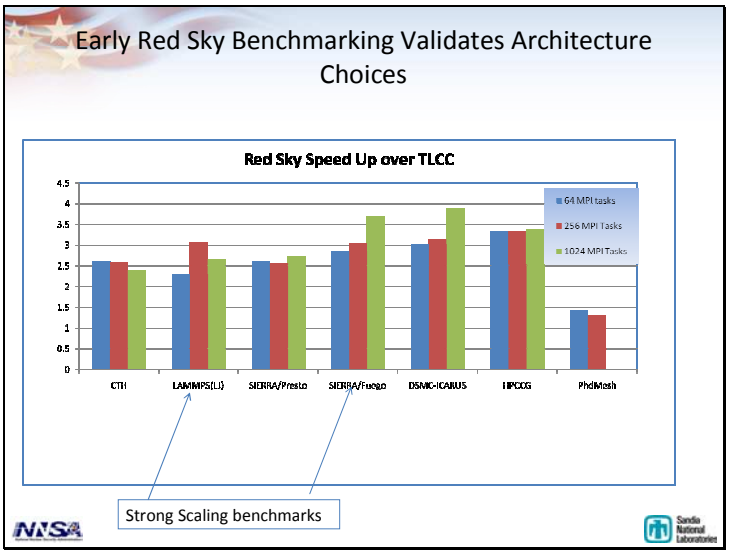
Early performance analysis of Red Sky is currently taking place, with performance being compared to the TLCC and Red Storm platforms to ensure that the platform is performing up to expectations. Without the expert knowledge of what performance is expected, it would be difficult to access performance issues and associated acceptance criteria.

Slide 15: Processor studies drove Red Sky architecture direction



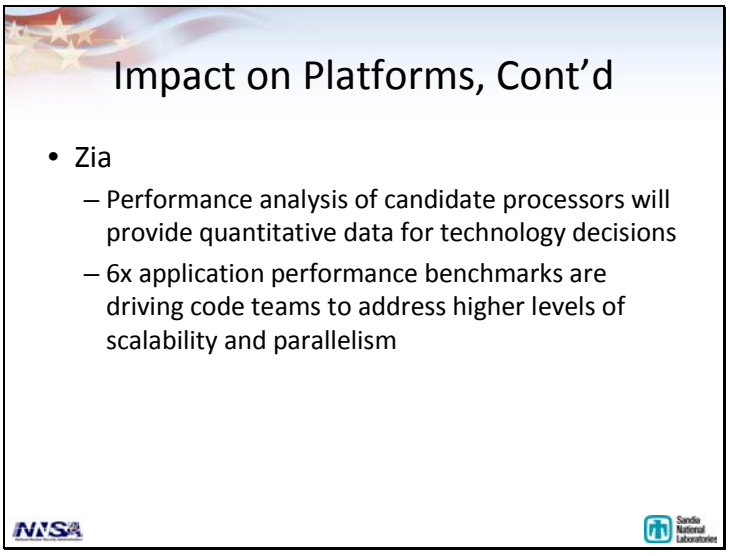
This is a summary of the processor performance study used for the Red Sky platform acquisition. CTH, phdMesh, HPCCG and LAMMPS were the applications used in the RFP, and hence analysis was focused on those applications also. Stream and GUPs are indicators of memory subsystem performance. The QPI/HT results for stream and GUPs are quantifying the performance of the memory subsystem when using the respective inter-socket interconnects. This analysis showed the significant performance potential of the Intel Nehalem process. This study used an early evaluation workstation from Intel and the higher CPU and memory subsystem clock rates of Red Sky were expected to provide even greater performance improvements.

Slide 16: Early Red Sky benchmarking



This slide summarizes some early application results on Red Sky, relative to TLCC. These early results help to validate the architectural choices made during the Red Sky acquisition. Performance is 2.5x to 3.5x that of TLCC for job sizes up to 1024 MPI ranks. It is expected that performance gains will be even larger at higher scales.

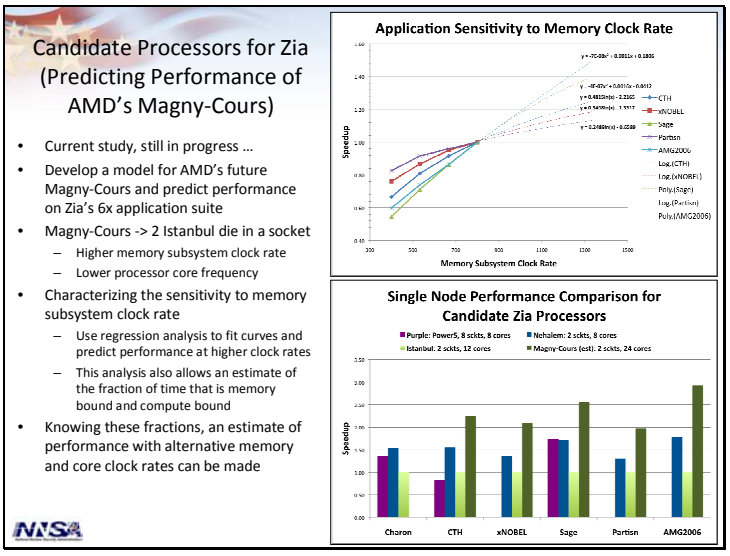
Slide 17: Impact on platforms, cont'd



Zia is the NNSA's next generation capability computing platform. A design goal of the Zia platform is to achieve a 6x to 8x performance improvement over the NNSA's Purple platform. Sandia and Los Alamos have formed an alliance, ACES, which has the responsibility of delivering and deploying the Zia platform. Achieving this level of performance improvement over Purple will require a multi-core architecture capable of scaling not only within a node, but also able to take full advantage of the high-speed interconnect and ensure that all processes (cores) within a node have sufficient memory bandwidth.

Current performance analysis efforts include quantifying the performance of Zia's 6x application suite on candidate processors for the Zia platform. This effort will evolve to a full system analysis that takes into account MPI performance and other inefficiencies that can affect platform scaling.

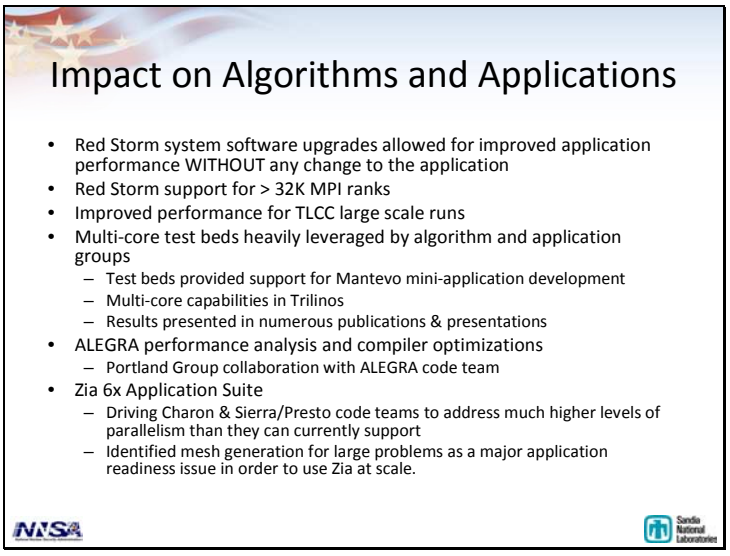
Slide 18: Candidate processors for Zia



This slide summarizes early application performance results for processors that are expected to be bid for Zia. The AMD Magny-Cours is not yet commercially available, so it is desirable to develop a predictive model of AMD's Istanbul processor for the Zia 6x application suite. The Magny-Cours is two Istanbul die in a single package, interconnected by Hypertransport. As such, a current 2 socket Istanbul workstation is a good surrogate for a single Mangy-Cours socket. However, the Magny-Cours will use DDR3 memory instead of DDR2, and the CPU clock rate will be lower than that of the Istanbul.

In this study, the performance for the Zia 6x applications is being measured on an Istanbul workstation using different memory subsystem clock rates. Regression analysis allows for a prediction of the performance for each application at a higher clock rate for the memory subsystem. This same analysis can be used to estimate the fraction of application time that is memory bound, and hence the fraction of time that is computationally bound. Knowing these fractions, it is possible to estimate Magny-Cours performance using a higher memory clock rate and a lower CPU clock rate.

Slide 19: Impact on algorithms and applications



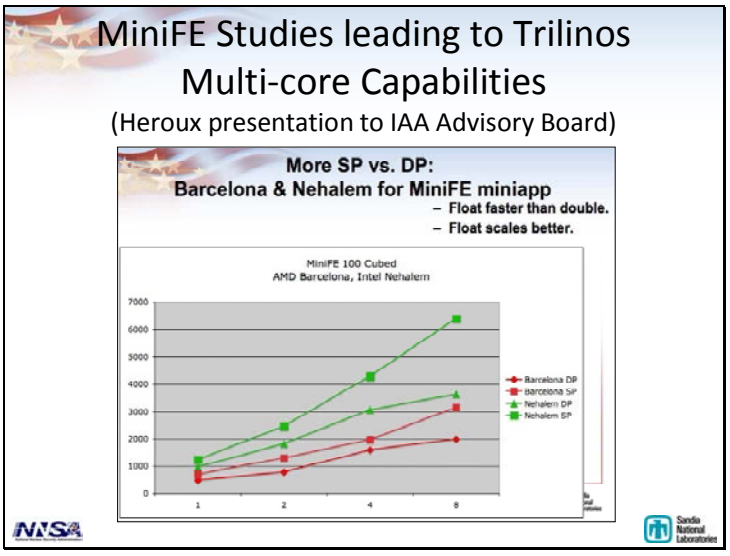
As was stated in an earlier slide, much of the system software efforts are transparent to the applications and analysts, but have an impact in improved performance and quicker turn around times.

There has been an excellent collaboration between the architecture, system software, algorithms and application teams in the effort to ensure a seamless transition to multi-core for the applications and analysts. The primary strategy for the architecture and performance analysis teams has been to work closely with the Trilinos group, as Trilinos is leveraged by numerous applications internal and external to Sandia.

In addition, the ALEGRA team has been collaborating directly with the Portland Group (PGI) compiler developers. PGI has been profiling and identifying areas of improvement for key ALEGRA kernels. This has been a two-way collaboration, in that performance issues and bugs with PGI's compiler are being feed back to their engineering team and being incorporated in future releases.

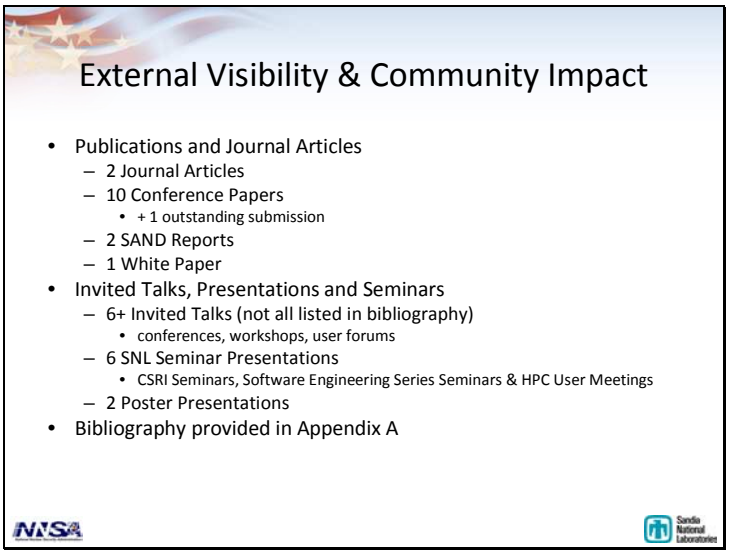
In defining the Zia 6x application suite problem sets, and collecting baseline data for the Purple platform, it has become evident that a major limiting factor in transitioning codes to Zia will be the ability to generate appropriate meshes of a size that will fill the Zia platform.

Slide 20: MiniFE studies leading to Trilinos multi-core capabilities



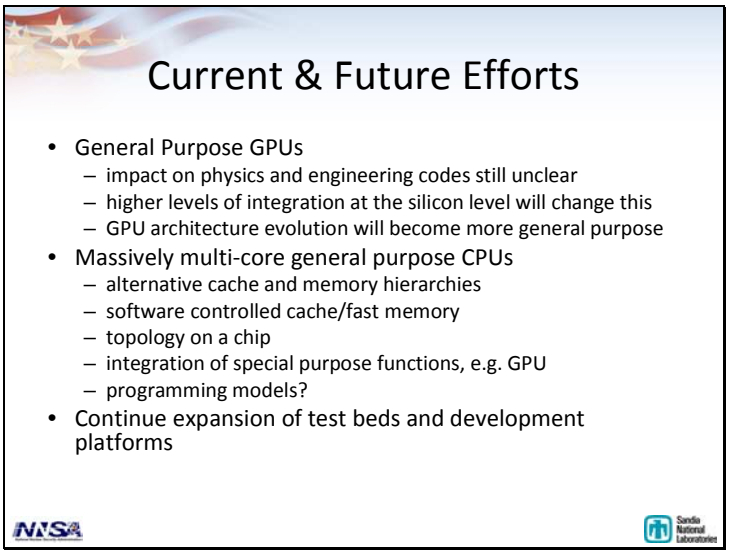
This study is an example of the collaboration between the architecture and algorithm groups. This analysis is looking at the performance difference between double-precision and single-precision floating-point operations.

Slide 21: External visibility & community impact



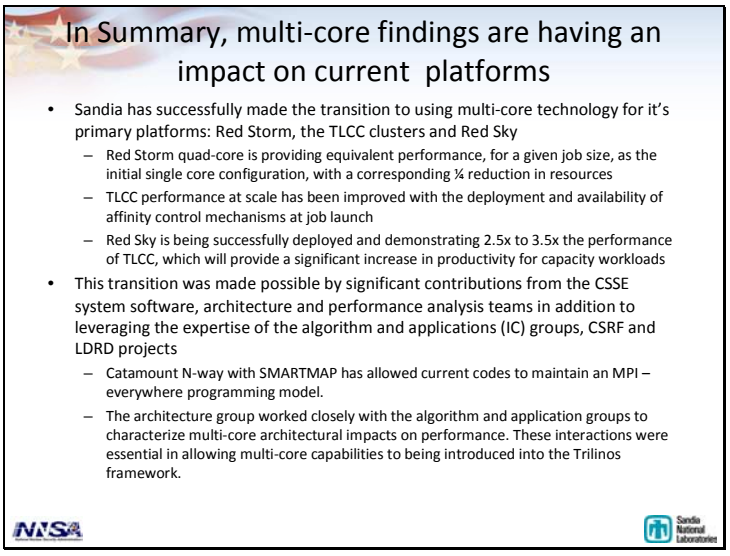
This slide summarizes the impact made in the form of publications and community interactions. All are documented as a bibliography in Appendix A.

Slide 22: Current and future efforts

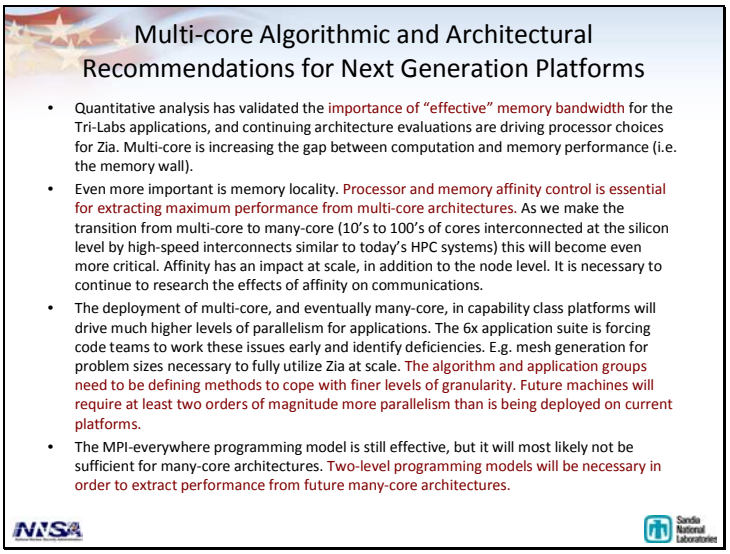


This effort does not end with the completion of this milestone. In fact, it's most likely only a beginning. The most challenging transition will be to massively multi-core processors which incorporate finer levels of granularity on chip and begin to introduce many of the architectural issues associated with full scale machines of only a few years ago.

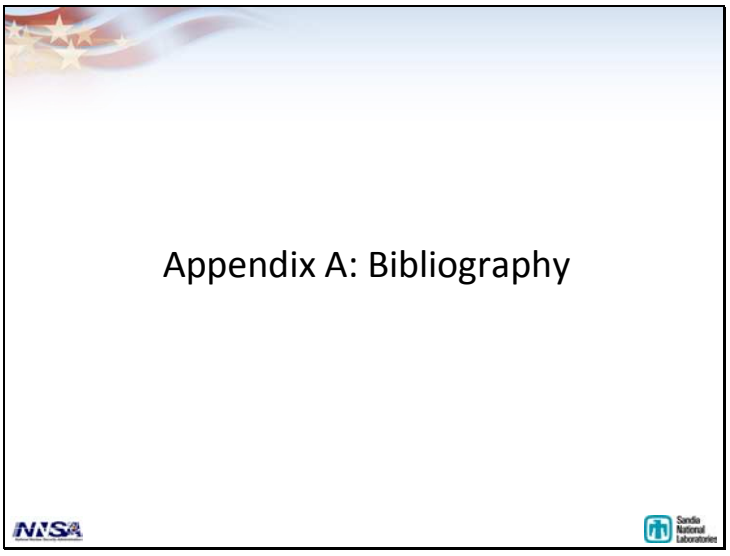
Slide 23: Summary



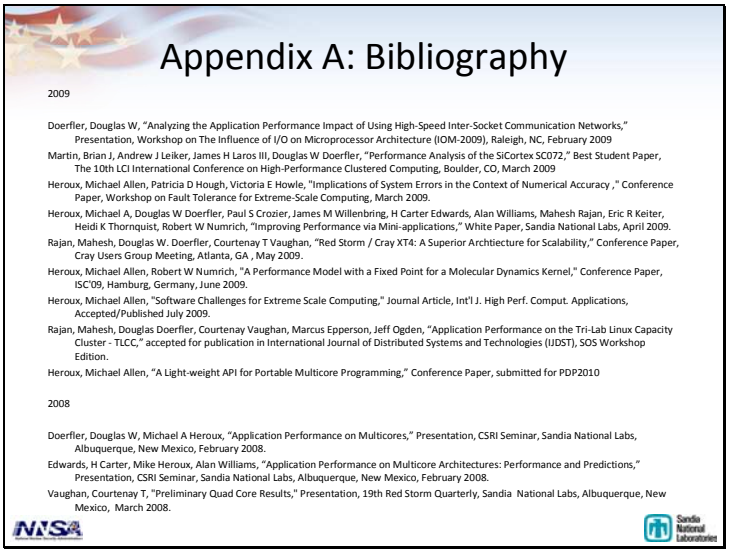
Slide 24: Multi-core algorithmic and architectural recommendations for NGS



Slide 25: Appendix A: bibliography



Slide 26: Appendix A: bibliography, cont'd



Appendix A: Bibliography

2009

Doerfler, Douglas W, "Analyzing the Application Performance Impact of Using High-Speed Inter-Socket Communication Networks," Presentation, Workshop on The Influence of I/O on Microprocessor Architecture (IOM-2009), Raleigh, NC, February 2009

Martin, Brian J, Andrew J Leiker, James H Laros III, Douglas W Doerfler, "Performance Analysis of the SiCortex SC072," Best Student Paper, The 10th LCI International Conference on High-Performance Clustered Computing, Boulder, CO, March 2009

Heroux, Michael Allen, Patricia D Hough, Victoria E Howle, "Implications of System Errors in the Context of Numerical Accuracy ," Conference Paper, Workshop on Fault Tolerance for Extreme-Scale Computing, March 2009.

Heroux, Michael A, Douglas W Doerfler, Paul S Crozier, James M Willenbring, H Carter Edwards, Alan Williams, Mahesh Rajan, Eric R Keiter, Heidi K Thornquist, Robert W Numrich, "Improving Performance via Mini-applications," White Paper, Sandia National Labs, April 2009.

Rajan, Mahesh, Douglas W. Doerfler, Courtenay T Vaughan, "Red Storm / Cray XT4: A Superior Archtiecture for Scalability," Conference Paper, Cray Users Group Meeting, Atlanta, GA , May 2009.

Heroux, Michael Allen, Robert W Numrich, "A Performance Model with a Fixed Point for a Molecular Dynamics Kernel," Conference Paper, ISC'09, Hamburg, Germany, June 2009.

Heroux, Michael Allen, "Software Challenges for Extreme Scale Computing," Journal Article, Int'l J. High Perf. Comput. Applications, Accepted/Published July 2009.

Rajan, Mahesh, Douglas Doerfler, Courtenay Vaughan, Marcus Epperson, Jeff Ogden, "Application Performance on the Tri-Lab Linux Capacity Cluster - TLCC," accepted for publication in International Journal of Distributed Systems and Technologies (IJDST), SOS Workshop Edition.

Heroux, Michael Allen, "A Light-weight API for Portable Multicore Programming," Conference Paper, submitted for PDP2010

2008

Doerfler, Douglas W, Michael A Heroux, "Application Performance on Multicores," Presentation, CSRI Seminar, Sandia National Labs, Albuquerque, New Mexico, February 2008.

Edwards, H Carter, Mike Heroux, Alan Williams, "Application Performance on Multicore Architectures: Performance and Predictions," Presentation, CSRI Seminar, Sandia National Labs, Albuquerque, New Mexico, February 2008.

Vaughan, Courtenay T, "Preliminary Quad Core Results," Presentation, 19th Red Storm Quarterly, Sandia National Labs, Albuquerque, New Mexico, March 2008.

Slide 27: Appendix A: bibliography, cont'd

Heroux, Michael Allen, "Challenges in Programming Next-generation Parallel Computer Systems," Presentation, SIAM Parallel Processing Conference, March 2008.

Rajan, Mahesh, Douglas Doerfler, Courtenay Vaughan, "A Preliminary Evaluation of Quad-Core Processors for Sandia Applications," Poster, HPCSW Workshop/Conference, Denver, CO, April 2008.

Vaughan, Courtenay T, "Preliminary Quad Core Results for Cray XT4 Systems," Poster, HPCSW Workshop/Conference, Denver, CO, April 2008.

Doerfler, Douglas, Kevin Pedretti, Mahesh Rajan, Carter Edwards, Courtenay Vaughan, Mike Heroux, "MPI Task Placement on Multicores," Presentation, CSRI Software Engineering Seminar Series, Sandia Labs, Albuquerque, New Mexico, April 2008.

Brightwell, Ron, "Application and Operating System Software Challenges in the Multi-core Era," Presentation, CSRI Software Engineering Seminar Series, Sandia Labs, Albuquerque, New Mexico, April 2008.

Rajan, Mahesh, Courtenay T Vaughan, Robert W Leland, Douglas W Doerfler, Robert E Benner, Jr., "Investigating the balance between capacity and capability workloads across large scale computing platforms," Conference Paper, 9th LCI International Conference on High-Performance Computing, Urbana, IL, April 2008.

Pedretti Kevin, "TLBs, DRAMs, and Other Scary Things, and Impact of Multi-core," Presentation, CSRI Software Engineering Seminar Series, Sandia National Laboratories, Albuquerque, New Mexico, April 2008.

Kelly, Suzanne M, John P VanDyke, Courtenay T Vaughan, "Catamount N-Way (CNW): An Implementation of the Catamount Light Weight Kernel Supporting N-cores Version 2.0," SAND Report, Sandia National Laboratories, Albuquerque, NM, June 2008

Heroux, Michael Allen, "Design Issues for Numerical Libraries on Scalable Multicore," Conference Paper, SciDAC 2008 Conference, Seattle, WA, July 2008.

Heroux, Michael Allen, "Scalable Algorithms for 1M Cores: What Might and Might Not Work, and Why," Presentation, Simulating the Future: 1M Cores and Beyond, September 2008.

Doerfler, Douglas, Courtenay Vaughan, "Adapting Codes for a Heterogeneous Multi-Core Red Storm," Presentation, Los Alamos Computer Science Symposium, Santa Fe, NM, Oct 2008.

Vaughan Courtenay T, "Level II ASC Milestone 3159, Red Storm 284 TeraFLOPS Upgrade – Final Report," SAND Report, Sandia National Laboratories, Albuquerque, NM, December 2008.

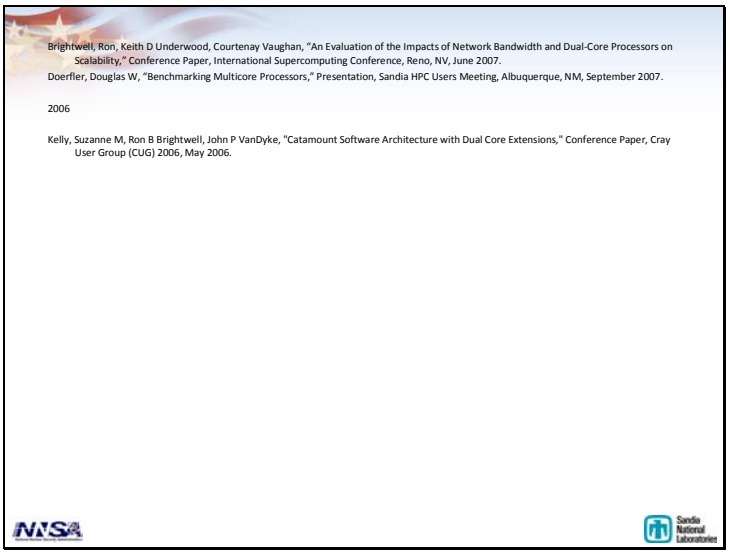
2007

Van Dyke, John P, Courtenay T Vaughan, Suzanne M Kelly, "Extending Catamount for Multi-Core Processors," Conference Paper, Cray User Group (CUG 07), May 2007.

Doerfler, Douglas, David Hensinger, Brent Leback, Douglas Miles, "Tuning C++ Applications for the Latest Generation x64 Processors with PGI Compilers and Tools," Conference Paper, Cray User Group (CUG) 2007, May 2007.

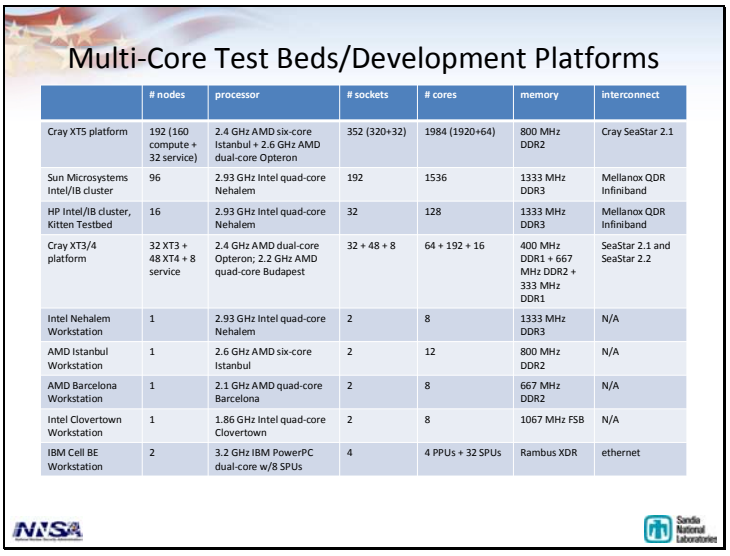
Heroux, Michael A, "Some Thoughts on Multicore," Conference Paper, Microsoft Manycore Workshop, Seattle, WA, June 2007.

Slide 28: Appendix A: bibliography, cont’d



Slide 29: Supporting slides

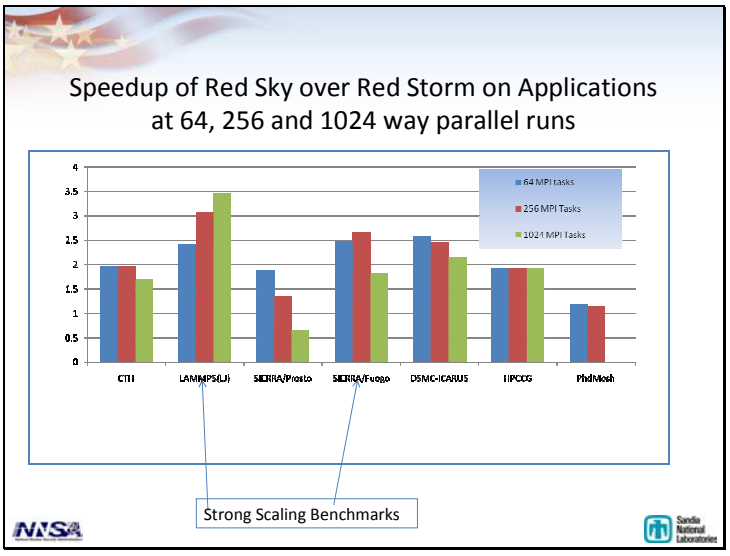
Slide 30: Multi-core test beds/development platforms



	# nodes	processor	# sockets	# cores	memory	interconnect
Cray XT5 platform	192 (160 compute + 32 service)	2.4 GHz AMD six-core Istanbul + 2.6 GHz AMD dual-core Opteron	352 (320+32)	1984 (1920+64)	800 MHz DDR2	Cray SeaStar 2.1
Sun Microsystems Intel/IB cluster	96	2.93 GHz Intel quad-core Nehalem	192	1536	1333 MHz DDR3	Mellanox QDR Infiniband
HP Intel/IB cluster, Kitten Testbed	16	2.93 GHz Intel quad-core Nehalem	32	128	1333 MHz DDR3	Mellanox QDR Infiniband
Cray XT3/4 platform	32 XT3 + 48 XT4 + 8 service	2.4 GHz AMD dual-core Opteron; 2.2 GHz AMD quad-core Budapest	32 + 48 + 8	64 + 192 + 16	400 MHz DDR1 + 667 MHz DDR2 + 333 MHz DDR1	SeaStar 2.1 and SeaStar 2.2
Intel Nehalem Workstation	1	2.93 GHz Intel quad-core Nehalem	2	8	1333 MHz DDR3	N/A
AMD Istanbul Workstation	1	2.6 GHz AMD six-core Istanbul	2	12	800 MHz DDR2	N/A
AMD Barcelona Workstation	1	2.1 GHz AMD quad-core Barcelona	2	8	667 MHz DDR2	N/A
Intel Clovertown Workstation	1	1.86 GHz Intel quad-core Clovertown	2	8	1067 MHz FSB	N/A
IBM Cell BE Workstation	2	3.2 GHz IBM PowerPC dual-core w/8 SPUs	4	4 PPUs + 32 SPUs	Rambus XDR	ethernet

A summary of test beds and development platforms that were acquired or leveraged as a part of this milestone. This is in addition to the production platforms of Red Storm, TLCC and Red Sky.

Slide 31: Speedup of Red Sky over Red Storm



This slide summarizes some early application results on Red Sky, relative to Red Storm. These early results help to validate the architectural choices made during the Red Sky acquisition. The performance of Red Sky is favorable to the Red Storm platform up to 1024 MPI tasks. SIERRA/Presto performance is an exception. For this application, the richer network and better MPI messaging performance of Red Storm contribute to better performance at scales above 256 MPI tasks.

Distribution list:

MS	Org.	Name	Copies
0807	9326	M. Rajan	1
0899	9536	Technical Library (electronic copy)	1
0823	9326	D. Pavlakos	1
1316	1412	D. Rintoul	1
1318	1414	K. Alvin	1
1319	1422	J. Ang	1
1319	1422	D. Barnette	2
1319	1422	R. Benner	1
1319	1423	R. Brightwell	1
1319	1422	D. Doerfler	1
1319	1423	S. Kelly	1
1319	1423	R. Oldfield	1
1319	1423	C. Vaughan	1
1320	1416	S. Collins	1
1320	1416	M. Heroux	1
1322	1420	S. Dosanjh	1
1323	1424	D. Rogers	1

Sandia National Laboratories