## SIAM Workshop on Exascale Applied Mathematics Challenges and Opportunities (EX14)

July 6, 2014
The Palmer House, Chicago, Illinois, USA
Co-located with the SIAM Annual Meeting July 7-11, 2014

## EX14 Program

**Sunday July 6, 2014**

8.00 - 10.15: Session 1

> 8:00 – 8:45: Invited Talk: What does Titan tell us about preparing for exascale supercomputers? *Jack Wells*, abstract, slides
> 8:45 – 9:30: Invited Talk: How the End of Moore's Law Scaling is Changing the Machines You Use, the Way You Code, and the Algorithms You Use, *Steve Oberlin*, abstract, slides
> 9:30 – 10:15: Invited Talk: Fault Tolerance Techniques for Sparse Matrix Methods, *Simon McIntosh-Smith*, abstract, slides

10.15-10.45: Break

10.45-noon: Session 2

> 10:45 – 11:30: Invited Talk: Time to Start over? Software for Exascale, *William Gropp*, abstract, slides
> 11:30 – noon: Resilient Solvers for Partial Differential Equations, *Miroslav Stoyanov and Eirik Endeve*, abstract, slides

Noon-1.25: Lunch

1.25-3.30: Session 3

> 1:25 – 1:50: Towards a Probabilistic Approach to Extreme-Scale Simulations under Uncertainty and System Faults, *Francesco Rizzi, Khachik Sargsyan, Karla Morris, Omar Knio, Paul Mycek, Olivier Le Maitre, Cosmin Safta, Habib Najm and Bert Debusschere*
> 1:50 – 2:15: Preliminary investigations on resilient parallel numerical linear algebra solvers, *Emmanuel Agullo, Luc Giraud, Pablo Salas Medina and Mawussi Zounon*, abstract, slides

2:15 – 2:40: Realizing Exascale Performance for Uncertainty Quantification, *Eric Phipps, H. Carter Edwards, Jonathan Hu and Clayton Webster*, abstract, slides
2:40 – 3:05: Driving Improvements to Algebraic Multigrid Through Performance Modeling, *Hormozd Gahvari, William Gropp, Kirk E. Jordan, Martin Schulz and Ulrike Meier Yang*, abstract, slides
3:05 – 3:30: A Reproducible Accurate Summation Algorithm for High-Performance Computing, *Sylvain Collange, David Defour, Stef Graillat and Roman Iakymchuk*, abstract, slides

3.30-3.55: Break

3.55-6.00: Session 4
3:55 – 4:20: Massive Asynchronous Parallelization of Sparse Matrix Factorizations, *Edmond Chow*, abstract, slides
4:20 – 4:45: Combinatorial Mathematics and Algorithms at Exascale: Challenges and Promising Directions, *Assefaw Gebremedhin and Alex Pothen*, abstract, slides
4:45 – 5:10: Sustained Petascale Performance of Seismic Simulations with SeisSol, *Michael Bader, Alexander Breuer, Alexander Heinecke, Sebastian Rettenberger, Alice-Agnes Gabriel and Christian Pelties*, abstract, slides
5:10 – 5:35: AmgX: Scalability and Performance on Massively Parallel Platforms, *Maxim Naumov, Marat Arsaev, Patrice Castonguay, Jonathan Cohen, Julien Demouth, Joe Eaton, Simon Layton, Nikolay Markovskiy, Nikolai Sakharnykh, Robert Strzodka and Zhenhai Zhu*, abstract, slides
5:35 – 6:00: The Graph BLAS effort and its implications for Exascale, *David Bader, Aydin Buluc, John Gilbert, Joseph Gonzalez, Jeremy Kepner and Timothy Mattson*, abstract, slides

* Underlined author was the speaker.

SIAM and the Workshop Organizing Committee wish to extend their thanks and appreciation to the U.S. National Science Foundation for its support of the Workshop.

SIAM and the Workshop Organizing Committee also wish to extend their thanks and appreciation to XSEDE for its partial support of the Workshop.

# Invited Talk Abstracts

What does Titan tell us about preparing for exascale supercomputers?, *Jack Wells*, Oak Ridge National Laboratory

Abstract: Significant advances in computational science and engineering have occurred over the past half-decade with the appearance of supercomputers with petascale performance capabilities, and beyond. And technology developments are occurring in addition to traditional CPU-based architectures in response to the growing energy requirements of these architectures, including many-core processors, graphical processing units (GPU), and related technologies. However, greater innovation in energy-efficient compute and data infrastructure is required to accelerate scientific breakthroughs and achieve mission-critical results in the exascale era. This talk will focus on the scientific drivers for exascale compute and data infrastructure and some of the challenges that we must overcome along the way. We will ground this discussion in lessons learned in deploying and operating Titan, the Department of Energy's Titan Cray XK7 supercomputer managed by the Oak Ridge Leadership Computing Facility, and in preparing applications to move from conventional CPU architectures to a hybrid, accelerated architectures. We will discuss implications for the research community as we prepare for exascale computational science and engineering within the next decade.

How the End of Moore's Law Scaling is Changing the Machines You Use, the Way You Code, and the Algorithms You Use, *Steve Oberlin*, NVIDIA

Abstract: The end of voltage scaling in CMOS is already upon us, limiting the generational bounty we've enjoyed from Moore's Law for decades, and soon (perhaps only two more silicon process shrinks, a mere half-dozen years) even transistor density increases will stall. While the massive economic forces behind the information technology market will almost certainly cause the creation of a successor to CMOS that will enable exponential technology growth to continue in the future, we are nonetheless experiencing the largest hiccup in that growth since the invention of the integrated circuit and it has already fundamentally altered the architectural strategy of processor and system manufacturers to emphasize parallelism over single thread speed. This talk will illuminate "the new normal" of CMOS technology and focus on the now-critical architecture response, describe the attributes of parallel architectures that are coming to underly every computing platform from supercomputers to mobile devices, and the implications for programming systems and algorithms.

Fault Tolerance Techniques for Sparse Matrix Methods, *Simon McIntosh-Smith*, University of Bristol

Abstract: High-performance computing systems continue to increase in size in the quest for ever higher performance. The resulting increased electronic component count, coupled with the decrease in feature sizes of the silicon manufacturing processes used to build these components, will result in future Exascale systems being more susceptible to soft errors than current HPC systems. Through the use of techniques such as hardware-based error-correcting codes (ECC) and checkpoint-restart, many of these faults can be mitigated, but at the cost of increased hardware overhead, run-time, and energy consumption that can be as much as 10–20%. For extreme scale systems, these overheads will represent megawatts of power consumption and millions of dollars of additional hardware cost, which could potentially be avoided with more sophisticated fault-tolerance techniques.

In this talk we present new software-based fault tolerance techniques that can be applied to one of the most important classes of software in HPC: sparse matrix solvers. Our new techniques enable us to move most of the error detection and correction burden into higher level software, exploiting knowledge of sparse matrix structures in such a way as to improve the performance, energy efficiency and fault tolerance of the overall solution.

Time to Start over? Software for Exascale, *Bill Gropp*, University of Illinois

Abstract: Exascale systems face many challenges in their design and implementation, and these are expected to impact the software for applications and for the system itself. This talk discusses the implications for programming systems and software of the major hardware issues, including the needs to limit data motion and power, handling performance irregularity and specialized processing elements, and concurrency and latency. Major Directions for software include simple evolution, addition to existing systems (such as MPI + X), and an all new approach. Pros and cons for each will be discussed, along with the merits of a blended strategy.

**Contributed Talk Abstracts**

Resilient Solvers for Partial Differential Equations, *Miroslav Stoyanov and Eirik Endeve*
Abstract: The exponential growth of computational power of the extreme scale machines over the past few decades has led to a corresponding decrease in reliability. Hardware faults are now the norm rather than the exception and as the supercomputers march towards the exaflop milestone, the occurrence of such faults is expected to intensify. In order to address this challenge and to ensure the reliability of the computed results, extensive measures must be taken at the level of hardware, software and mathematical algorithms. Most hardware is build with various types of error detection and correction, however, the ever increasing number of components in modern supercomputers offers unprecedented opportunity for failure. Due to constraints on cost and power consumption, it is infeasible to create hardware so reliable that it can guarantee no faults over millions of components and billions of CPU hours. Thus, extreme scale software must be build with the ability to withstand a wide variety of possible (and likely) hardware faults.
Our research focuses on the mathematical challenges presented by the silent hardware faults that can perturb the result of computations in an inconspicuous way (also called soft-transient faults), namely we study the new algorithms that are needed in order to detect, reject and correct silent faults.

Towards a Probabilistic Approach to Extreme-Scale Simulations under Uncertainty and System Faults, *Francesco Rizzi, Khachik Sargsyan, Karla Morris, Omar Knio, Paul Mycek, Olivier Le Maitre, Cosmin Safta, Habib Najm and Bert Debusschere*
Abstract: We present a novel approach for solving PDEs, using a probabilistic representation of uncertainty in the PDE solution due to incomplete convergence and the effect of system faults. Using domain decomposition, the problem is reduced to solving the PDE on subdomains with uncertain boundary conditions. An iterative approach is then applied to solve this problem in a resilient and scalable way, using subdomain computations for sampled values of the subdomain boundary conditions. Preliminary tests using a 1D elliptic problem show promising convergence results, even under synthetically-generated bit-flips.

Preliminary investigations on resilient parallel numerical linear algebra solvers, *Emmanuel Agullo, Luc Giraud, Pablo Salas Medina and Mawussi Zounon*
Abstract: The advent of extreme scale machines will require the use of parallel resources at an unprecedented scale, probably leading to a high rate of hardware faults. Handling fully these faults at the computer system level may have a prohibitive cost. High performance computing applications that aim at exploiting all these resources will thus need to be resilient, i.e., be able to compute a correct solution in presence of core crashes. In this work, we investigate possible remedies in the framework of numerical linear algebra problems such as the solution of linear systems or eigen-problems that are the inner most numerical kernels in many scientific and engineering applications and also ones of the most time consuming parts. More precisely, we present recovery techniques followed by

restarting strategies. In the framework of Krylov subspace linear solvers the lost entries of the iterate are interpolated using the available entries on the still alive cores to define a new initial guess before restarting the Krylov method. In particular, we consider two interpolation policies that preserve key numerical properties of well-known linear solvers, namely the monotony decrease of the A-norm of the error of the conjugate gradient or the residual norm decrease of GMRES. We extend these interpolation ideas in the context of some state of the art eigensolvers where these recovery approaches are applied to reconstruct a meaningful search space for restarting. We assess the impact of the recovery method, the fault rate and the number of processors on the robustness of the resulting numerical linear solvers. Finally, we will expose some preliminary investigations in the context of soft-error detections for the conjugate gradient implemented in the Parallel Ocean Program of the Community Earth System Model in the framework of the G8-ECS project.

Realizing Exascale Performance for Uncertainty Quantification, *Eric Phipps, H. Carter Edwards, Jonathan Hu and Clayton Webster*

Abstract: Exascale computing promises to address many scientific and engineering problems of national interest by facilitating computational simulation of physical phenomena at tremendous new levels of accuracy, fidelity, and scale, as well as unprecedented capabilities for high-level analysis such as uncertainty quantification for today's petascale computational simulations. However realizing these uncertainty quantification capabilities on exascale architectures will be challenging due to the mismatch between traditional methods of implementing uncertainty quantification algorithms in computational simulations and likely exascale architectures. In this paper we propose embedded uncertainty propagation approaches to overcome these challenges by reducing memory usage, improving data motion, and exposing greater levels of fine-grained parallelism. We highlight recent progress on developing these ideas as well as further mathematical, computational, and software challenges that must be overcome.

Driving Improvements to Algebraic Multigrid Through Performance Modeling, *Hormozd Gahvari, William Gropp, Kirk E. Jordan, Martin Schulz and Ulrike Meier Yang*

Abstract: Multilevel solvers, such as Algebraic Multigrid (AMG), are expected to play a big role on the massively parallel machines of the future, owing to superior computational complexity. AMG in particular allows for the fast solution of unstructured grid problems. It is not without its challenges, however. Since there is no assumption about the structure of the problem, the hierarchy of grids needs to be explicitly set up. The stencils of the resulting coarse grid problems increase in size, leading to increased interprocessor communication and posing a challenge to scalability.

On emerging parallel machines, we have found that this increased communication combined with the increasingly multicore nature of the underlying architecture has in fact led to an observable degradation in scalability. The urgent need for a remedy is made clear by projections of exascale systems having hundreds or even thousands of cores per node. The space of possible remedies is, however, wide, spanning changes in the algorithm and changes in the programming model. Thus, there is also a need for a means of readily arriving at an effective remedy without excessive time and effort on the part of the end users.

In this talk, we highlight the importance of an approach based on performance modeling, which has helped us in all phases of our work. We developed a performance model for the AMG solve cycle that starts with basic network parameters and then builds on those to consider network contention, issues with multicore nodes, simultaneous multithreading, and the mixture of shared memory and distributed memory programming models. This enabled us to pinpoint that it was increased communication on coarse grids combined with stresses put on the network by multicore nodes that were causing the performance degradation we were seeing. We then used this model to help us survey a large space of algorithmic innovations in multigrid methods, designed with massive parallelism in mind but largely set aside when drastic increases in single-core speeds were occurring, and selected which were the most promising for further investigation. The end result was a data redistribution algorithm that gathers data onto a subset of the machine, trading communication for computation. This algorithm, which was guided by the performance model at runtime, enabled us to realize speedups when running AMG on modern parallel machines.

A Reproducible Accurate Summation Algorithm for High-Performance Computing, *Sylvain Collange, David Defour, Stef Graillat and Roman Iakymchuk*

Abstract: On modern multi-core, many-core, and heterogeneous architectures, floating-point computations, especially reductions, may become non-deterministic and thus non-reproducible mainly due to non-associativity of floating-point operations. We introduce a solution to compute deterministic sums of floating-point numbers efficiently and with the best possible accuracy. Our multi-level algorithm consists of two main stages: a filtering stage that uses fast vectorized floating-point expansions; an accumulation stage based on superaccumulators in a high-radix carry-save representation. We present implementations on recent Intel desktop and server processors, on Intel Xeon Phi accelerator, and on AMD and NVIDIA GPUs. We show that the numerical reproducibility and bit-perfect accuracy can be achieved at no additional cost for large sums that have dynamic ranges of up to 90 orders of magnitude by leveraging arithmetic units that are left underused by standard reduction algorithms.

Massive Asynchronous Parallelization of Sparse Matrix Factorizations, *Edmond Chow*

Abstract: A vital challenge for exascale computing is the massive concurrency required in scientific and engineering algorithms in order to run efficiently on exascale architectures. We present a completely new approach for parallelizing certain sparse matrix factorizations, proposed for the regime where one has more cores than one knows how to utilize. Our approach is to express the matrix factorization as a large number of scalar bilinear equations, and then solving these equations via an asynchronous iterative method. The concurrency is very fine-grained, which is necessary for massive amounts of parallelism. The approach is driven by computer architectural considerations, rather than modifying the best existing algorithms, and is conceivable only in the case of massive amounts of parallelism within a node. The paradigm of transforming a problem into a more

parallel, iterative one could be applied in many situations. The matrix algorithms that benefit the most are those where an approximation is suitable, e.g., in preconditioning or when the matrix data has redundancies or noise. The paradigm, in effect, trades accuracy for concurrency. Our main example will be incomplete factorizations, but we will also discuss how the ideas extend to other sparse matrix problems such as matrix completions and the solution of sparse triangular systems.

Combinatorial Mathematics and Algorithms at Exascale: Challenges and Promising Directions, *Assefaw Gebremedhin and Alex Pothen*
Abstract: Developing numerical algorithms and implementations that fully harness the potential of many-core and accelerated architectures brings on daunting challenges at different levels, but these challenges are so much more exasperated in the case of combinatorial algorithms. Yet combinatorial models and algorithms play a crucial and growing role both as ``embedded'' and ``forefront"" enablers in advanced computing. We discuss a selection of contexts in which combinatorial models arise, outline the challenges, discuss a sample of approaches we found promising in recent activities, and point out potentially fruitful avenues for future exploration.

Sustained Petascale Performance of Seismic Simulations with SeisSol, *Michael Bader, Alexander Breuer, Alexander Heinecke, Sebastian Rettenberger, Alice-Agnes Gabriel and Christian Pelties*
Abstract: We present our current work on optimizing SeisSol, a software package for high-order fully-adaptive seismic simulations, for current petascale supercomputers and manycore platforms. SeisSol's ADER-DG approach has been demonstrated to be especially successful for simulations that need to account for highly complicated geometrical structures, with applications ranging from seismology via exploration industry to earthquake physics.

It's high computational intensity and comparably simple communication pattern make it an attractive candidate for simulations on exscale and beyond. We will present recent results on the SuperMUC machine (1.08 sustained performance, 20% effective peak efficiency) and on HPC platforms using Intel Xeon Phi.

AmgX: Scalability and Performance on Massively Parallel Platforms, *Maxim Naumov, Marat Arsaev, Patrice Castonguay, Jonathan Cohen, Julien Demouth, Joe Eaton, Simon Layton, Nikolay Markovskiy, Nikolai Sakharnykh, Robert Strzodka and Zhenhai Zhu*
Abstract: In this presentation we discuss the AmgX iterative solve framework, which includes implementation of two forms of Algebraic Multigrid (AMG), Krylov subspace methods, as well as parallel variants of typical smoothers and preconditioners.
The goal of AmgX is to fully utilize the hardware capabilities of large GPU-accelerated clusters, especially for solving elliptic PDEs. While parallel versions of many suitable numerical methods exist, it is challenging to find algorithms which fully exploit both fine-grained parallelism (“scale-in”) within a single node, as well as the coarse-grained parallelism (“scale-out”) between nodes. In particular, there are only a few choices when it

comes to fault tolerant numerical schemes that retain the robust and fast convergence properties of serial methods when scaling on massively parallel computer platforms.

The Graph BLAS effort and its implications for Exascale, *David Bader, Aydin Buluc, John Gilbert, Joseph Gonzalez, Jeremy Kepner and Timothy Mattson*

Abstract: The graph abstraction provides a natural way to represent relationships among complex fast-growing scientific data sets. Graph algorithms are hard to parallelize, making their performance suboptimal on high-performance architectures. Many graph computations, however, contain sufficient parallelism that can be uncovered by using the right primitives. This position paper lays out the opportunities of defining a standard for linear-algebraic primitives for graph algorithms. Specifically, we emphasize the implications of having a Graph BLAS standard for the accessibility and availability of high-performance graph algorithms in exascale architectures.

SIAM and the Workshop Organizing Committee wish to extend their thanks and appreciation to the U.S. National Science Foundation for its support of the Workshop.

SIAM and the Workshop Organizing Committee also wish to extend their thanks and appreciation to XSEDE for its partial support of the Workshop.