Final report for "Extreme-scale Algorithms and Solver Resilience"

PI: William D. Gropp, University of Illinois at Urbana-Champaign
DOE # DE-FG02-13ER26138/DE-SC0010049
Period Covered: June 15, 2013 – June 16, 2017
Date: June 30, 2017
Product Number DOE-UIUC-10049

Overview

This is a joint project with PIs at ORNL, SNL, UC Berkeley, and the Univ. of Tennessee. Our part of the project involves developing performance models for highly scalable algorithms and the development of latency tolerant iterative methods. During this project, we extended our performance models for Multigrid and conducted experiments with highly scalable variants of conjugate gradient methods that avoid blocking synchronization. In addition, we worked with the other members of the project on alternative techniques for resilience and reproducibility. We also presented an alternative approach for reproducible dot-products in parallel computations that performs almost as well as the conventional approach by separating the order of computation from the details of the decomposition of vectors across the processes.

Results

Evaluating and developing algorithms, particularly at extreme scale, requires an effective performance model. By effective, we mean one that is simple enough to use to guide development of algorithms but accurate to predict which algorithms will perform better than others. We have been developing such methods, including terms for memory contention on multicore nodes and (relatively simple and coarse) approximations for network contention. Recent work has looked at algebraic multigrid because (a) AMG is optimal for a large class of problems and (b) there are promising approaches for resilience to transient memory errors for AMG based on making use of values on the other grids and defining the appropriate interpolation or restriction operation. First steps were published in (Calhoun, Olson et al. 2015). Recent results were presented at HPDC this year (Calhoun, Snir et al. 2017).

We also looked at resilience and reproducibility by considering the problem as one of scheduling or graph labeling. Work in this area has looked at lightweight scheduling mechanisms for fine grain operations such as stencil application, including performance models to guide parameter selection in the scheduling mechanism. This provides the basis for an overdecomposition approach that we evaluated for providing a general and scalable approach to reproducibility across a range of process counts for parallel applications. A presentation on this was given at the Workshop on Batched, Reproducible, and Reduced Precision BLAS, held in May 2016 at the University of Tennessee, and is available online at http://wgropp.cs.illinois.edu/bib/talks/tdata/2016/fixed-schedule.pdf , which included results from an implementation restricted to distributions of data with the same number of items on each process. These results showed that the overdecomposition approach could provide bit-reproducible results without sacrificing performance, even in parallel computations.

We have results for several variants of conjugate gradient that rearranges the operations to provide better memory locality as well as permitting the use of nonblocking Allreduce. Runs on Blue Waters, the large Cray XE6 at Illinois, show better performance in PETSc than other versions of CG. A performance model is under development to explain the results. Unfortunately, none of the available systems, including the Blue Gene/Q system at ANL, supports an efficient nonblocking allreduce. We hope to provide evidence to the vendors that providing a good implementation will provide higher performance to important applications. These results were presented at SC16 (Eller and Gropp 2016).

Presentations

"Tying Iterative Solvers to Performance Models: Experiences with Algebraic Multigrid," Hormozd Gahvari, William Gropp, Kirk E. Jordon, Martin Schulz, Ulrike Meier Yang, presentation at Copper Mountain, April, 2014.

"Improving Scalability of Parallel Algebraic Multigrid," Amanda Bienz, Rob Falgout, William Gropp, Luke Olson, Jacob Schroder, presentation at Copper Mountain, April, 2014.

"Driving Improvements to Algebraic Multigrid Through Performance Modeling," Hormozd Gahvari, William Gropp, Kirk E. Jordon, Martin Schulz, Ulrike Meier Yang, presentation at the SIAM EX14 Workshop, July 2014

"Non-blocking conjugate gradient methods for extreme scale computing." P.R. Eller and W. Gropp, presentation at 17th Copper Mountain Conference on Multigrid Methods, Copper Mountain, CO, April, 2015.

"Non-blocking conjugate gradient methods for extreme scale computing." P.R. Eller and W. Gropp, poster presentation at SC15, Austin, TX, November, 2015.

"Reproducibility of Computations and Distributed Data Structures," William Gropp, presentation at Workshop on Batched, Reproducible, and Reduced Precision BLAS, May 18-19, 2016, Knoxville, TN.

Bibliography

Calhoun, J., et al. (2015). Towards a more fault resilient multigrid solver. Proceedings of the Symposium on High Performance Computing. Alexandria, Virginia, Society for Computer Simulation International: 1-8.

Calhoun, J., et al. (2017). Towards a More Complete Understanding of SDC Propagation. Proceedings of the 26th International Symposium on High-Performance Parallel and Distributed Computing. Washington, DC, USA, ACM: 131-142.

Eller, P. R. and W. Gropp (2016). Scalable non-blocking preconditioned conjugate gradient methods. Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. Salt Lake City, Utah, IEEE Press: 1-12.