6th European Conference on Computational Mechanics (ECCM 6)
7th European Conference on Computational Fluid Dynamics (ECFD 7)
11-15 June 2018, Glasgow, UK

Fault-tolerant parallel multigrid

M. Altenbernd*[1], D. Göddeke[2]

[1] University of Stuttgart, Allmandring 5b, 70569 Stuttgart,
mirco.altenbernd@ians.uni-stuttgart.de, http://www.ians.uni-stuttgart.de/cmcs/
[2] University of Stuttgart, Allmandring 5b, 70569 Stuttgart,
dominik.goeddeke@ians.uni-stuttgart.de, http://www.ians.uni-stuttgart.de/cmcs/

Keywords: *Fault-Tolerance, Resilience, Multigrid, High-Performance Computing, MPI*

Fault-tolerance is one of the major challenges towards extreme-scale computing. There is broad consensus that future leadership-class machines will exhibit a substantially reduced mean-time-between-failure (MTBF). The resilience challenge can best be summarised that at scale, faults and failures are likely to become the norm rather than the exception: Any simulation run will be compromised without inclusion of resilience techniques into the underlying software stack and system. Especially the rising number of cores, which is expected on the way to exascale computing, has a great impact.

We introduce an algorithm-based fault-tolerance scheme to detect and repair soft transient faults (SDC, bitflips) in multigrid solvers [2]: By applying the full approximation scheme (FAS) variant of multigrid to linear systems, we use invariants that enable fault detection and correction, and ultimately lead to a black-box protection of the smoothing stage. A statistical analysis for a wide range of prototypical problems demonstrates the efficiency of our approach, especially compared to full checksum protection. In particular, the overhead of our new method is negligible in the fault-free case, since we only employ readily available quantities. Furthermore the Full Approximation Scheme directly enables the application of a hierarchically compressed checkpointing to counteract node-loss scenarios [1]. Therefore we are developing, in cooperation with the University of Münster, a high-level C++ approach to manage local errors, asynchrony and faults in MPI applications which will integrate seamlessly with the upcoming MPI-ULFM [3] standard.

REFERENCES

[1] D. Göddeke, M. Altenbernd and D. Ribbrock, Fault-tolerant finite-element multigrid algorithms with hierarchically compressed asynchronous checkpointing. *Parallel Computing*, Vol. **49**, pp. 117–135, 2015.

[2] M. Altenbernd and D. Göddeke, Soft fault detection and correction for multigrid. *The International Journal of High Performance Computing Applications*, 2017.

[3] Bosilca, George and Bouteiller, Aurelien and Guermouche, Amina and Herault, Thomas and Robert, Yves and Sens, Pierre and Dongarra, Jack, Failure Detection and Propagation in HPC Systems. *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, pp. 72:1–27:11, 2016.