

Resilience versus Performance in Numerical Linear Algebra (REPEAL)

Abstract

Concepts for achieving fault tolerance are becoming more important in many application areas of computer systems because various factors (including the rapidly growing complexity of these systems and efforts to improve energy efficiency) lead to an increase of the frequency of various types of faults and failures in computer systems. This phenomenon can be observed in many types of computer systems, ranging from networked systems such as "smart grids" or "smart metering" over the Internet of things to high performance computing (HPC) systems. While there is already a long history of concepts for fault tolerance at the system level, it is currently an active area of research to investigate the potential of resilience and fault tolerance against various types of faults at the algorithmic level. This includes the analysis of the overhead and of the cost associated with resilience and fault tolerance at the algorithmic level. In the course of the REPEAL project we developed, analyzed and evaluated new fault tolerant algorithms for selected problems from numerical linear algebra. Our new algorithms yield results of comparable accuracy despite unreported faults during their execution. In particular, we investigated new algorithmic approaches for matrix multiplication, for summation and averaging in parallel or distributed systems, for large linear least squares problems and for large sparse linear systems. The problems considered are central building blocks for numerical simulations in many application areas, from classical models based on differential equations to modern data-based models. There are many types of faults which can influence the execution of an algorithm, ranging from the failure of hardware components over unreliable communication in parallel or distributed systems (e.g., leading to message loss) to spontaneous silent data corruption (SDC, also called "bit flips") in memory blocks. In the REPEAL project, we concentrated mostly on two types of faults: In the first half of the project the focus was on algorithmic resilience against SDC. In the second half of the project the focus was on parallel algorithms which are resilient against unreported node failures. The project team successfully developed and analyzed new resilient algorithms for these selected problem settings. Moreover, we also evaluated them experimentally based on prototypical implementations. This way, we could show that our new algorithms are significantly more resilient than the state-of-the-art and that the cost to be paid for this increased resilience (in terms of runtime overhead) is very low even in fault-free scenarios.

Scientific disciplines:

102023 - Supercomputing (60%) | 101014 - Numerical mathematics (25%) | 102025 - Distributed systems (15%)

Keywords:

fault tolerant matrix computations, algorithm-based fault tolerance, gossip-based aggregation

Principal Investigator: Wilfried Gansterer

Institution: University of Vienna



Further links about the involved persons and regarding the project you can find at

https://archiv.wtf.at/programmes/information_communication/ICT15-113