Mixed Precision Sparse Triangular Solves

Fritz Göbel[1] Tobias Ribizel[2] Thomas Grützmacher[3]

Many scientific applications require the solution of sparse systems of linear equations. In a lot of cases, Krylov subspace methods are the tool of choice. These iterative solvers naturally yield a lot of parallelism and are therefore well suited for the task. Most of the time, especially for ill-conditioned problems, the convergence behavior of these iterative solvers can be significantly improved by the use of a preconditioner. A popular and often very effective choice are incomplete factorization preconditioners like incomplete LU (ILU) or Cholesky (IC) factorizations approximating the exact factorization of the matrix on a given sparsity pattern. These preconditioners require the solution of two triangular systems per iteration. Since on modern hardware architectures, the solution of sparse systems of linear equation is a memory-bound problem with performance limited by memory bandwidth rather than compute power, reducing the memory traffic in the preconditioner application can significantly reduce its cost at the trade-off of some accuracy. However, since the preconditioner itself is already only a rough approximation to the inverse of the matrix this can be an attractive choice and yields a reduction in overall solver runtime in many cases. With the *Memory Accessor*, the open-source library Ginkgo provides a tool that easily allows us to decouple the arithmetic from the storage precision and therefore enables us to store our preconditioners in a lower precision format in main memory while still performing all arithmetic precisions in full IEEE double precision. This talk discusses results obtained with this decoupling of precision formats, either to reduce memory traffic during the preconditioner application or to extend the sparsity pattern of the approximate factorization while keeping the memory traffic roughly constant.

[1]Karlsruher Institut für Technologie, Steinbuch Center for Computing
fritz.goebel@kit.edu

[2]Karlsruher Institut für Technologie, Steinbuch Center for Computing
tobias.ribizel@kit.edu

[3]Karlsruher Institut für Technologie, Steinbuch Center for Computing
thomas.gruetzmacher@kit.edu