GPU Accelerated Algebraic Multigrid

Gundolf Haase[1] Manfred Liebmann[2] Gernot Plank[3]

The race for higher and higher clock frequencies hit a physical road block five years ago, forcing the big processor companies to focus on multi–core architectures instead. One major direction in semiconductor industry is the many–core GPU architecture of Nvidia and AMD with hundreds of simple scalar processors with only a small amount of on–chip memory and a very high bandwidth memory interface. The current generation of GPUs deliver 1TFLOPS peak in single precision arithmetic and 100 GFLOPS in double precision. The high performance scalar processors on the GPU are fed by a very high bandwidth memory interface to the on-board DRAM with up to 150GB/s throughput. This is more than one order of magnitude faster than on a typical CPU based server.

The easy–to–use to CUDA-interface for programming NVidia cards encouraged us to check whether many–core GPUs are useful in scientific computing for solving systems of equations with sparse unstructured matrices from 3D applications. First results on the NVIDIA Geforce 8800 GT showed an acceleration of 10 in comparison to a high-end CPU for the sparse matrix–vector product in single precision. Based on this matrix-vector product we accelerated our Algebraic Multigrid (AMG) solver also by a factor of 10 on the GPU. The recent tests on a pre–release of the NVIDIA GTX 280 confirmed the measured accelerations also for double precision arithmetics. This allows together with our modular parallelization toolbox a very fast and efficient parallel AMG solver on a rather cheap parallel computer based on GPUs from the shelf.

[1]University of Graz, Institute for Mathematics and Scientific Computing, Heinrichstr. 36, 8010 Graz, Austria,
gundolf.haase@uni-graz.at
[2]University of Graz, Institute for Mathematics and Scientific Computing,
manfred.liebmann@uni-graz.at
[3]Medical University of Graz, Institute of Biophysics,
gernot.plank@meduni-graz.at