GPU Computing: Übung #9

Abgabe am Donnerstag, 15. Januar 2014

Günther Schindler, Alexander Schnapp, Klaus Naumann

Inhaltsverzeichnis

Reading: High-Performance Code Generation for Stencil Computations on GPU Architectures

In this paper the method of overlapped tiling is presented in order to accelerate stencil codes implementatition on GPUs. After giving an short overview over the properties and challenges of GPUs and also stencil codes they explain the method of overlapped tiling in detail. The idea behind it is not to synchronize the so called halo regions of values that are needed by different parts of the grid, but simply redundandly recalculate them in order to avoid this traffic. By this they are also giving an tutorial how to implemnt it. In the last part of the paper they are analysing the performance of the method. The main outcome was that global memory access is the limiting factor for smaller time tile sizes, and computational overhead is the limiting factor for larger time tile size.
It is positiv that this can be applied well vor many diffrent GPUs and different stencils. Furthermore they are very concrete in telling how to actually apply and implement this method. So I really like the paper.