GPU-Programming with CUDA

M.Sc. Marius Brehler

Within the last decade, graphic processing units (GPUs) have evolved from specialized processors for computer graphics and video gaming into a general purpose computing architecture for high parallel codes, especially in the area of scientific computing. Researchers have demonstrated substantial speedups with GPU-accelerated codes in many different application domains. However, the possible performance benefits come with the challenge to master a programming model with literally thousands of simultaneously active threads.

This course teaches GPU-programming with CUDA from scratch. Prior to the GPU-related topics, the course starts with a brief introduction on programming in C. Topics in GPU-programming include CUDA basics, the hardware and software model of GPUs, and memory optimizations. The course further covers the usage of debugging and profiling analysis tools.
Furthermore, the course gives advices on how to use CUDA with MATLAB. We will briefly review the MATLAB Parallel Computing Toolbox, MEX files, and also the possibility to read/write MAT files within a C/C++ application.

As the course includes hands-on programming, the participants should bring their own laptops with an SSH (Secure Shell) client installed. Windows users are recommended to use PuTTY[1]. Access to a GPU will be provided by the University of L'Aquila.

---

[1] `https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html`