Performance and Toolchain of a Combined GPU/FPGA Desktop

Bruno da Silva, *Erasmus University College, Brussels*
An Braeken, *Erasmus University College, Brussels*
Erik H. D'Hollander, *Ghent University, Ghent, Belgium*
Abdellah Touhafi,*Erasmus University College,Brussels*
Jan G. Cornelis, *Free University of Brussels, Belgium*
Jan Lemeire, *Free University of Brussels, Belgium*

Low-power, high-performance computing nowadays relies on accelerator cards to speed up the calculations. Combining the power of GPUs with the flexibility of FPGAs enlarges the scope of problems that can be accelerated [2, 3]. We describe the performance analysis of a desktop equipped with a GPU Tesla 2050 and an FPGA Virtex-6 LX240T. First, the balance between the I/O and the raw peak performance is depicted using the roofline model [4]. Next, the performance of a number of image processing algorithms is measured and the results are mapped onto the roofline graph. This allows to compare the GPU and the FPGA and also to optimize the algorithms for both accelerators. A programming toolchain is implemented, consisting of OpenCL for the GPU and several High-Level Synthesis compilers for the FPGA. Our results show that the HLS compilers outperform handwritten code and offer a performance comparable to the GPU. In addition the FPGA compilers reduce the development time by an order of magnitude, at the expense of an increased resource consumption. The roofline model also shows that both accelerators are equally limited by the input/output bandwidth to the host. A well-tuned accelerator-based codesign, identifying the parallelism, the computation and data patterns of different classes of algorithms, will enable to maximize the performance of the combined GPU/FPGA system [1].

References

[1] Asanovic, K. et al. 2009. A view of the parallel computing landscape. Communications of the ACM. 52, 10 (2009), 56–67.

[2] Bauer, S. et al. 2010. FPGA-GPU architecture for kernel SVM pedestrian detection. Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on (2010), 61–68.

[3] Inta, R. et al. 2012. The "Chimera": An Off-The-Shelf CPU/GPGPU/FPGA Hybrid Computing Platform. International Journal of Reconfigurable Computing. January 2012, Article 2, 10 pages (2012).

[4] Williams, S. et al. 2009. Roofline: an insightful visual performance model for multicore architectures. Communications of the ACM. 52, 4 (2009), 65–76.