

HEIDELBERG UNIVERSITY
INSTITUTE FOR COMPUTER ENGINEERING (ZITI)

MASTER OF SCIENCE COMPUTER ENGINEERING
GPU COMPUTING

Exercise 6

gpucomp03

Benjamin Maier
Daniel Barley
Laura Nell

Due date January 12th, 09:00

6.1 Reading

Roofline: An Insightful Visual Performance Model for Multicore Architectures

To address the increasing variety in microprocessors, the Roofline model was introduced to offer cross-platform understandable performance guidelines using bound and bottleneck analysis. It ties together floating-point performance, operational intensity and memory performance in a 2D-graph. After explaining the function of the Roofline model, the paper demonstrates its utility on four diverse multicore computers, then optimizing four floating-point kernels taken from the Seven Dwarfs before clearing up some fallacies.

The Roofline model sets an upper bound on performance depending on the kernel's operational intensity, thus showing if performance of the kernel is compute-bound or memory-bound and, by adding ceilings, which optimizations may be necessary.

Finally, the ridge point is a better predictor for performance than clock rate or peak performance and indicates when a computer is imbalanced.

The Roofline model is still used today to check for a kernel's peak performance, although one has to especially consider to choose the right metric. Therefore, we accept the given paper and its content.

6.2 Reduction - CPU sequential version

6.3 Reduction - GPU parallel initial version

6.4 Reduction - GPU parallel optimized version

6.5 Reduction - Profiling

6.6 Willingness to present

Hereby, we declare our will to present the results presented in the former sections.