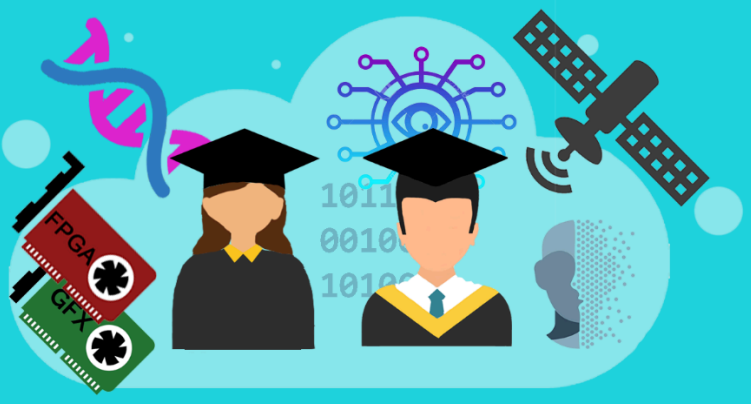
Diploma Thesis

Microproccessors and Digital Systems Laboratory



RTL-Informed GPU Performance and Power Modeling

Graphics Processing Units (GPUs) are becoming increasingly dominant in the accelerator landscape with their extensive use in data analytics and machine learning applications. Along with peak performance, energy efficiency—measured as performance per watt—has become a vital metric for assessing GPU designs. As a result, GPU architects need comprehensive tools to model both the performance and power consumption of modern GPUs.

Accel-Sim and Accelwattch are the cycle-accurate state-of-the-art tools to this end. Accel-Sim is a configurable simulation framework designed for detailed, validated GPU performance modeling, across different GPU generations. AccelWattch is a GPU power modeling tool, integrated with Accel-Sim that captures most aspects of contemporary GPU architectures such as separate dynamic, static and constant power modeling as well as complex power gating patterns. Vortex is an open-source, RISC-V-based soft GPU designed to support both GPGPU tasks, offering a scalable platform for a variety of General Purpose and Machine Learning applications. Implemented on FPGA, Vortex supports OpenCL, providing a fully functional software and hardware stack for GPU research and development.

A fundamental purpose of the above GPU simulation frameworks is to enable reliable Design Space Exploration . However, they are tuned to specific GPU architectures and remain accurate only when sweeping specific architecture parameters or modeling adjacent architecture generations. Moreover, specific components completely lack a power model, introducing errors that are amortized over the whole GPU. **The purpose of this thesis is to produce hybrid analytical/ML performance and power models** trained on FPGA measurements from a set of differently configured GPUs produced by Vortex. This would enable reliable Design Space Exploration of drastically different architectures.

PREREQUISITES:

Familiarity with: C++, Python, Bash.

Desirable qualifications: Verilog, OpenCL.

RELATED MATERIAL:

https://paragon.cs.northwestern.edu/papers/2021-MICRO-AccelWattch-Kandiah.pdf,
https://ieeexplore.ieee.org/document/9138922,
https://dl.acm.org/doi/pdf/10.1145/3466752.3480128

CONTACT INFORMATION:

Asst. Prof. Sotirios Xydis (sxydis@microlab.ntua.gr)

Konstantinos Iliakis (kiliakis@microlab.ntua.gr)

Panagiotis Eleftherakis (pelef@microlab.ntua.gr)