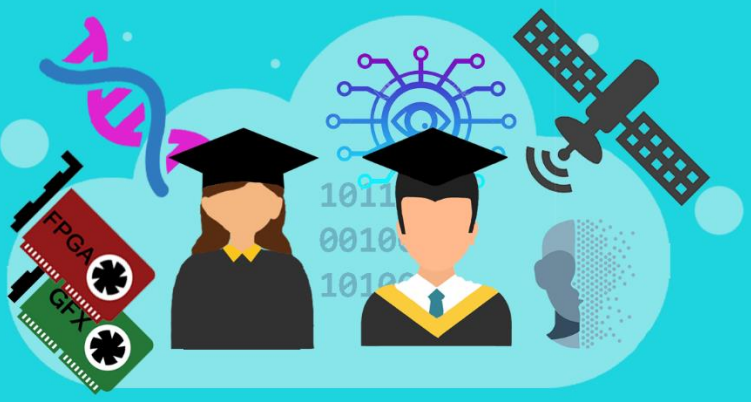
Diploma Thesis

Microproccessors and Digital Systems Laboratory



Exploring multi-application execution on integrated CPU-GPU heterogeneous many-core systems

In order to overcome the performance limitations, a recent trend in heterogeneous systems is to integrate CPUs and GPUs on the same die forming integrated (fused) CPU-GPU systems. Under this architecture principle, CPUs and GPUs share memory space, memory controller, and a level of cache all of which greatly reduce data transfer time. Large industrial companies endorse this vision of such and have started to experiment and commercialize integrated heterogeneous systems: Intel Sandy Bridge, Intel Ivy Bridge, AMDL-lano, AMD's accelerated processing units, NVIDIA's Echelon project etc. However, the design space for such heterogeneous platforms is much more complex than discrete and homogeneous systems, and thus, more sophisticated and efficient resource management concepts are required.

The overall goal of the thesis is to build a systematic methodology for such heterogeneous systems, that overcomes the problem of destructive resource sharing and maximizes system's throughput under multi-application execution. In order address these problems, CPU-to-GPU and GPU-to-GPU interefence must be quantified and analyzed.

PREREQUISITES:

Good knowledge of C/C++, computer architecture, CUDA

READING MATERIAL:

1. Throughput optimization and resource allocation on GPUs under multi-application execution
2. Co-Scheduling on Fused CPU-GPU Architectures with Shared Last Level Caches
3. Airavat: Improving energy efficiency of heterogeneous applications.

CONTACT INFORMATION:

Prof. Iraklis Anagnostopoulos: (iraklis.anagno@siu.edu)

Prof. Dimitrios Soudris: (dsoudris@microlab.ntua.gr)