Boyang Zhang

https://foolock.github.io/

Email : bzhang523@wisc.edu
Mobile : +1-908-705-7875
Linkedin: www.linkedin.com/in/boyangzhangb49820190

Research Summary

- GPU kernel design and optimization for memory-bound applications, leveraging shared memory, tiling, and pipeline scheduling. • Task Graph Partitioning. • Static Timing Analysis.

Education

- **University of Wisconsin-Madison** — Madison, WI
  Ph.D. in Electrical and Computer Engineering — *May 2027*
- **Rutgers University** — New Brunswick, NJ
  Master of Science in Electrical and Computer Engineering — *May 2022*
- **South China University of Technology** — Guangzhou, China
  Bachelor of Science in Electronic Engineering — *May 2020*

Projects

- **PASTA**: PASTA is a fast task-graph partitioner for static timing analysis (STA), with both a parallel CPU version (C-PASTA ISPD'24) and a GPU version (G-PASTA DAC'24). PASTA aims to reduce task-scheduling overhead by partitioning the original task graph into a smaller graph without affecting too much its parallelism.
- **iTAP**: iTAP is an incremental task graph partitioner for STA built on top of PASTA (ASP-DAC'25). iTAP aims to further reduce the partitionig runtime by incrementally updating the partitioned task graph, avoiding fully repartition the entire graph when only a small portion of the graph changes.
- **FDTD Simulation**: Implemented an efficient GPU kernel for Finite-Difference Time-Domain (FDTD) simulation that applies diamond tiling to exploit temporal data reuse in shared memory, achieving 40% speedup over a state-of-the-art implementation on a 4M-cell problem.
- **G-STAR**: G-STAR is a GPU-accelerated statistical static timing analysis algorithm using level-by-level replication. G-STAR aims to enable efficient levelized data propagation for memory-bound workloads whose full level list cannot fit in GPU memory at once.

Work Experience

- **Software Intern at Cadence** — San Jose, CA
  Worked on GPU-accelerated Statistical Static Timing Analysis — *May 2024 - Aug. 2024*

Selected Publications

- **Boyang Zhang**, Che Chang, Cheng-Hsiang Chiu, Dian-Lun Lin, Yang Sui, Chih-Chun Chang, Yi-Hua Chung, Wan-Luan Lee, Zizheng Guo, Yibo Lin, and Tsung-Wei Huang, "iTAP: An Incremental Task Graph Partitioner for Task-parallel Static Timing Analysis", *ASP-DAC*, 2025
- **Boyang Zhang**, Dian-Lun Lin, Che Chang, Cheng-Hsiang Chiu, Bojue Wang, Wan Luan Lee, Chih-Chun Chang, Donghao Fang, and Tsung-Wei Huang, "G-PASTA: GPU-Accelerated Partitioning Algorithm for Static Timing Analysis", *DAC*, 2024.
- Tsung-Wei Huang, **Boyang Zhang**, Dian-Lun Lin, and Cheng-Hsiang Chiu, "Parallel and Heterogeneous Timing Analysis: Partition, Algorithm, and System", *ISPD*, 2024
- **Boyang Zhang**, Yang Sui, Lingyi Huang, Siyu Liao, Chunhua Deng, Bo Yuan, "Algorithm and Hardware Co-design for Deep Learning-powered Channel Decoder: A Case Study", *ICCAD*, 2021

Skills

- **Programming Language**: C, C++, Verilog, VHDL, SystemVerilog, Python
- **Programming Model**: Taskflow, CUDA, OpenMP, oneTBB
- **Testing & Profiling**: doctest, Nsight Compute, Valgrind
- **DevOps & Tooling**: GitHub, CMake