BOYANG ZHANG

bzhang523@wisc.edu

https://foolock.github.io

EDUCATION

University of Madison, Wisconsin   Seq. 2023 - now
Ph.D. student in Electrical and Computer Engineering

Rutgers, The State University of New Jersey   Seq. 2019 - May. 2021
M.S. in Electrical and Computer Engineering

South China University of Technology   Sep. 2016 - May 2020
B.S. in Electronic Engineering

RESEARCH INTERESTS

Parallel and Heterogeneous Computing, Graph Partitioning, GPU Computing, Physical Design Automation

EXPERIENCE

University of Madison, Wisconsin, Madison, WI, USA   Seq. 2023 - now
Research Assistant

Worked on task graph partitioning for Static Timing Analysis.

Cadence, San Jose, CA, USA   May. 2024 - Aug. 2024
Software Intern

Worked on GPU-accelerated Statistical Static Timing Analysis.

SKILLS

Programming Languages	CUDA, C/C++, Verilog, VHDL, SystemVerilog, Python, BASH
Unit Test	doctest
Profiler	NVIDIA Nsight Systems
Programming Model	Taskflow, oneTBB, CUDA, OpenMP, Pytorch

PUBLICATIONS

Conference Papers

· Wan-Luan Lee, Shui Jiang, Dian-Lun Lin, Che Chang, **Boyang Zhang**, Yi-Hua Chung, Ulf Schlichtmann, Tsung-Yi Ho, and Tsung-Wei Huang, "iG-kway: Incremental k-way Graph Partitioning on GPU," *ACM/IEEE Design Automation Conference (DAC)*, San Francisco, CA, 2025

· Cheng-Hsiang Chiu, Wan-Luan Lee, **Boyang Zhang**, Yi-Hua Chung, Che Chang, and Tsung-Wei Huang, "A Task-parallel Pipeline Programming Model with Token Dependency," *Workshop on Asynchronous Many-Task Systems and Applications (WAMTA)*, St. Louis, MO, 2025

· **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," *IEEE/ACM Asia and South Pacific Design Automation Conference (ASP-DAC)*, Tokyo, Japan, 2025

- Che Chang, **Boyang Zhang**, Cheng-Hsiang Chiu, Dian-Lun Lin, Yi-Hua Chung, Wan-Luan Lee, Zizheng Guo, Yibo Lin, and Tsung-Wei Huang, "PathGen: An Efficient Parallel Critical Path Generation Algorithm," *IEEE/ACM Asia and South Pacific Design Automation Conference (ASP-DAC)*, Tokyo, Japan, 2025
- Cheng-Hsiang Chiu, Chedi Morchdi, Yi Zhou, **Boyang Zhang**, Che Chang, and Tsung-Wei Huang, "Reinforcement Learning-generated Topological Order for Dynamic Task Graph Scheduling", *IEEE High-performance and Extreme Computing Conference (HPEC)*, virtual, 2024
- Chih-Chun Chang, **Boyang Zhang**, and Tsung-Wei Huang, "GSAP: A GPU-Accelerated Stochastic Graph Partitioner," *ACM International Conference on Parallel Processing (ICPP)*, Gotland, Sweden, 2024
- **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," *ACM/IEEE Design Automation Conference (DAC)*, San Francisco, CA, 2024
- Che Chang, Cheng-Hsiang Chiu, **Boyang Zhang**, and Tsung-Wei Huang, "Incremental Critical Path Generation for Dynamic Graphs," *IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, Knoxville, Tennessee, 2024
- Tsung-Wei Huang, **Boyang Zhang**, Dian-Lun Lin, and Cheng-Hsiang Chiu, "Parallel and Heterogeneous Timing Analysis: Partition, Algorithm, and System," *ACM International Symposium on Physical Design (ISPD)*, Taipei, Taiwan, pp. 51-59, 2024
- Donghao Fang, **Boyang Zhang**, Hailiang Hu, Wuxi Li, Bo Yuan, and Jiang Hu, "Global Placement Exploiting Soft 2D Regularity," *ACM International Symposium on Physical Design (ISPD)*, New York, NY, 2022.
- Lingyi Huang, Xiao Zang, Yu Gong, **Boyang Zhang**, and Bo Yuan, "VLSI Hardware Architecture of Neural A* Path Planner," *Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, 2022
- **Boyang Zhang**, Yang Sui, Lingyi Huang, Siyu Liao, Chunhua Deng, and Bo Yuan, "Algorithm and Hardware Co-design for Deep Learning-powered Channel Decoder: A Case Study," *IEEE/ACM International Conference On Computer Aided Design (ICCAD)*, Munich, Germany, 2021

PROJECTS

GPU-Accelerated Task Graph Partitioning for Static Timing Analysis

- This project develops G-PASTA, a fast GPU-accelerated partitioning method for large task dependency graphs (TDGs) used in modern static timing analysis (STA). While recent STA engines rely on TDG parallelism to speed up graph-based and path-based analyses, the scheduling overhead of very large TDGs has become a major bottleneck. G-PASTA addresses this challenge by providing a lightweight, GPU-powered partitioner that drastically reduces scheduling cost while improving overall STA performance. In experiments on large designs, G-PASTA achieves up to $41.8\times$ faster partitioning than state-of-the-art CPU methods and boosts end-to-end STA runtime by up to 43%.
- Publications: *DAC'24*

Incremental Task-Graph Partitioning for Static Timing Analysis

- This project develops iTAP, an incremental task-dependency graph (TDG) partitioner designed for timing-driven static timing analysis (STA). While existing TDG partitioners reduce scheduling overhead, they lack incremental capabilities—an essential requirement for STA tools that repeatedly update timing information during optimization. iTAP addresses this gap by supporting efficient, incremental updates to TDG partitions without rebuilding them from scratch. This enables practical integration into timing-driven workflows and significantly improves end-to-end performance. Compared to G-PASTA, iTAP delivers up to $2.97\times$ speedup in overall STA runtime.
- Publications: *ASP-DAC'25*