

TSUNG TAI YEh

ttyeh@cs.nycu.edu.tw

1001 University Road, Hsinchu 300, Taiwan

Personal Website: <https://www.cs.nycu.edu.tw/~ttyeh/>

EDUCATION

Purdue University, USA 2012-2020

Ph.D. in Electrical & Computer Engineering
Area: Computer Architecture and Parallel Computing
Advisor: Timothy G. Rogers

National Tsing Hua University, Taiwan 2009

Master of Science in Institute of Information & System Application
Thesis: Efficient Parallel Algorithm for Nonlinear Dimensionality Reduction on GPU

POSITION

National Yang Ming Chiao Tung University, Taiwan 2020/08-
Assistant Professor

- Research topics: Domain-specific accelerator, GPU, Systems on AI

Advanced Micro Devices (AMD) Inc., USA 2018/01-2018/06
Research Intern

- Developed GPU architecture for machine learning workloads.
- Advisors: Bradford M. Beckmann, Matthew Sinclair.

Purdue University, USA 2012-2018, 2019-2020
Research Assistant

- NEESCom, ITAP, NSF-funded project

Academia Sinica, Taiwan 2009-2011
Research Assistant

- Advisor: Professor Hsu, Tsan-Sheng

RESEARCH GRANTS

Curriculum Plan: Edge AI Accelerator for TinyML Models 4/2022 - 3/2023
Intelligent Chip and Applications Alliance, MOE

Designing an In-storage Accelerator for Deep Neural Networks 3/2022 - 3/2023
Phison

Accelerating Ray Tracing on CGRA-based Architecture 1/2022 - 6/2023
MediaTek MARC

Evaluating memory allocation on AI accelerator 3/2021 - 6/2022
Skymizer Inc.

Optimizing Domain-Specific Accelerator Hardware Resource Utilization 11/2020 - 10/2022
109-2222-E-009-009-MY2, MOST

HONOR & AWARD

Lynn Fellowship, Purdue University, West Lafayette, IN 2012-2013

- Award recipient of Lynn fellowship is for students who work in the Computational Science and Engineering and supports research in any of high performance computing or heterogeneous programming or area that is related to computational science program.

CONFERENCE PAPERS

[C1] **Deadline-Aware Offloading for High-Throughput Accelerators** **Tsung Tai Yeh**, Matthew D. Sinclair, Brad Beckmann, Timothy G Rogers in the 27th IEEE International Symposium on High-Performance Computer Architecture (HPCA), 2021. Acceptance rate: $(63/258) = 24.4\%$

[C2] **Dimensionality-Aware Redundant SIMT Instruction Elimination** **Tsung Tai Yeh**, Roland Green, Timothy G. Rogers in the 25th ACM International Conference on Architectural Support for Programming Languages and Operating Systems,(ASPLOS), 2020. acceptance rate $86/476 = 18.0\%$

[C3] **Optimizing GPU Cache Policies for ML Workloads** Johnathan Alsop, Matthew D. Sinclair, Anthony Gutierrez, Srikant Bharadwaj, Xianwei Zhang, Bradford Beckmann, Alexandru Dutu, Omur Kayiran, Michael LeBeane, Brandon Potter, Sooraj Puthoor, **Tsung-Tai Yeh**, in IEEE International Symposium on Workload Characterization (IISWC), 2019

[C4] **Pagoda: Fine-Grained GPU Resource Virtualization for Narrow Tasks** **Tsung Tai Yeh**, Amit Sabne, Putt Sakdhnagool, Rudolf Eigenmann, Timothy G. Rogers In Proceedings of the 22nd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PPoPP), 2017 [**Best Paper Nomination**], acceptance rate $29/132 = 22.0\%$

[C5] **CacheRAID: An Efficient Adaptive Write Cache Policy to Conserve RAID Disk Array Energy** Tseng-Yi Chen, **Tsung-Tai Yeh**, Hsin-Wen Wei, Yu-Hsun Fang, Wei-Kuan Shih and Tsan-Sheng Hsu in proceedings 2012 IEEE/ACM International Conference on Utility and Cloud Computing, 2012

JOURNAL PAPER

[J1] **Pagoda: A GPU Runtime System For Narrow Tasks** **Tsung Tai Yeh**, Amit Sabne, Putt Sakdhnagool, Rudolf Eigenmann, Timothy G. Rogers In ACM Transactions on Parallel Computing (Accepted, May, 2019)

[J2] **An Energy-Efficient and Reliable Storage Mechanism for Data-Intensive Academic Archive Systems** Chen, Tseng-Yi and Wei, Hsin-Wen and **Yeh, Tsung-Tai** and Hsu, Tsan-Sheng and Shih, Wei-Kuan In IEEE Transaction on Storage, Vol 11, 2015

PATENT

LAXITY-AWARE, DYNAMIC PRIORITY VARIATION AT A PROCESSOR, **Tsung Tai Yeh**, Bradford M. Beckmann, Sooraj Puthoor, Matthew D. Sinclair, United States Patent Application Number: 16200503, Nov. 23, 2018. Assignee: Advanced Micro Devices Inc.

TEACHING

National Yang Ming Chiao Tung University, Taiwan

IOC 5009: Accelerator Architectures for Machine Learning. (Fall 2020, 2021)

- DCP 1244: Discrete Mathematics. (Spring 2021)
- IOC 5226: Operating System Capstone. (Spring 2022)

SERVICES

Session Chair for IEEE International Symposium on VLSI Design, Automation and Test (VLSI-DAT), 2022

Track Chair for IEEE International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoc), 2022

External Reviewer for USENIX Annual Technical Conference (ATC), 2022

Invited reviewer for IEEE Transactions on Signal Processing, 2022

Program Committee for IEEE International Symposium on Workload Characterization (IISWC), 2021

External Reviewer for IEEE Transactions on Computers (TOC), 2021

External Reviewer for Transactions on Architecture and Code Optimization (TACO), 2021

External reviewer for International Symposium on Microarchitecture (MICRO), 2021

STUDENTS

PhD Student Supervision

- 2022
- Ravi Krishna Sundaram (Cadence Inc.)

Master Student Supervision

- 2021
- Chih-Ning Hsi
- Mu-Le Lee
- Kai-Chun Chang
- Ching-Jui Lee
- Guan-Wu Lin
- Wei-Qi Chen (co-advised with Prof. Cheng-chung Lin)
- Hong-Yen Tai (co-advised with Prof. Yung Sung Wu)
- Cherie Hsieh (PUFsecurity)
- 2020
- Meng-Hsien Lin (First Job: MediaTek)
- Chia-Wei Liu (First Job: SiFive)
- Keng-Ting Lin
- Bo-Cheng Chen

Undergraduate Student Supervision

- 2021
- Heng-Chun Hung (Funded by undergraduate student plan, MOST)
- Yu-Fang Hu
- Yu-Ren Liu
- Yu-Yuan Liu