Bringing over 13 years of extensive experience across GPU computing, parallel computation, distributed systems, data center virtualization, embedded systems, Deep Learning model optimization, and AI chip compiler enhancements. I've held roles as a CUDA technology instructor, earned recognition in CUDA competitions, and made notable contributions to GPU virtualization, data center virtual machine migration, and fault tolerance. Proficiency extends across foundational hardware and software techniques, currently concentrating on advanced compiler-related methods to maximize AI model performance on embedded devices.

EXPERIENCE

• Mediatek

AI Compiler Enginner

May 2021 - Present

• APU Compiler development: Involved in spearheading AI Compiler development, specifically designing and implementing the tiling mechanism. Significantly boosted performance metrics (power & FPS) bridging the gap between ISP and APU. Spearheaded the implementation of GPU Neuron Network Related Backend. Innovated in the enhancement of SMP and layer pipeline performance within the APU architecture. Achieved notable advancements in benchmark scores, including ETHZ and MLPerf.

• Industrial Technology Research Institute

Hsinchu, Taiwan

Research and Development Engineer

Mar 2018 - Mar 2021 (3 years)

Email: coldfunction@gmail.com

Mobile: +886-932-651-215

- Cuju: An open source project for Virtualization-Based Fault Tolerance; primarily used for preventing the downtime causing by hardware failures.
- Patents: Holder of patents related to Fault-Tolerant Systems and their Control Methods: P52080043TW, P52080043US
 P52080043CN

• Lecturer Taiwan

GPU Computing and Virtualization Technique Lecturer

2010 - 2018 (8 years)

o Parallel, GPU and Virtualization Technique: Design and teaching courses for the industry and academic.

• Tsinghua University

Beijing, China

Teaching Assistant

Summer in 2013 & 2015

• Parallel Programming: Design the course and teach students to implement the parallel computing algorithm by pthread, MPI and CUDA for solving the practical problem.

EDUCATION

• National Tsing Hua University

Hsinchu, Taiwan

Ph.D in Department of Computer Science

Jul. 2012 - Dec. 2017

 Doctoral Dissertation: A New API Remoting Policy for General-Purpose Computing on GPU Virtualization and its Application on Biological Tools

• Chang Gung University

Taoyuan, Taiwan

Master in Computer Science & Information Engineering

Aug. 2010 - July. 2012

• Masters Dissertation: Efficient Constrained Multiple Sequence Alignment Tool and Parallel RSA Decryption Algorithm for Many-core GPUs with CUDA

AWARDS & CERTIFICATES

- Outstanding Research Award, Industrial Technology Research Institute, 2020: Research on low latency virtualization-based fault tolerance technique.
- Certificate of CUDA: Certificate Of NVIDIA CUDA Engineers (Certificate No. : 201309-01-0003).
- CUDA programming competition 2012 & 2011: The Second Place Award and Honorable Mention in the CUDA programming competition 2012 & 2011 in Taiwan.