Cao Gao

caogao@umich.edu • (734) 834-3274
web.eecs.umich.edu/~caogao

Summary

Interested in computer architecture, performance analysis, mobile systems and machine learning. Currently working on designing mobile architecture for future machine learning algorithms.

Education

- ***University of Michigan, Ann Arbor, MI*** *January.2014 – May.2017 (expected)*
  Ph.D., *Computer Science and Engineering*
  Area of Specialization: Computer Engineering --- Hardware
  Advisor: Prof. Trevor Mudge
- ***University of Michigan, Ann Arbor, MI*** *September.2012 – December.2013*
  M.S., *Computer Science and Engineering*
  Overall GPA: 4.0/4.0

- ***Zhejiang University, Hangzhou, Zhejiang, China*** *September.2008 – June.2012*
  B.Eng., Major: Electronic and Information Engineering Minor: English
  Member of Chu Kochen Honors College Overall GPA: 3.91/4.0

- ***National Cheng Kung University, Tainan, Taiwan, R.O.C*** *September.2010 – January.2011*
  Exchange student in Electrical Engineering and Computer Science Overall GPA: 93.7/100

Professional Experience

- ***ARM Ltd., Austin, TX*** *June.2014 – August.2014*
  R&D Intern at the Mobile System Group. Worked on Android workload characterization.

Major Projects

- An ultra-low power non-uniform memory accelerator for wearable devices

Developed the architecture, ISA, and compiler for the accelerator, which has been fabricated as a chip
Designed a framework that automatically generates optimal memory layout based on target applications

- Accelerating deep learning algorithms on mobile platforms

Analyzed the characteristics of Deep Neural Network workloads on mobile / server GPUs
Participated in designing a framework which intelligently partitions workloads between mobile and server

- Graph analytics processing accelerator

Proposed an accelerator architecture for billion-edge scale graph applications
Led four grad students to characterize applications, explore algorithm and architecture choices

- User quality-of-experience metrics for android applications *(at ARM)*

Identified a set of metrics that measures the user experience of Android applications
Implemented a framework in Android to automate workload execution and metrics collection

- A study of mobile device utilization

Evaluated the CPU and GPU utilization of a wide range of commonly used mobile applications
Identified the diminishing return of increasing core counts and suggested a more flexible system

Publications

- Yiping Kang, Johann Hauswald, Cao Gao, Austin Rovinski, Trevor Mudge, Jason Mars, Lingjia Tang. The Neural Edge: Collaborative Intelligence Between Cloud and the Mobile Edge. Architectural Support for Programming Languages and Operating Systems (ASPLOS), April 2017.
- Suyoung Bang, Jingcheng Wang, Ziyun Li, Cao Gao, Yejoong Kim, Qing Dong, Yen-Po Chen, et.al. A 288μW Programmable Deep Learning Processor with 270kB On-chip Weight Storage Using Non-Uniform Memory Hierarchy for Mobile Intelligence. 2017 IEEE international Solid-State Circuits Conference (ISSCC), February 2017.
- Qi Zheng, Cao Gao, Trevor Mudge, and Ronald G. Dreslinski. Leveraging Mobile GPUs for Flexible High-speed Wireless Communication. The 3rd International Workshop on Parallelism in Mobile Platforms (PRISM-3), June 2015.
- Cao Gao, Anthony Gutierrez, Madhav Rajan, Ronald G. Dreslinski, Trevor Mudge, and Carole-Jean Wu. A Study of Mobile Device Utilization. 2015 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), March 2015.
- Cao Gao, Anthony Gutierrez, Ronald G. Dreslinski, Trevor Mudge, Kristian Flautner, and Geoffery Blake. A Study of Thread Level Parallelism on Mobile Devices. 2014 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), March 2014.

Skills

Programming: experienced in C/C++, Python, CUDA, familiar with Verilog, Matlab
Environments: Linux, Android, shell scripting, git, ARM streamline, Keil uVision, nvprof, Caffe
Languages: Fluent in English, native Mandarin speaker

Related Courses

EECS 583 -- Advanced Compilers Grade: A
Project: Design a loop-distribution compiler optimization technique in LLVM
EECS 570 -- Parallel Computer Architecture Grade: A
Project: Fine-grain reconfigurable energy-efficient Network-on-Chip router architecture
EECS 578 -- Computer-Aided Design Verification of Digital Systems Grade: A+
Project: BugCalc: a quantitative evaluation of bug effects and characteristics
EECS 470 -- Computer Architecture Grade: A
Project: 2-way superscalar out-of-order processor. Ranked 2nd overall, 1st in clock frequency.

Awards

University of Michigan EECS Departmental Fellowship	2012-2013
ISCA conference travel grant	2015
Rackham travel grant	2014, 2015
Scholarship for Excellence in Research and Innovation	2009, 2010
Scholarship for Outstanding Students	2008, 2009, 2010