Zhuoran Zhao Email: zhuoran@utexas.edu Personal Website: https://zoranzhao.github.io Mobile: +1-512-751-1819

GitHub: https://github.com/zoranzhao

LinkedIn: https://www.linkedin.com/in/zoranzhao

SUMMARY

My research interests mainly include efficient runtime/middleware for distributed deep learning, system-level performance modeling and software/hardware co-design.

SKILLS

• Programming languages: C/C++, Python

- Tools and frameworks: LLVM, Caffe2, PyTorch, Darknet, OMNeT++, lwIP.
- Project experiences with operating system kernel programming, multi-threaded programming (POSIX pthreads) and network programming (Socket Programming in C).

EDUCATION

• Ph.D. in Electrical and Computer Engineering;

University of Texas at Austin;

Advisor: Prof. Andreas Gerstlauer

• M.S. in Electrical and Computer Engineering:

University of Texas at Austin; GPA: 3.93/4.00

• B.S. in Electrical Engineering;

Zhejiang University; GPA: 3.95/4.00

Dec. 2014 - May 2019

Austin, Texas

Aug. 2012 – Dec. 2014 Austin, Texas

Sep. 2008 – Jun. 2012

Zhejiang, China

Honored Minor: Advanced Honor Class of Engineering Education (ACEE)

EXPERIENCE

Facebook

Research Scientist

Menlo Park, CA Oct. 2019 - Present

• University of Texas at Austin

Graduate Research Assistant/Teaching Assistant

Austin. Texas Aug. 2012 - May 2019

- o DeepThings: A portable and lightweight runtime framework for locally distributed CNN/DNN inference in resource-constrained IoT edge clusters, developed in C [1].
- o NoSSim: A source-level network/system co-simulation framework for rapid embedded/mobile system prototyping, developed in C++ with LLVM, OMNeT++ and SystemC framework [2].
- HCSim: A fast full-system simulation platform with abstract models of real-time operating systems (RTOS) and high-level multi-core processor models, developed in C++ with SystemC framework [2].
- RBA: A compile-time profiling and instrumentation tool for source-level system performance evaluation, developed in C++ and python with gcc and LLVM framework [3].

• FutureWei Technologies

Research Intern

Research Intern

Plano, Texas

May 2014 - Aug. 2014

Manager: Weizhong Chen

o Architectural Description Language (ADL) framework prototype for digital signal processors (DSP), developed in C++ and Python.

• NXP Semiconductors

Austin, Texas

May 2013 - Aug. 2013

Manager: Mark Bader

• Automatic microprocessor performance calibration framework between RTL and cycle-accurate simulator, developed in C++ with ADL/uADL framework.

• University of California, Los Angeles

Summer Exchange Program

Advisor: Prof. Vwani P. Roychowdhury

Los Angeles, California Jul. 2011 - Sep. 2011

• Interactive complex network visualization in mobile applications and browsers, developed in JavaScript, Java and Python with AJAX technique and Django framework.

SELECTED PUBLICATIONS

- [1] Zhuoran Zhao, K. Mirzazad and A. Gerstlauer, "DeepThings: Distributed Adaptive Deep Learning Inference on Resource-Constrained IoT Edge Clusters," CODES+ISSS, special issue of IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD), 2018.
- [2] Zhuoran Zhao, V. Tsoutsouras, D. Soudris, A. Gerstlauer, "Network/System Co-Simulation for Design Space Exploration of IoT Applications," Proceedings of the International Conference on Embedded Computer Systems: Architectures, Modeling and Simulation (SAMOS), 2017.
- [3] Zhuoran Zhao, A. Gerstlauer and Lizy K. John, "Source-Level Performance, Energy, Reliability, Power and Thermal (PERPT) Simulation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, 2017.
- [4] Zhuoran Zhao, D. Lee and A. Gerstlauer, "Host-Compiled Reliability Modeling for Fast Estimation of Architectural Vulnerabilities," In Silicon Errors in Logic, System Effects Workshop (SELSE), 2015
- [5] S. Chakravarty, Zhuoran Zhao, A. Gerstlauer, "Automated, Retargetable Back-Annotation for Host-Compiled Performance and Power Modeling," Proceedings of the IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS), 2013.
- [6] L. Guckert, M. O'Connor, S. K. Ravindranath, Zhuoran Zhao and V. J. Reddi, "A Case for Persistent Caching of Compiled JavaScript Code in Mobile Web Browsers," In Workshop On Architectural And Microarchitectural Support For Binary Translation (AMAS-BT), 2013

Relevant Graduate Coursework

- EE382V Dynamic Compilation
- EE382V Advanced Programming Tools
- EE380L Engineering Programming Languages
- EE382C Multicore Computing
- EE382N Computer Architecture
- EE382N Embedded System Design and Modeling
- EE382M System-on-a-Chip Design

Professional Service

• Reviewer:

- o IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD), 2017
- o Design, Automation and Test in Europe (DATE) Conference, 2018
- IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS), 2018

• Teaching:

- o Teaching Assistant: EE382N Embedded System Design and Modeling, 2016
- Teaching Assistant: EE319K Introduction to Embedded System, 2012

Honors and Awards

- Best in Session Award for the presentation "Automated, Retargetable Back-Annotation for Host-Compiled Power and Performance Modeling," in Semiconductor Research Corporation (SRC) TECHCON, Sep 11, 2013
- National Scholarship in China (2%), 2009, 2010