Wenzhi Cui

Google LLC cwz920716@gmail.com

SUMMARY

I am interested in building high-performance, energy-efficient computer systems. My past projects focus on developing advanced compiler and runtime support for emerging programming models and hardware platforms, such as GPGPU and Pixel Visual Core (PVC).

EDUCATION

The University of Texas at Austin *2014-2018*
M.S., Computer Science

Nanjing University, Nanjing, China *2010-2014*
B.S., Software Engineering

SKILLS

- Programming Languages: C/C++, Cuda, Java, JavaScript, Python
- System Software: Halide, LLVM, Node.js, Coq
- Hardware Description Language: Verilog, Bluespec Verilog

WORK EXPERIENCE

Google Full-time 06/2018 - present
- Compiler/runtime support for new hardware/language features on Pixel Visual Core
- Add compiler features to assit application teams in porting emerging applications to PVC
- Improve PVC compiler infrastucture reliability and performance via fixing critical bugs and refactoring compiler code base.

Graduate Research Assistant (GRA) 2017 - 2018
- LLVM based instrumentation Framework for CUDA device kernels and host applications
- Profile conditional branch divergence and memory divergence analysis on GPGPU
- Profile GPU/CPU/Battery power on drone applications on NVIDIA Tegra TX2

Google Intern Summer 2017
- Compiler/runtime support for advanced hardware features on Pixel Visual Core (PVC)
- Halide library for image applications (Transpose, Rectification, etc.) on PVC

Graduate Research Assistant (GRA) 2016 - 2017
- Develope a program profiling technique called the "Event Dependence Graph" (EDG) to deconstruct the server response time of event-driven applications such as Node.js
- Use EDG to characterize the tail latency of Node.js application and identified JavaScript garbage collector as a dominant root cause for the Node.js tail
- Demonstrate how to alleviate the tail latency by applying frequency boosting during garbage collection and carefully tuning garbage collector parameters

IBM Research Austin Summer 2015
- Analyze the HTTP request latency distribution of web servers in data centers
- Design an end-host based load balancing scheme by offloading HTTP requests
- Implemente a prototype in OpenVSwitch

PUBLICATIONS

Conference Papers

- Behzad Boroujerdian, Hasan Genc, Srivatsan Krishnan, Wenzhi Cui, Aleksandra Faust, Vijay Janapa Reddi
  MAVBench: Micro Aerial Vehicle Benchmarking
  MICRO 2016
- Nadav Chachmon, Daniel Richins, Robert Cohn, Magnus Christensson, Wenzhi Cui, Vijay Janapa Reddi
  Simulation and Analysis Engine for Scale-Out Workloads
  ICS 2016
- Wenzhi Cui, Chen Qian
  Scalable and Load-balanced Data Center Multicast
  Globecom 2015
- Wenzhi Cui, Chen Qian
  DiFS: Distributed Flow Scheduling for Adaptive Routing in Hierarchical Data Center Networks
  ANCS 2014

Patents

- Kanak B. Agarwal, Wenzhi Cui, Wesley Felter, Yu Gu, Eric Rozner
  Job assignment using artificially delayed responses in load balanced groups
- Kanak B. Agarwal, Wenzhi Cui, Wesley Felter, Yu Gu, Eric Rozner
  Tail latency-based job offloading in load-balanced groups
- Kanak B. Agarwal, Wenzhi Cui, Wesley Felter, Yu Gu, Eric Rozner
  Offloading at a virtual switch in a load-balanced group

HONORS & RECOGNITIONS

- UT Austin Microelectronics and Computer Development Fellowship, 2014-2017
- Google Scholarship, 2013

COURSEWORK PROJECTS

Misc.

- Verify the correctness of (simplified) mark-sweep garbage collector using Coq
- Implemente an analysis pass in the LLVM compiler to determine C/C++ pointer bounds and integrate with Softbound, a compiler transformation pass for enforcing spatial safety of C/C++ pointers
- OpenGL Based Cloth Simulation using mass-spring model
- Utilize parallel execution and asynchronous IO to speed up the recursive copy operation on SSD
- Implement a multi-cycle DLX (simplified MIPS) microprocessor using Verilog

TEACHING EXPERIENCE

Teaching Assistant

- Undergraduate: Programming Languages(Honor), Principles of Computer Systems, Computer Networks
- Graduate: Code Generation and Optimization

COURSEWORK

Computer Architecture, Compilers, Advanced Operating Systems, Algorithms: Techniques/Theory, Formal Verification and Semantics, Natural Language Processing, Program Verification, Hardware Verification, Programming Languages, Computer Graphics, Physical Simulation