ZHUANHAO WU

37-249 Cedarbrae Ave. ⋄ Waterloo, Ontario, N2L 4S8 Canada
+86 156-815-73728 (mainland China) / +1 519-573-6243 (Canada) ⋄ zhwu95@gmail.com

EDUCATION

University of Waterloo — Waterloo, Canada
Ph.D. candidate in Electrical and Computer Engineering — Fall. 2019 - Spring. 2025
Research area: computer architecture, cache coherence, real-time systems
Supervisor: Prof. Hiren Patel

University of Waterloo — Waterloo, Canada
MASc in Electrical and Computer Engineering — Fall. 2017 - Spring. 2019
Department of Electrical and Computer Engineering
Supervisor: Prof. Hiren Patel

Nankai University — Tianjin, China
B.E. in Computer Science and Technology — Sept. 2013 - Jul. 2017

PUBLICATIONS

Conference Publications

- Exclusive Hierarchies for Predictable Sharing in Last-Level Cache
  Xinzhe Wang, **Zhuanhao Wu**, Rodolfo Pellizzoni, Hiren Patel
  RTAS 2024
- SCCL: An open-source SystemC to RTL translator
  Zhuanhao Wu, Maya Gokhale, Scott Lloyd, Hiren Patel
  FCCM 2023
- Ditty: Directory-based Cache Coherence for Multicore Safety-critical Systems
  Zhuanhao Wu, Marat Bekmyrza, Nachiket Kapre, Hiren Patel
  DATE 2023, nominated for Best Paper Award
- ZeroCost-LLC: Shared LLCs at No Cost to WCL
  Zhuanhao Wu, Anirudh Kaushik, Hiren Patel
  RTAS 2023
- Predictable sharing of last-level cache partitions for multi-core safety-critical systems
  Zhuanhao Wu, Hiren Patel
  DAC 2022
- ZHW: A Numerical CODEC for Big Data Scientific Computation
  Michael Barrow, **Zhuanhao Wu**, Scott Lloyd, Maya Gokhale, Hiren Patel, Peter Lindstrom
  FPT 2022
- A Hardware Platform for Exploring Predictable Cache Coherence Protocols for Real-time Multicores
  Zhuanhao Wu, Anirudh Kaushik, Paulos Tegegn, Hiren Patel
  RTAS 2021
- CARP: A Data Communication Mechanism for Multi-core Mixed-Criticality Systems
  Anirudh Kaushik, Paulos Tegegn, **Zhuanhao Wu**, Hiren Patel
  RTSS 2019
- Strengthening PUFs using Composition
  Zhuanhao Wu, Hiren Patel, Manoj Sachdev, Mahesh Tripunitara
  ICCAD 2019

Workshop Publications

- PASoC: A Predictable Accelerator-rich SoC
  Susmita Tadepalli, **Zhuanhao Wu**, Hiren Patel
  TCRS'23

Journal Publications

- High Performance and Predictable Shared Last-level Cache for Safety-Critical Systems
  Zhuanhao Wu, Anirudh Kaushik, Hiren Patel
  TECS 2024
- Enhancing Strong PUF Security With Nonmonotonic Response Quantization
  Kleber Stangherlin, **Zhuanhao Wu**, Hiren Patel, Manoj Sachdev
  TVLSI 2023

RESEARCH PROJECT EXPERIENCE

Memory-model directed cache coherence protocol Feb. 2023 - current
Sole developer *Computer Architecture and Embedded System Group, University of Waterloo*
Keywords: memory model, cache coherence, formal methods, gem5

· Developed two approaches to reduce the worst-case latency for TSO systems, decoupling the MSHR entry size (up to 97% reduction for 32 entries system).
· Correctness of the technique proven in Lean 4.

Predictable directory-based coherence protocol May. 2022 - current
Developer in a group of 2 *Computer Architecture and Embedded System Group, University of Waterloo*
Keywords: cache coherence, HLS, hardware prototype

· Developed a directory-based coherence protocol and hardware with WCL guarantees, where the coherence protocol includes coherence state changes to a baseline MSI protocol
· Integrated of designs from multi-language (the real-time interconnect is implemented in SystemVerilog, the cache design is implemented in HLS C++, the glue logic is implemented in Scala/SpinalHDL)
· Synthesized and run the design in AWS FPGA F1 platform

Ensuring worst-case latency with ZeroCost-LLC May. 2022 - December. 2022
Sole developer *Computer Architecture and Embedded System Group, University of Waterloo*
Keywords: gem5, C++, micro-architectural simulator

· Implemented the ZIV mechanism to eliminate back-invalidations in inclusive LLC
· Implemented an invariant-based mechanism on top of ZIV to eliminate write-backs caused due to LLC replacement, which lowers the WCL bound.

systemc-clang SystemVerilog translation backend Sept. 2019 - Dec. 2023
Developer in a group of 4 *Computer Architecture and Embedded System Group, University of Waterloo*
Keywords: SystemC, Python, hardware prototyping, compiler

· Translated the frontend generated intermediate representation (IR) into SystemVerilog with Python parser
· Synthesized and run a ZHW floating point number encoder/decoder generated from SystemC on a Zynq UltraScale+ FPGA board
· Synthesized and run a systolic array design generated from SystemC on a Zynq UltraScale+ board

Predictable Cache Coherence Prototyping Platform Feb. 2020 - Sept. 2021
Sole developer *Computer Architecture and Embedded System Group, University of Waterloo*
Keywords: heterogeneous cache coherence, RISC-V, FPGA, linux device driver, qemu

- Implemented a QEMU device, enabling co-simulation of host applications and the actual hardware design
- Implemented a host device driver, allowing the host to access the FPGA memory coherently with RISC-V cores
- Set up CI environment for verifying the core and verified the single core implementation against riscv-tests with iverilog and verilator
- Implemented a set-associative cache design
- Implemented a DSL in Chisel/Scala for describing hardware coherence protocols
- Synthesized RV32IA to Xilinx and Altera FPGAs and verified the synthesized core against riscv-tests
- Implemented the Predictable MSI cache coherence
- Implemented the predictable TDM bus arbitration schemes with various flavors, some leading to improvement of WCL bound over previous works
- Implemented the system call emulation mechanism, enabling the core to handle file I/O and to spawn threads without an operating system

Physical Unclonable Functions (PUFs) Apr. 2018 - Sept. 2022
Developer in a group of 2 *Computer Architecture and Embedded System Group, University of Waterloo*

- Implemented a framework in Python for evaluating the resilience of different PUF architectures to machine learning-based modeling attacks.
- Validated the machine learning resistance of a new architecture with evolutionary strategies and logistic regression.
- The framework leads to the publication of two PUF architectures, PoP and NMQ-PUF.

Wavenet implementation on FPGA Mar. 2018 - Apr. 2018
Course Project

- Implemented a text-to-speech neural net architecture, the Wavenet, on FPGA using high-level synthesis.
- Cross-validated the result with tensorflow implementation.

WORK EXPERIENCES

Infrastructure R&D Intern Jun. 2023 - Nov. 2023
Alibaba Group *Beijing, China*

- Worked on the full-system simulation infrastructure and its the calibration, with a specific focus on translation-look-aside buffer.

AWARDS AND HONORS

Engineering Graduate Scholarship Fall 2017, Winter 2021, Fall 2022
ACM/ICPC Asia Regional Onsites, 2 Silver Medals 2016

RELEVANT COURSES

Register-Transfer-Level Design
Reconfigurable Computing
Reinforcement Learning
Computer Organization
Algorithm Design and Analysis
Safety-critical Embedded Software

TECHNICAL STRENGTHS

Languages & DSLs	C++, Python, Verilog, Bash, Scala (Chisel/SpinalHDL), SystemC, Lean 4 and mechanized proofs
Technologies	Vitis/Vivado, Tensorflow/PyTorch, Git, Linux