Zhantong Qiu

studyztp[1] | Zhantong Qiu[2] | studyztp@gmail.com[3] |

Introduction

Computer architecture researcher focusing on computer system performance evaluation, simulation methodology, and hardware-software co-design for emerging applications. I build novel frameworks for fast, accurate, and portable computer system performance evaluation using sampled simulation. I tackle robotic applications on tiny robots and HPC workloads in datacenters through hardware-software co-design. I am a strong programmer and active contributor to production open-source simulators (e.g., gem5). I seek challenges in constrained systems and solve them with solutions that span multiple layers of the computer system stack.

Education

Master of Science in Computer Science Fall 2023 – Spring 2026 (expected)
University of California, Davis Davis, California

Visiting Student in Electrical and Computer Engineering Fall 2025 – Spring 2026
Cornell University Ithaca, New York

Bachelor of Science in Computer Science and Engineering Fall 2020 – Spring 2023
University of California, Davis Davis, California

Publications

- **Nugget: Portable Program Snippets[4] (Accepted to HPCA 2026)**
  Authors: **Zhantong Qiu**, Mahyar Samani, Jason Lowe-Power.

  Nugget introduces an LLVM IR-based sampling framework that turns long-running workloads into portable program snippets ("nuggets") which can be analyzed once, then reused across binaries, ISAs, and microarchitectures. It replaces simulator-based interval analysis with compiler instrumentation on real hardware, cutting sampling overhead by up to hundreds× versus functional simulation, while keeping intervals binary-independent via an IR-level unit of work. Using large SPEC CPU2017, NPB, and LSMS workloads, the work shows how Nugget enables fast, hardware-based validation of sampling methods and helps diagnose simulator modeling errors in gem5's Arm memory system.

- **Accelerating the Simulation of Parallel Workloads using Loop-Bounded Checkpoints (Under review at ACM TACO)**
  Authors: Alen Sabu, **Zhantong Qiu**, Harish Patil, Changxi Liu, Wim Heirman, Jason Lowe-Power, Trevor E. Carlson.

  Proposed *LoopPoint*, a synchronization-agnostic loop-based sampling methodology that enables fast, accurate simulation of multi-threaded workloads via loop-bounded checkpoints, achieving up to 801× speedup on SPEC CPU2017 with ∼2.3% average runtime error, and introduced *ROIperf* for validating representative regions directly on silicon. This work extends the prior *LoopPoint* paper published at HPCA.

[1] https://github.com/studyztp
[2] https://www.linkedin.com/in/zhantong-qiu-60165a165/
[3] mailto:studyztp@gmail.com
[4] https://arxiv.org/abs/2509.02873

Talks and Tutorials

- **HPCA 2026** – Paper talk: *Nugget: Portable Program Snippets* (Main Conference, Sydney, Australia).
- **CGO 2026** – Technique talk: Nugget with LLVM infrastructure (LLVM–CGO 2026 Workshop, Sydney, Australia).
- **gem5 Bootcamp 2024** – Week-long hands-on training on the gem5 simulator (UC Davis, CA).
- **HPCA 2023** – LoopPoint Tools: Sampled Simulation of Complex Multi-threaded Workloads using Sniper and gem5 (LoopPoint Tutorial, Montreal, Canada).
- **ISCA 2023** – Technique talk on full-system sampling support in gem5 (gem5 Workshop, Orlando, FL).

Research Experience

Graduate Researcher June 2023 – Present
DArchR Lab, UC Davis
Advisor: Jason Lowe-Power

- **Nugget (LLVM IR sampling)**
  - Designed and implemented a cross-platform, architecture-independent sampling framework at the LLVM IR level; supports rapid interval analysis, native execution for validation, and cross-ISA simulation.
  - Evaluated Nugget on real hardware platforms with diverse performance characteristics (e.g., Ampere Altra) and in the gem5 simulator to validate fidelity and portability.
  - Analyzed a diverse set of workloads (SPEC CPU2017, NPB, LSMS) to demonstrate robustness across CPU, HPC, and scientific applications.
- **Accelerating the Simulation of Parallel Workloads using Loop-Bounded Checkpoints**
  - Co-led methodology and implementation to validate LoopPoint samples on gem5; under journal review.
- SPEAR: Scalable Power and Energy Analysis and Performance Tools for Frontier
  - Working with researchers from Oak Ridge National Laboratory and University of
- **gem5 contributions**
  - Ongoing upstream contributions since 2022 to full-system sampling support and related simulation features in gem5; author of 50 commits[5] (public history).

Visiting Student August 2025 – Present
Computer Systems Laboratory, Cornell University
Advisor: Christopher Batten

- **Agile Robotic Hardware-Software Co-Design**
  - Developing a closed-loop, multi-robot evaluation framework integrating robotics simulators and architecture simulators; enables parallel per-robot simulation and decouples from a specific ISA/simulator to broaden studies (e.g., ISA extensions).
- **Accurate STM32G4 Board in gem5**
  - Building an accurate STM32G4 MCU model in gem5 (Arm M-profile). Implemented the ART (Adaptive Real-Time) accelerator[6] to improve Flash access latency; planning to upstream core M-profile components (e.g., NVIC).

[5] https://github.com/gem5/gem5/commits?author=studyztp
[6] https://github.com/studyztp/gem5/tree/studyztp/ART-cache

Undergraduate Researcher June 2022 – June 2023
DArchR Lab, UC Davis
Advisor: Jason Lowe-Power

- **SimPoint**
  - Implemented SimPoint support in the gem5 stdlib and ran SPEC CPU2006 experiments using SimPoint.
- **LoopPoint**
  - Implemented LoopPoint support in system-emulation and full-system modes in gem5; added a generalized PC-execution counter to track occurrences of specific PCs.

Course Projects

CXL Shared Memory Filesystem Optimization[7]
Investigated application needs in a CXL shared-memory system and improved the FAMFS framework with more efficient allocation/deallocation and zero-copy operations to avoid redundant copies.
Limitations of Disaggregated Memory and Innovations[8]
Analyzed limitations of disaggregated memory and proposed a hardware-software co-designed page management approach.
CUDA Microbenchmarks[9]
Developed configurable microbenchmarks to measure shared-memory latency and memory-scaling behavior on GPUs.
RISC-V Operating System
Built a simple RISC-V OS on the UC Davis RISC-V Console Simulator for a course project.

Teaching Experience

WQ 2024 ECS 154B: Computer Architecture Teaching Assistant
Led weekly discussion sections and office hours; created assignment material on Chisel-based CPU model (DINO CPU).

Organizations and Service

Founder of Computer Systems Seminar at UC Davis[10]
Organized weekly speaker series with 12+ talks from academia and industry (still on-going).

HPCA 2026 Artifact Evaluation Reviewer
Helped reviewing the artifact of the accepted paper.

Skills

Programming Languages: C/C++, Python, Bash, CUDA, assembly, Latex
Hardware Description Languages: Chisel
System Evaluation Tools: gem5, QEMU
Profiling & Instrumentation: LLVM passes, DynamoRIO, Valgrind, Linux perf, PAPI
Other Linux Related Tools: cpuset, CRIU, cgroups, Docker
Compilers: GCC/G++, GFortran, Clang/LLVM
Languages: English, Cantonese, Mandarin

[7] https://github.com/orgs/ECS-289D/repositories
[8] https://github.com/studyztp/my-course-paper/blob/main/Limitations_of_Disaggregated_Memory_and_Innovations.pdf
[9] https://github.com/studyztp/CUDA_microbenchmarks.git
[10] https://arch.cs.ucdavis.edu/events/computer-system-seminar