Reese Kuper

reesekuper.com | reese.kuper@gmail.com

EDUCATION

Masters – University of Illinois at Urbana-Champaign (3.97/4.00) 2021 - 2023
Department of Electrical and Computer Engineering
Focus: Computer Architecture

Bachelors – University of Wisconsin-Madison (3.89/4.00) 2017 - 2021
Degrees in Computer Engineering and Computer Science

Relevant Coursework

- Parallel Computer Architecture
- Computer Microarchitecture
- Artificial Intelligence
- Advanced Computer Architecture
- Digital System Design & Synthesis
- Algorithms
- Memory and Storage Systems
- Operating Systems
- System-on-Chip Design

SKILLS

Hardware Description and Programming Languages System Verilog, C/C++, Python, Bash Scripting, MySQL
Development and Software Tools Vivado, Vivado HLS, ModelSim, Quartus, Vim, Git

EXPERIENCE

Arm 2024 - Present
Graduate Hardware Engineer – Systems Interconnect *Austin, TX*
- Improved and automated portions of interconnect model simulation flow for system analysis work.

Intel Inc. 2022 - 2023
Graduate Research Intern – On-Chip Accelerators *Remote*
- Conducted performance analysis on DSA – an on-chip accelerator found on Sapphire Rapids and later Xeon processors.
- Explored use cases for DSA to take advantage of cache pollution mitigation and higher throughput for memory operations.
- Submitted patents for improving memory deduplication techniques using DSA.
- Prepared and Presented a tutorial for Intel's on-chip accelerators at ISCA 2023.

Samsung Semiconductor Inc. 2022
Graduate Research Intern – Memory Systems Laboratory *San Jose, CA*
- Investigated the effectiveness of an L1 cache in a CXL device for DLRM offloading.
- Built a functional cache simulator to evaluate both hit rates and cache occupancy rates of real DLRM memory traces.
- Analyzed the characteristics of real DLRM data for locality patterns to aid in the design of the memory system.
- Simulated DRAM and cache designs via Ramulator to obtain bandwidth, hit rate, and other metrics used in analysis.

Arm Ltd. 2021
Hardware Engineering Intern – CPU Memory System Verification *Austin, TX*
- Debugged failing signatures for the CPU memory system testbench.
- Implemented new statistical coverage (SCOV) workflow:
  - Used scoreboard listener functions to implement SCOV events for assessing stimulus coverage.
  - Coded macro for ease of adding additional SCOV events within the UVM testbench.
  - Developed Python script to parse generated simulation log files during regression tests to send to a MySQL database.

Arm Ltd. 2020
Hardware Engineering Intern – Systems Interconnect Verification *Austin, TX*
- Developed internal Python tool to analyze the use of all plusargs within the UVM testbenches.
- Fixed UVM register definition auto-generation for more flexible RAL models.
- Programmed module for modeling transactions between a master device to interconnect return nodes in SystemC.
- Formally verified round robin and LSB priority arbiters using System Verilog assertions.

Qualcomm Inc. 2019
Software Engineering Intern – Linux Kernel Memory System *San Diego, CA*
- Improved kernel ION allocation memory speeds by around 10%.
- Analyzed the efficiency of IOVA's use of caching and compared it with MMAP's gap searching RBTree.
- Created internal Python tool for parsing Linux RAM dump binaries.
- Worked towards shifting mmap allocations to use the mempool API.

PROJECTS

Relevant Projects and Research

- Investigated finer-grained GPU scheduling for improved GPU utilization in machine learning workloads.
- Self-balancing Segway written in Verilog for a DE0-Nano FPGA board.
- Synthesized a 5-stage pipelined CPU and a dataflow processor written in Verilog.
- Implemented prefetcher designs and age matrix issueing schemes for BOOM RISC-V processor on an FPGA.

PUBLICATIONS, PATENTS, AND CONFERENCES

- **A Quantitative Analysis and Guidelines of Data Streaming Accelerator in Modern Intel Xeon Scalable Processors**
  R. Kuper, I. Jeong, Y. Yuan, R. Wang, N. Ranganathan, N. Rao, J. Hu, S. Kumar, P. Lantz, N. S. Kim
  International Conference on Architectural Support for Programming Languages and Operating Systems (**ASPLOS**), 2024
- **Efficiently Merging Non-Identical Pages in Kernel Same-Page Merging (KSM) for Efficient and Improved Memory Deduplication and Security**
  R. Kuper, Y. Yuan, R. Wang
  US Patent App 18/369,090 (**Patent**), 2024
- **Method and Appratus for Batching Pages for a Data Movement Accelerator**
  R. Kuper, Y. Yuan, R. Wang
  US Patent App 18/477,628 (**Patent**), 2024
- **Demystifying CXL Memory with Genuine CXL-Ready Systems and Devices**
  Y. Sun, Y. Yuan, Z. Yu, **R. Kuper**, C. Song, J. Huang, H. Ji, S. Agarwal, J. Lou, I. Jeong, R. Wang, J. H. Ahn, T. Xu, N. S. Kim
  International Symposium on Microarchitecture (**MICRO**), 2023
- **On-chip Accelerators in 4th Gen Intel® Xeon® Scalable Processors: Features, Performance, Use Cases, and Future!**
  R. Kuper, I. Jeong, Y. Yuan, J. Hu, R. Wang, N. Ranganathan, N. S. Kim
  [TUTORIAL] International Symposium on Computer Architecture (**ISCA**), 2023
- **STYX: Exploiting SmartNIC Capability to Reduce Datacenter Memory Tax**
  H. Ji, Y. Sun, M. Mansi, Y. Yuan, J. Huang, **R. Kuper**, M. Swift, N. S. Kim
  The USENIX Annual Technical Conference (**ATC**), 2023
- **Improving GPU Utilization in ML Workloads Through Finer-Grained Synchronization**
  R. Kuper, S. Pati, M. Sinclair
  Young Architect Workshop (**YArch**), 2021