Fei Gao

https://fei-g.github.io feig@princeton.edu +1 609-933-7188
Engineering Quadrangle J301, 41 Olden St, Princeton, NJ 08544

Education

Ph.D. Candidate in Electrical and Computer Engineering *May 2019 - Present*
Princeton University || Department of Electrical and Computer Engineering || Princeton, NJ, USA
Advisor: Prof. David Wentzlaff

M.S. in Electrical Engineering *Sep 2017 - May 2019*
Princeton University || Department of Electrical and Computer Engineering || Princeton, NJ, USA

B.S. in Microelectronics *Sep 2013 - Jul 2017*
Tsinghua University || School of Integrated Circuits || Beijing, China
Rank: 2/25, GPA: 94.1/100, Minor in Business Administration

Work Experience

Research Intern *May 2022 - Aug 2022*
Microsoft || Cloud Accelerated Systems and Technologies (CAST) group

- Improved peak performance of the Lightweight Transport Layer (LTL) by adding out-of-order support and packet spraying.
- Finished RTL design, timing optimization, and scale-out experiments with more than 100 endpoints.

Publications

2022

Fei Gao, Ting-Jung Chang, Ang Li, Marcelo Orenes-Vera, Davide Giri, Paul J. Jackson, August Ning, Georgios Tziantzioulis, Joseph Zuckerman, Jinzheng Tu, Kaifeng Xu, Grigory Chirkov, Gabriele Tombesi, Jonathan Balkind, Margaret Martonosi, Luca Carloni, David Wentzlaff
"DECADES: A 67mm2, 1.25TOPS, 47 Giga Cache-Coherent 64-bit RISC-V Inst/s Heterogeneous Manycore SoC with 109 Tiles including Accelerators, Intelligent Storage, and eFPGA in 12nm FinFET"
in submission

Ting-Jung Chang, Ang Li, **Fei Gao**, Tuan Ta, Georgios Tziantzioulis, Yanghui Ou, Moyang Wang, Jinzheng Tu, Kaifeng Xu, Paul J. Jackson, August Ning, Grigory Chirkov, Marcelo Orenes-Vera, Shady Agwa, Xiaoyu Yan, Eric Tang, Jonathan Balkind, Christopher Batten, David Wentzlaff
"CIFER: A 12nm, 16mm2, 22-Core SoC with a 1541 LUT6/mm2, 1.92 MOPS/LUT, Fully Synthesizable, Cache- Coherent, Embedded FPGA"
in submission

Jinzheng Tu, **Fei Gao** and David Wentzlaff
"REEDRAM: A RISC-V-based, End-to-End Framework for In-memory Computation with Off-the-Shelf DRAM Chips"
in submission

Fei Gao, Georgios Tziantzioulis, and David Wentzlaff
"FracDRAM: Fractional Values in Off-the-Shelf DRAM"
In Proceedings of the 55nd International Symposium on Microarchitecture (MICRO-55)

Marcelo Orenes Vera, Aninda Manocha, Jonathan Balkind, **Fei Gao**, Juan Luis Aragn, David Wentzlaff and Margaret Martonosi
"Tiny but Mighty: Designing and Realizing Scalable Latency Tolerance for Manycore SoCs"
In Proceedings of the 49th International Symposium on Computer Architecture (ISCA '22)

2021

Georgios Tziantzioulis, Ting-Jung Chang, Jonathan Balkind, Jinzheng Tu, **Fei Gao** and David Wentzlaff
"OPDB: A Scalable and Modular Design Benchmark"
IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2021

2020

Jonathan Balkind, Ting-Jung Chang, Paul J. Jackson, Georgios Tziantzioulis, Ang Li, **Fei Gao**, Alexey Lavrov, Grigory Chirkov, Jinzheng Tu, Mohammad Shahrad, and David Wentzlaff
"OpenPiton at 5: A Nexus for Open and Agile Hardware Design"
IEEE Micro, July-August 2020, pp. 22-31, vol. 40.

Jonathan Balkind, Katie Lim, Michael Schaffner, **Fei Gao**, Grigory Chirkov, Ang Li, Alexey Lavrov, Tri M. Nguyen, Yaosheng Fu, Florian Zaruba, Kunal Gulati, Luca Benini, and David Wentzlaff
"BYOC: A "Bring Your Own Core" Framework for Heterogeneous-ISA Research"
In Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS20), March 2020, Lausanne, Switzerland

2019

Fei Gao, Georgios Tziantzioulis, and David Wentzlaff
"ComputeDRAM: In-Memory Compute Using Off-the-Shelf DRAMs"
In Proceedings of the 52nd International Symposium on Microarchitecture (MICRO-52), October 2019, Columbus, Ohio, USA
Honorable Mention in IEEE Micro Top Picks From the 2019 Computer Architecture Conferences (top 26 papers in the year)

Jonathan Balkind, Michael Schaffner, Katie Lim, Florian Zaruba, **Fei Gao**, Jinzheng Tu, David Wentzlaff, and Luca Benini
"OpenPiton+Ariane: The First Open-Source, SMP Linux-booting RISC-V System Scaling From One to Many Cores"
Workshops at Third Workshop on Computer Architecture Research with RISC-V (CARRV'19), June 2019, Phoenix, AZ, USA.

Research Experience

In-Memory Compute Using Off-the-Shelf DRAMs *Jul 2018 - Present*

- **First work demonstrating row copy and bit-wise logical AND/OR with unmodified, off-the-shelf, commercial DRAM**
- Proposed issuing DRAM commands with shortened timing interval to trigger a charge sharing on the bit-line, which can be utilized to build logical operation
- Characterized the capabilities and robustness of the proposed in-memory compute operations across multiple DDR3 modules from all major DRAM vendors
- Presented an algorithmic method to perform arbitrary computations based on non-inverting operations, since only AND and OR can be constructed in unmodified DRAM

Storing Fractional Value in Off-the-Shelf DRAMs *May 2021 - Present*

- Proposed primitive operations to store the same fractional value in the entire row, or store a mixture of zeros, ones and half value in a row, and demonstrate that with off-the-shelf DRAM
- Utilized half value to build PUFs (Physical Unclonable Functions) with off-the-shelf DRAM module, reaching the same throughput and reliability as the state-of-the-art, which in contrast requires hardware modification to the existing DRAM design
- Utilized half value to build majority-of-three from the charge sharing among four rows, which enables more DRAM modules to perform in-memory compute operations

DECADES Test Chip *Feb 2019 - Mar 2021*

8mmx8mm heterogeneous many-core processor taped-out with GF12 process, including 44 RISC-V Ariane tiles, 18 "intelligent storage" tiles, 18 Maxtrix-Multiplication/2d-Convolution accelerator tiles, and a controller tile. Designed for fast graph processing with decoupled access/execute architecture, near-memory bit-serial computing, and accelerators.

- **Led a six-people team working on the hierarchical back-end design**, including synthesis, floor-plan, place-and-route, static timing analysis (STA) and sign-off check
- Designed the cache system, based on the OpenPiton Network-on-Chip, to enable different coherence level among heterogeneous tiles
- Optimized the global clock network, and fine-tuned the timing budget across different hierarchies to push the maximum frequency to 1GHz and meet hold time constraints
- Conducted gate-level simulation after place-and-route to verify system functionality

CIFER Test Chip *Mar 2020 - Oct 2020*

4mmx4mm heterogeneous 8-core processor with an on-chip FPGA integrated, taped-out with GF12 process. Designed for real-time, fine-grained, on-chip FPGA configuration.

- Worked on system-level RTL integration, and verified the functionality with FPGA prototyping
- Designed the back-end of the cache system and the Network-on-Chip from synthesis to sign-off check
- Optimized critical paths in L2 cache and successfully pushed the maximum frequency to 1.1GHz

OpenPiton+Ariane: The RISC-V Hardware Research Platform *Apr 2019 - Present*

- Add internal instructions inside the cache system of OpenPiton to support RISC-V atomic operations, in order to integrate OpenPiton with RISC-V cores
- Maintained the open-source project, resolving questions and pull-requests from the community.

Patents

Fei Gao, David Wentzlaff, and Georgios Tziantzioulis, **"Storing fractional values in DRAM and applicaions"**, in application.

David Wentzlaff, **Fei Gao**, and Georgios Tziantzioulis, **"System and method for in-memory compute"**, US Patent No. 11,043,259, issued June 22, 2021.

Awards & Honors

Honorable Mention in IEEE Micro Top Picks *2020*

Yan Huo *94 Graduate Fellowship in Electrical Engineering *Sep 2019*

TP-LINK Scholarship *Oct 2016*

Scholarship of Academic Excellence, Tsinghua University *2014, 2015*

First Prize of the 31st National Undergrad. Physics Contest *Dec 2014*

Teachings & Tutorials

ELE/COS 206 Contemporary Logic Design *Fall 2022*
Teaching Assistant, Princeton University

"OpenPiton with RISC-V Cores: A Hands-On Tutorial with the Open Source Manycore Processor" *Oct 2019*
Tutorial in MICRO-52, Columbus, Ohio.

"OpenPiton+Ariane: The RISC-V Hardware Research Platform" *Jun 2019*

Tutorial in ISCA/FCRC 2019, Phoenix, Arizona.

"OpenPiton+Ariane: The RISC-V Hardware Research Platform" *Jun 2019*
Tutorial in Week of Open Source Hardware, ETH Zurich, Switzerland.

ELE/COS 475 Computer Architecture *Fall 2018*
Teaching Assistant, Princeton University

Professional Skills

Language: Native Speaker of Mandarin, English.

Programming: Verilog HDL, Python, C/C++, Matlab, LaTeX, bash, tcl.