Yue Liang

Email: liangy@cqu.edu.cn Add.: Chongqing China

EDUCATION

Chongqing University *M.S.* in Computer Technology **GPA:** 3.5/4 *Sep 2021 - Jun. 2024*

Chongqing University *B.S.* in Computer Science **GPA:** 3.17/4 *Sep 2017 - Jun 2021*

PUBLICATIONS

[1] ***Yue Liang***, Di Mou et al. DISC: Exploiting Data Parallelism of Non-Stencil Computations on CGRAs via Dynamic Iteration Scheduling, in *2024 ACM/IEEE International Conference on Computer-Aided Design*. (ICCAD) 2024.

[2] Minghao Tian, ***Yue Liang*** et al. CoSpMV: Towards Agile Software and Hardware Co-design for SpMV Computation, *IEEE Transactions on Computers*. (under review)

PROJECT EXPERIENCE

Research on Software and Hardware Co-design for SpMV Computation[2] *Jul 2021 - Jan2022*

- Addressed challenges in the inflexibility of unified architecture, resolving issues such as poor resource utilization or poor bandwidth utilization for different SpMV kernels after software preprocessing.
- Designed a scalable compressed data format as the bridge between software and hardware such that design space could be significantly reduced, and constructed a scalable and highly-pipelined hardware template that can be instantly instantiated to obtain the corresponding accelerators.
- Implemented an agile software and hardware co-design framework for SpMV acceleration on FPGA. With efficient design space exploration, a high-performance and resource-efficient accelerator could be quickly achieved for a specific SpMV kernel. The paper on this work is currently under review by *IEEE Transactions on Computers*, and I am the second author.

Research on Exploiting Data Parallelism of Irregular Programs on CGRA[1] *Jan 2022 - Nov 2023*

- Addressed challenges in memory partitioning for non-stencil computation with irregular memory access patterns on SRAM, resolving issues such as high memory access conflicts, low computational parallelism, and poor data reuse.
- Designed a dynamic iteration scheduling method that can select conflict-free iterations from an iteration buffer for parallel data access to dynamically exploit data parallelism from non-stencil computations.
- Implemented CGRA with dynamic iteration scheduling (DISC) and verified that it provides significant performance improvement with modest hardware overhead. A paper about this work was accepted by *ICCAD 2024*.

Research on Accelerating Indirect Memory Access by Near-Data Processing *May 2024 - Present*

- Addressed challenges in indirect memory accesses of sparse data formats, resolving issues such as little locality for caches and poor performance of conventional prefetchers.
- By offloading operations that cause cache misses to buffer chip of DIMM using near-memory processing, and reducing bandwidth waste through memory access merging, locality is improved.
- Still in the design and implementation phase, this project has provided an opportunity to learn about near-data processing and related tools, such as Gem5 and Ramulator2.

SKILLS

Programming: C/C++, Python, Chisel, Verilog.

Tool: Systemc, DC, Vivado, Gem5, Zsim, Ramulator2.