Dedong Xie

Website: ddxxdd-code.github.io Email: dedongx@cs.washington.edu LinkedIn: dedong-xie-547a501a8 GitHub: github.com/ddxxdd-code

EDUCATION

University of Washington PhD

University of Toronto

Honours Bachelor of Science

- Majors in computer science and mathematics, minor in statistics
- Cumulative GPA (cGPA): 3.99/4.0
- 90 or above in all Computer Science courses
- 96 or above in all programming courses

PUBLICATION

- Dedong Xie, Theano Stavrinos, Kan Zhu, Simon Peter, Baris Kasikci, and Thomas Anderson. "Can Storage Devices be Power Adaptive?" *Proceedings of the 16th ACM Workshop on Hot Topics in Storage* and File Systems. HotStorage '24. Santa Clara, CA, USA: Association for Computing Machinery, 2024, pp. 47–54. DOI: 10.1145/3655038.3665945.
- [2] Dedong Xie, Zhen Jia, Zili Zhang, and Xin Jin. "Optimizing Half Precision Winograd Convolution on ARM Many-Core Processors". Proceedings of the 13th ACM SIGOPS Asia-Pacific Workshop on Systems. APSys '22. Virtual Event, Singapore: Association for Computing Machinery, 2022, pp. 53–60. DOI: 10.1145/3546591.3547529.

Research Experience

Research Assistant

Syslab & EfesLab. Paul G. Allen School of Computer Science and Engineering, University of Washington. Advisors: Prof. Simon Peter and Prof. Baris Kasikci

- Research project: power adaptive storage system.
- Conducted a power measurement study of power consumption patterns of datacenter storage devices including HDDs and SSDs.
- Proposed how the datacenter storage devices can work to make storage system power adaptive, the ability to be able to adjust to dynamically changing power budgets.
- Published a first-author paper in ACM Workshop on Hot Topics in Storage and File Systems (HotStorage 2024).
- <u>slides</u> of my presentation in HotStorage 2024.
- Follow-up work on designing a power adaptive storage server.
- Designed and implemented a power adaptive storage all-flash server for disaggregated storage.
- Thoroughly studied the power dynamic range and performance of the system over various microbenchmarks and application benchmarks of file systems and database system.
- Achieved widest power dynamic range (94% of marginal power over idle power) and superior performance (about $10 100 \times$) compared with best performing baselines in dimensions of power dynamic range and performance.
- Paper under submission.

Seattle, WA, USA Sep. 2023–Jun. 2028 (expected)

> Toronto, ON, Canada Jun. 2023

> > Sep. 2023–Present

Research Assistant

SysNet Lab. Department of Computer Science. University of Toronto. Supervisor: Prof. Eyal de Lara

- Participated in IBM CAS Canada project 1153 Reducing JVM memory costs in the cloud https://www-40.ibm.com/ibm/cas/canada/projects?projectId=1153
- $-\,$ Sole developer of the run-time memory profiler of OpenJ9 JVM JIT-compiler.
- Proposed instrumenting dynamic memory allocation logger in OpenJ9's memory allocator.
- Proposed visual illustration of memory usage over time to find source of peak usage.
- Implemented the memory allocation logger, post-process pipeline, and visualizer with 3,000 lines of code in C++ and Python.
- Found external fragmentation and late release of memory to be main causes of memory inefficiencies.
- Proposed using program slicing to identify memory that could have a shorter lifetime.
- Currently working on identifying the scope of each allocated memory.
- Video of my presentation, and slides of the presentation.

Research Intern

AI Lab. Amazon Web Services (AWS).

Supervisors: Dr. Zhen Jia (AWS) and Prof. Xin Jin (Peking University)

- Sole developer of HAWC, a half-precision Winograd convolution system for Amazon Graviton-2 ARM architecture chips.
- Proposed customized memory layout for Amazon Graviton-2 chips, ARM-specific matrix multiplication kernel generator, and minimal multi-threading scheduler to accelerate Winograd convolution.
- Implemented 3000 lines of code using C++ and ARM assembly.
- Studied and adapted baseline systems for comparison.
- Extensively tested on a variety of representative convolution layers against state-of-the-art solutions.
- Achieved on average $11 \times$ and up to $28 \times$ speedup.
- Generated matrix multiplication kernels exploited up to 89% of theoretical maximum TFLOPS of the hardware.
- Published a first-author paper in ACM Asia-Pacific Workshop on Systems (APSys 2022).
- -<u>Video</u> and <u>slides</u> of my presentation in APSys 2022.

AWARDS AND ACHIEVEMENTS

•	Dean's List Scholar	Jun.	2021, Jun.	2022, Ju	n. 2	2023
	- Faculty of Arts and Science, University of Toronto					
•	Dr. James A. & Connie P. Dickson Scholarship In Science & Mathematic	cs	Sept.	2022, Sej	ot. 2	2023
	- "Given to the best students enrolled in science and mathematics programs."					
	- University College, University of Toronto					
•	Department of Computer Science Undergraduate Research Award			Μ	ay 2	2022
	- Department of Computer Science, University of Toronto					
•	Galois Awards in Mathematics			O	ct. 2	2021
	– "Given to the best students enrolled in a mathematics specialist program."					
	- University College, University of Toronto					
•	The Faculty of Engineering Dean's Award				4	2020
	- "For the best performance in year 1, 2 or 3."					
	- University of New South Wales					

Jun. 2021–Jul. 2022

• COMP1511 (Programming Fundamentals) Hall of Fame

- "A list of students who have achieved great distinction and honour by completing large amounts of extra work."
- http://web.cse.unsw.edu.au/~cs1511/hall_of_fame/
- COMP1511 Teaching Team, University of New South Wales

SKILLS

- LANGUAGES
- **Programming languages:** C/C++, Python, Java, Racket, and Haskell.
- Assembly programming: MIPS, ARM Aarch64, Intel x86 instruction sets.
- Database management systems: Microsoft Access, MySQL.
- Mathematical computation and data analysis: R, Mathematica, MATLAB.
- English: proficient
- **IELTS:** Overall 8.0 (Aug. 2019)
- Chinese: native speaker