

Kai Wu

+1 517-763-1599 | kwu42@ucmerced.edu | kaikylewu.com

RESEARCH INTERESTS

My research interests span Computer System, Heterogeneous computing and High Performance Computing, with a focus on building runtime, programming model, and system (OS and file system) support for emerging *non-volatile memory*.

EDUCATION

University of California, Merced

Ph.D., in Electrical Engineering and Computer Science

Merced, CA

June 2016 – Present

Michigan State University

M.S., in Computer Science and Engineering

East Lansing, MI

August 2014 – May 2016

Harbin Normal University

B.S., in Digital Media Technology

Harbin, China

September 2010 – June 2014

REFEREED CONFERENCE PUBLICATIONS

[FAST'21] Kai Wu, Jie Ren, Ivy Peng and Dong Li. “ArchTM: Architecture-Aware, High Performance Transaction for Persistent Memory”. In 19th USENIX Conference on File and Storage Technologies, 2021.

[HPCA'21] Jie Ren, Jiaolin Luo, Kai Wu, Minjia Zhang, Hyeran Jeon and Dong Li. “Efficient Tensor Migration and Allocation on Heterogeneous Memory Systems for Deep Learning”. In The 27th IEEE International Symposium on High-Performance Computer Architecture, 2021.

[PACT'20] Kai Wu, Ivy B. Peng, Jie Ren and Dong Li. “Ribbon: High Performance Cache Line Flushing for Persistent Memory”. In 29th International Conference on Parallel Architectures and Compilation Techniques, 2020.

[IPDPS'20] Ivy B. Peng, Kai Wu, Jie Ren, Dong Li and Maya Gokhale. “Demystifying the Performance of HPC Scientific Applications on NVM-based Memory Systems”. In 34rd IEEE International Parallel and Distributed Processing Symposium, 2020.

[CLUSTER'20] Jie Ren, Kai Wu and Dong Li. “Exploring Non-Volatility of Non-Volatile Memory for High Performance Computing Under Failures”. In 22th IEEE Cluster Conference, 2020.

[SC'18] Kai Wu, Jie Ren and Dong Li. “Runtime Data Management on Non-Volatile Memory-based Heterogeneous Memory for Task-Parallel Programs”. In 30th ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis, 2018.

[ICPP'18] Kai Wu, Wenqian Dong, Qiang Guan, Nathan Debardeleben and Dong Li. “Modeling Application Resilience in Large Scale Parallel Execution”. In 47th International Conference on Parallel Processing, 2018.

[SC'17] Kai Wu, Yingchao Huang and Dong Li. “Unimem: Runtime Data Management in Non-Volatile Memory-based Heterogeneous Main Memory”. In 29th ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis, 2017.

[CLUSTER'17] Shuo Yang, Kai Wu, Yifan Qiao, Dong Li and Jidong Zhai. “Algorithm-Directed Crash Consistence in Non-Volatile Memory for HPC”. In 19th IEEE Cluster Conference, 2017.

[NAS'17] Wei Liu, Kai Wu, Jialin Liu, Feng Chen and Dong Li. “Performance Evaluation and Modeling of HPC I/O on Non-Volatile Memory”. In 12th International Conference on Networking, Architecture, and Storage, 2017.

RESEARCH EXPERIENCE

Graduate Student Researcher

University of California Merced

June 2016 – Present

Advisor: Prof. Dong Li

- Ph.D. Dissertation: Runtime Data Management on Non-Volatile Memory-Based High Performance Systems, invited talk in SC'20 Doctoral Showcase
- Data placement and migration on non-volatile memory-based heterogeneous memory [SC'17, SC'18, HPCA'21]
- System resilience and reliability on non-volatile memory [FAST'21, PACT'20, CLUSTER'20, MCHPC'18]
- Performance characterization and system optimization for non-volatile memory devices (e.g, Intel Optane DC Persistent Memory and Optane SSD) [IPDPS'20, NAS'17]

Research Intern

Bytedance (System Infrastructure Lab), USA

May 2020 – November 2020

- Worked on hybrid transactional/analytical processing (HTAP) system; designed an adaptive data compaction/compression algorithm; lead the performance benchmarking and optimization of HTAP store; coordinated with other team members to complete prototype development of a new large-scale HTAP system.

Research Intern

May 2018 – August 2018

Lawrence Livermore National Laboratory, USA

Mentor: Dr. Maya Gokhale

- Worked on memory caching mechanism; explored pre-fetch and eviction optimizations using the user faulted approach for efficient access to persistent memory for data-intensive applications such as out-of-core sorting, graph application (i.e., process Graph500) and asteroid detection application (i.e., process imaging data sets of the sky generated by the optical cameras like the Dark Energy camera in Chile); achieved up to 2.5x speedups [MCHPC'19]

Research Intern

May 2017 – August 2017

Los Alamos National Laboratory, USA

Mentor: Dr. Nathan DeBardeleben, Dr. Qiang Guan

- Worked on system resilience and reliability for large-scale parallel HPC applications; built an analytical model to predict the fault injection result of the application running in large-scale based on fault injection results of the application running in small-scale and serial [ICPP'18, SC'17 poster]

REFEREED WORKSHOP PUBLICATIONS

[MCHPC'19] Ivy B. Peng, Marty McFadden, Eric Green, Keita Iwabuchi, **Kai Wu**, Dong Li, Roger Pearce and Maya Gokhale. "UMap: Enabling Application-driven Optimizations for Page Management". In Workshop on Memory Centric High Performance Computing held in conjunction with SC'19.

[MCHPC'18] Jie Ren, **Kai Wu** and Dong Li. "Understanding Application Recomputability without Crash Consistency in Non-Volatile Memory". In Workshop on Memory Centric High Performance Computing held in conjunction with SC'18.

REFEREED JOURNAL PUBLICATIONS

[JCST] **Kai Wu** and Dong Li. "Unimem: Runtime Data Management on Non-Volatile Memory-based Heterogeneous Main Memory for High Performance Computing". The International Journal of Computer Science and Technology, Special Section on Memory-Centric System Research for HPC, 2021.

POSTERS

[SC'20] **Kai Wu** and Dong Li. "Runtime Data Management on Non-Volatile Memory-based High Performance Systems". Doctoral Showcase in 31th ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis, 2020.

[SC'20] Jiaolin Luo, Luanzheng Guo, Jie Ren, **Kai Wu** and Dong Li. "Enabling Faster NGS Analysis on Optane-based Heterogeneous Memory". Poster in 31th ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis, 2020.

[NVMW'19] **Kai Wu**, Jie Ren and Dong Li. "Architecture-Aware, High Performance Transaction for Persistent Memory". In 10th Workshop on Non-Volatile Memory, 2019.

[NVMW'19] **Kai Wu**, Jie Ren and Dong Li. "Runtime Data Management on Non-Volatile Memory-based Heterogeneous Memory for Task-Parallel Programs". In 10th Workshop on Non-Volatile Memory, 2019.

[NVMW'19] Jie Ren, **Kai Wu** and Dong Li. "EasyCrash: Exploring Non-Volatility of Non-Volatile Memory for High Performance Computing Under Failures". In 10th Workshop on Non-Volatile Memory, 2019.

[NVMW'18] **Kai Wu** and Dong Li. "Unimem: Runtime Data Management in Non-Volatile Memory-based Heterogeneous Main Memory". In 9th Workshop on Non-Volatile Memory, 2018.

[NVMW'18] Jie Ren, **Kai Wu** and Dong Li. "Algorithm-Directed Crash Consistence in Non-Volatile Memory for HPC". In 9th Workshop on Non-Volatile Memory, 2018.

[SC'17] **Kai Wu**, Qiang Guan, Nathan DeBardeleben and Dong Li. "Characterization and Comparison of Application Resilience for Serial and Parallel Executions". In 29th ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis, 2017.

WORK UNDER SUBMISSION

Jie Ren, **Kai Wu** and Dong Li. "Relaxation of Crash Consistency on Non-Volatile Main Memory for HPC Under Failures".

Jie Ren, Jiaolin Luo, Ivy Peng, **Kai Wu** and Dong Li. "Performance Analysis and Optimization of Electromagnetic Particle-In-Cell Method on Emerging Persistent Memory-based Platform".

PREPRINTS

Kai Wu, Frank Ober, Shari Hamlin, Qiang Guan and Dong Li, “Early Evaluation of Intel Optane Non-Volatile Memory with HPC I/O Workloads”. arXiv:1708.02199v2, 2017.
Yingchao Huang, **Kai Wu** and Dong Li. “High Performance Data Persistence in Non-Volatile Memory for Resilient High Performance Computing”. arXiv:1705.00264, 2017.

SERVICES

Technical Program Committee: Eurosys’21 (shaow PC), HPCC’20
Reviewer: IPDPS’21, IEEE Access, NPC’20, CLUSTER’20, NPC’19, ICPP’19, SC’18, IPDPS’17, CLUSTER’17, HPCC’17, NAS’17
Student Volunteer: SC’20, SC’19, SC’18, SC’16
Graduate Student Representative, UC Merced EECS, January 2020 - December 2020

AWARDS

Conference Student Grant: OSDI’20, NVMW’20, SC’19, NVMW’19, SC’18, OSDI’18, ASPLOS’18, NVMW’18, CLUSTER’17, NVMW’17, SC’16
UC Merced Graduate Travel Fellowship: 2018 & 2020
UC Merced Bobcat Graduate Research Fellowship: 2017
China National Scholarship: 2013
First Class Scholarship of Harbin Normal University: 2011 & 2012 & 2013 & 2014

PROGRAMMING SKILLS

C/C++, Java, Python, Fortran, JavaScript
Linux kernel programming
PMDK(Persistent memory programming), MPI, OpenMP, GPU(CUDA)

RELEVANT TEACHING EXPERIENCE

Teaching Assistant

University of California Merced

- CSE15 - Discrete Mathematics 2020 Spring
- CSE20 - Introduction to Computing I (Java Programming) 2016 Summer & 2018 Fall & 2019 Fall
- CSE179 - Parallel Computing 2020 Spring

REFEREES

Available on request