

Kai Wu

+1 517-763-1599 | kaiwu0126@gmail.com | kaikylewu.com

SUMMARY

- 5+ years of research and work experience in designing and developing high-performance and scalable systems
- Research interests: Computer systems, Programming models, Runtimes, High Performance Computing (HPC)
- Specifically, I have worked on:
 - Memory management on heterogeneous memory systems for HPC and AI
 - Cache and graph analytic platform on emerging memory/storage hardware (e.g., persistent memory)
 - Next-gen cloud memory systems (e.g., disaggregated memory)
- Programming skills: C, C++, Java, Python, Fortran, SQL, Shell scripting and JavaScript

EDUCATION

University of California, Merced <i>Ph.D., in Electrical Engineering and Computer Science</i>	Merced, CA June 2016 – Dec 2020
Michigan State University <i>M.S., in Computer Science and Engineering</i>	East Lansing, MI August 2014 – May 2016
Harbin Normal University <i>B.S., in Digital Media Technology</i>	Harbin, China September 2010 – June 2014

EXPERIENCE

Researcher/Engineer <i>ByteDance/TikTok (System Infrastructure Lab), Mountain View, CA</i>	Feb 2021 – Present
<ul style="list-style-type: none">• Collaboratively designed and developed a new distributed graph analytic platform, supporting hundreds of trillions of edge graph processing and mining, and providing a vertex-centric programming interface; led the design and implementation of a high-performance, large-capacity, fault-tolerant graph data store subsystem on PMEM• Collaboratively designed and developed a new multi-tier cache system (based on DRAM, PMEM and SSD) to support multiple data services (e.g., OLTP/OLAP/serving workloads); led the design and implementation of PMEM-aware cache (e.g., log-structure allocator, NUMA optimization, and concurrency optimization)• Led investigation of memory disaggregation technologies (i.e., based on CXL and RDMA) for internal use cases	
Graduate Student Researcher <i>University of California Merced, Merced, CA</i>	June 2016 – Dec 2020 <i>Advisor: Prof. Dong Li</i>
<ul style="list-style-type: none">• Ph.D. Dissertation: Runtime Data Management on Non-Volatile Memory(NVM)-Based High Performance Systems, invited talk in SC'20 Doctoral Showcase• Designed and implemented data/page management mechanisms on heterogeneous memory systems for (MPI/OpenMP-based) HPC and deep learning workloads [HPCA'21, SC'18, SC'17]• Designed and implemented runtimes and programming models to provide fault tolerance/crash consistency for data-intensive workloads on NVM-based systems [FAST'21, PACT'20, CLUSTER'20, MCHPC'18]• Characterized and optimized HPC applications on NVM devices (e.g, Intel Optane DC Persistent Memory and Optane SSD) [IPDPS'20, NAS'17]	
Research Intern <i>ByteDance/TikTok (System Infrastructure Lab), Mountain View, CA</i>	May 2020 – November 2020
<ul style="list-style-type: none">• Worked on a new hybrid transactional/analytical processing (HTAP) system; coordinated with other team members to complete prototype development; designed and implemented adaptive data compaction and garbage collection modules [InSub'22]	
Research Intern <i>Lawrence Livermore National Laboratory, Livermore, CA</i>	May 2018 – August 2018 <i>Mentor: Dr. Maya Gokhale</i>

- Worked on application specific page cache management; explored caching optimizations using the user faulted approach for efficient access to a large file (on NVM) for data-intensive applications such as out-of-core sorting, graph processing and asteroid detection (i.e., process imaging data sets of the sky generated by the optical cameras like the Dark Energy camera in Chile) applications; achieved up to 2.5x speedups [MCHPC'19]

Research Intern

May 2017 – August 2017

Los Alamos National Laboratory, Los Alamos, NM

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]

SELECTED REFEREED PUBLICATIONS

- [InSub'22] Jianjun Chen, Yonghua Ding, Ye Liu, Fangshi Li, Li Zhang, Mingyi Zhang, Kui Wei, Lixun Cao, Dan Zou Yang Liu, Lei Zhang, Rui Shi, Wei Ding, **Kai Wu**, Shangyu Luo, Jason Yang Sun and Yuming Liang. "ByteHTAP: ByteDance's HTAP System with Fresh and Consistent Data". In Submission to 48th International Conference on Very Large Data Bases, 2022.
- [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.
- [ICS'21] 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". In 35th International Conference on Supercomputing, 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.
- [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, 2019.
- [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.

PATENTS

- [In application] Hongzhi Chen, Peng Lin, Lin Ma, **Kai Wu**, Haiyang Shi, Xiao Liu, Wei Xu, Dawei Gong, Shuai Zhang and Jianjun Chen. "Serverless and scalable graph data processing and mining with hierarchical persistent storage".
- [In application] Jianjun Chen, Yonghua Ding, Ye Liu, Fangshi Li, Li Zhang, Mingyi Zhang, Kui Wei, Wei Ding, **Kai Wu** and Jason Yang Sun. "Storage Engine for Hybrid Data Processing".

SERVICES

Program Committee: Eurosys'21 (shaow PC), HPCC'20 & 2021

Reviewer: DOE SBIR/STTR, IEEE Transactions on Parallel and Distributed Systems, Future Generation Computer Systems, IEEE Access, IEEE IPDPS'21, NPC'20, IEEE CLUSTER'20, NPC'19, ICPP'19, ACM/IEEE SC'18, IEEE IPDPS'17, IEEE CLUSTER'17, IEEE HPCC'17, IEEE 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