Kai Wu

CONTACT INFORMATION

5200 N Lake Road SE2 Room 213C, Merced, CA, 95348
517-763-1599 kwu42@ucmerced.edu http://kaikylewu.com

RESEARCH INTERESTS

My research broadly falls into general areas of High Performance Computing (Large-Scale Parallel Systems). Specially, I focus on the following areas: (i) Resource Management in Heterogeneous Computing (Non-volatile memory); (ii) Parallel programming models and runtime; (iii) Performance optimization and modeling; (iv) Resilience and Consistency.

EDUCATION

University of California, Merced, CA Jun 2016 – Now
Ph.D., in Electrical Engineering and Computer Sciences
Advisor: Dong Li

Michigan State University, East Lansing, MI Aug 2014 - May 2016
M.S., in Computer Science and Engineering
Advisor: Yiying Tong

Harbin Normal University, Harbin, CHINA Aug 2010 - Jun 2014
B.S., Digital Media Technology

EXPERIENCE

Lawrence Livermore National Laboratory
Research Intern May 2018 – Aug 2018
Mentor: Maya B Gokhale

- **Project: Pre-fetch and eviction policies design for UMap**
  Design, implement and evaluate pre-fetch and eviction optimizations using the user faulted approach for efficient access to mapped files in persistent memory.

Los Alamos National Laboratory
Research Intern May 2017 – Aug 2017
Mentor: Nathan DeBardeleben and Qiang Guan

- **Project: Characterizing and Modeling Application Resilience Difference Between Serial and Parallel Executions (See ICPP'18 paper and SC'17 poster)**
  We study a new methodology to evaluate application resilience in the large scale. Instead of injecting faults into the application in the large scale, we use fault injection in a small-scale execution and the serial execution to model and predict application resilience in the large scale.

UC Merced
Graduate Student Researcher with Prof. Dong Li Jun 2016 – Now

- **Project 1: Runtime Data Management on Non-Volatile Memory-based Heterogeneous Memory for Task-Parallel Programs (see SC'18 Paper)**
  We introduce a runtime system to enable efficient data management (data migration between NVM and DRAM) on NVM-based HMS. Our runtime leverages persistence of NVM and task parallelism to do

checkpoint, enabling resiliency for task-parallel programs while significantly reducing checkpoint overhead.

- **Project 2: Runtime Data Placement of MPI-based Parallel Applications on Heterogeneous Memory (NVM/DRAM) System (See SC'17 paper)**
  We introduce a lightweight runtime solution that automatically and transparently manage data placement on HMS without the requirement of hardware modifications and disruptive change to applications. Leveraging online profiling and performance models, the runtime characterizes memory access patterns associated with data objects and minimizes unnecessary data movement.
- **Project 3: High Performance Data Persistence in Non-Volatile Memory for Resilient High Performance Computing (remove checkpoint)**
  We explore how to build resilient HPC with emerging NVM. Then we introduce new schemes and optimization techniques, and explore how to leverage high performance and non-volatility of NVM to establish a consistent data status as the traditional checkpoint mechanism.
- **Project 4: Algorithm-Directed Crash Consistence in Non-Volatile Memory for HPC (See Cluster'17 paper)**
  We introduce an algorithm-based method to establish crash consistence in NVM for HPC applications. We slightly extend application data structures or sparsely flush cache blocks, which introduce ignorable runtime overhead.
- **Project 5: Performance Implications of Persistent Memory on HPC Applications (See NAS'17 paper)**
  We study the implication of NVM (as a block device) on HPC applications. We focus on measuring and comparing the different performance of HDD, SSD and PMBD (NVM simulator) in three directions: POSIX I/O vs MPI I/O, Independent I/O vs Collective I/O, and page cache.
- **Project 6: Early Evaluation of Intel Optane Non-Volatile Memory with HPC I/O Workloads**
  We analyze the performance of I/O intensive HPC applications with Optane as a block device from three perspectives: (1) basic read and write bandwidth of Optane, (2) a performance comparison study between Optane and HDD, including checkpoint workload, MPI individual I/O vs. POSIX I/O, and MPI individual I/O vs. MPI collective I/O, and (3) the impact of Optane on the performance of a parallel file system, PVFS2.
- Others: I am maintaining a computing cluster ALPHA.

UC Merced
Teaching Assistant Jun 2016 – Aug 2016

- CSE 20 - Introduction to Computing I

PUBLICATION

[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.

[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.

[SC'17] **Kai Wu**, Qiang Guan, Nathan DeBardeleben and Dong Li, "**Characterization and Comparison of Application Resilience for Serial and Parallel Executions**". Poster 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 Persistent Memory for High Performance Computing**". In 19th IEEE Cluster Conference.

[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.

[TR] **Kai Wu**, Frank Ober, Shari Hamlin, Qiang Guan and Dong Li, "**Early Evaluation of Intel Optane Non-Volatile Memory with HPC I/O Workloads**". Technical Report, PASA Lab.

Kai Wu, Yingchao Huang and Dong Li. **High Performance Data Persistence in Non-Volatile Memory for Resilient High Performance Computing (In submission)**

PROFESSIONAL ACTIVITES

External reviewers: SC'18, IPDPS'17, CLUSTER'17, HPCC'17, NAS'17, etc.
Student Volunteer SC'18, SC'16

AWARDS

Student travel grant for ASPLOS'18	2018
Student travel grant for NVMW'18	2018
NSF student travel fund for Cluster'17	2017
UC Merced Bobcat scholarship	2017
Student Travel Grant for NVMW2017	2017
ACM/IEEE Travel Grant for SC'16	2016
First-Prize, 'LanQiao Cup' National Software & Information Technology Professional Talents Competition, C/C++ group	2013
Third-Prize, International Mathematics and Computer Programming Olympiad of RF and PCR university students	2013
Bronze Medal, ACM/ICPC International Collegiate Programming Contest China Tonghua Invitational Contest	2013
Third-Prize, ACM/ICPC International Collegiate Programming Contest China Hei Longjiang Province Contest	2012
China National Scholarship	2013
First Class Scholarship of Harbin Normal University	2011- 2014

SKILLS

C/C++, Python, Java, Fortran
MPI, OpenMP, GPU (CUDA)
Linux kernel programming
Hadoop, Pig, Hive, Spark, Weka and AWS.
Web Development (PHP and JavaScript).

REFERENCE

Dong Li
Assistant Professor
University of California, Merced
Email: dli35@ucmerced.edu
Maya B Gokhale
Distinguished Member of Technical Staff
Lawrence Livermore National Laboratory
Email: maya@llnl.gov
Nathan Debardeleben
Scientist
Los Alamos National Laboratory
Email: ndebard@lanl.gov
Qiang Guan
Assistant Professor
Kent State University
Email: qguan@kent.edu