

Feng Li

755 Lockefield ST APT C, Indianapolis, IN 46202

<https://fenggli.github.io/> | 317-772-3886 | li2251@purdue.edu

Education

Purdue University – Ph.D. Indianapolis, IN 2015 – 2022 (expected)

Research Areas: in-situ workflow, high-performance computing, scientific workflow and its applications.

Huazhong University of Science and Technology – B.Eng. Wuhan, China 2011 – 2015

GPA: 3.8/4.0

Work Experience

Research Intern – IBM Research Almaden, CA May 2019 – Aug 2019

- Designed and implemented a unified file system interface (KVFS) for multiple key-value store backends, so that file operations are translated into key-value store put/get operations.
- Used FUSE to implement KVFS, and designed mechanisms to handle the mappings between file abstractions and data objects.
- Code base in IBM Comanche: <https://github.com/IBM/comanche/tree/unstable/src/fuse>.

Research Intern – IBM Research Almaden, CA May 2018 – Aug 2018

- Designed and implemented NVMeStore, which is a key-value store that uses NVMe SSDs as data storage while keeping critical metadata in persistent memory.
- Added continuous integration for the Comanche project using Linux containers and Travis-CI.

Research Intern – IBM Research Almaden, CA May 2017 - Aug 2017

- Designed and implemented an NVMe-backed light-weight memory service – CO-PAGER (Collaborative Paging).
- CO-PAGER captures virtual memory page faults and performs paging operations on NVMe SSDs using fast userspace I/O.
- A paper [3] is published in HP3C'19.

Skills & Abilities

- Programming languages (C/C++/Java/Scala/Python)
- Performance analysis tools (Intel Vtune, Linux Perf, TAU)
- Cloud/container solutions (Google Cloud Platform, Amazon AWS, Openstack, Docker, Kubernetes)
- Mathematical-modelling tool (CPLEX)
- Big Data/Deep learning frameworks (Tensorflow/Pytorch/Apache Spark)
- Storage related: Redis, spdk, dpdk, pmdk, fuse, fio
- CI/Build tools(cmake, Apache Maven, Travis CI, Google Gtest)

Selected Research Projects

(More projects are described on my homepage at <https://fengggli.github.io/>)

Cross-environment in-situ workflow management

Jan 2020 — Now

- Design, and prototype a framework that can be used to launch in-situ workflows across HPC and Cloud systems.
- Formalize the scheduling problem for in-situ workflows, design and evaluate heuristic-based algorithms to improve workflow metrics such as throughput and latency.
- Part of this work is published in PASC'21 [2].

Performance Analysis of In-situ Methods in HPC

April 2017 — Nov 2017

- Design and implement the first-of-its-kind in-situ workflow benchmark that supports flexible configurations of workflows, transport libraries, and execution environments.
- Artifacts are available at: <https://github.com/IUPU-HPC/workflow-bench>.
- This work contributes to part of the HPDC'18 paper [6].

Services & Awards

- Reviewer: BIGCOM'16, PEARC'18, BDCAT'18, PEARC'21.
- Best Student Paper Award: PEARC'17.

Publications

- [1] **Feng Li**, Ranran Chen, Yuankun Fu, Fengguang Song, Yao Liang, Isuru Ranawaka, Sudhakar Pamidighantam, Daniel Luna, and Xu Liang. 2021. Accelerating complex modeling workflows in CyberWater using on-demand HPC/Cloud resources. In Proceedings of 2021 IEEE eScience21.
- [2] **Feng Li**, Dali Wang, Feng Yan, and Fengguang Song. 2021. X-composer: enabling cross-environments in-situ workflows between HPC and cloud. In Proceedings of ACM PASC '21.
- [3] **Feng Li**, Daniel G. Waddington, and Fengguang Song. 2019. Userland CO-PAGER: boosting data-intensive applications with non-volatile memory, userspace paging. In Proceedings of ACM HP3C'19.
- [4] **Feng Li** and Fengguang Song. 2019. Building a scientific workflow framework to enable real-time machine learning and visualization. Concurrency Computat Pract Exper. 2019; 31:e4703.
- [5] Yuankun Fu, **Feng Li**, Fengguang Song, and Luoding Zhu. 2018. Designing a Parallel Memory-Aware Lattice Boltzmann Algorithm on Manycore Systems. In Proceedings of SBAC-PAD'18.
- [6] Yuankun Fu, **Feng Li**, Fengguang Song, and Zizhong Chen. 2018. Performance analysis and optimization of in-situ integration of simulation with data analysis: zipping applications up. In Proceedings of ACM HPDC'18.
- [7] **Feng Li** and Fengguang Song. 2017. A Real-Time Machine Learning and Visualization Framework for Scientific Workflows. In Proceedings of ACM PEARC'17.
- [8] Xiao Bian, **Feng Li**, and Xia Ning. 2016. Kernelized Sparse Self-Representation for Clustering and Recommendation. In Proceedings of SIAM SDM'16.
- [9] Dan Luo, Jiguang Wan, Yifeng Zhu, Nannan Zhao, **Feng Li**, and Changsheng Xie. 2016. Design and Implementation of a Hybrid Shingled Write Disk System. IEEE TPDS, Volume:27, Issue:4 (April 2016), 1017-1029.