Fahim Tahmid Chowdhury

Mobile: +1(786)-406/2617
Skype: fahim.tahmid.chowdhury
E-mail: fchowdhu@cs.fsu.edu
Website: http://ww2.cs.fsu.edu/~fchowdhu
LinkedIn: http://www.linkedin.com/in/fahimtahmidchowdhury

EDUCATION	
	Florida State University, Tallahassee, Florida
	Ph.D. candidate, Computer Science, currently enrolled, *CGPA* **4.00**/4.00 (Intended grad: **January 2022**)
	Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
	BSc., Computer Science and Engineering, February 2013, *CGPA* **3.54**/4.00

RESEARCH INTERESTS

- **High Performance Computing (HPC) Systems:** Domain-specific Systems Design, HPC I/O Optimization, Heterogeneous Storage Stack, HPC File Systems, HPC Workflow, Performance Analysis
- **Artificial Intelligence:** Deep Learning(DL) at Scale, Reinforcement Learning

RESEARCH EXPERIENCE

Department of Computer Science, Florida State University

Graduate Research Assistant **August 2017 - Present**
- Ph.D. student researcher at *Computer Architecture and SysTems Research Lab (CASTL)* supervised by *Professor Dr. Weikuan Yu*, specializing in domain-specific distributed systems design

Center for Applied Scientific Computing (CASC), Lawrence Livermore National Laboratory (LLNL)

Student Intern **May 2019 - August 2019**
- Worked on optimizing I/O strategies in HPC application workflows like Cancer Moonshot Pilot 2 in the Data Analysis Group at CASC. Achieved 84.7% latency improvement by using burst buffers on Lassen.

National Energy Research Scientific Computing Center (NERSC), Lawrence Berkeley National Laboratory (LBNL), Berkeley, California

Student Assistant(Summer intern) **May 2018 - August 2018**
- Worked in the Data Analytics and Services group at NERSC. Analyzed scalable data pipeline for distributed DL atop TensorFlow and Horovod. Determined I/O bottleneck of upto 11.04% in DL training time.

NERSC, LBNL

LBNL Affiliate **August 2018 - August 2019**
- Enhanced the summer internship project on determining I/O bottlenecks in distributed DL applications.

RESEARCH PROJECTS

- **HPC Workflow I/O Optimization:** Built an Emulator during internship at CASC to analyze different HPC I/O patterns, e.g., Deep Learning Training I/O, Checkpoint/Restart, Producer-Consumer, etc. Developing a middleware to implement optimization strategies for application workflows.
- **BeeGFS Performance Evaluation:** Published a research paper on the performance evaluation of BeeGFS parallel cluster file system using IOR and MDTest, and deep learning applications atop TensorFlow, Horovod and LBANN. Currently, working on analyzing the fitness of BeeGFS On Demand (BeeOND) as an ephemeral burst buffer file system.
- **Scalable Data Pipeline for Distributed Deep Learning:** Analyzed and profiled I/O behavior posed by cutting-edge deep learning applications at scale by using a logging framework developed during internship at NERSC. Pinpointed I/O bottlenecks caused by metadata overhead in deep learning training.

PUBLICATIONS

- **F. Chowdhury**, F. Di Natale, A. Moody, E. Gonsiorowski, K. Mohror, and W. Yu. "Understanding I/O Behavior in Scientific Workflows on High Performance Computing Systems," in Proceedings of the *International Conference on High Performance Computing, Networking, Storage and Analysis 2019 (SC19)*, Regular Poster, Nov. 2019.
- **F. Chowdhury**, Y. Zhu, T. Heer, S. Paredes, A. Moody, R. Goldstone, K. Mohror, and W. Yu, "I/O Characterization and Performance Evaluation of BeeGFS for Deep Learning," in Proceedings of the *48th International Conference on Parallel Processing (ICPP 2019)*, Research Paper, Aug. 2019.
- **F. Chowdhury**, J. Liu, Q. Koziol, T. Kurth, S. Farrell, S. Byna, Prabhat, and W. Yu, "Initial Characterization of I/O in Large-Scale Deep Learning Applications," in *SC18, 3RD Joint International Workshop on Parallel Data Storage & Data Intensive Scalable Computing Systems (PDSW-DISCS 2018)*, Work-in-progress (WIP) Abstract, Nov. 2018.
- Y. Zhu, **F. Chowdhury**, H. Fu, A. Moody, K. Mohror, K. Sato, and W. Yu, "Entropy-Aware I/O Pipelining for Large-Scale Deep Learning on HPC Systems," in *IEEE International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS 2018)*, Research Paper, Sep. 2018.

TECHNICAL SKILLS

- Programming Languages: **C/C++**, **Python**, C#, MATLAB, Java, Javascript
- Libraries: **MPI**, **HDF5**, BSD sockets, WinSock, OpenGL, Boost, Windows API, Google Test
- Frameworks: **TensorFlow**, **Horovod**, **LBANN**, Qt Framework, MFC, .NET Framework
- Distributed File Systems: **BeeGFS**, Lustre, BurstFS, UnifyCR