Xiangfeng(Allen) Zhu

zxfeng@umich.edu · xzhu27.me · 650-660-0918 · github.com/Romero027 · xzhu

Education

University of Michigan, Ann Arbor — Expected: Dec 2020
Bachelor of Science, Computer Science — GPA:3.82/4.0

Experiences

Databricks San Francisco, CA — May 2020 - Aug. 2020
Incoming Software Engineer Intern

- Will be working in Databricks as a Software Engineer Intern

Software Systems Lab University of Michigan — Nov. 2018 - Now
Research Assistant Advisor: Prof. Mosharaf Chowdhury

- Co-developed a general-purpose execution engine, Sol, that can adapt to diverse network conditions on top of Apache Spark.
- Improved SQL, machine learning, and streaming jobs by 4.2× and 16.4× on average, respectively, in offline and online settings compared to Apache Spark in resource-constrained networks.

Disorderly Lab UC Santa Cruz — Mar. 2018 - Now
Research Assistant Advisor: Prof. Peter Alvaro

- Designing a tracer and a fault injector using system-level provenance for unmodified distributed systems
- Developed a debugging approach for distributed systems based on analysis of provenance data obtained during system executions
- Evaluated our approach on the TaxDC collection of real-world bugs from four large-scale distributed systems(Cassandra, Hadoop, HBase and ZooKeeper).

Dropbox San Francisco, CA — May 2019 - Aug. 2019
Software Engineer Intern Filesystem Team

- Worked on the next-generation distributed filesystem for Dropbox
- Designed and implemented an asynchronous system to unmount namespaces that a user loses access to
- Rearchitected our MapReduce framework to be more efficient and fault-tolerant using RocksDB and gRPC

♡ Projects

Distributed Debugger Using Provenance Graph (Go) — Mar. 2018 - Aug. 2018

- Implemented a lineage-driven distributed debugger that can analyze the given program and give suggestions to the programmer how and where to correct the program

Fault-tolerant Scalable Key-Value Store (Python) — Jan. 2018 - Mar. 2018

- Built a distributed, fault-tolerant, highly available and eventually consistent key-value store that can store the amount of data that cannot fit into one single machine.

Skills

- Language: Java, C, C++, Python, Go, Scala, Bash, SQL, HiveQL, HTML, CSS, JavaScript, LaTeX
- Tool: Perf, Valgrind, Git, Vim, GDB, Docker, Xcode, Flask, Pytorch
- Data: Oracle, MySQL, Neo4j, Hadoop, Hive, Spark, Flink, Zookeeper