PEIQIN ZHAO

peiqinz@andrew.cmu.edu♦(412)616-6652♦github.com/hbgxsm95

EDUCATION

Carnegie Mellon University, *M.S. in Information Networking* Pittsburgh, Aug. 2017 – Dec. 2018

- Courses: Computer Network, Distributed Systems, Database Systems, Storage Systems

Sun Yat-sen University, *B.E. in Electronic Engineering* Guangzhou, Aug. 2013 – Jun. 2017

- Courses: Data Mining, Operating Systems, Introduction to Cloud Computing

WORK EXPERIENCE

Google, Software Engineering Intern, Cloud Infrastructure Sunnyvale, May. 2018 – Aug. 2018

- Built a metric collection system for distributed deployment on IoT under OpenCensus and *Stackdriver* in **Go**.
- Designed and implemented a robust and low-overhead protocol for the measurement transmission in the above system to support efficient coordination between any IoT device and the master node under *OpenCensus*.
- Fixed and updated the tracing and metrics exporter for *Stackdriver* under *OpenCensus.*

PROJECT EXPERIENCE

Disk-Oriented Storage Manager for SQLite DBMS (C++), CMU October. 2018

- Implemented a thread-safe buffer pool manager by building up an extendible hash table, which supports moving physical pages back and forth from main memory to disk based on LRU page replacement policy.
- Built a B+ Tree dynamic index structure, which supports concurrent operations including search, update and index iterator based on the index concurrency control.
- Designed a lock manager which tracks the tuple-level locks issued to transactions and supports shared & exclusive lock grant and release based on 2PL and strict 2PL, as well as deadlock prevention and detection.

SSD Flash Translation Layer (C), CMU Sep. 2018

- Implemented the system for managing mapping relationship between physical and logical addresses in SSD.
- Developed the garbage collection system to compress valid pages and erase stale blocks when needed.

Heterogeneity-aware Job Scheduler (C++), CMU May. 2018

- Built a scheduler based on *Apache YARN* to maximize utility for jobs with different types and flow patterns.
- Scheduled jobs based on a normalized quantitative indicator related to factors like estimated running time.

Haystack File System (Python), CMU Apr. 2018

- Implemented a prototype of haystack distributed file system by deploying a directory server based on Cassandra.
- Developed the cache cluster based on Redis cluster to reduce I/O overhead on the backend servers.

Data Processing on AWS using Apache Spark (Python), CMU Feb. 2018

- Built a distributed program that ETL preprocesses 500G WET files crawled from Common Crawl in 36 minutes.
- Optimized 3x ~ 5x on model learning using SGA on data of eight-hundred-million features with inverted indices.

Implementation of RMI (JAVA), CMU Feb. 2018

- Developed remote-method-invocation library functions including ***Stub*** and ***Skeleton*** which forwards the methods calls to the remote objects and implemented a multi-threaded TCP server on the remote sides.
- Validated the performance of RMI library with PingPongClient and PingPongServer in **Java**.

BitTorrent-like File Transfer Application (C), CMU Oct. 2017

- Built a file transfer application that downloads files concurrently from multiple peers like BitTorrent.
- Implemented TCP-like congestion control protocol including Slow Start and Congestion Avoidance.

A Multi-user Web Server called Lisod (C), CMU Sep. 2017

- Used the Berkley Sockets API to build a web server which could handle concurrent clients simultaneously. The server supports pipeline requests with GET, HEAD and POST methods based on the HTTP 1.1 protocol.
- Implemented the Common Gateway Interface and SSL server-side to support HTTPS via TLS.

SKILLS

Programming Language: Expert in Go, C/C++, Java; Familiar with Python, Swift, SQL, JavaScript
Framework and tools: Linux/Unix, Shell, Django, Hadoop, Spark, OpenCensus, Stackdriver, Vim, Git