Pufan Jiang

jiangpufan@gmail.com | Mountain View, CA, USA | 650-704-0357

Education

Carnegie Mellon University (M.S. in Computer Engineering)	09/2015 - 12/2016
- Teaching Assistant in Cloud Computing	
Shanghai Jiao Tong University (B.S. in Information Engineering)	09/2011 - 07/2015

Work Experience

Google Ads & Commerce Payment Team (SDE Intern)   05 - 08/2016

- Prevented and alerted abnormal payments and found ~10 bugs in current payment system by analyzing complex payment business logic and one year's historical payment data and building a validation and monitor framework.
- Tested the monitor framework on production data (~10TB) using FlumeJava (Google's internal data processing framework).
- ~10K lines of code submitted. Exceed expectation review (mid term).

High Performance Web Server Design (Cloud Computing Teaching Assistant)   10 - 12/2016

- Developed multiple ETL processes on 1TB data with $10 budget using MapReduce and Spark framework.
- Built a read-only web service handling ~8K RPS from ~20M records(~70G data) with $1/hour budget on AWS, by fine-tuning HBase database, data compression, data balance and well-designed schema etc.
- Built a read/write mixed web service handling ~10K RPS from ~20G data with $1/hour budget on AWS, by fine-tuning MySQL database, batching technique and well-designed concurrency model etc.
- Identified and eliminated performance bottlenecks, achieving ~10x higher throughput compared with naive design, with tracing tools(perf), monitoring tools(top, vmstat, sar, Ganglia), visualization tools (flame graph) etc.

Selected Projects

Hybrid File System on SSD and AWS S3 with FUSE library   10 - 12/2016

- Implemented a size-aware file system storing large data on cloud (AWS S3) and small data/cache locally.
- Implemented a file de-duplication layer to minimize storage and data transfer cost.

Object-oriented Design under Formal Design Process   10 - 11/2016

- Designed and implemented a Scrabble game supporting extensible rules and having a GUI, following formal design process i.e. domain description, sequence diagram, object diagram.

Computer Network Modules - Reinvent the Wheels   09 - 12/2016

- Implemented a robust HTTP server, supporting static files, CGI requests and HTTP persistent requests.
- Implemented a TCP-like reliable network module on top of datagram socket system calls.
- Implemented adaptive video streaming and locality-aware DNS services.

Distributed System Playground   03 - 05/2016

- Implemented a Chord-like distributed hash table, supporting consistent hashing, gossip protocol, failure detection, data replication and tunable consistency level.
- Implemented a communication framework supporting reliable multicast communication.

Skills

Programming Language: Java, C/C++, Go, Python, Shell, Javascript, SQL
Platform and Tool: MapReduce/Hadoop, AWS(EC2, S3, ELB, Elastic MapReduce), Spark, Node.js, Git