Ziming WANG

Mountain View, CA, 94043
zimingwang@cmu.edu • +1 (650) 417-8700 • https://github.com/ZimingWang • https://www.linkedin.com/in/zimingwang93

WORK EXPERIENCE

Data Engineer Intern, Sparkline Data Inc — May 2016 – Aug 2016

- Wrote code to validate the correctness of data-frames generated by SparkSQL query and transformed Druid Query.[Spark, Druid, Scala, JVM]
- Designed performance tests against the server clusters to test the Real-time performance of Sparkline Accelerator.[TPC-H, BI, Jmeter, RTAP]
- Contributed bug-free code to the Sparkline Accelerator every week, and added new features to its SparkUI.[Druid, Hive, SparkUI]

RESEARCH EXPERIENCE

Research Assistant, Jilin-Georgia University Bioinformatics Research Center — Dec 2012 – Nov 2014

- Created an original Entropy-PageRank Model which could make the process of disease related genes prediction automatically in the GB-scaled data sets. Wrote 2 academic papers with first author's identity.

SKILLS

- Language: Java, Scala, Python, Javascript, Ruby, Bash Script
- Cloud Computing: AWS, Spark, Druid, Hadoop, Kafka, Samza, Zookeeper, Docker, Mesos
- Database: MySQL, Cassandra, Neo4j, PostgreSQL, HBase, MongoDB, Hive, Impala, RedShift
- Web/Mobile: AngularJS, Node.js, Meteor, Rails, Play, Servlet, Android

A fast-learner and enthusiastic coder who performs well in team as well as individually.

EDUCATION

Carnegie Mellon University, USA

- Master of Engineering (M.E.) in Software Engineering — Sep 2015 – Dec 2016
  - Course: Cloud Computing, Practical Data Science, Data Intensive Workflow, Advanced Database, Distributed System, Service Oriented Computing, Information Security, Java Smartphone Dev, Foundation of Software Engineering

Jilin University, China

- Bachelor of Engineering (B.E.) in Software Engneering — Sep 2011 – Jun 2015
  - Course: Operating System, Principle of Database, Computer Network, Data Structure, Algorithm Analysis, Software Architecture, C/C++ Programming, J2EE Design, Mobile Phone Software Programming

PROJECTS

Real Time Stock Data Processing Platform, Kafka, Cassandra, Spark Streaming, Mesos, Docker

- Implemented a real time data processing platform to analyze stock data from Yahoo Finance.
- Developed a high performance data ingestion layer using Apache Kafka (200k msg/s).
- Built a data stream processing pipeline using Apache Spark Streaming and a data storage layer using Apache Cassandra to store time series data.
- Created a scalable cloud deployment environment using Docker and Apache Mesos(job dispatch).
- Developed a visualization web app using Redis, Node.js and D3.js.

Twitter Big Data Analysis Web Service, Cloud Computing(AWS), HikariCP, Hbase, MySQL

- Developed a fault-tolerant and scalable web-server based on servlet and used load balancer which could handle six different requests concurrently(14000-25000 rps within 1 dollar per hour).
- Used ETL Pipeline to normalize, tokenize and parse over 1TB Tweet data(500 million rows).
- Optimized database performance by indexing,schema design and memory caching.
- Applied JDBC connection pool like HikariCP to handle highly concurrent user read requests.

App Web Crawler, Python, Scrapy, Scrapyjs, Meteor, Node.js

- Crawled 100,000 apps from Xiaom App Store using Scrapy, Scrapyjs, Proxy, and stored result in MongoDB.
- Used Cosine-Similarity to recommend top 10 related apps according to user's download history.
- Built front-end to display details of apps using Meteor.

Survivable Social Network, Node.js, Angular JS, Scrum, BeagleBone Black, Software Design

- Implemented a social network web application with features such like Real-time chatting, Information searching ,Role-based Users and Performance measuring based on BeagleBone Black Board.
- Applied Hybrid Scrum-Kanban development method in a four-member team.
- Developed front-end and back-end using Angular JS, Jade, passport.js, Bootstrap, Node.js, socket.io and sqlite.
- Used Mocha, Superagent and Selenium to test the project.
- Used Bower, Shippable and Grunt to develop, integrate and maintain the project.

Guuber FreeRide Application, Android, Google API, Socket, AWS

- Developed a freeride android app which provides passengers with free rides based on their location.
- Used Google Map API to implement location identification and navigation.
- Applied Socket to build communication between driver and passenger on the server which is deployed on AWS.