JASON (BAO) SUN

Phone: (86) 1536-107-1315 E-mail: arthursunbao@gmail.com

SUMMARY

3+ years' professional software development experience in Java, Python and Golang

SKILLS

Languages: Java,Python,Golang, JavaScript, Shell Scripting, SQL, C#
Database: Cassandra**,** MySQL, MongoDB, Oracle(SQL, PL/SQL)
Middleware: Kubernetes, Docker, Hadoop, Spark, Zookeeper, Kafka
Framework: Spring, Hibernate, Django, Flask
Version Control: Git, Subversion

EDUCATION

Master of Computer Engineering
University of British Columbia, Vancouver, Canada 2015 – 2018

Bachelor of Computer Engineering
Shanghai University, Computer Science Department, Shanghai, China 2009 – 2013

PROFESSIONAL/RESEARCH EXPERIENCE

Software Engineer, Sensetime Research, Shenzhen Sep 2018 - Now

- Lead a team of five software engineers to develop massive kubernetes-based distributed Spark deep learning feature clustering service with Cassandra, Kafka as data ingress, Ceph as object backend storage, Django as RESTFul controller, MySQL as metadata storage, supporting batch feature clustering of 0.8 billion at one time within an hour using Python, Java and Golang
- Developed a K8S-based microservice high-dimensional feature database with Cassandra as persistent storage, minio as feature snapshot, GPU as feature cache with distributed master-worker architecture, which can persist 0.1 billion features and 50 1:N QPS per GPU

Software Engineer, CA Technologies Canada May 2018 – Sep 2018

- Worked in an agile team to develop a SaaS based API management portal based on Kubernetes, Docker and Spring Boot, Hibernate in Java and was responsible for backend feature development, debugging and customer support

Software Engineer Intern, Baidu ECOM Research Nov 2016 – Nov 2017

- Developed a distributed Hadoop-based search relevance batch, near real-time streaming service which can support 20 million core ads keyword pair matching an hour with one server using Word2Vec as text similarity model, MapReduce as core pipeline, MySQL cluster as persistence using Java, Python
- Designed and implemented a machine learning model called PARIGQ using DNN and GBDT to improve current overall wise AUC from 76.8% to 77.2% with PaddlePaddle, Spark, C++ and Python

Software Engineer Intern, Intel May 2016 – Nov 2016

- Worked on distributed search engine index development and optimization and utilized machine learning methods to improve encoding, decoding and compress time on CPU and GPU platform using Java, C++, Python
- Developed SIMD version Intel Caffe multi-node version and contributed to open-source Caffe using C++

IT Specialist, IBM, Shanghai, China Sep 2013 – July 2015

- Developed front-end web interface for IBM enterprise storage web authentication system as well as back-end DS8000 automation storage testing framework, customized plugin for customer with Java, Python
- Developed PoC demos for various customers by requirement based on IBM products