

XUSHENG JI

Email: Porter.Ji.Porter@gmail.com • Github: <https://github.com/PorterPorterUS>

Education Background

University of Rochester **ROCHESTER, NY**
Master of Computer Science, GPA 4.0 2019.9- 2020.12
● Coursework: Operating System, Algorithm, Compiler, Parallel Computing, Database System

Imperial College London **LONDON, ENGLAND**
Phd candidate in Computer Science (QUIT) 2019.1-2019.3
● Thesis: Efficient Large-Scale Data processing on the Edge Computing Platform

University of Leeds **LEEDS, ENGLAND**
Master of Advanced Computer Science, GPA 3.7 2017.7-2018.12
● Coursework: Cloud Computing, Distributed System, Data-Intensive system, Distributed System

South-Central University for Nationalities **WUHAN, CHINA**
Bachelor of Computer Science, GPA 3.5 2008.9-2012.6
● Coursework: Java, C++, Data Structure, Computer Networks, Computer Organization, Data Mining

SKILLS

Programing Language: Java, Golang, C++, Python, SQL, shell, C
FrameWork: Hadoop, Flink, Spark, Kafka, Flume, Presto, Impala, Zookeeper, Spanner, Druid
Code Management Tool: Jira, Git, Confluence
DataBase System: MYSQL, Oracle, MongoDB, HBASE, ETCD, Cassandra, Redis

Industry Experience

Baidu Inc. **BEIJING, CHINA**
Crawler DNS, Big Data System Engineer 2012.7-2013.12
● Developed a distributed service that was dismantled into three processes after reconstruction: internal query service, recursive resolution service, public network resolution service.
● Participated in the development and open source work of tera, an open source project of Baidu, which is a Golang Hbase-like distributed column storage, led by T8 architects.

Hisense Key Lab **QINGDAO, CHINA**
Intelligent Elastic Containers Cloud Platform, Big Data System Engineer 2019.3-2019.9
● Implemented automatic scaling and intelligent elastic spark containers services which can support billions users access services simultaneously by modifying Spark on Kubernetes source code.
● Reduced latency of Spark Streaming tasks from 300ms to around 100ms with Optimized Dizzle, which is a new double-layer scheduler model that can schedule tasks in advanced according to graph model.
● Developed service discovery and service health checks system with Apache Zookeeper and etcd for the internet cloud platform
● Optimized Emitter framework by modifying source code to access MQTT services for smart devices

Hisense R&D Center **QINGDAO, CHINA**
HDFS Cloud ETL system, Big Data System Engineer 2016.1-2017.3
● Designed, implemented an Apache Flume Service based on Restful architecture and extracting user Logs in real- time From Apache kafka
● Designed, implemented an ETL service that can complete ETL tasks to log data from Apache Kafka and this service can support ETL for 83 different business topics simultaneously

Hisense R&D Center **QINGDAO, CHINA**
KV Storage System Based on Redis, Big Data System Engineer 2014.4-2015.1
● Optimized Redis Gossip protocol to reduce fail over delay about 200ms and improve availability when the cluster size reaches above TB.
● Improved I/O response time about 246ms for KV storage system using kernel-bypass and DPDK techniques.
● Implemented strong consistency for read request in KV storage system by combining Raft protocol and Read Index strategy together.

Research Experience

University of Rochester **ROCHESTER, NY**
Real time point Cloud system for Deep learning (Supervised by Dr.Yuhao Zhu) 2019.11-2019.12
● Optimized and designed a read time point cloud database system for collecting LiDAR data

University of Leeds **LEEDS, ENGLAND**
Pareto analysis to optimized VMs allocation strategy (Supervised by Dr.Jie Xu) 2018.05-2018.12
● Designed services that can collect monitoring data from Zabbix and execute Hadoop WordCount job simultaneously, and it can also apply different VMs allocation method to finish Hadoop jobs.
● Clustered the running time, CPU usage and VM allocation strategies on clustering algorithms