Haoxiang Lin

linhaoxiang32@gmail.com • +86 15501243326 • Blog • Github

Education

Central South University – Bachelor degree in Automation   2013 - 2018
Skill: Java/Python/C/Scala/Golang
Data structure, algorithm, operating system, databases
Three-years experiences in Linux/OS X
Using Git/Vim/Sublime/IntelliJ to complete development

Internship Experience

ByteDance – Infrastructure   2017.3 – Present
Online Storage Team
- Keep up the redis, springdb, abase cluster.
- Develop the ssdb in a new develop branch in Bytedance, add a new way of master-slave replication.
- Improve the abase performance, which is a K/V database written by Bytedance.

DiDichuxing – DiDi Reasearch   2016.9 – 2017.2
Realtime Monitor Team
- Finish the pressure test for the whole real time monitor system, Optimize the Druid's query efficiency and system stability.
- Build a log monitor system, collect the Druid's log to Kafka every day and ETL the Exception and Error in log using Spark streaming and using Spark SQL to offer a universal interface to query detail logs.

Alibaba Group – Ant Financia   2016.7 – 2016.9
Data Platform Team
- Build a DeepQA System, which is for BD and operators, help them to see the data index and ask the data which they want to know.
- Build a linear model to predict the data index trends, save the relations of different data indexes in knowledge graph.

Alibaba Group – Ant Financial   2015.6 - 2015.11
Alipay+1.0 Team
- Complete part of Alipay+1.0's data requirement, write SQL script to complete ETL task.
- Use sofa framework, to complete a data visualization system, which can easily for operator to see the data's change in real time for Alipay+1.0.The back-end using Java to complete MVC design and complete front-end using A-charts, which is a chart lib, canvas and d3.js.

Cheetah Mobile – Cloud Business Unit   2014.6 - 2014.10
Offline Compute Team
- Use Python to call hive-SQL to complete Clean Master's usual data requirement.
- Take part in a new Data Computing Framework's Development. Improved computational efficiency of Hadoop MapReduce. make the framework to be more flexible, Let jobs to be a DAG graph to be beneficial to Chain Job.

Personal Project

Mars–a simple script language
- A simple script language written by Python, Implement Scanner, Syntax, Symbol Table, Semantic analyze. Refer to CPython's bytecode to complete code generate.
- Implement a stack-based Virtual Machine to parse the bytecode to simulate CPU'stack

BenzeDB–a time series database
- A time series database written by golang, using LevelDB to be the storage engine.
- Implement a compress algorithm refer to facebook's Gorilla.
- Now adding raft to make data consistency.