

饶舸璇

📞 18271321721 · ✉ r776379772@gmail.com · 🔗 AZHELLRGX · 🌐 azhell

👤 个人信息

- 男, 1996年9月出生
- 求职意向: 数据仓库开发工程师
- 工作经验: 3.5 年
- 期望薪资: 25k~30k

🎓 教育经历

- 学士, 三峡大学, 计算机科学与技术专业, 2014.9~2018.7

💼 工作经历

- 深圳名通科技服务有限公司, 网优大数据部门, 大数据开发工程师, 2019.7~至今
 1. 负责DataX的二次开发工作, 完成多服务器关系型数据库之间的数据同步
 2. 负责“虚拟路测定位系统”中“常驻用户”模块的需求开发、定位精度优化、数据维护工作
 3. 负责通用SparkSQL工具”的开发、迭代、优化工作; 并完成程序设计文档、操作文档的编写
 4. 参与联通集团“PMCM数据底座”项目的架构设计、模型设计; 使用Flink完成流式数据处理与指标计算

🔧 项目经历

• DataX二次开发

DataX、Java多线程、SpringBoot

1. **项目功能:** 本项目主要完成多台、多种关系型数据库之间的数据自动同步
2. **担任角色:** 负责了基于开源的DataX的二次程序开发工作
3. **解决问题:** 替代了过去繁琐的跨服务器“数据库脚本+FTP”同步方式, 在MSSQL的基础上添加了对Oracle、Mysql、SyBase等数据库的支持, 解放了运维与数据库开发人员
4. **解决方式:** 阅读了DataX的部分源码, 结合DataX的设计思想以及具体业务需求, 整合出一款更易使用、更稳定、扩展性更强的工具

• “常驻用户”项目

MapReduce、Spark、Hive

1. **项目介绍:** 通过运营商客户提供的数据模拟用户常驻地信息, 协助基站信号质量监测、市场营销等
2. **MapReduce阶段:** 本项目属于“4G虚拟路测定位”系统的一部分, 初始阶段使用MapReduce进行需求开发和部署, 计算数据量大约为5T【170G * 30天, 各省人口不同略有变化】
3. **Spark阶段:** 因为涉及到业务算法变更, 部门决定开发Spark版本; 我参与规划了新的业务计算流程, 以及协助拆分子程序, 并负责核心模块的开发、调优工作
4. **Hive数据优化:** 客户为提高市场营销成功率, 要求提高常驻用户的定位精度; 我负责将新定型的优化算法转化为HiveSQL【快速迭代, 短期使用】, 并编写Shell脚本辅助执行、进行HiveSQL的调优

• “联通集团PMCM数据底座”项目

Flink、Spark、HBase、Kafka、Hive

1. **项目介绍:** 联通集团基站运行产生的PMCM数据解码、指标计算
2. **担任角色:** 参与了项目结构的模型设计、业务架构设计、技术架构设计; 并绘制架构图和流程图

3. **数据介绍**：全国31省份总体解码结果数据量大约为40T、每15分钟【数据粒度】计算指标大约为600个
4. **具体实施**：将通信服务商提供的基站运行过程中的参数、性能数据进行解码，并使用Flink流式计算指标，并将部分指标结果通过Kafka共享；解码详情数据、全量指标结果则存至Hive

✂ 技能清单

- ★★★★★ Java
- ★★★ 数据结构与算法、多线程、JVM
- ★★★ Scala、Spark、Hive、数据仓库
- ★★★ MySQL、Redis、ES、SpringBoot
- ★★☆☆ Flink、HBase、Kafka、Presto
- ★★ Shell