

## 个人信息

基本信息	张伦/男/1996	南昌大学(18年毕业)	电子信息 (本科211)
联系方式	13647956188	zhanglun.me@gmail.com	2529450174(QQ)
博客 Github	<a href="http://blog.zhanglun.me">blog.zhanglun.me</a>	<a href="https://github.com/mrzhangboss">github.com/mrzhangboss</a>	

## 工作经历

### 浙江普树信息科技

#### 数据工程师(18年7月-至今)

##### Flink实时计算SQL管理系统(19年6月-9月)

原始计算平台依赖SQL存储过程，吞吐性低、开发维护困难，经过调研各种实时计算平台，最终决定基于Flink搭建一套集开发以及监控的计算平台

- 在迁移历史项目的时候发现，一个单一的Flink Job开发涉及到代码编写、代码编译、包上传服务器、查看日志运行情况，开发周期以周为单位，针对这个问题，基于Flink SQL Client和terminado开发了一套SQL实时调试系统，支持在web直接编写SQL开发和调试，支持实时返回运行结果，将Flink Job开发调试周期压缩到日，减少了90%的调试时间
- 随着平台跑的Flink Job越来越多，经常出现Flink资源不足导致任务提交失败，只能通过人肉运维远程增加配置，针对这种情况我们将Flink平台迁入Docker Swarm平台，通过监控Flink Master负载情况，编写资源管理模块定时扩增或缩减TaskManager数量，避免资源不足导致集群性能下降
- 针对SQL任务由于数据或者逻辑错误导致的任务失败，设计了一套监控预计系统，实时监控MasterManager Rest API接口，针对异常退出的任务，及时发送预警邮件或短信

##### 服务回调监控系统(19年4月-6月)

某些业务系统涉及到系统多个服务串式调用，由于新的模块不断添加以及部分第三服务不稳定，当业务方发现产生数据异常或者长时间未获取到数据，排除问题的时候需要涉及多个部门，耗时周期长且效率低下，故基于Flink CEP模块开发一套服务回调监控系统

- 针对服务架构不同，采用Filebeat采集各个子系统的服务日志，并针对大流量无用日志进行过滤优化，全链服务日志实时性控制在2秒之内
- 针对CEP规则由于各个服务修改导致的服务重启问题，采用Broadcast State实现线上规则动态更新，有效的减少服务重启次数
- 针对数据指标监控实时修改与展示，使用ElasticSearch作为数据的Sink，使用Kibana做数据展示，支持线上服务业务指标规则统计的动态修改以及实时展示

#### 其他项目

- 风控大数据系统
  - 负责编写一套数据导入+数据处理+数据同步程序，使用Sqoop、Airflow、Spring Boot、Spark等技术
- 文书检索项目
  - 使用Scala编写一套Spark批处理程序将历史数据以及增量更新数据导入到ElasticSearch中，并使用Flask提供接口实现动态DSL查询ElasticSearch

## 技能清单

- 擅长Java Scala Python
- 熟悉Hadoop生态区，熟悉Flink Hive HBase Spark等组件
- 熟悉MySQL MongoDB Redis等主流数据库
- 掌握主流机器学习模型算法
- 三年Linux使用经验，掌握Linux平台常用软件使用
- 熟悉英语，能够流畅阅读英文文档

## 开源项目

- [Json2DB](#)

可以存贮任意Json字符串到主流数据库（支持MySQL，PostgreSQL，MSSQL，SQLite等），并支持在数据库中创建删除数据表，提供定制生成数据表结构的接口，编写了一套前端页面支持动态修改数据库字段设计以及复杂映射关系，[体验地址](#)

- [行政规划预测](#)

可以根据任意地名推测出最可能的完整行政规划名称，[项目地址](#)