

# 张作峰 大数据架构师

18320872958 [Call me](tel:18320872958) | [featzhang@apache.org](mailto:featzhang@apache.org) [Mail to me](mailto:featzhang@apache.org) | [<github.com/featzhang>](https://github.com/featzhang)

---

## 专业技能

1. Apache InLong committer, Apache hudi/Flink contributor
  2. 腾讯 Flink Oteam PMC 、腾讯金融实时计算平台负责人、实时数仓负责人
  3. 9年Java大数据开发, 百亿级实时流处理和PB级批处理开发经验
  4. 熟悉Hadoop、Storm、Spark、HBase、Flink、Impala等大数据系统的原理, 有应用开发、系统性能优化及故障排除能力; 经历过完整的设计、应用开发和集群运营周期.
  5. 参与腾讯Flink内核、Connector、Format、ETL工具等的协同开发.
  6. 熟悉大数据基础架构平台的监控、资源管理、数据流管理.
  7. 熟悉离线批处理、预计算、即时计算和实时流计算等常用技术.
  8. 大数据中台、实时OLAP系统和大数据基础平台开发经验.
- 

## 教育经历

- 2010.09 - 2013.06 湘潭大学 控制理论与控制工程(智能计算) 硕士研究生
  - 2006.09 - 2010.06 青海大学 自动化 本科
- 

## 工作经历

- 2017.06 - 至今 腾讯
  - 2016.06 - 2017.06 平安科技
  - 2013.07 - 2016.06 创维软件
- 

## 项目经验

### 腾讯金融实时计算平台 2019年10月~至今

平台架构师、实时数仓负责人 带领9人

一站式实时开发和治理平台, 提供元数据管理和任务管控能力, 提供了 SQL 、 画布 、 Jar 和 拖拽式开发 API , 无缝适配 Flink 、 Spark 、 MR 和 微服务 等计算引擎. 第三方的平台可以通过SDK快速接入, 只需要按照规范和协议提交参数就可以方便的创建任务, 不需要关注数据流转. 根据对数据的实时性要求可以选择不同的计算引擎. 用户开发的实时任务经过编译、封装, 保存下来整个任务的逻辑, 并结合任务的元数据和 UDF 然后进入解析层. 经过编译、执行计划、代码生成等, 解析参数带有要使用的计算引擎, 路由到不同的引擎进行调度. 此外, 基于 Flink 调度能力, 构建实时ETL工具, 支撑业务50T/天的数据量.

金融级数据保障:

链路血缘分析, 基于 Flink SQL 流批一体对账和数据回补, 状态数据分析, 数据染色校验, 自动压测, 双链路, 任务灰度发布

#### 智能运维:

延迟监控, 实时指标体系, 健康度评分, 自动诊断, 自动修复

#### 增量数仓:

在实时数仓的基础上, 基于 Apache hudi 实现流批一体增量数仓, 解决离线任务计算瓶颈

---

## 标签工厂 2019年4月~2019年9月

项目负责人 带领7人

基于Spark SQL和Flink CEP实现了画像标签的挖掘、分析、更新和应用的配置化和一键部署. 目前接入30个业务, 30亿×2万维标签, 链路每天处理数据量超过10PB.

#### 拖拽式标签生产:

通过SQL适配层下发到不同的计算引擎(Spark或Flink), 利用OLAP引擎实现数据探索, 探索结果自动转换为Spark SQL.

#### 标签整合:

按照主题将标签SQL自动合并并优化, 配置化Spark SQL任务将主题表整合成画像宽表

#### 标签上线:

使用Spark将宽表压缩转换成HFile, bulkload到HBase表, 用作点查询场景; 利用Flink batch将orcFile转换为Parquet, 写入CDH集群, 在Impala中做多维分析;

#### 链路治理:

① 血缘分析与链路自动寻优: 根据标签字段在数仓中寻找代价最小路径; ② 数据染色监控, 异步标签质量监控

---

## 腾讯亚秒级OLAP数据分析平台 2017年5月~2017年9月

结合 Tableau 灵活的可视化能力和 Kylin (开源版)多维CUBE预处理能力, 构建的大数据分析工具. 在此过程中, 通过突破 Tableau 鉴权体系, 实现内部OA多用户和角色的管理; 通过请求伪造等方法整合 Tableau viz. 优化 Kylin 开源版引擎, 提升性能. 目前已经支持20余个业务和日用户增量5亿条的数据分析. Kylin 和 Tableau 多个优化已反馈开源社区和官方.

职责: 项目架构设计, Tableau 鉴权体系突破和展示视图的整合, kylin 优化

---

## 平安“渠道联盟”广告推荐系统 2016年5月~2017年5月

渠道联盟广告推荐系统通过打通平安公司内外部的各个渠道, 引入外部ADX流量, 对接平安系近2000款 APP和PC的广告位, 借助人工智能和大数据来, 提高广告转化率的广告营销平台. 在项目中主要解决了业务自动化接入、在线预测工程化实现、算法服务的动态部署、画像整合与快速上线等难点. 通过Storm计算实时推荐效果. 曝光量超过10亿次, 累计用户超过5亿.

职责：架构设计, 推荐大数据平台设计、搭建和应用开发, 后台开发, 10亿级数据写入, 实时效果统计

ZALPHA-智能优化算法平台 2011.08 - 2013.04

ZALPHA 是为解决算法设计过程中仅能依靠理论分析和无法可视化的痛点, 开发的一套算法设计工具. ZALPHA 采用插件式的结构, 将算子（选择、交叉和变异算子）和测试问题模块化, 使用配置或插件即可构建算法, 并实时呈现算法计算过程中的结果和评估值. 有效避免了理论分析错误, 算法执行过程的分析 和理解更加方便. ZAL

