

滴滴实时计算平台架构与实践

梁李印

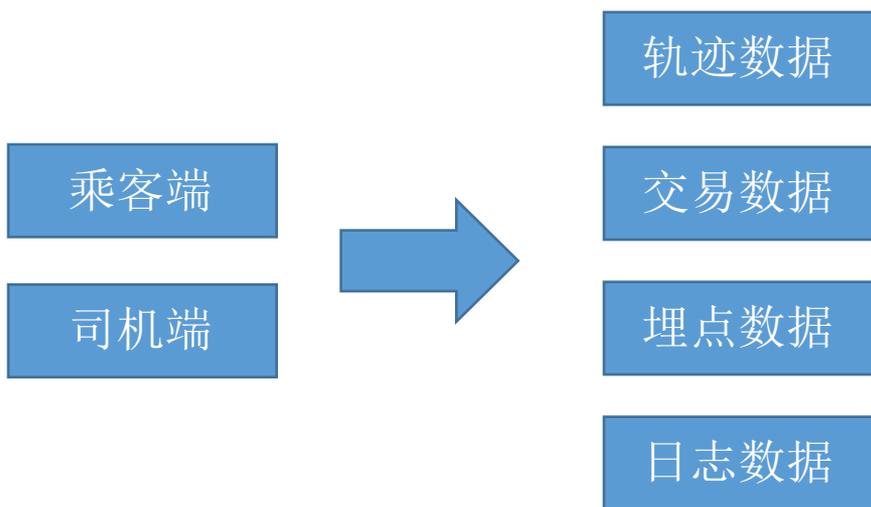
滴滴出行-大数据架构部

自我介绍

1. 2010~2014年 阿里巴巴Hadoop集群（云梯）负责人之一
2. 2014~2016年 负责阿里云分布式图计算框架研发
3. 2017年加入滴滴，目前负责实时计算

滴滴有着丰富的实时数据及计算场景

1. 滴滴核心业务是一个在线交易系统
2. 滴滴有丰富的实时轨迹数据
3. 滴滴率先完成了数据链路实时化改造



缺乏一站式实时计算平台是业务痛点之一

1. 单机版实时计算

- 不可扩展
- 性能受限于单机物理资源
- 容错难度大
- 复杂逻辑开发难度大

2. 自建实时计算集群

- 预先采购大量机器
- 监控报警等体系缺失
- 维护难度大，牵涉大量精力
- 缺乏技术支持

3. 难以做到成本最低，机器资源浪费

目录

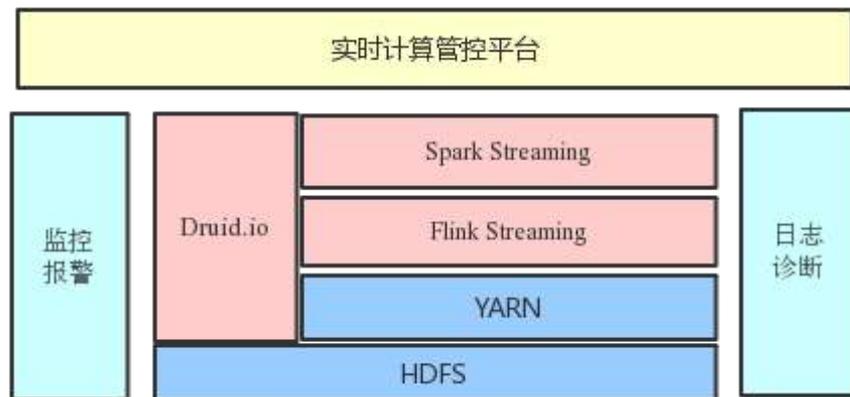
CONTENTS

- 01 实时计算服务体系
- 02 实时计算平台简介
- 03 实时计算应用案例
- 04 实时计算发展规划

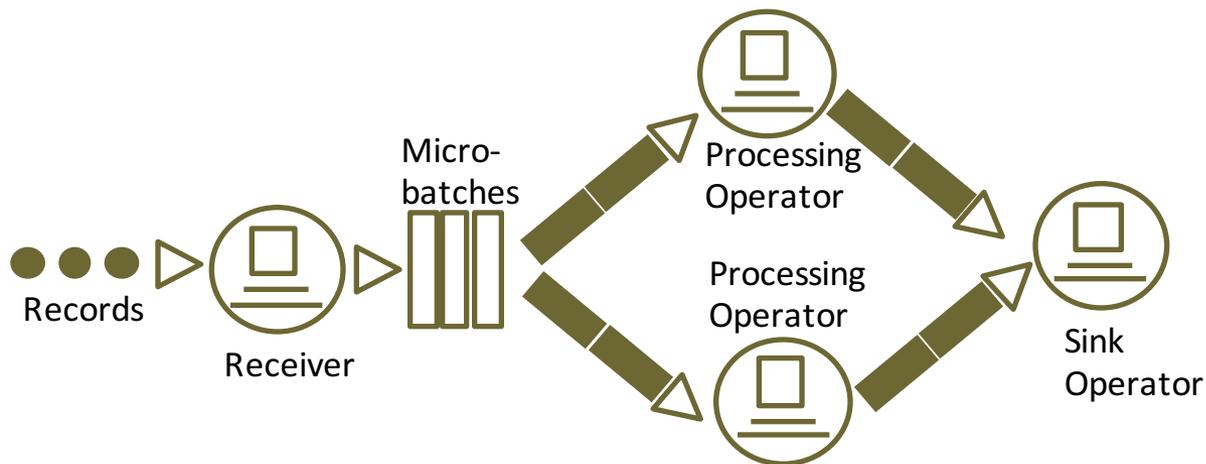
01 实时计算服务体系

实时计算服务体系

1. 流计算+实时存储or消息队列
2. 流计算+实时OLAP

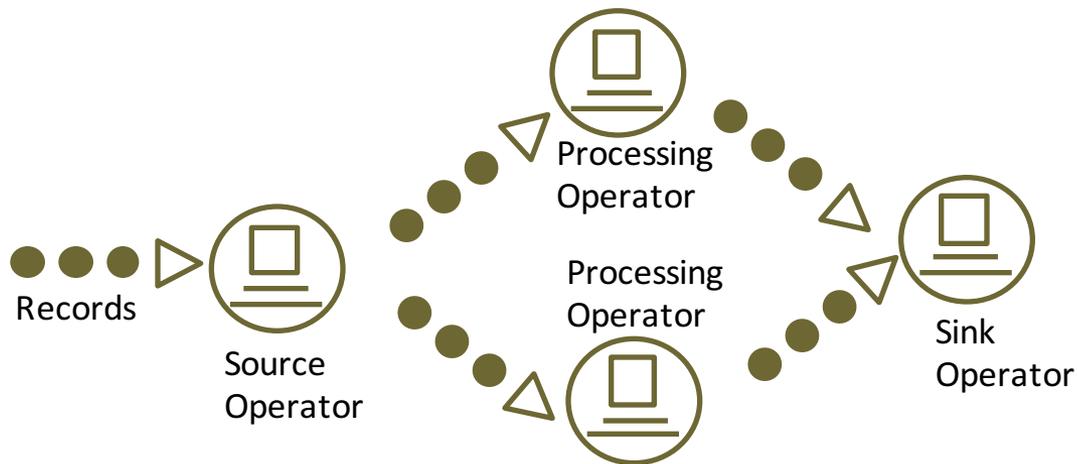


Spark Streaming模型



1. Micro-Batch模式、秒级延时
2. 使用场景：数据清洗、数据通道、数据ETL等
3. 对于熟悉Spark批处理的RD非常容易上手

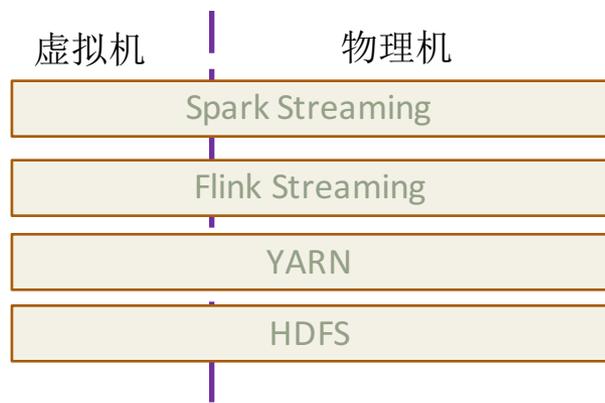
Flink Streaming模型



1. Native Streaming模式、毫秒级延时
2. 使用场景：事件驱动、实时风控、在线业务等
3. 受到喜欢新技术的RD欢迎

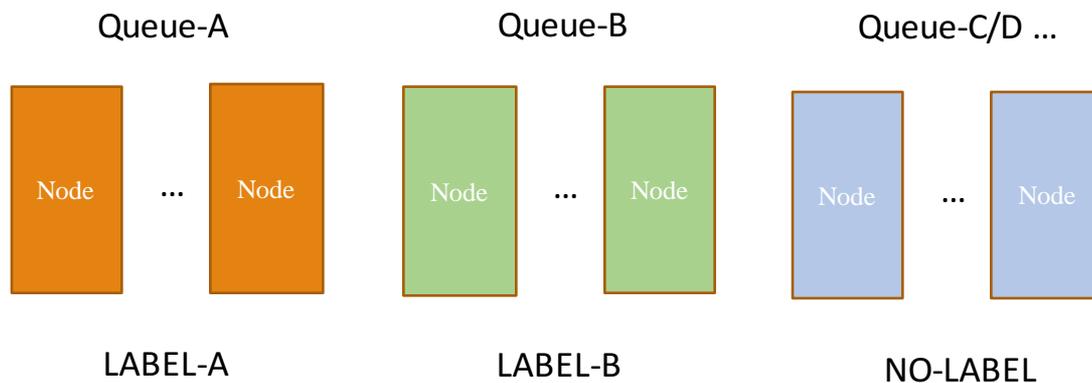
独立构建流计算集群

1. 避免离线业务干扰
2. 与在线业务、Kafka、MQ等相同机房，延时低
3. 提供部分虚拟机作为公共测试任务机器
4. 按照计算队列（queue）分配资源和计费



流计算资源隔离机制

1. 核心业务独占机器（基于NodeLabel机制）
2. 普通业务混用机器（基于CGroup隔离）



Web化管控流计算任务

1. 通过网页提交及管理任务
2. 没有客户机的烦恼
3. 内置参数调优及最佳实践

The screenshot shows a web form for submitting a task. The form includes the following fields and options:

- 所属集群:** 生产集群 (Production Cluster)
- 所属项目:** [Input field with text]
- 公共账户:** [Input field with text]
- 任务名称:** [Input field]
 示例: 客服中心-数据可视化大屏项目 (Example: Customer Service Center - Data Visualization Dashboard Project)
- 任务类别:** flink1.3.2-102 (Task Category)
 注: 切换任务类型时若存在已上传依赖资源会被清除 (Note: When switching task types, if there are already uploaded dependencies, they will be cleared)
- 主程序:** 已上传资源 (Already uploaded resources) | 上传资源 (Upload resources)
 主程序有且只有一个, 允许上传的文件类型: jar/.py (Main program has only one, allowed file types: jar/.py)
- checkpoint工作目录:** [Input field with text]
- 依赖资源:** [Input field] + 增加依赖资源 (Add dependencies)
 注意: 资源不可重复上传 (Note: Resources cannot be uploaded repeatedly)

流计算任务级监控大盘

1. 重要指标实时可视化监控
2. 存活报警、基于指标的报警

每秒流入数据条数(inDataRate)

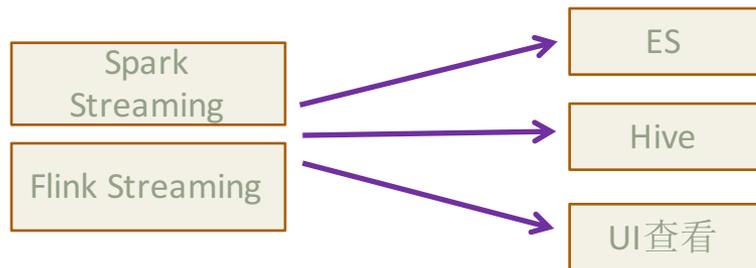


数据延迟度(latency)



流计算任务诊断体系

1. 通过Spark或Flink UI查看日志（定制优化版本）
2. 通过ES检索日志（日志实时上报ES）
3. 通过Hive或Spark分析日志（日志已上报HDFS）
4. 未来提供在线Local调试模式

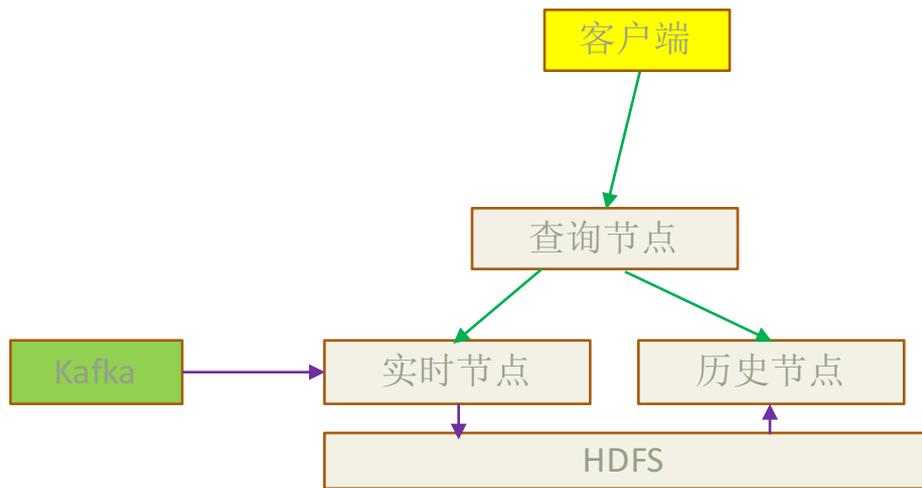


流计算业务规模

1. 集群规模：约500台机器，平均CPU利用率：20%
2. 流计算任务数：约600
3. 服务于快车、专车、顺风车、安全部、运维部等各个业务部门

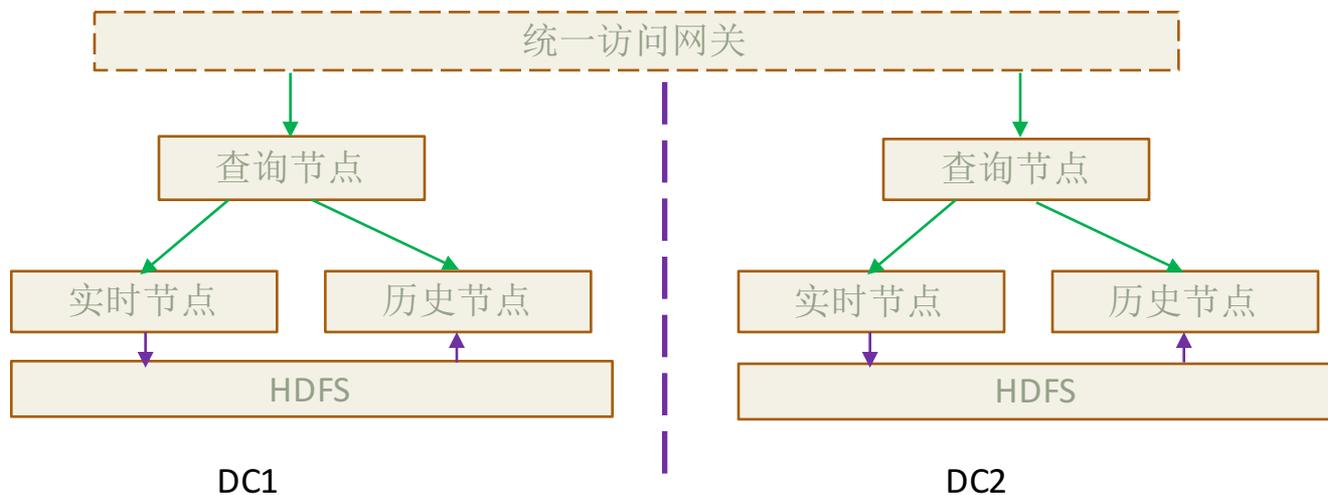
Druid模型

1. 实时数据存储及多维度聚合计算引擎
2. 低延时数据写入（直接消费kafka），数据实时可见
3. 快速交互式查询（99%查询1秒内返回）
4. 支持SQL及DSL两种查询语言
5. 适用场景：实时报表、实时监控



Druid集群建设

1. 异地双活部署
2. 通过网关建设实现多集群对业务透明



Druid数据接入体系

1. Web化操作

2. 选择kafka集群、选择kafka topic、选择维度列、指标列和时间列

Datasource 构建

✓ 基本信息 ———— ✓ 配置数据源信息 ———— ✓ 解析数据 ———— ✓ 编辑字段 ———— 5 匹配字段信息

时间戳列: time 格式: UTC auto + 自定义时间格式
注: 时间字段的格式为时间戳时, 请务必使用UTC (时间戳已包含时区信息)

维度列:

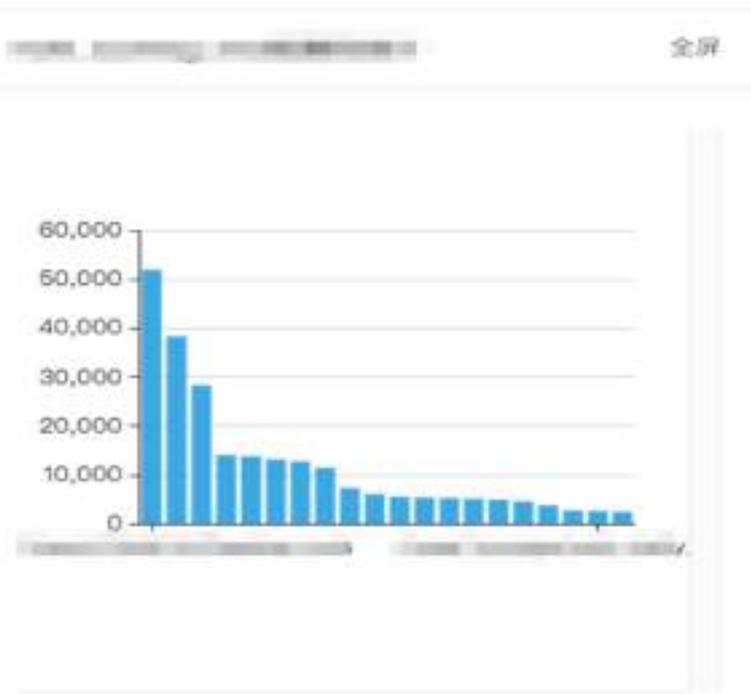
appid	sourceName
appid	Strings
sourceN...	Strings

指标列:

value
value doubleSum as sum

Druid数据查询体系

1. 支持Web化构建监控指标及监控视图
2. 支持拖拽式分析数据
3. 支持Web写Druid SQL分析数据
4. 支持通过RESTful API查询Druid (业务方自建平台)



Druid业务规模

1. 集群规模：120台（双活集群）
2. 接入数据源：250+
3. 每天查询量：百万级别
4. 支撑了公司核心的业务监控

实时计算服务体系总结

1. 独立建设流计算集群
2. 支持NodeLabel资源隔离体系
3. Spark Streaming应用于高吞吐场景
4. Flink Streaming应用于低延时场景
5. Druid异地双活集群，支撑了核心业务监控与报警

02 实时计算平台简介

平台定位

1. 一站式实时计算开发套件
2. 流计算任务开发及管控
3. 开放实时报表/监控能力，用户自助服务
4. 实时数据分析及可视化能力



流计算任务开发

1. 任务管理

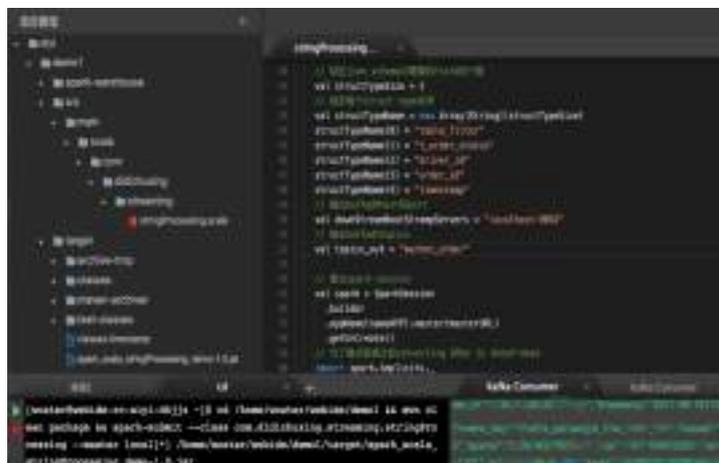
- 本地开发流计算任务
- web化提交及管理

2. 数据ETL

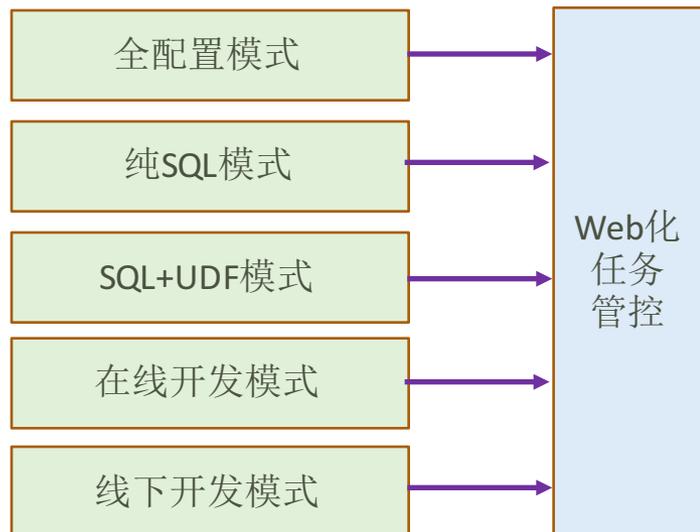
- 配置+模板类+UDF模式
- web化开发流计算任务

3. WebIDE

- 在线调试
- 小白用户



流计算任务开发规划



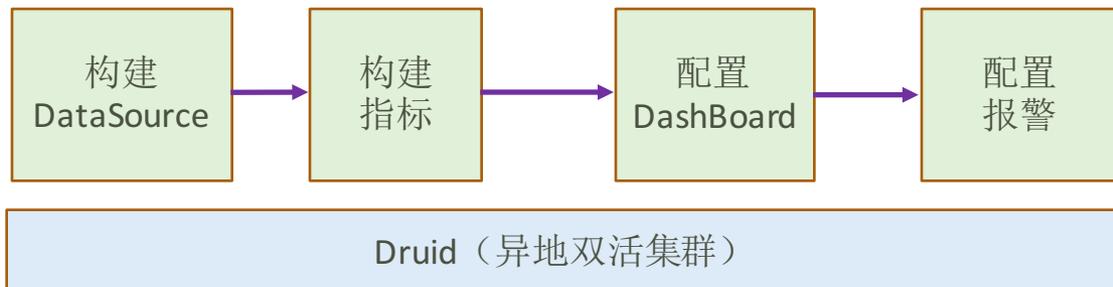
实时报表及监控开发

1. 指标构建

- DataSource构建 (Kafka->Druid)
- 指标构建 (查询模板服务化)

2. Dashboard及报警

- 监控视图自助配置
- 报警服务自助配置



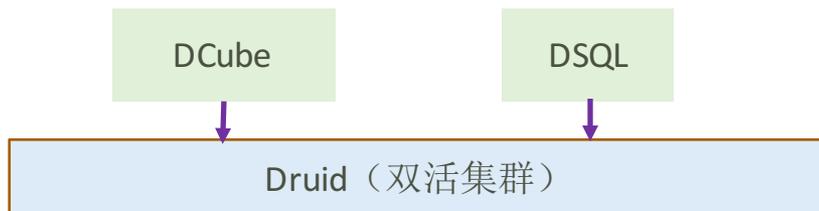
实时数据分析及可视化

1. DCube

- 拖拽式数据分析及可视化

2. DSQL

- Druid SQL开发
- 结果数据可视化



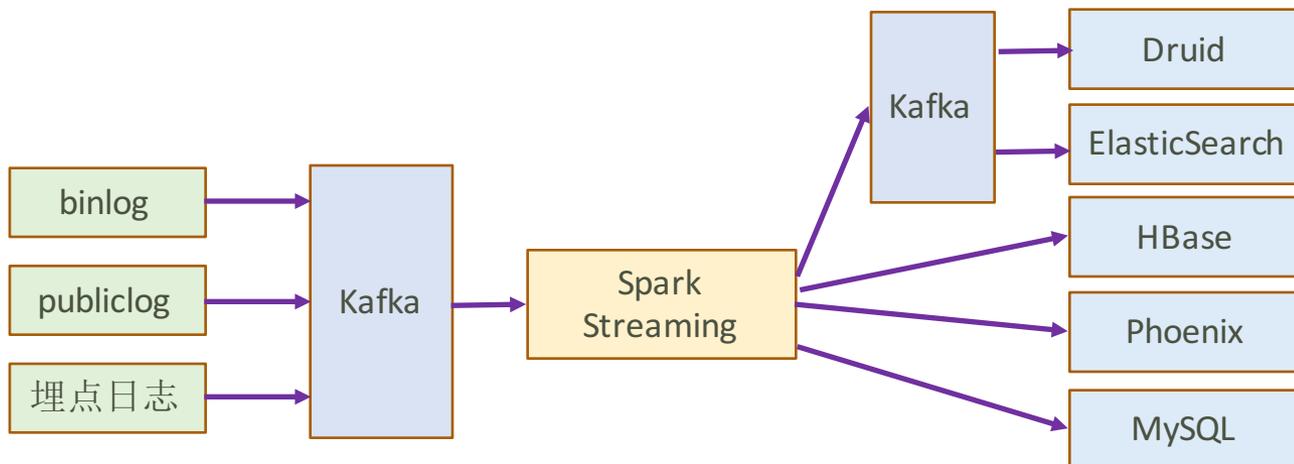
实时计算平台总结

1. 一站式实时计算开发套件
2. 流计算开发、监控与构建、实时数据分析与可视化
3. 实时计算服务唯一入口

02 实时计算应用案例

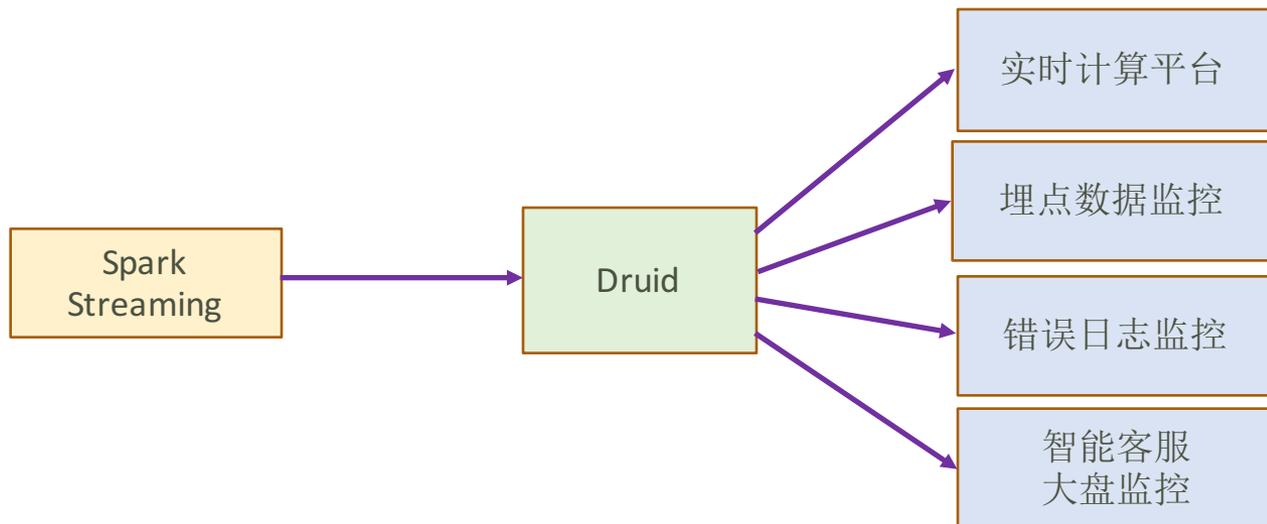
实时数据清洗&通道

1. 实时数据清洗
2. 适配实时存储格式
3. 实时多流join



业务实时监控与报表

1. 清洗后的数据进入Druid
2. 80%查询通过Druid SQL
3. 通过DSL描述更复杂的计算，如PostAggregator、分位数、直方图、Javascript等



智能客服大屏

在线数据

2017-05-03
17:08:18

总UV
400,002

在线分流比
7.70
+2.02%



FAQ UV 点赞率 4.00%

自助 UV 使用率 50.14%

在线聊天
机器人UV 51,570
人工UV 61,570
人工接起率 14%
转人工率 38.2%
转人工量 20

热线 使用量 998

FAQ TOP5

- 2000127
- 30108004_并前30天...
- 30185300_并前30天...
- 2000255
- 2000253

机器人关键词 TOP5

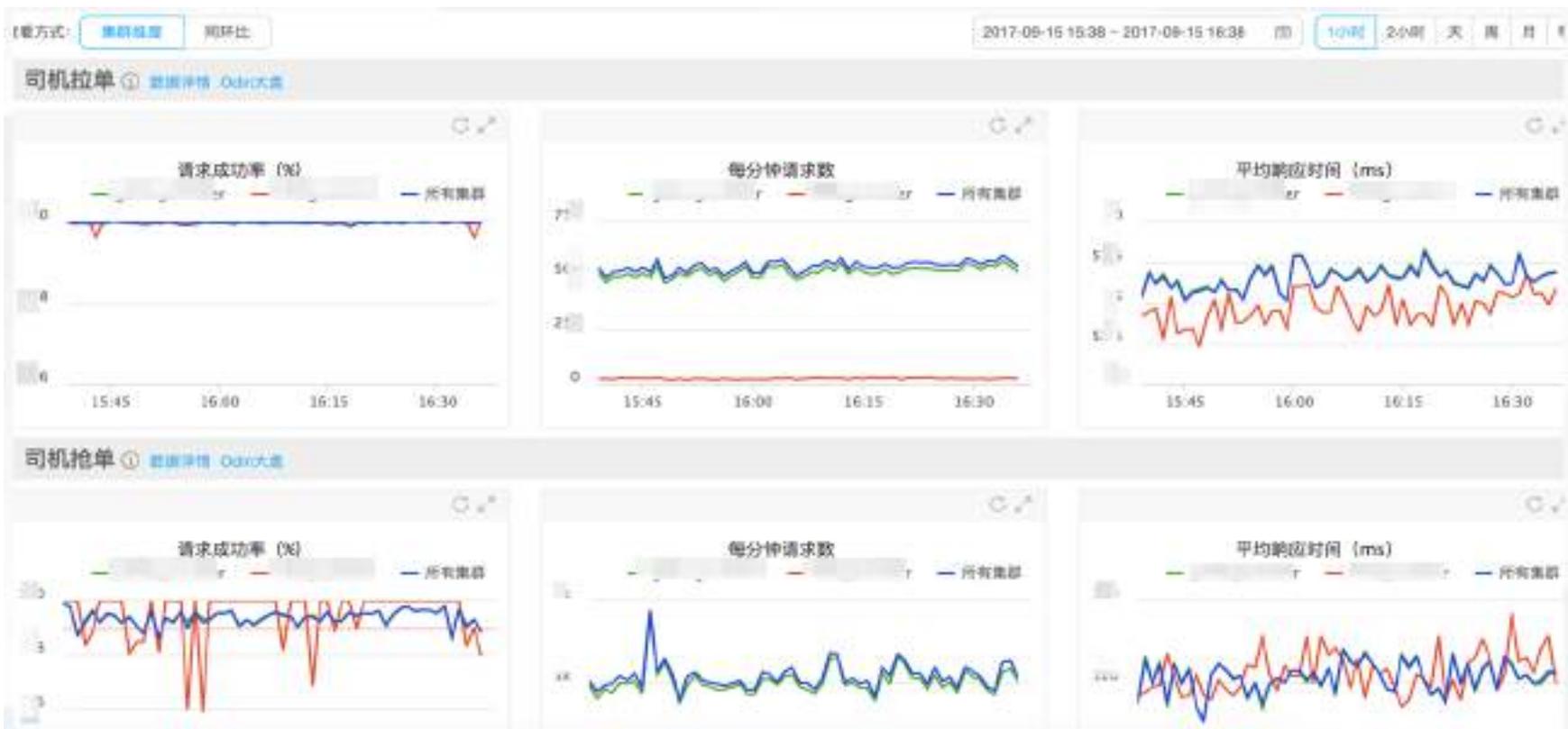
- 显示中文乘客
- 全屏

采用流计算实时监控

1. 在流计算中执行聚合计算
2. 适用于较复杂的聚合逻辑
3. 应用于网关请求监控、过车指标分析等场景

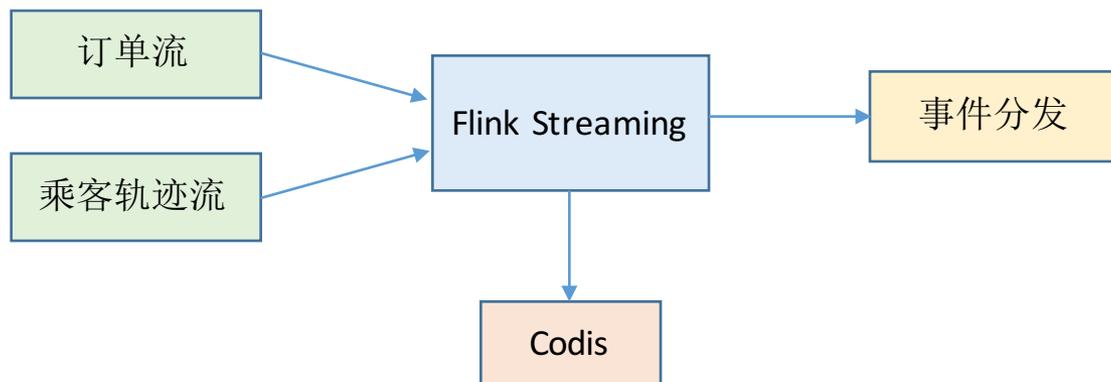


网关日志监控



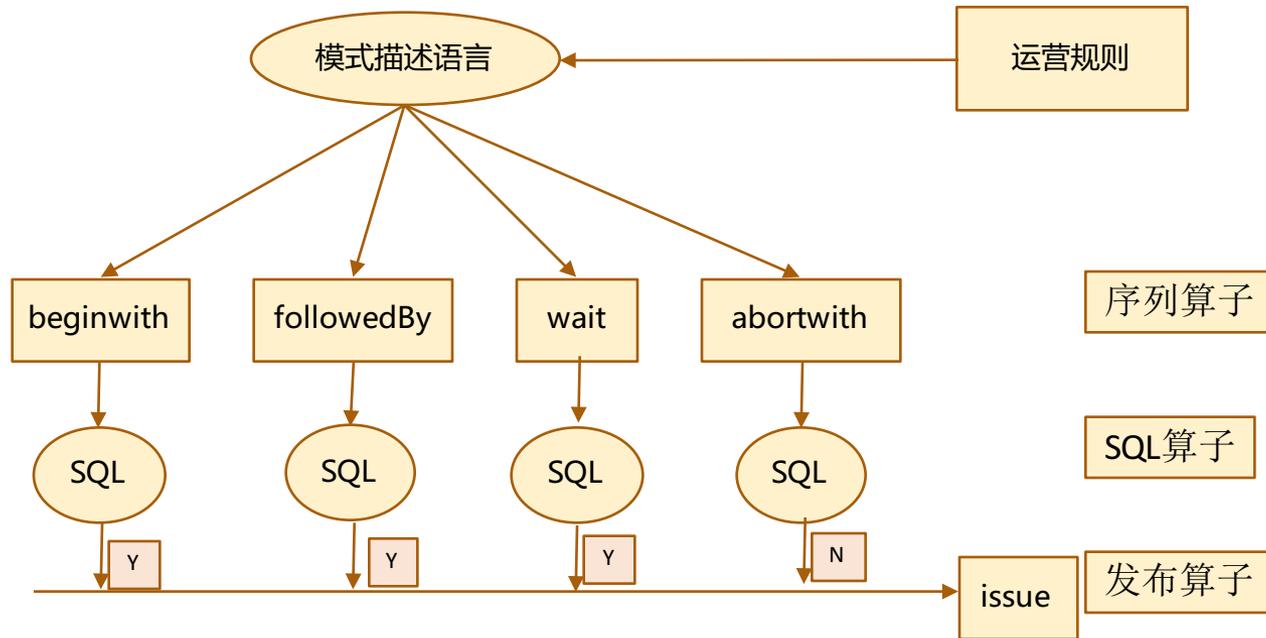
实时业务-乘客位置语义推送

1. 订单流和乘客流实时join
2. 乘客位置变化推送给司机端，准确率94%
3. 减少司乘沟通成本，提升接驾效率



实时规则引擎

- 1. 实时精准运营
- 2. 精准发放优惠券
- 3. 研发中



实时计算应用案例总结

1. 实时数据清洗、数据通道、数据ETL
2. 核心业务指标实时报表及监控
3. 实时日志监控及报表
4. 实时业务

04 实时计算发展规划

技术规划

1. Streaming SQL

- 开发流计算像写Hive SQL/Spark SQL一样简单

2. 流计算任务平滑升级技术

3. 资源隔离技术

- Druid-on-YARN
- CPU资源弹性隔离及调度

业务规划

1. 实时规则引擎
 - 实时精准营销
 - 实时风控
2. 核心在线业务
3. 实时机器学习
4. IoT&边缘计算

谢谢大家！

GIAC | 全球互联网架构大会
GLOBAL INTERNET ARCHITECTURE CONFERENCE

GIAC

全球互联网架构大会

GLOBAL INTERNET ARCHITECTURE CONFERENCE



扫码关注GIAC公众号

2017.thegiac.com