

# StarRocks 在携程的应用

魏宁

使用介绍

生态建设

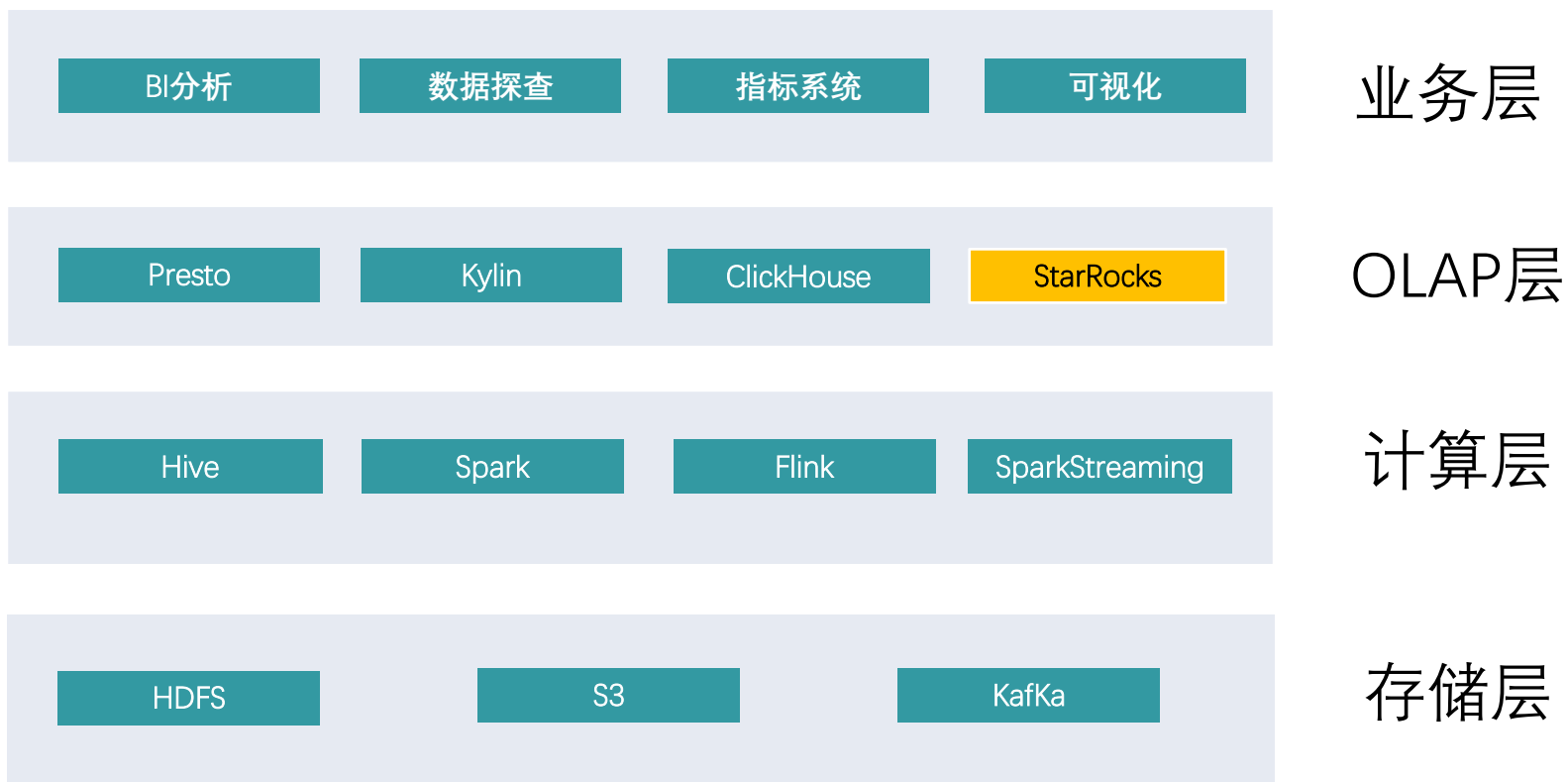
稳定性保障

具体应用

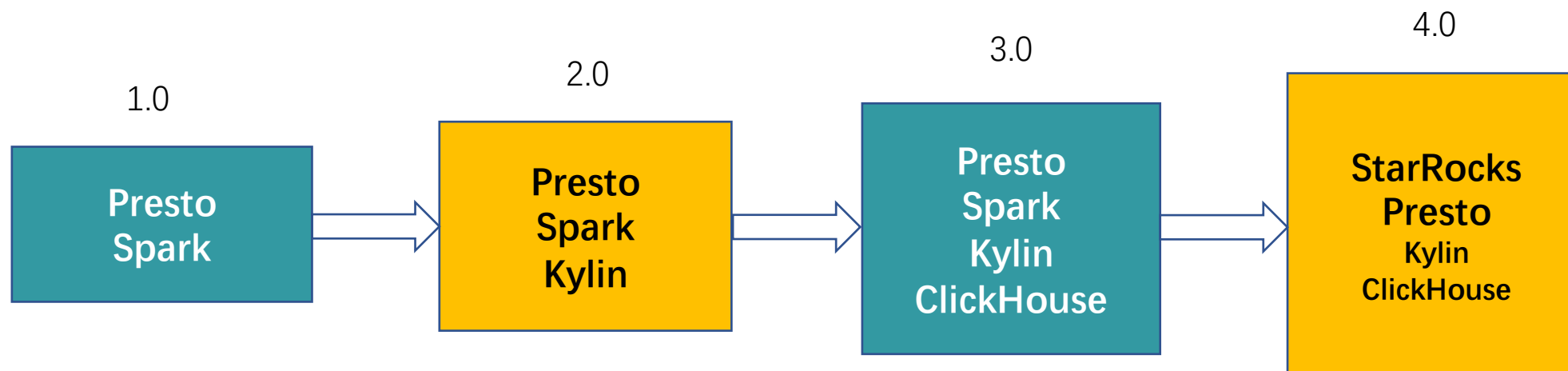
后续

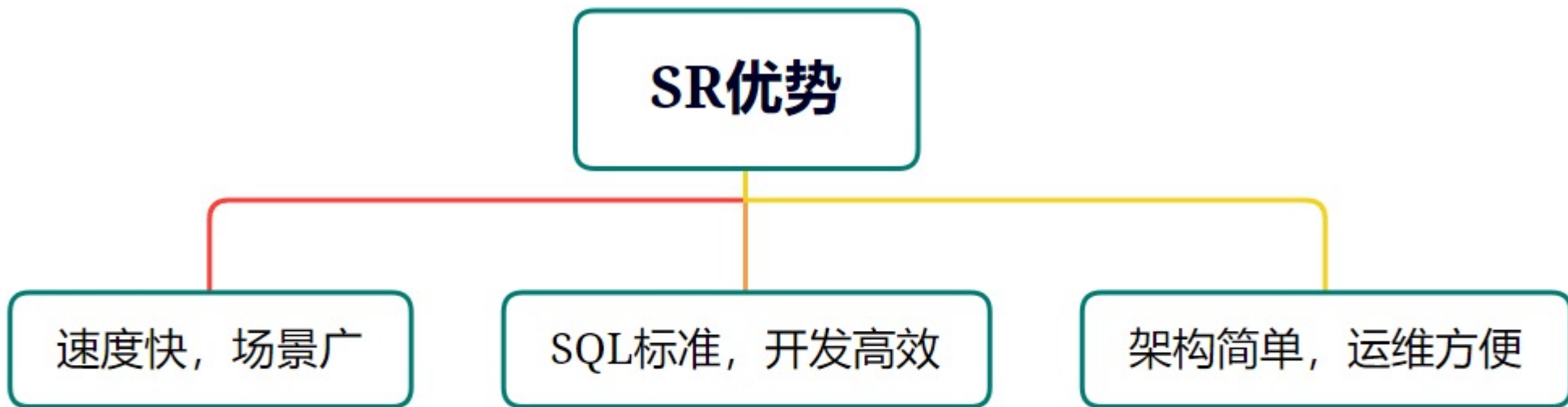
# 使用介绍

# 生态定位

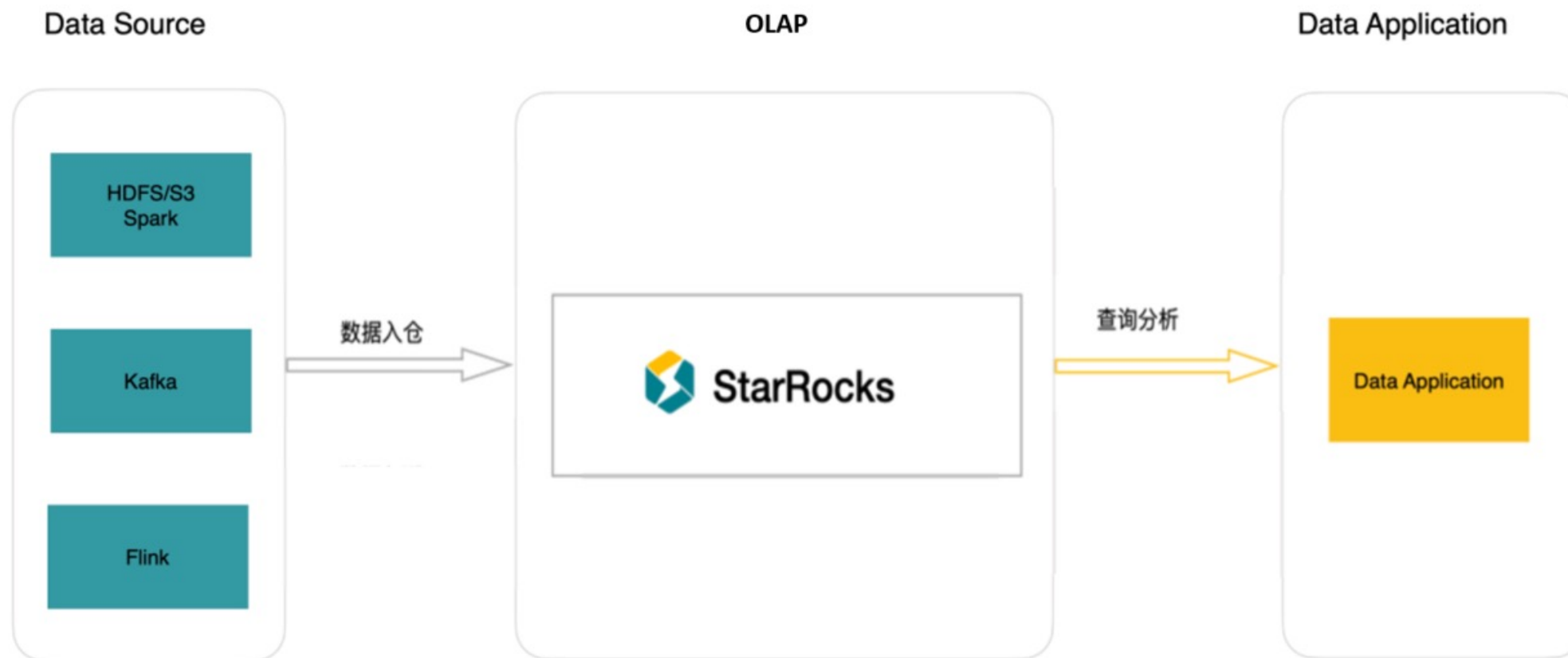


# OLAP系统演进





# 应用架构



## 应用案例

酒店实时业绩  
数据

机票fltinsight

TI标签系统

酒店调价

火车票指标  
平台

ArtNova

商旅在线  
报告

酒店核算

商旅日志数据

Hdata/ebk

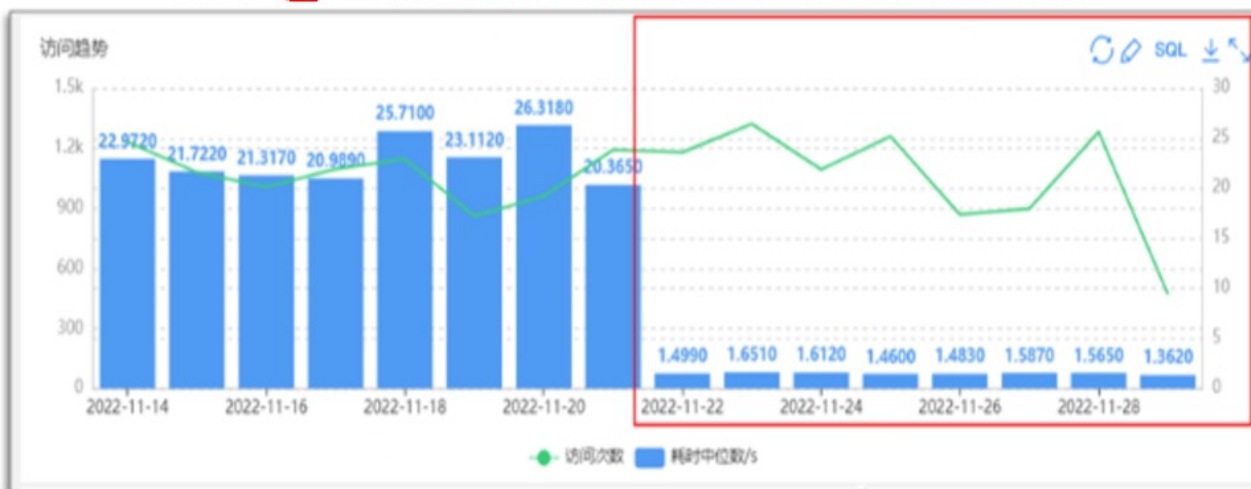
AI平台emoss  
日志

.....

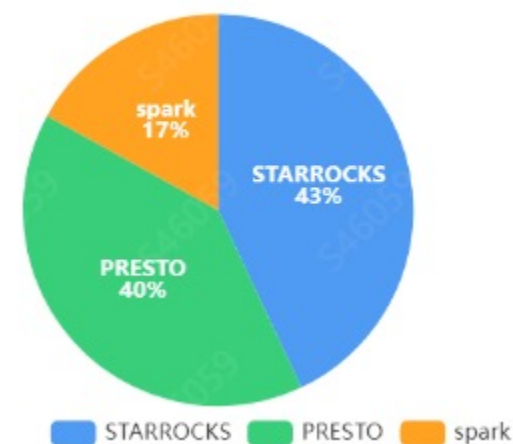


# 应用案例

数据集《~~携程集团2022年11月14日至2022年11月28日~~》查询耗时, 从25.49s 降低到 1.47s

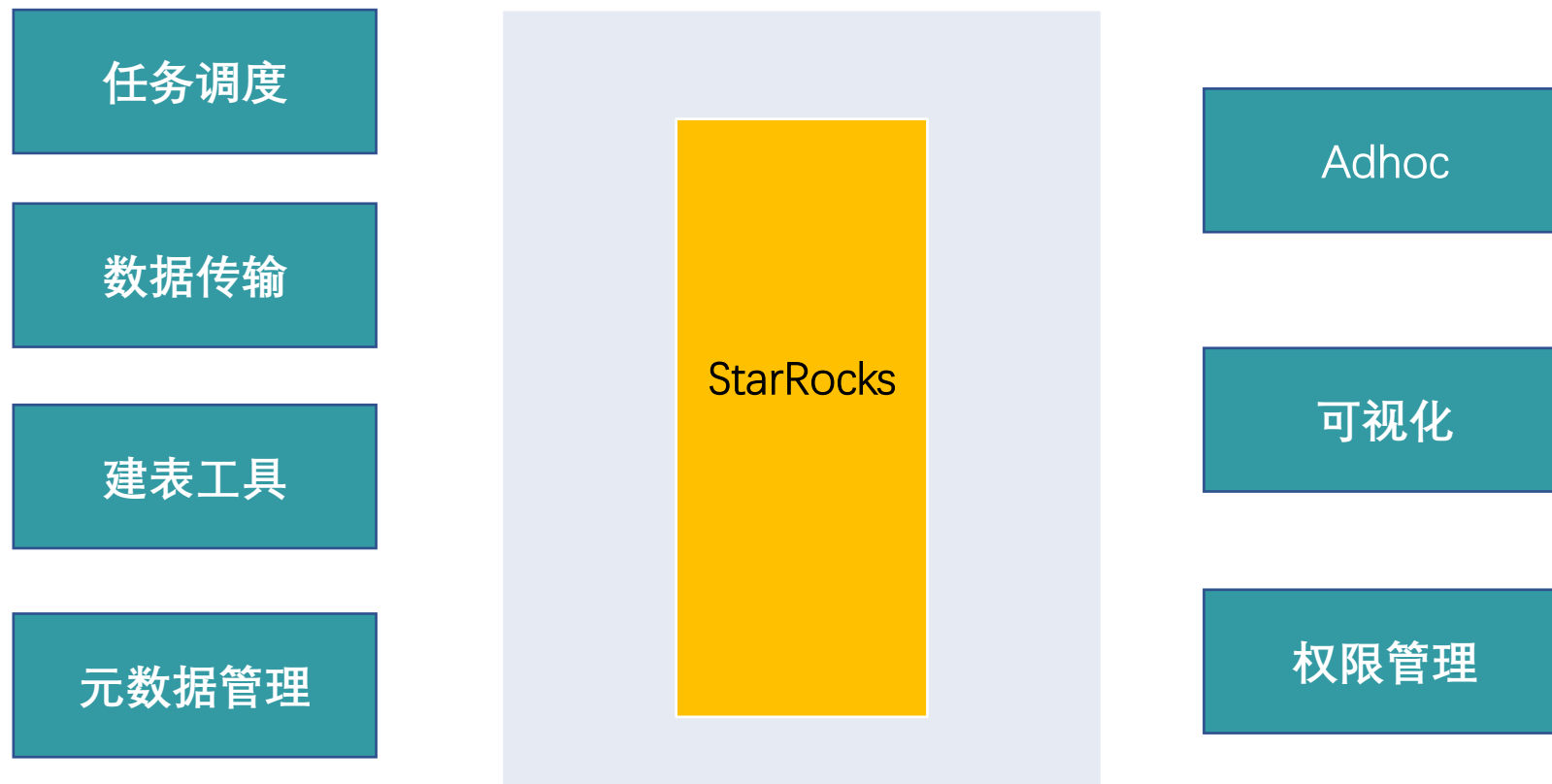


artnova报表Top100中的引擎分布

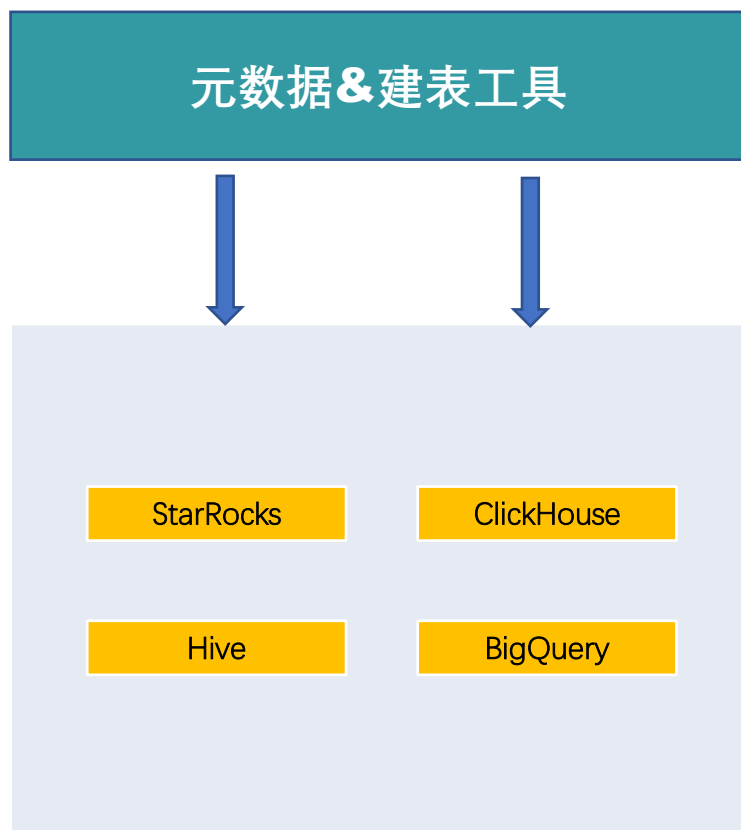


# 生态建设

# 系统建设

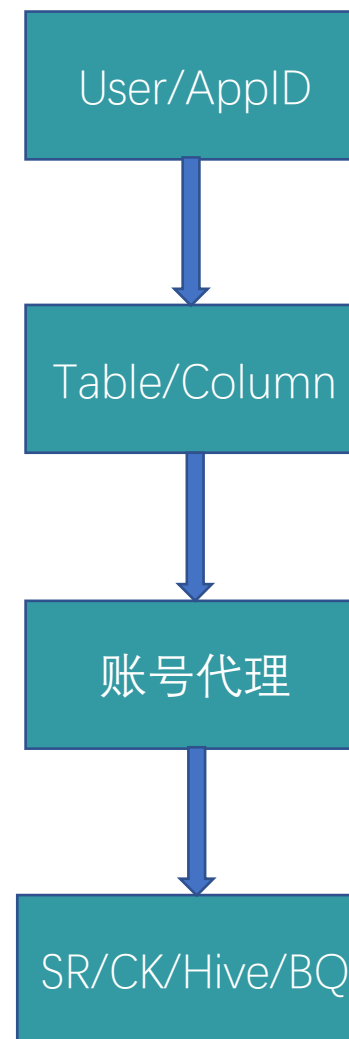
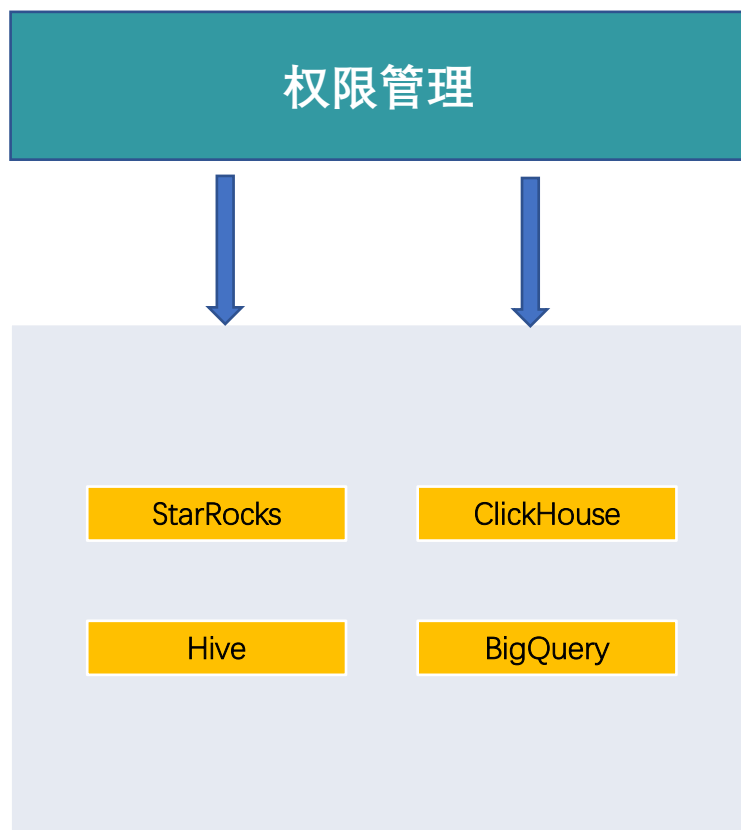


# 元数据&建表



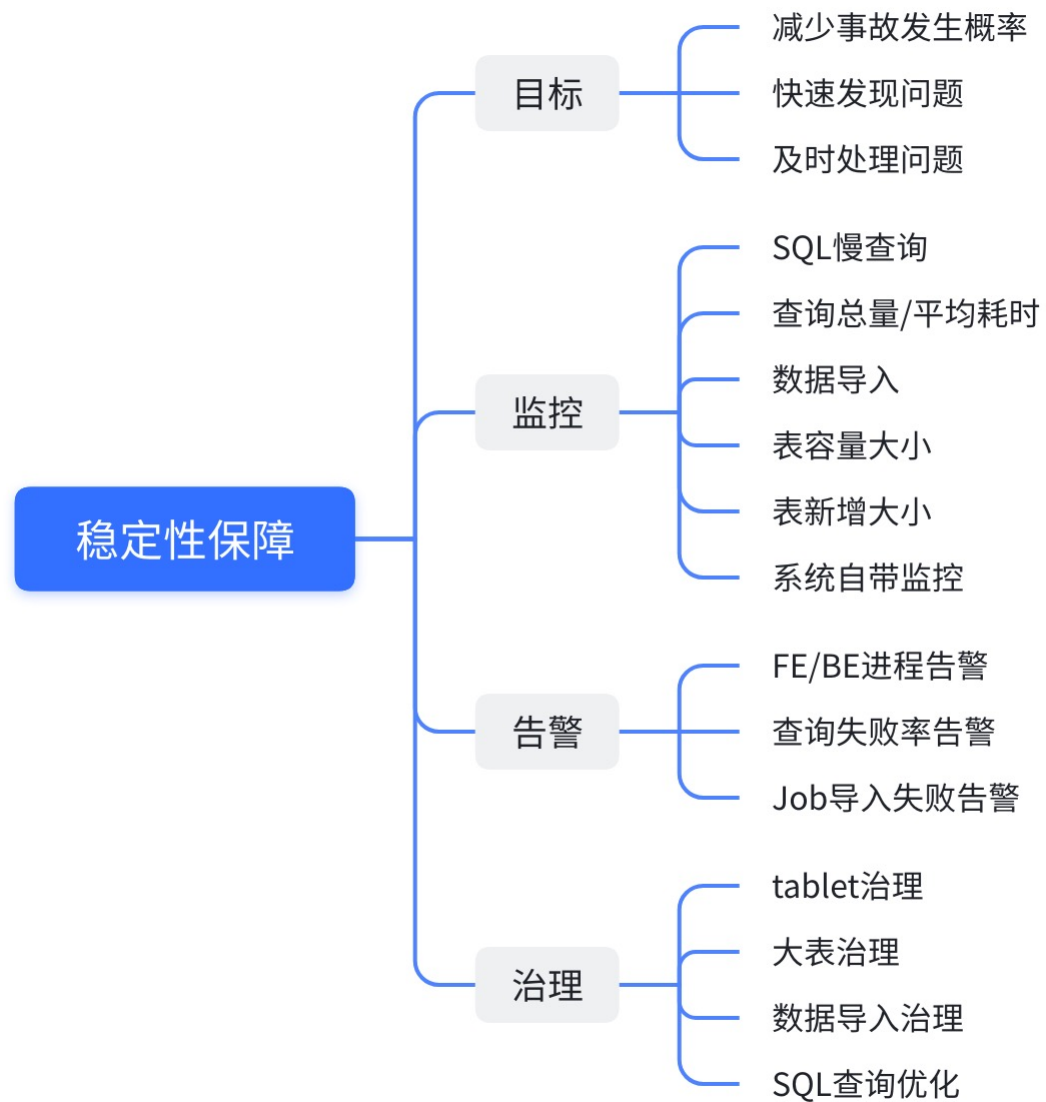
1. 建表工具, 表结构变更留有记录,可审计
2. 元数据, 方便服务API调用
3. 方便查看元数据信息,例如表结构, 大小, 分区信息, 数据owner, 业务任务等

# 权限管理

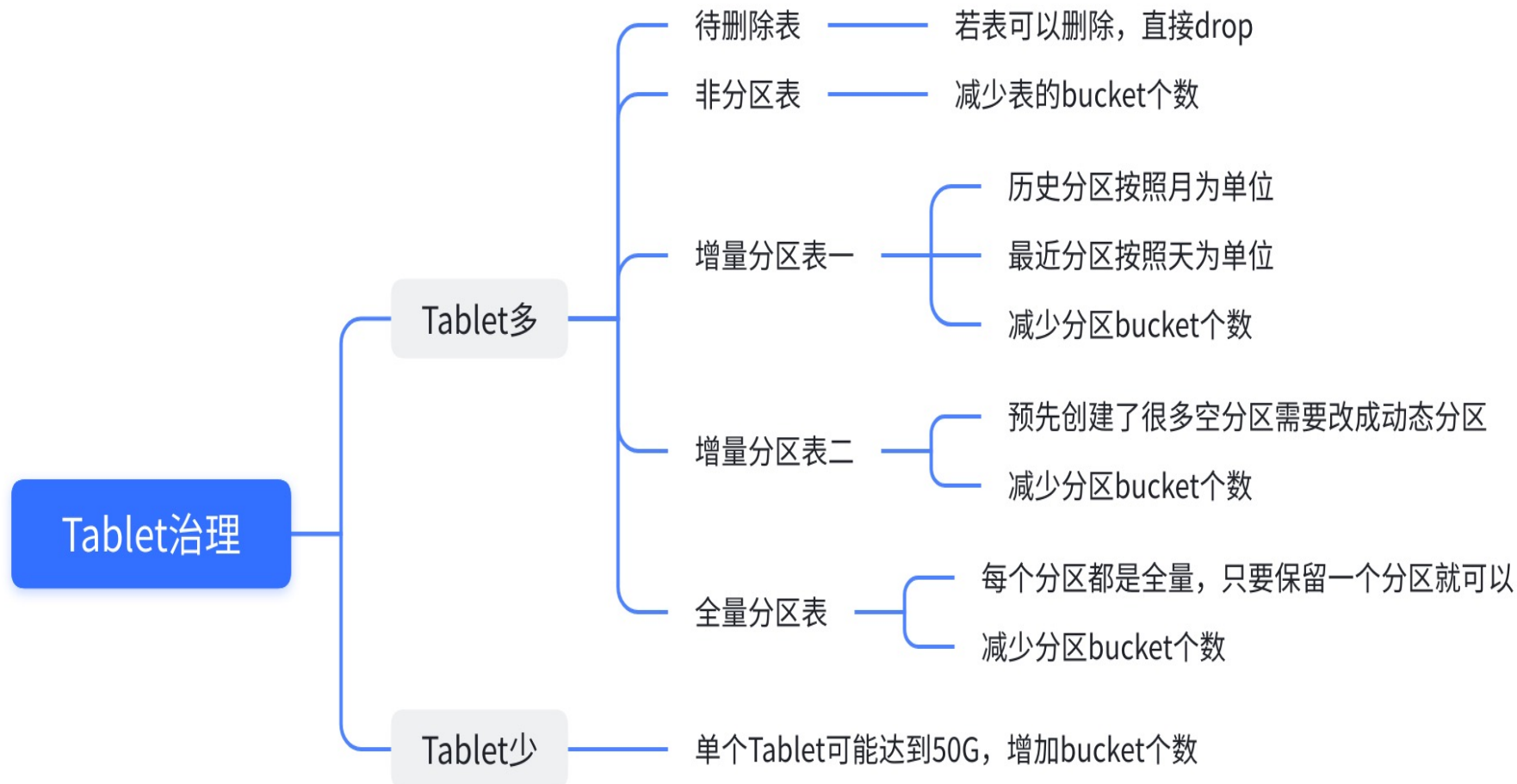


稳定性保障

# 稳定性保障

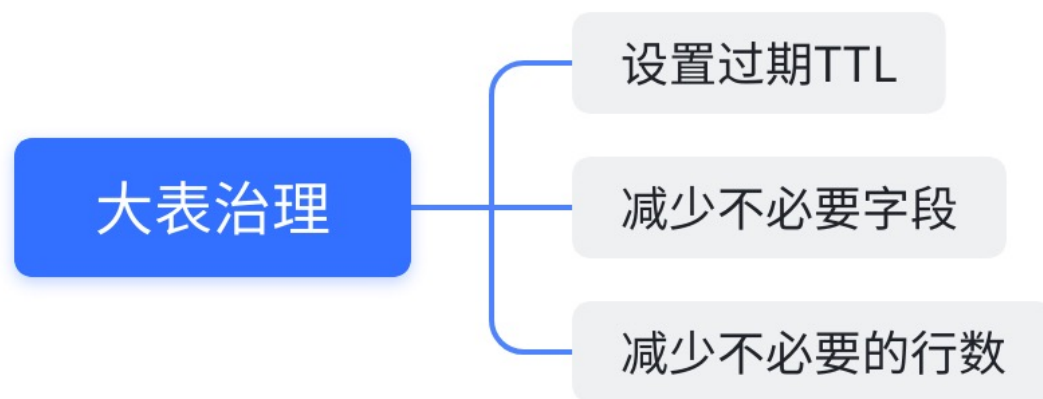


# Tablet治理

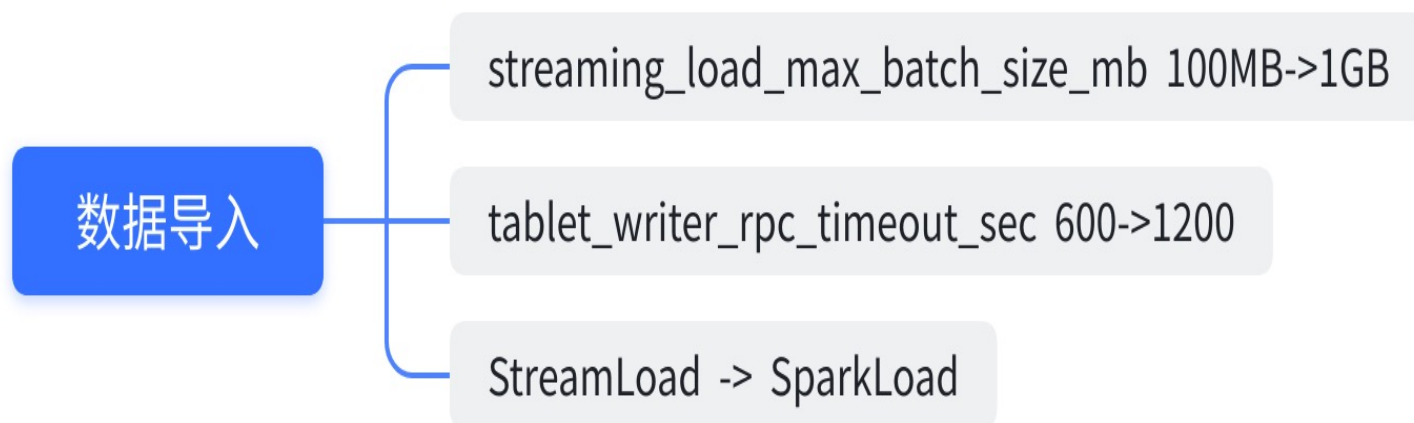




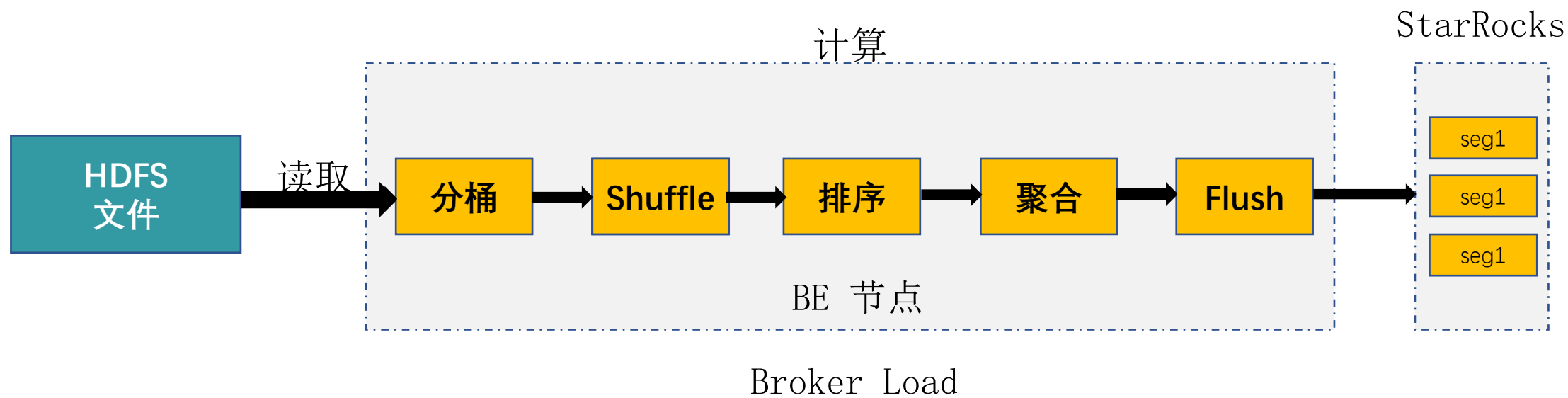
# 大表治理

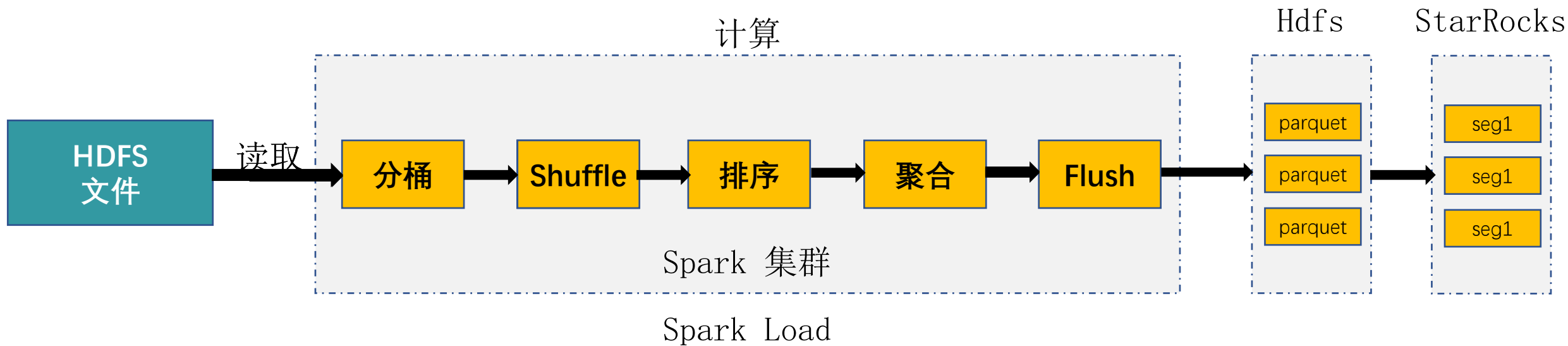


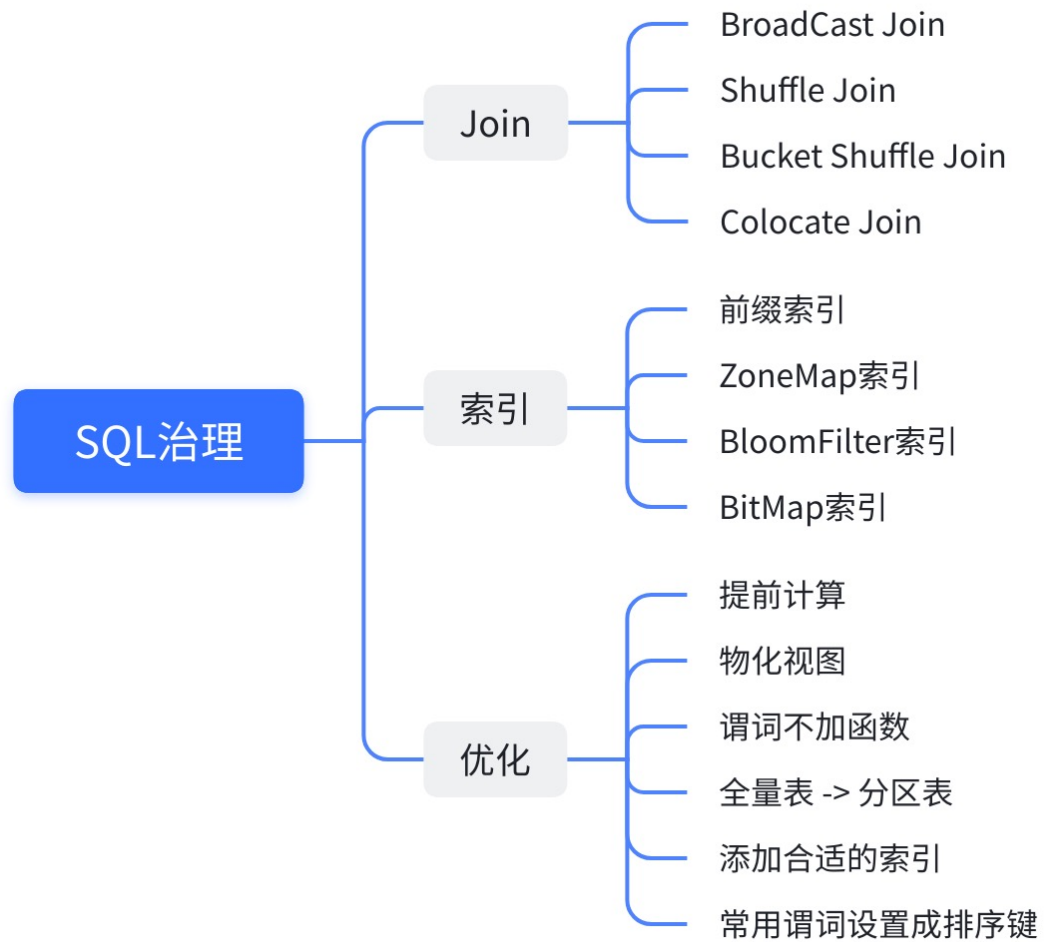
# 导入治理



# Broker Load







# 案例

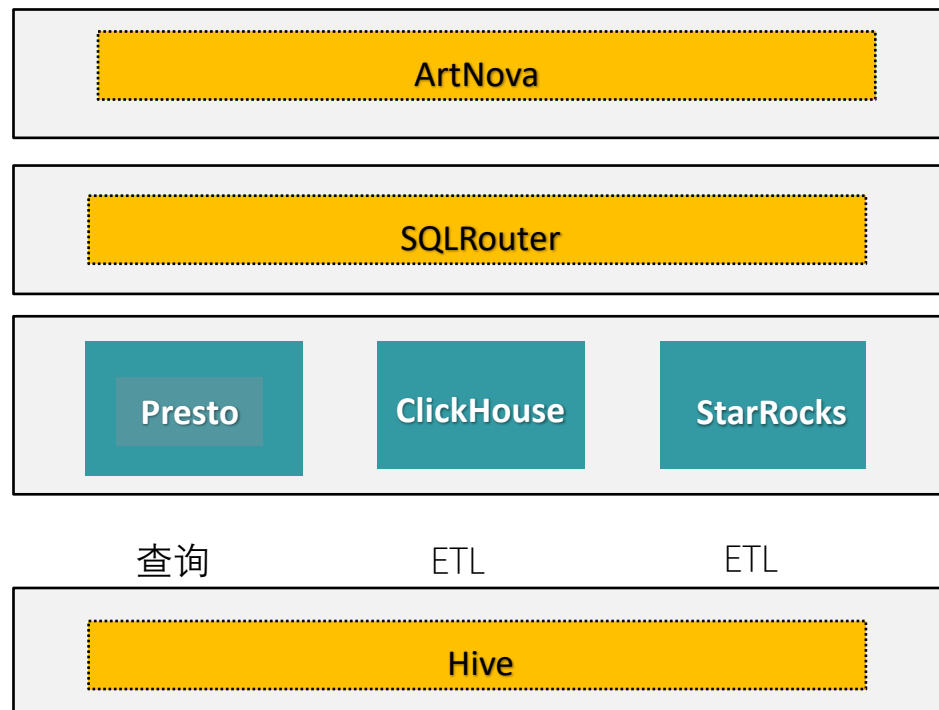
1. 分区粒度Day -> Hour <https://github.com/StarRocks/starrocks/pull/13073>
2. SR两级存储，分区和tablet。compaction的范围是tablet，随着时间的累加 读写放大严重，数据量大时，问题也显著放大



## 具体应用

# ArtNova

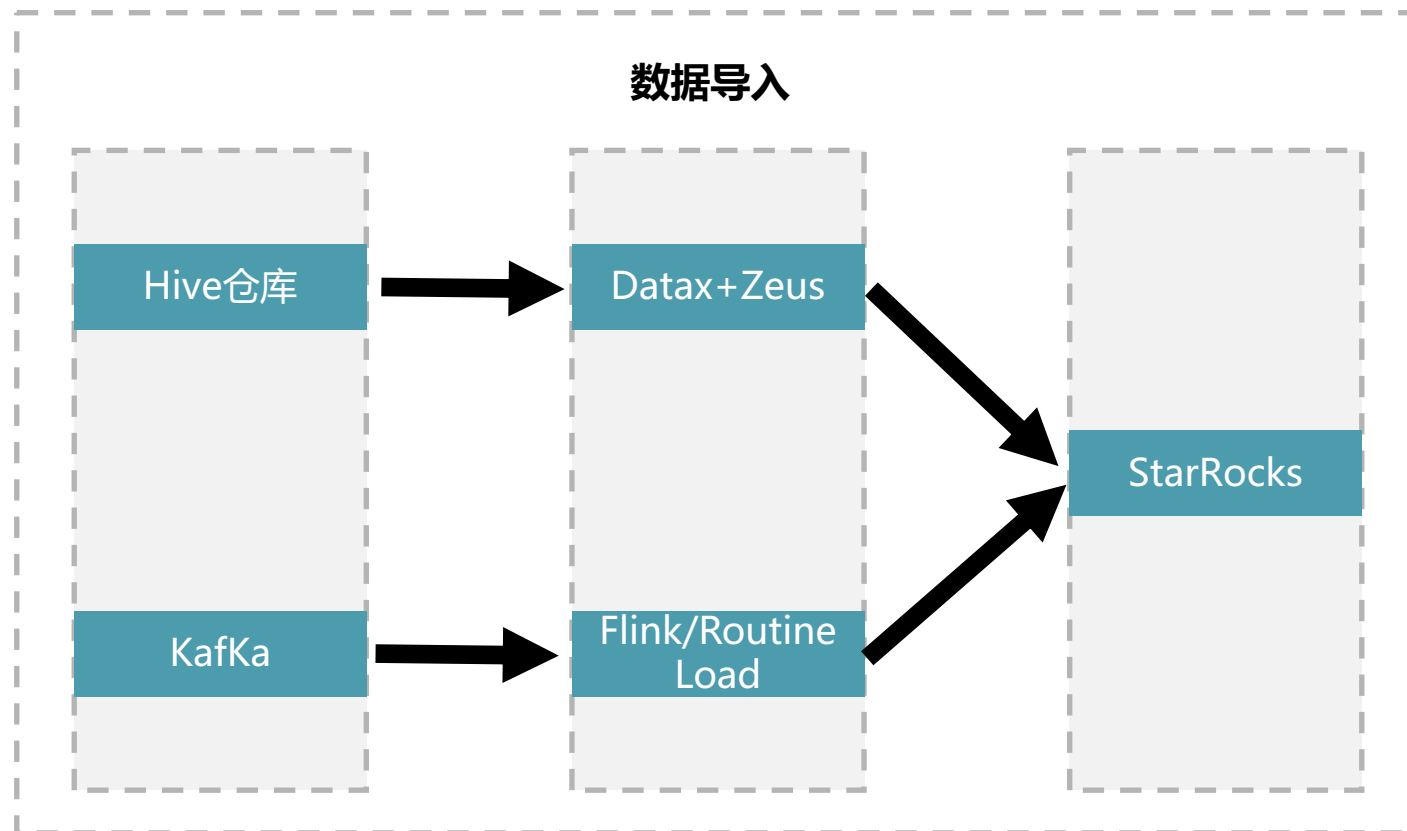
整体架构



## 挑战

1. SQL复杂
2. 整点流量大
3. DashBoard图表多
4. 查询耗时要求高





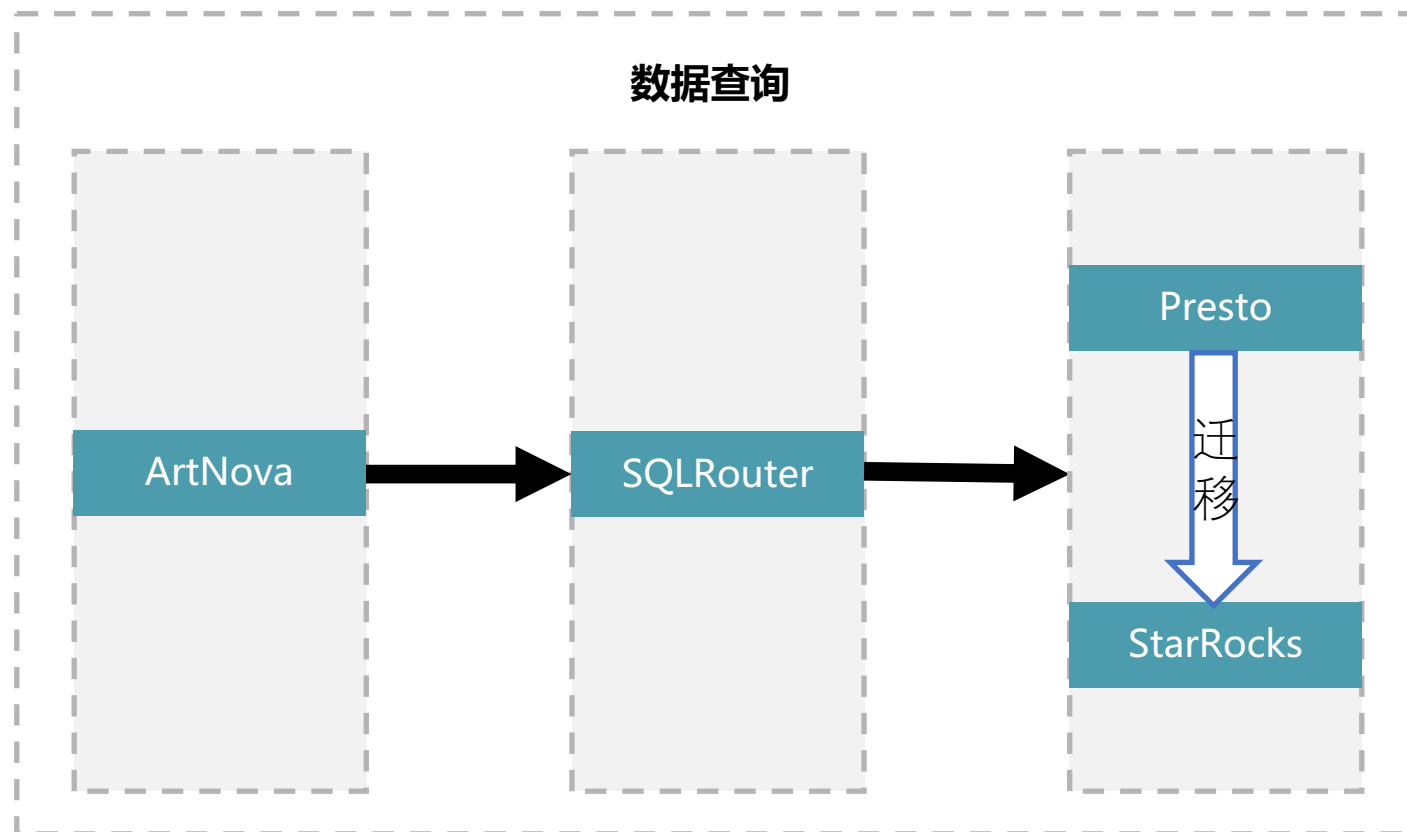
1

离线部分 通过zeus调度Datax从Hive仓库拉取数据导入到StarRocks

2

实时部分 通过Flink/RoutineLoad消费KafKa导入到StarRocks

## ArtNova



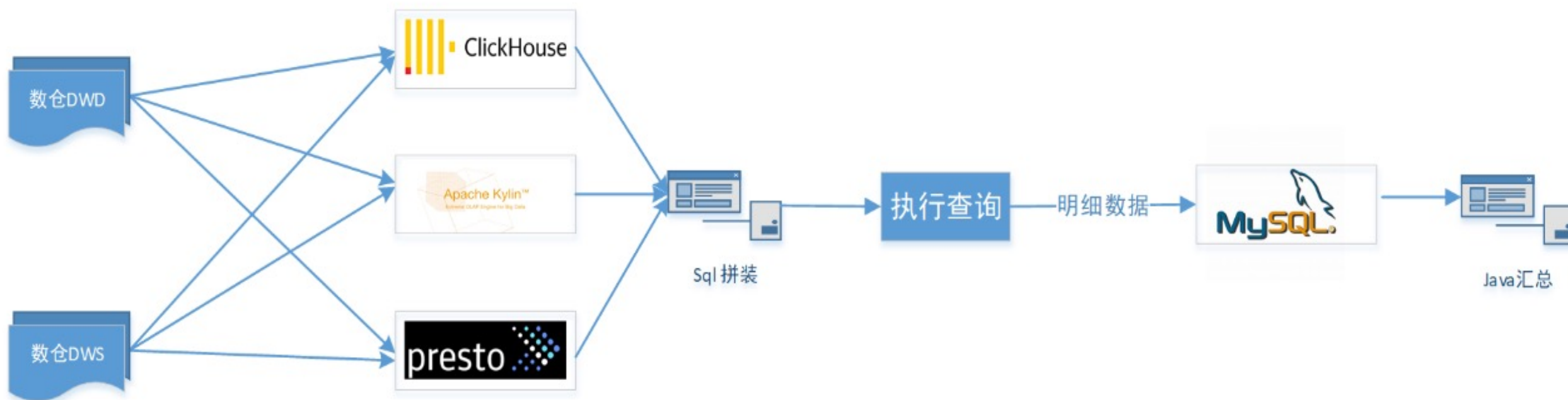
1

Top慢的报表从Presto引擎迁移到StarRocks

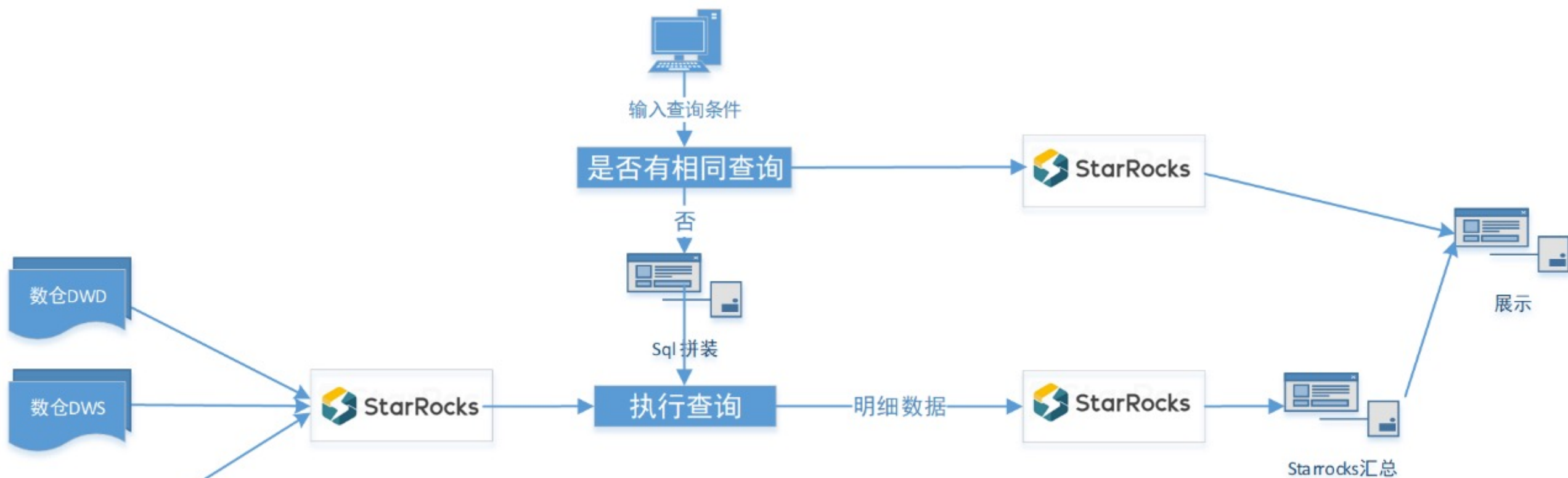
2

已迁移的报表 性能提升10倍+，平均耗时1.2秒

# 火车票指标平台



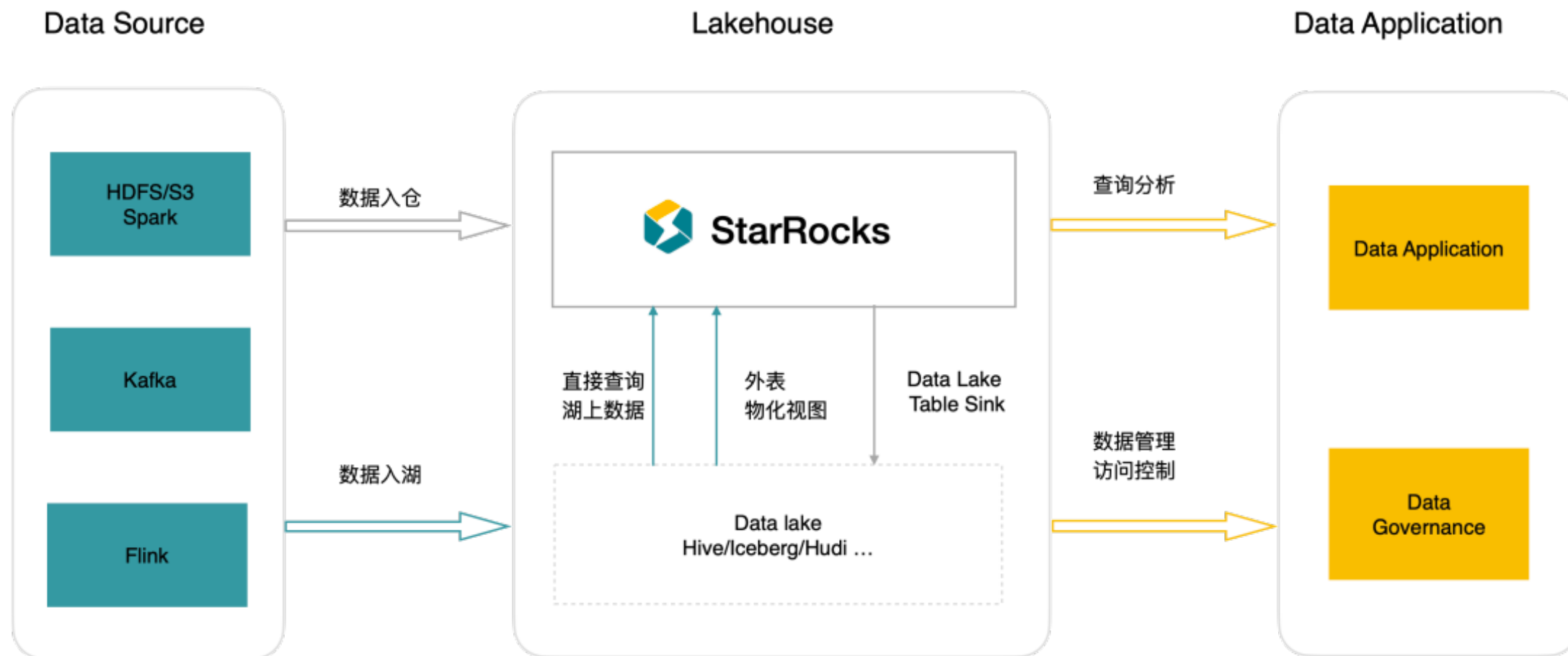
# 火车票指标平台



1. 复杂查询30min -> 10s
2. P99 2秒
3. 用户使用量翻倍

后续

# 湖仓融合





# StarRocks & Friends

StarRocks Meetup 上海站

活动时间：2023/06/17 13:30-17:00



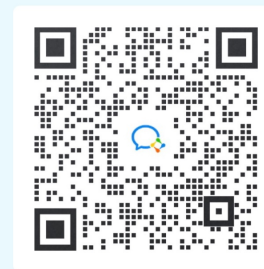
StarRocks 公众号



镜舟公众号



扫码填写问卷



User Group 上海

<https://github.com/StarRocks/starrocks>