



# 字节跳动用户反馈中台的大数据技术演进





# 讲师简介



唐浪飞  
大数据架构师

履历：曾就职于腾讯，vivo等互联网公司，主要负责大数据实时、离线、OLAP平台开发。

工作：目前在字节跳动，主要从事反馈中台的大数据体系建设，基础组件建设和分布式系统优化等基础技术实践。



曹铭斌  
大数据架构师

履历：曾就职于腾讯，OPPO等互联网公司，主要负责大数据实时、离线、OLAP平台开发。

工作：目前在字节跳动，主要从事反馈中台的大数据与智能化体系建设，大数据基础平台建设和OLAP引擎优化等基础技术实践。





# 目录

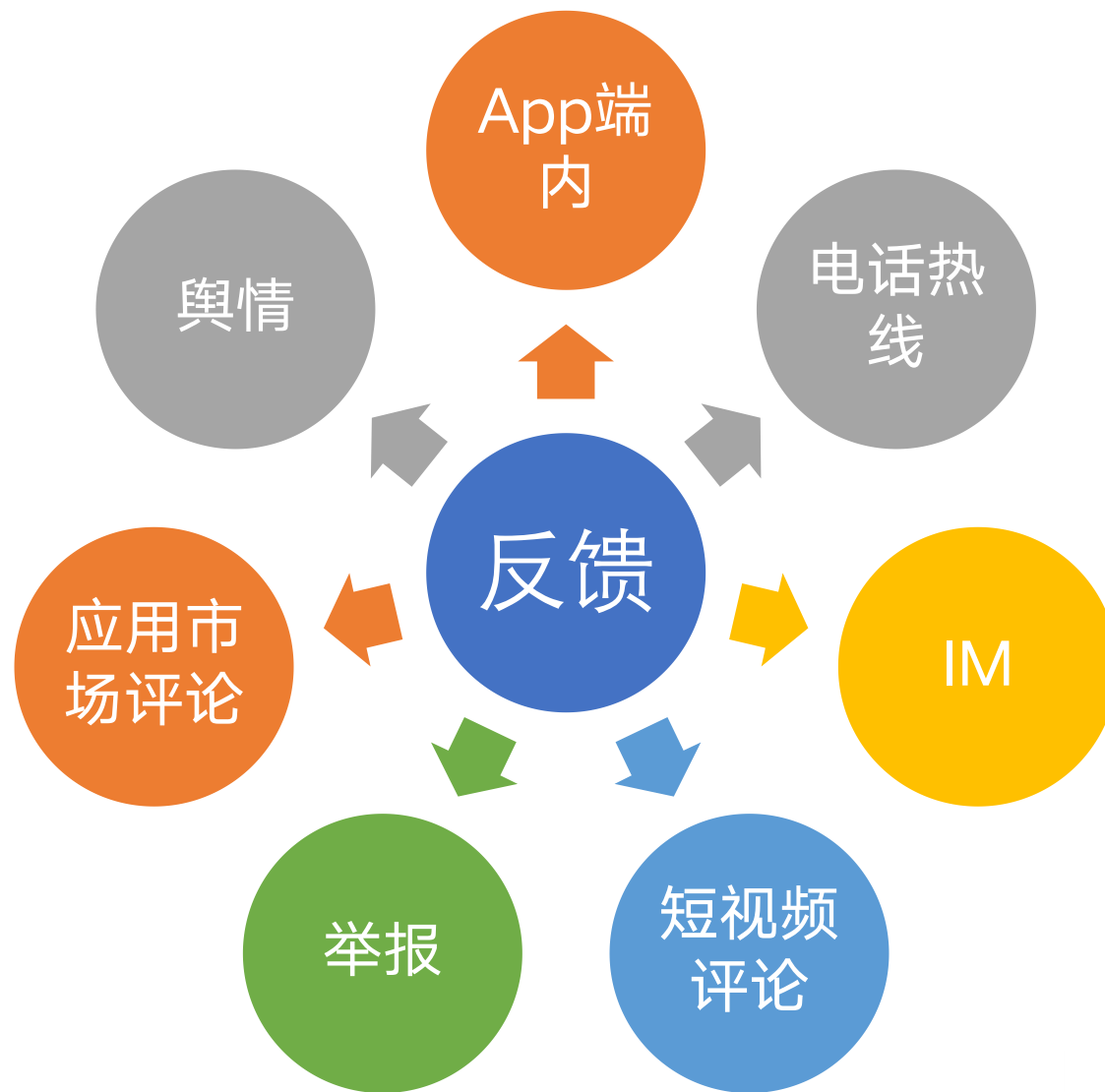
1. 前言
  1. 什么是用户反馈
  2. 用户反馈的价值
2. 业务背景
3. 字节跳动反馈中台实时架构演进
  1. 泛反馈大数据平台
  2. 基于NLP分类的标签和关键词挖掘
  3. 通用的高吞吐实时数据接入引擎
  4. PB级ElasticSearch存储实践
  5. 架构演进-OLAP引擎ClickHouse
  6. ClickHouse优化经验
4. 未来规划
  1. 平台架构
  2. 架构下一次演进





# 什么是用户反馈

字节跳动作为全球月活超19亿(截至2021年6月)的互联网公司，日均用户反馈超**百亿级**，在这场景下，快速识别**用户意图**、挖掘**用户反馈价值**是对平台能力的极大挑战与价值体现。





## 抖音用户反馈产品思考



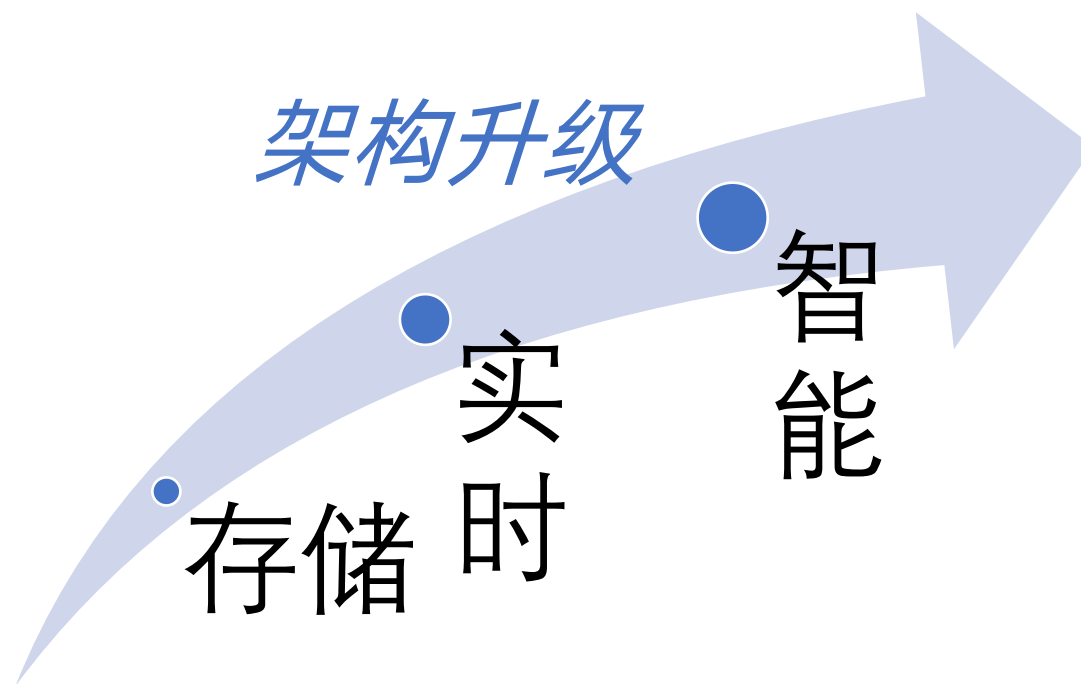


# 业务驱动架构升级

字节跳动用户反馈中台历经数据规模由万级到百亿级，业务线多达90多个，收到如此多的反馈，人工离线处理已经无法满足业务发展。

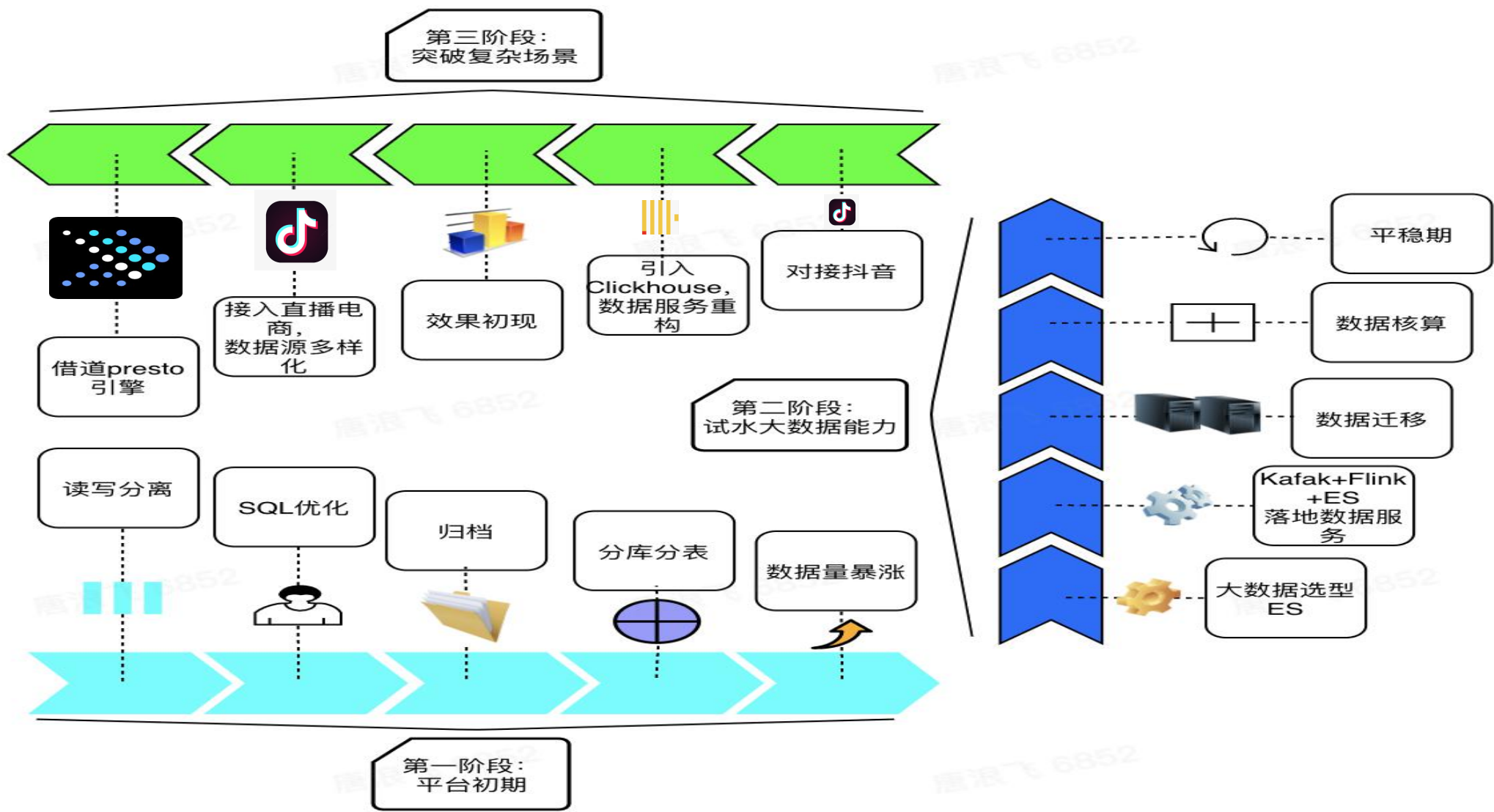
痛点：

1. 数据量大
2. 业务线多
3. 无效反馈占大多数
4. 分析工具缺乏
5. 分析耗时长
6. 问题发现时效低





# 泛反馈大数据平台架构演进历程





引入Kafka, Flink, ES等大数据组件, 通过构建全球的实时数据流, 快速处理用户反馈数据

存储

- PB级数据存储

实时

- 高吞吐, 低延迟

扩展

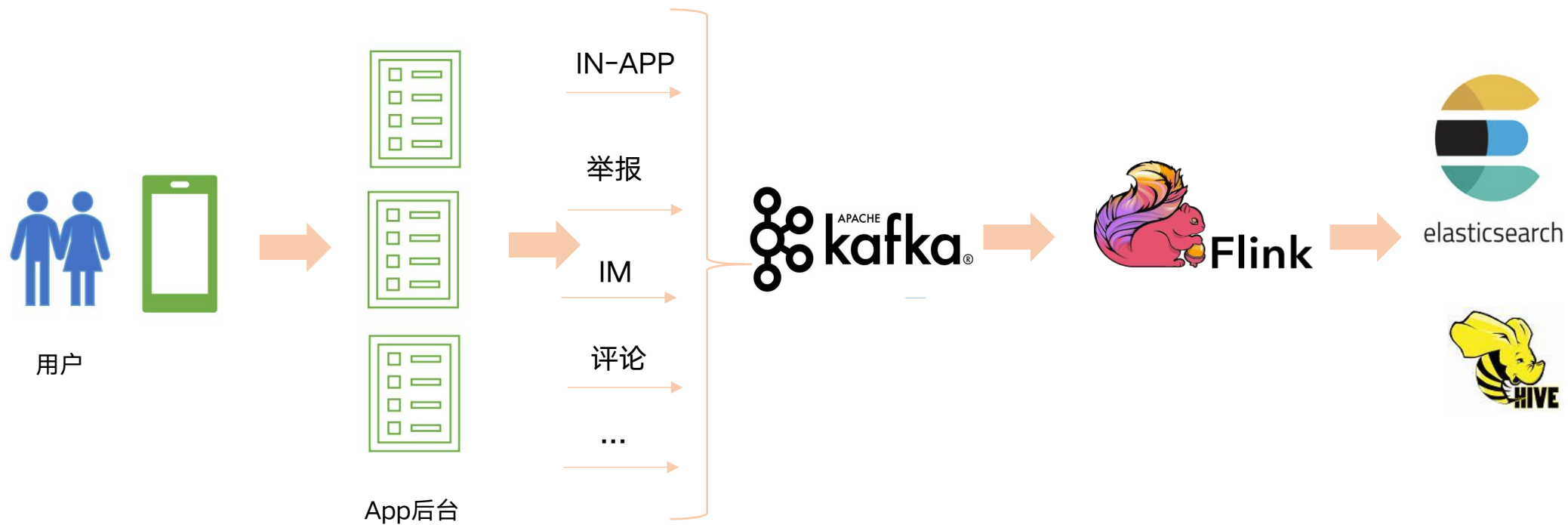
- 适配多业务线的异构数据

智能化

- 基于NLP分类的标签和关键词挖掘

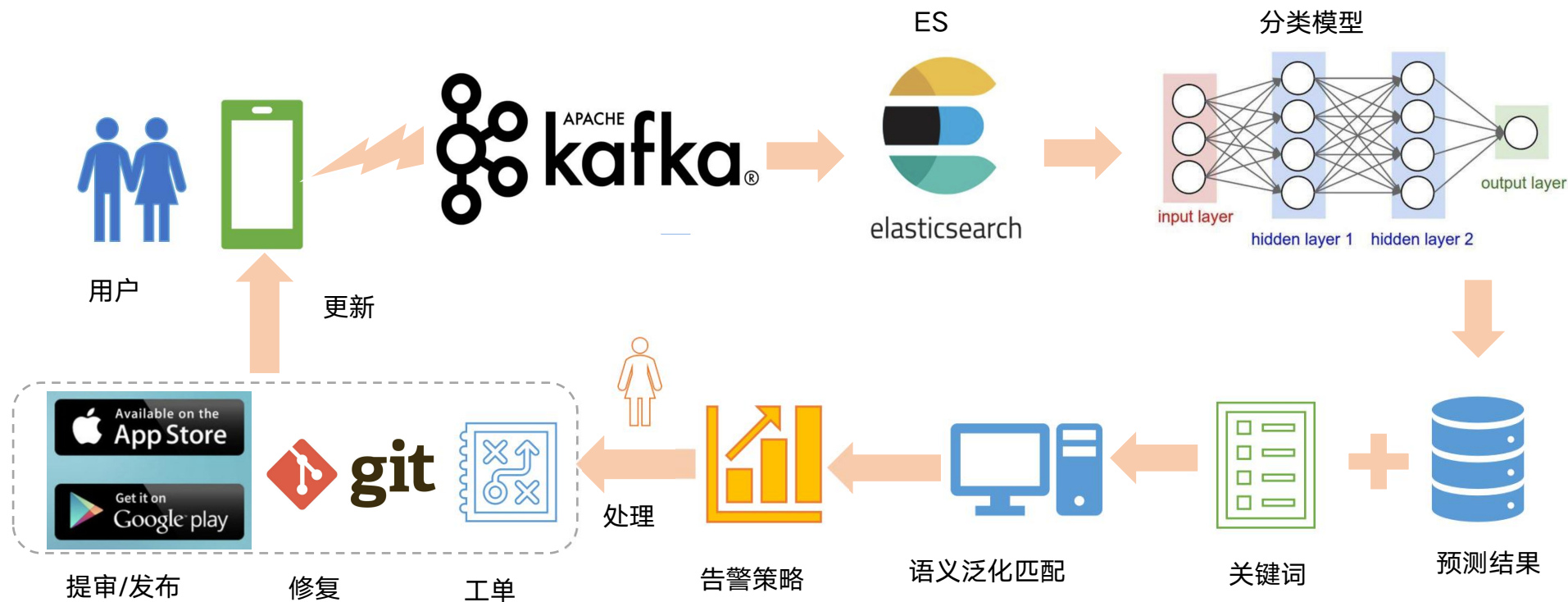








# 基于NLP分类的标签和关键词挖掘



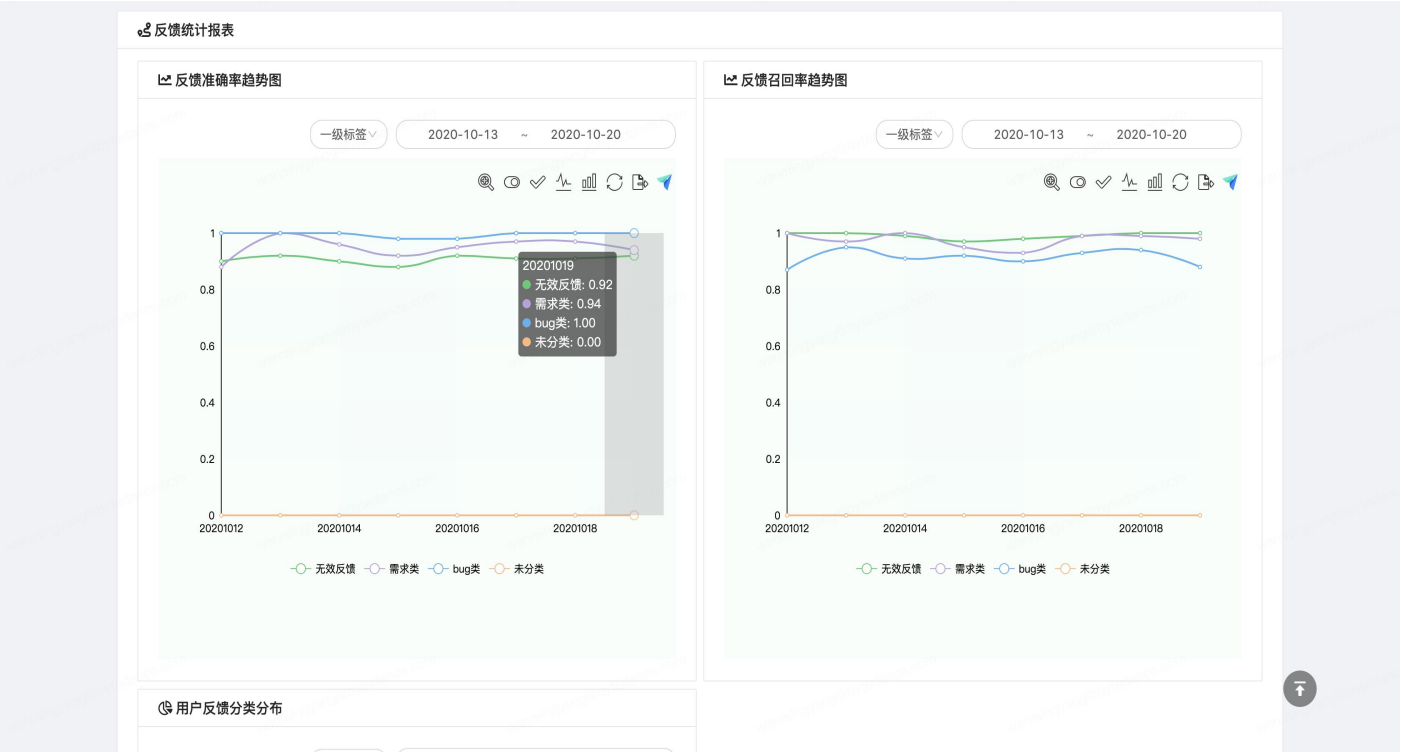


## 落地效果



### 三级分类：

- 一级BUG类准确率95%以上
- 二级BUG类准确率92%以上
- 三级BUG类准确率91%以上





# 通用的高吞吐实时数据接入引擎

平台已经接入多种反馈渠道，90多个数据源。

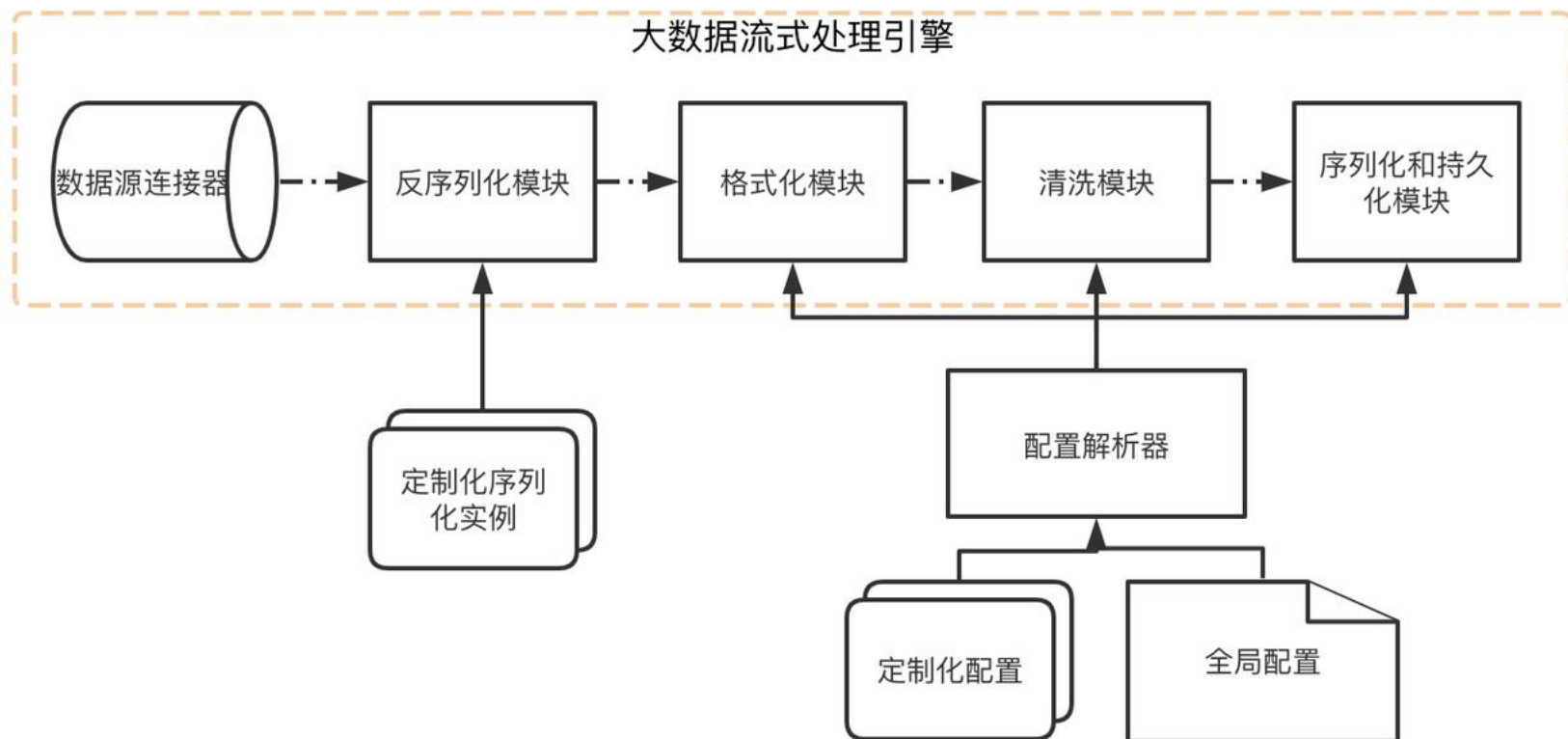
基于Flink框架开发的通用的数据接入引擎，使用统一的数据接入框架，适配多业务线的异构数据，将极大地提高开发、运维的效率





通过业务抽象，一项数据接入任务一般可分为以下几个子步骤：

- 反序列化
- 格式化
- 清洗
- 序列化
- 持久化。

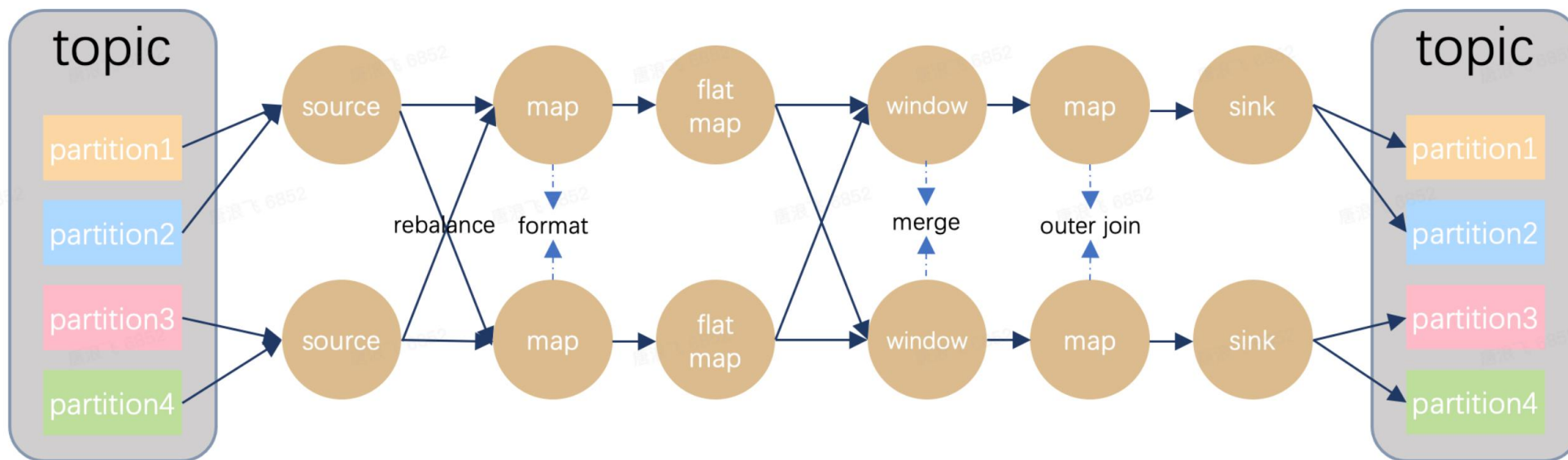




# 通用的高吞吐实时数据接入引擎

高吞吐优化：

- 原始数据进行shuffle，分发给多个窗口进行处理，每个窗口处理完成后都将数据放入内存缓存，再批量写入目标存储，从而达到短时间的高并发效果
- 结果数据消重





## 数据合并：精准去重

### MapState

- MapState 是 Flink 中 KeyedState 的状态类型，将主键保存在 MapState，实现精准去重

### HyperLogLog

- HyperLogLog 算法也就是基数估计统计算法，结合redis，预估一个集合中不同数据的个数

### bitmap

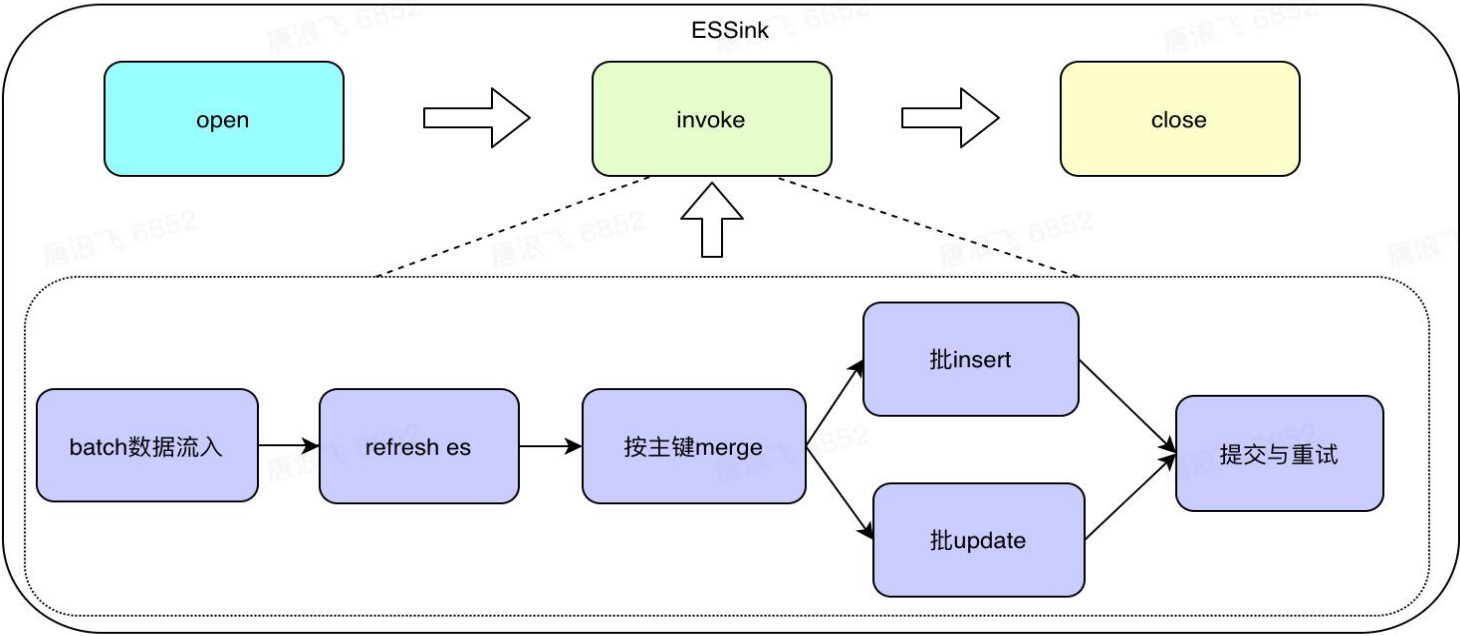
- 优点是精确去重，占用空间小(在数据相对均匀的情况下)。缺点是只能用于数字类型(int或者long)，不过可以通过hash方式进行转换解决





## flink sink ES端优化:

自研 es connector, 支持异常重试, 数据merge, es refresh等功能, 处理逻辑如下:

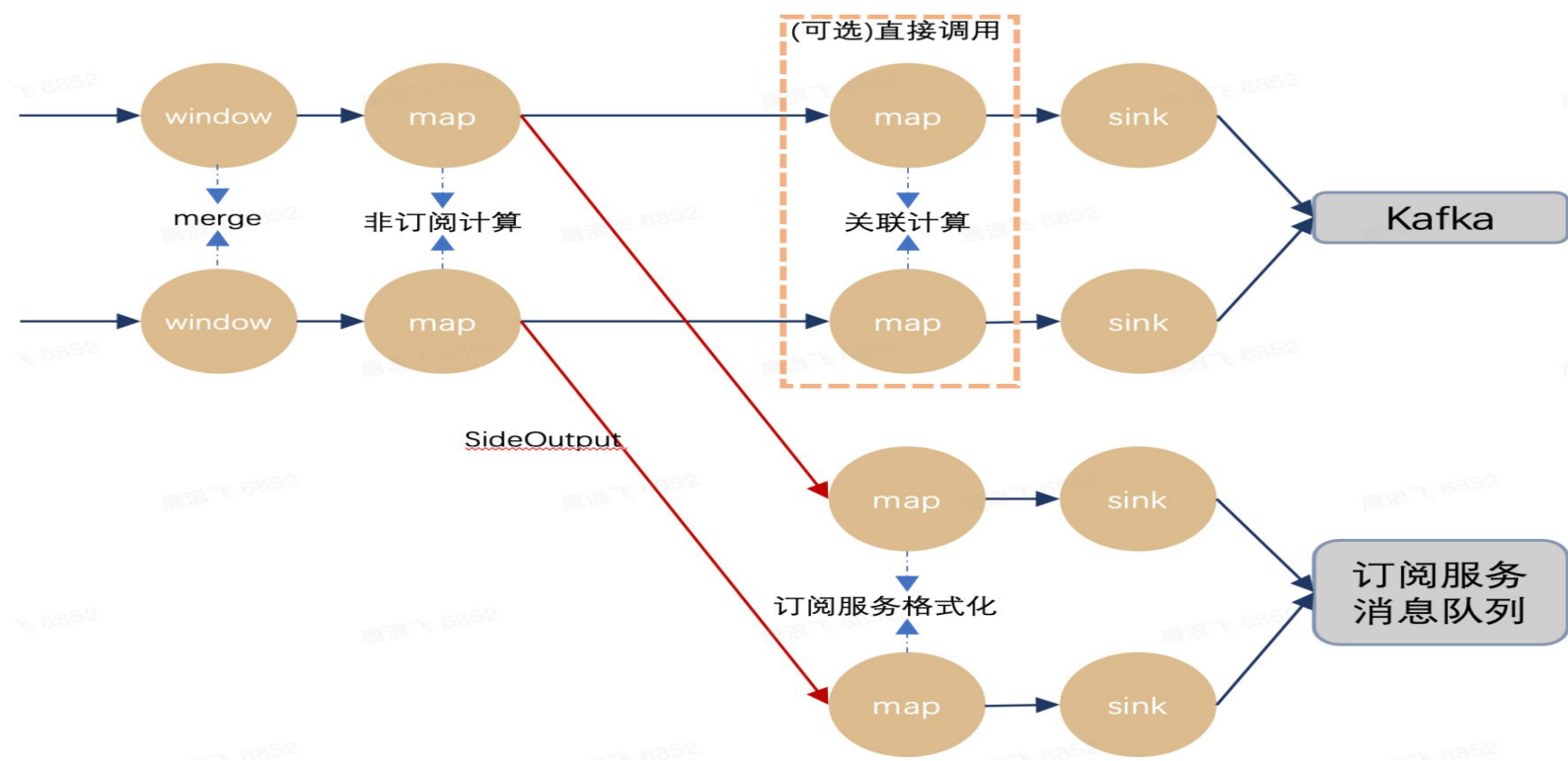






## 大数据与算法能力结合：

算法模型调用的QPS与大数据计算自带冲突，但是泛反馈平台是依托大数据+算法能力构建，才能挖掘出异常的实时问题和及时告警处理，缺一不可。





# 通用的高吞吐实时数据接入引擎

高吞吐：  
采用改进后的数据处理系统，200个Partition，50Taskmanager(4C8G)，  
每分钟能够处理达到200万的数据，并且延迟下降到毫秒级别

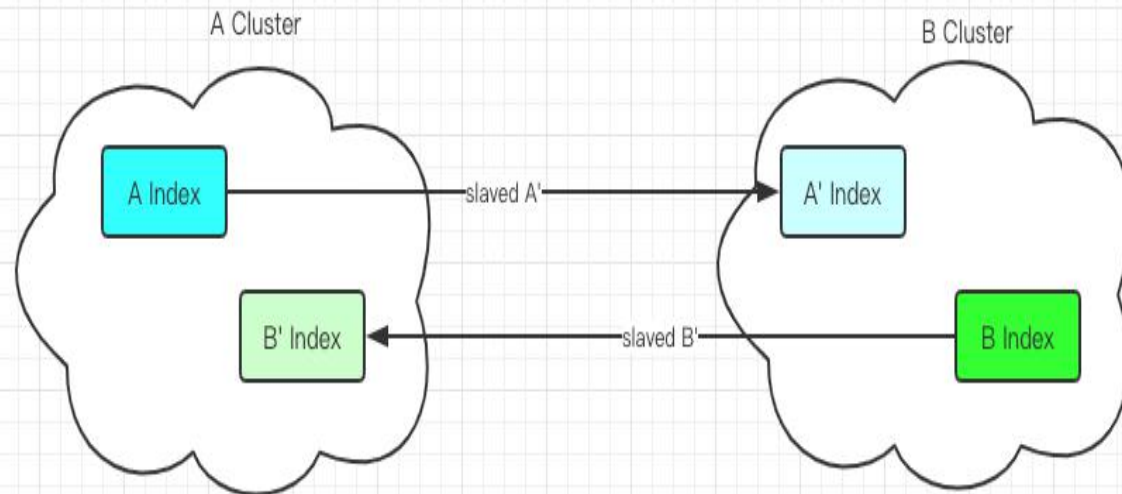
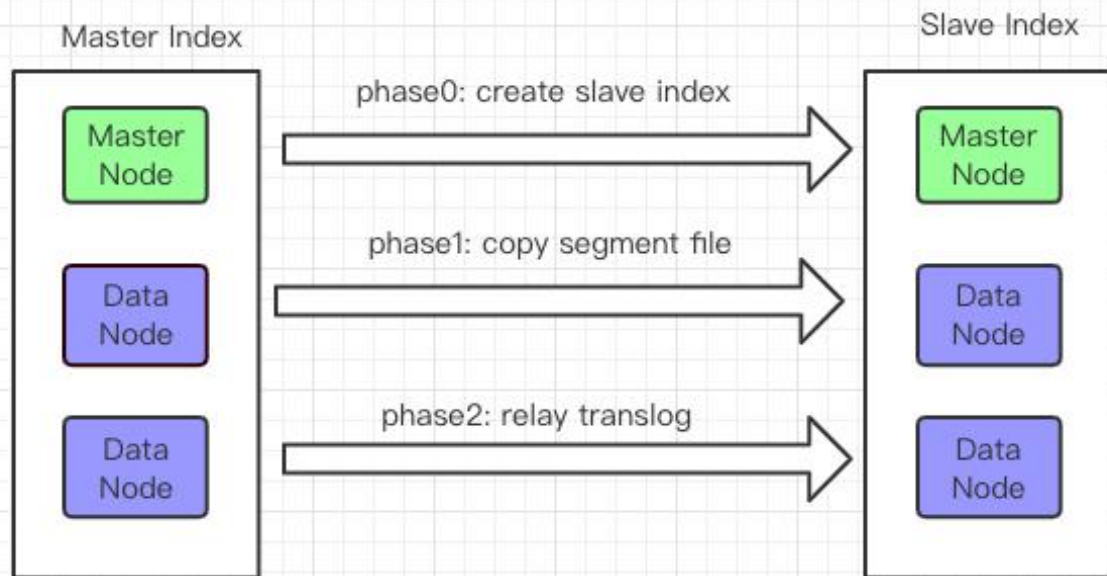




# PB级ElasticSearch存储实践

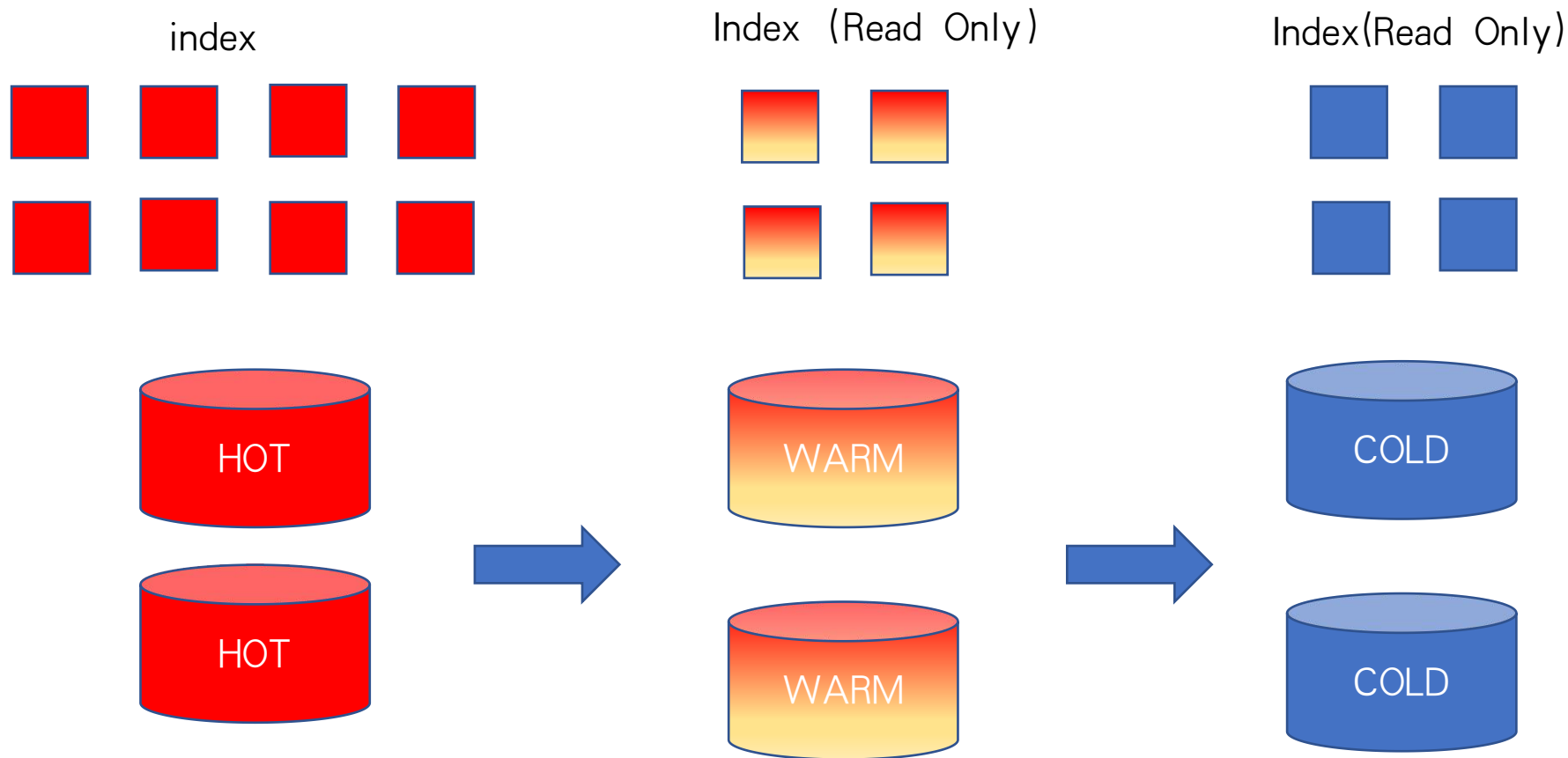
读写分离

- freeCCR（自研，非官方）同步存量数据
- 双机房集群部署





ILM索引生命周期管理-冷热模式，滚动，单分片不超过50G，shrink





## 落地效果

- 热机20+T， 温机+冷机 1.5P
- 13台物理机： CPU:96Core RAM:384GB SSD-2T\*1 HDD-16T\*24
- 每台2个实例
- 冷机10个实例
- 写入QPS可以达到3w/s

基本信息

集群

运营情况

治理

集群 ID

状态

机房

部署信息

更新时间

22881

运行中



类型：物理机 [服务树](#)  
实例个数：26

2021-01-15 06:18:45



## ES遇到问题与业务挑战

- 20w+/s的大数据量写入，写入延迟高，偶尔发生reject
- 业务需求原因，Flink sink es需要强制refresh。
- 聚合查询慢，甚至部分节点超时导致结果不全。
- 语法局限，不能支持丰富的聚合算子，大部分聚合算子支持性能较差。





# 架构演进-OLAP引擎ClickHouse

## OLAP技术选型

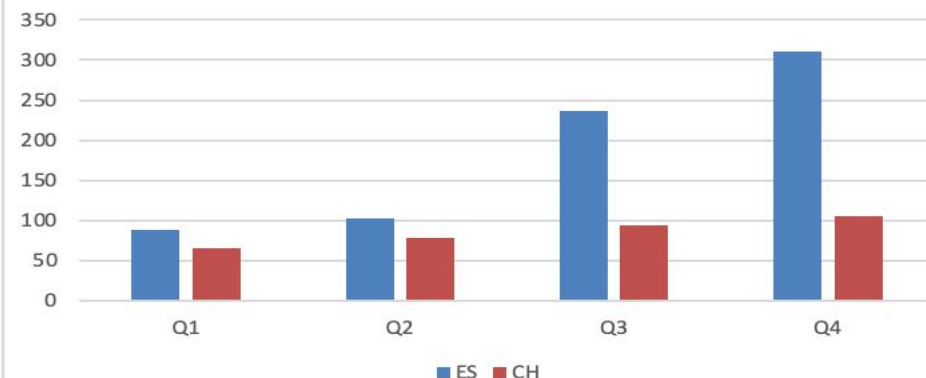
- ROLAP (ClickHouse / ES / Presto / Doris)
- MOLAP (Druid / Kylin)

## CH与ES性能对比:

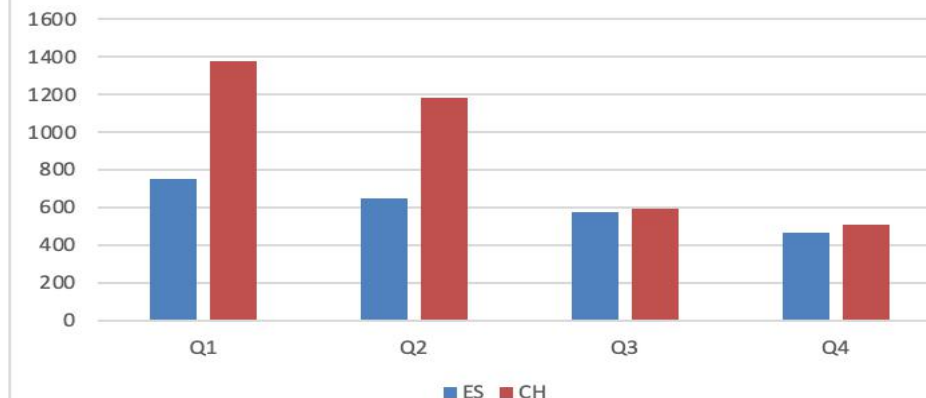
在响应时间和高并发方面, CH性能明显优于ES。

所以选择Clickhouse

查询单并发测试ms



查询高并发测试(并发个数)





## 技术挑战

- 抖音评论数据量20+w/s。
- 业务要求数据按key全局去重并且取最新时间的一条。
- 平台架构支持多引擎的写入与查询。



groupId Nullable(Int64)	updateTime DateTime
7011395954217585951	2021-09-29 07:16:05
7011395954217585951	2021-09-29 08:34:08
7011395954217585951	2021-09-29 12:04:49
7011395954217585951	2021-09-29 14:38:42
7011395954217585951	2021-09-29 15:24:01
7011395954217585951	2021-09-29 16:27:03
7011395954217585951	2021-09-29 17:57:28
7011395954217585951	2021-09-29 18:33:14
7011395954217585951	2021-09-29 22:54:19
7011395954217585951	2021-09-29 23:38:33



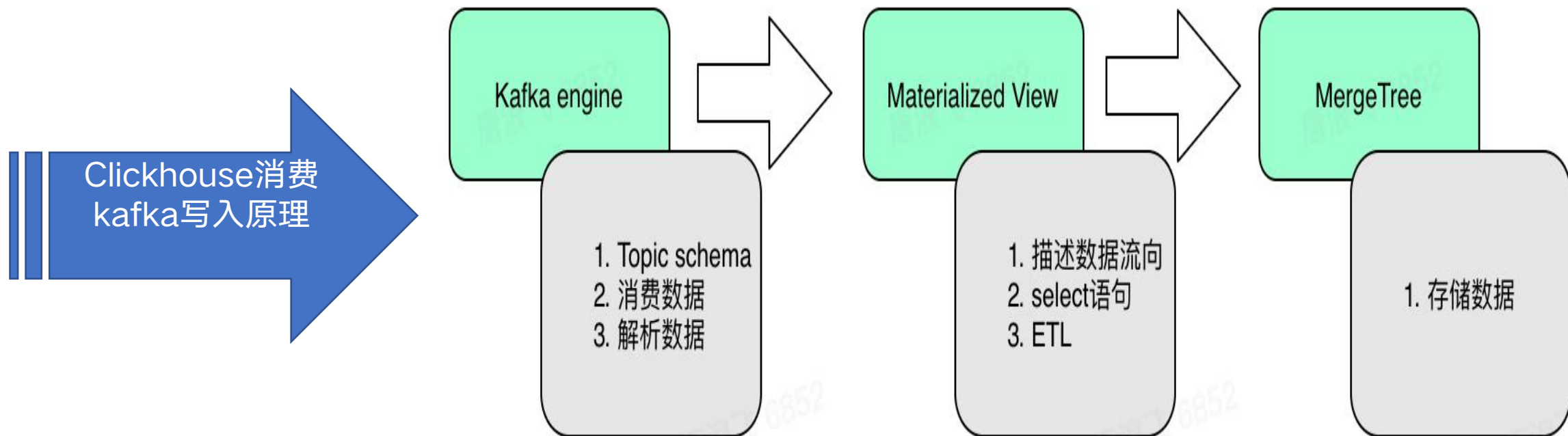




# 架构演进-OLAP引擎ClickHouse

解决方案: 引入高性能Clickhouse

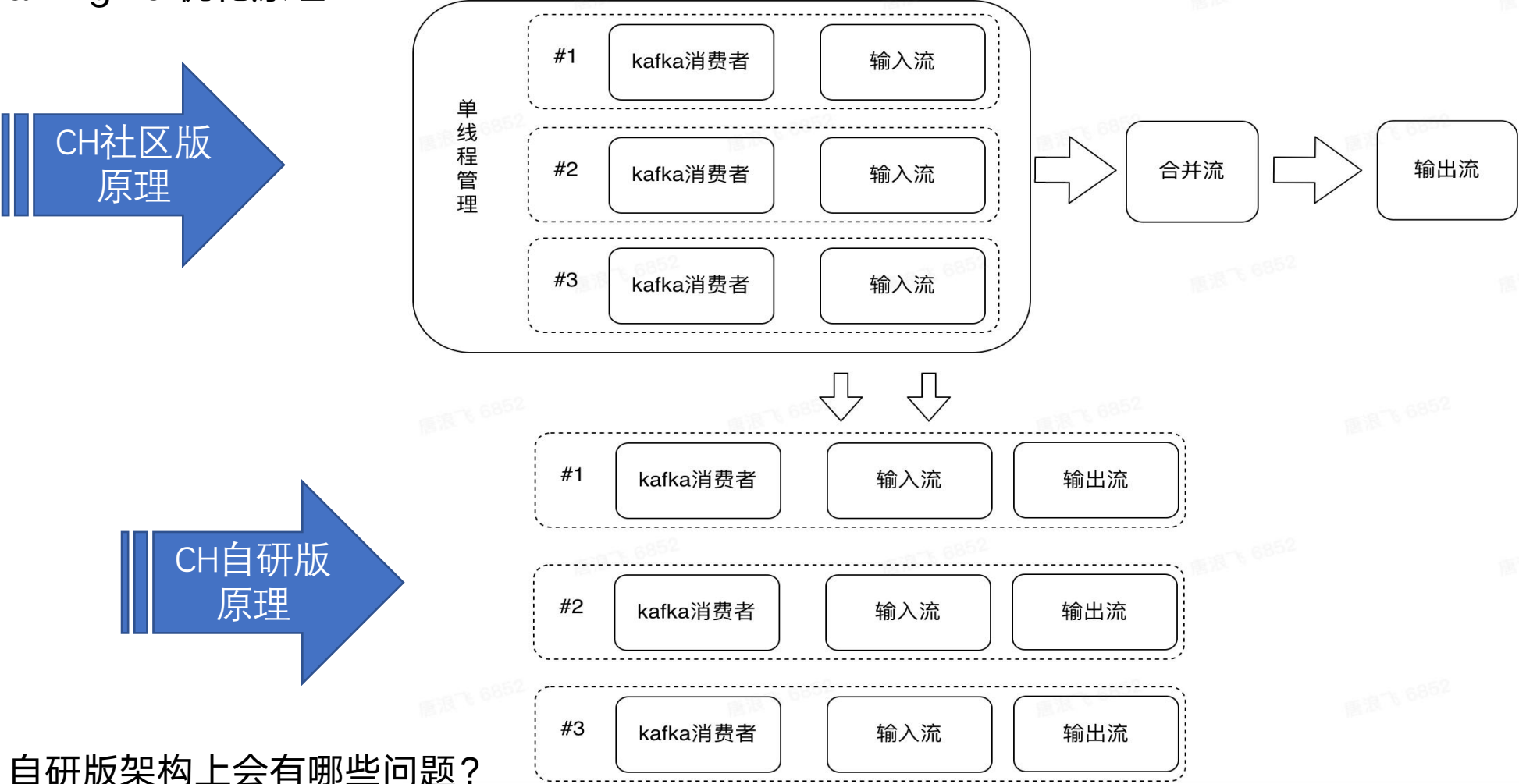
- 支持单机5w/s写入
- 支持精准去重HaUniqueMergeTree引擎，而不是ReplacingMergeTree引擎
- 抽象数据服务SDK，支持多引擎插拔式扩展能力。





# 架构演进-OLAP引擎ClickHouse

Kafka Engine 优化原理:



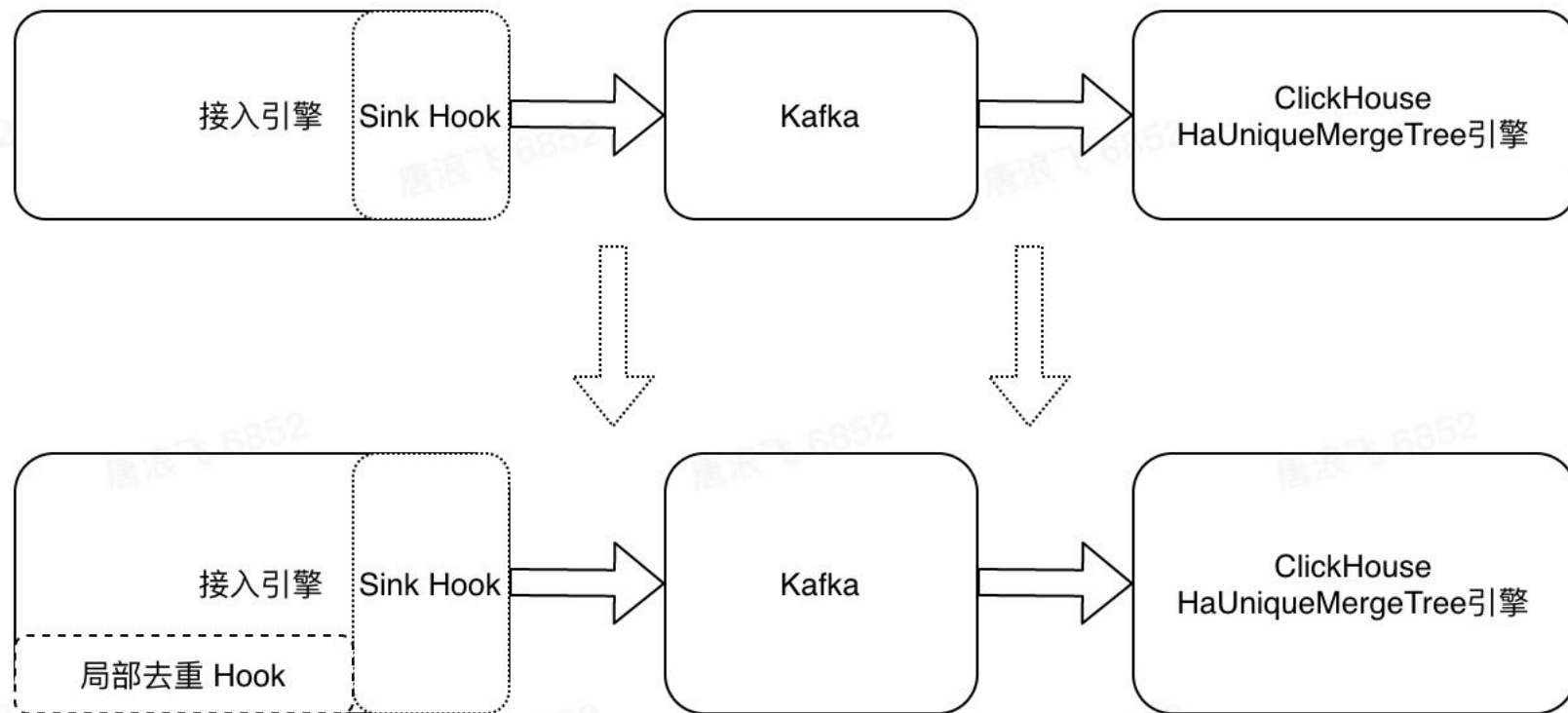
思考：自研版架构上会有哪些问题？





# 架构演进-OLAP引擎ClickHouse

业务需求：精确去重，需要对相同主键的数据进行清洗合并

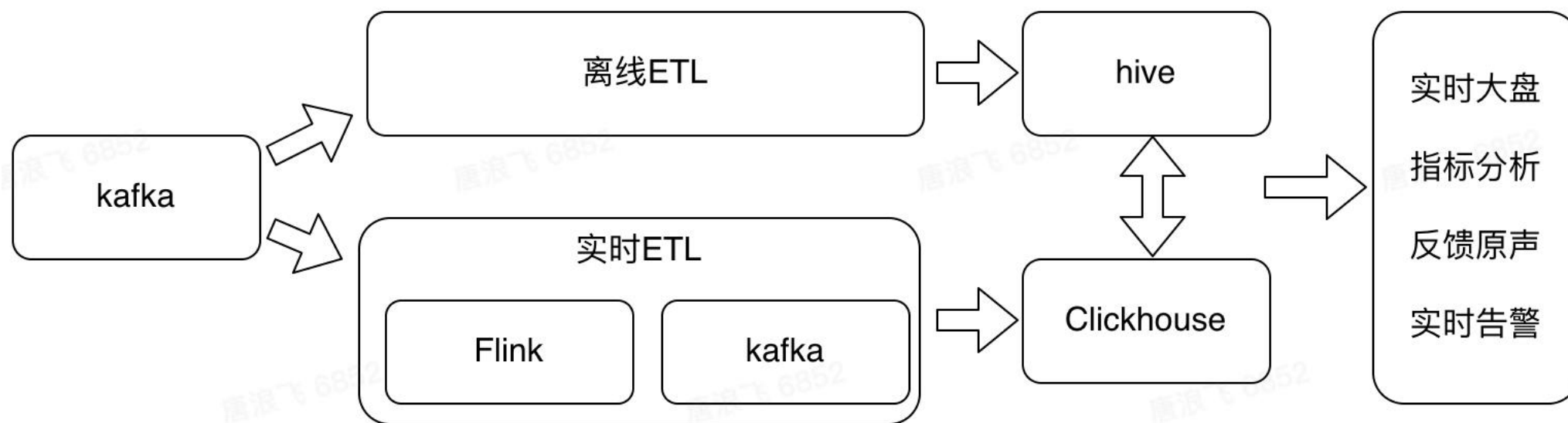


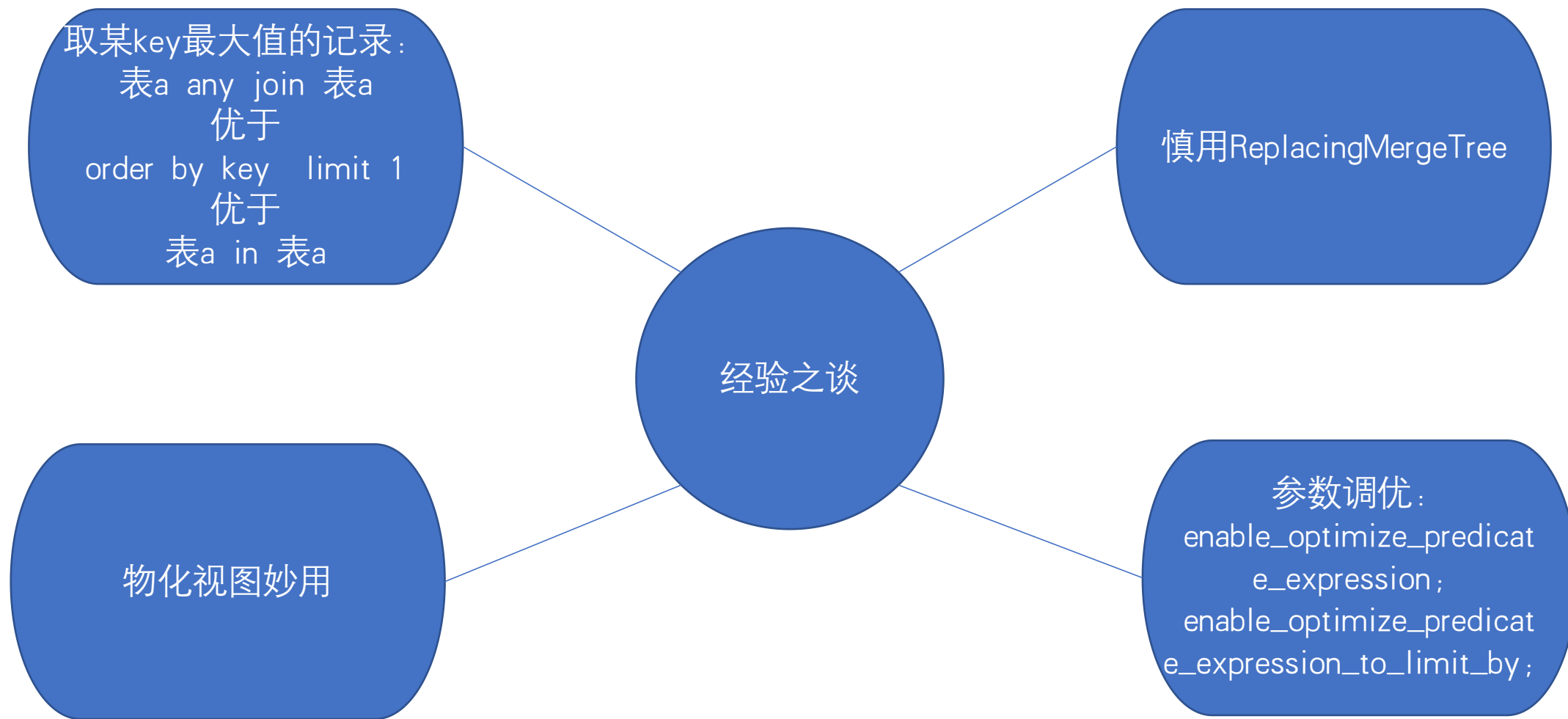
业务要求相同  
key取最新的  
记录，CH去重  
压力大

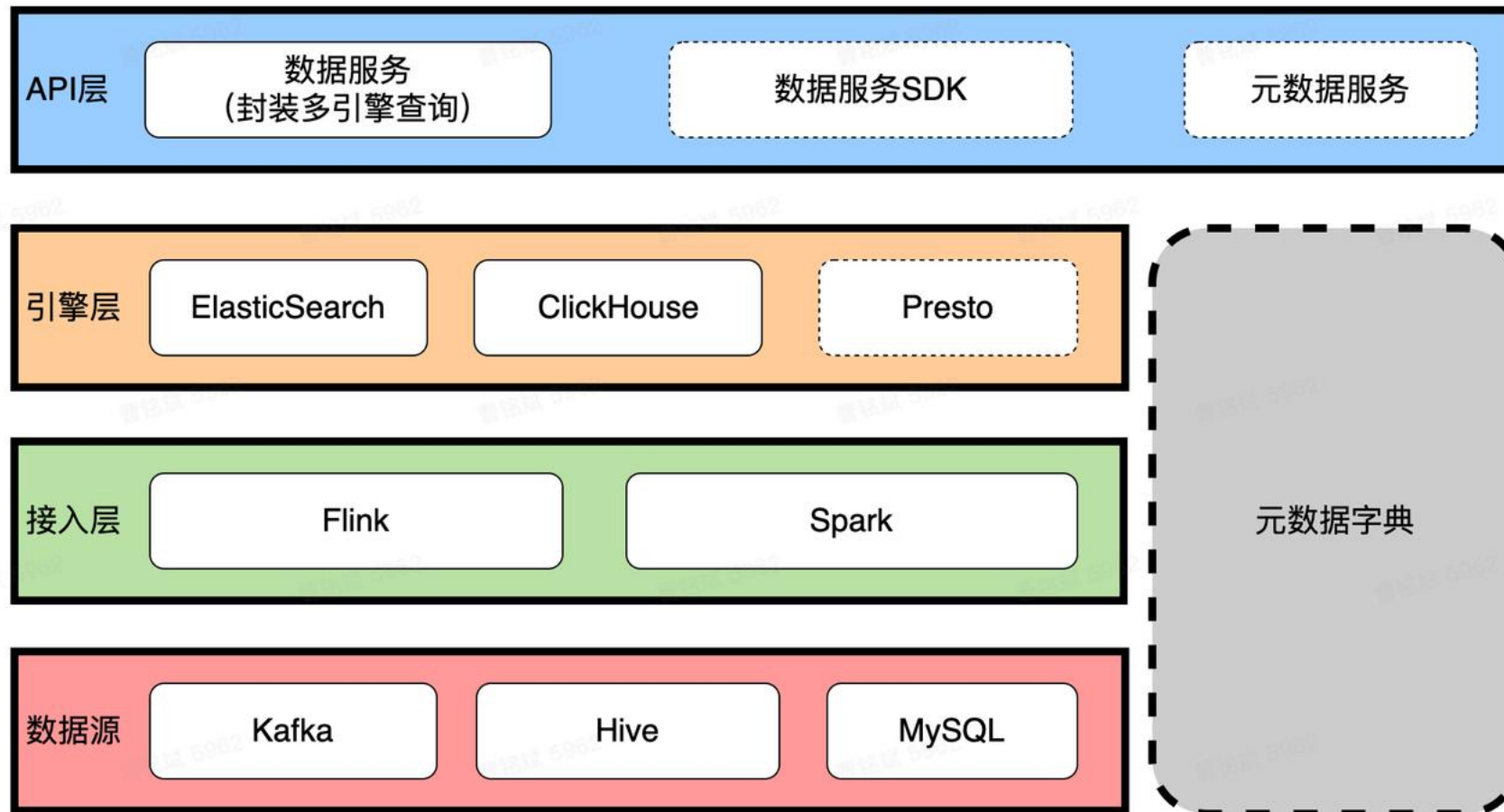




效果图:









# 未来规划--架构下一次演进(Presto)

- 业务目标：治理买卖双方、达人、平台矛盾点，净化电商平台交易环境。
- 业务背景：当前问题发现维度比较单一，需要一个多维度视角快速挖掘低质买卖家和达人。

## 平台挑战：

1. 电商数据源多样化(买卖家hive数仓、反馈评论实时数据、客服域离线反馈数据)
2. 电商数据庞大，并且多表join
3. 需要快速查询，支持下钻、去重





关注msup公众号  
获取更多AI落地实践

麦思博(msup)有限公司是一家面向技术型企业的培训咨询机构，携手2000余位中外客座导师，服务于技术团队的能力提升、软件工程效能和产品创新迭代，超过3000余家企业续约学习，是科技领域占有率第1的客座导师品牌，msup以整合全球领先经验实践为己任，为中国产业快速发展提供智库。