龙玺 / AIRBNB爱彼迎 / 中国基础架构组（CHINA INFRASTRUCTURE）

AirTrain: Airbnb的通用数据产品平台

airbnb

关于Airbnb爱彼迎

成立于2008年8月，爱彼迎总部位于加利福尼亚州旧金山市。爱彼迎是一个值得信赖的社区型市场，在这里人们可以通过网站、手机或平板电脑发布、发掘和预订世界各地的独特房源。

房客总数
超过
150,000,000个

城市
超过65,000个

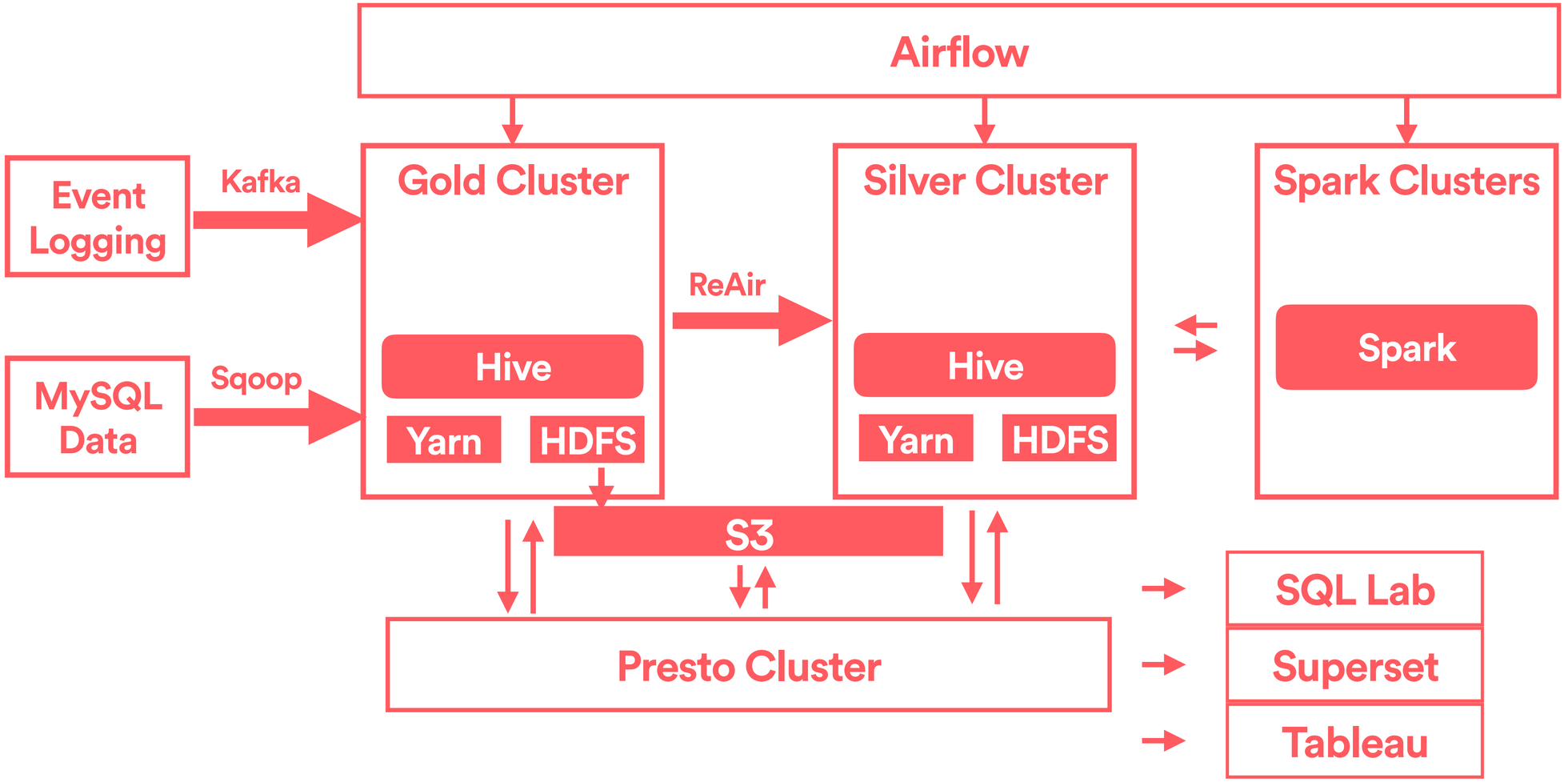
城堡
超过1,400个

国家
超过191个

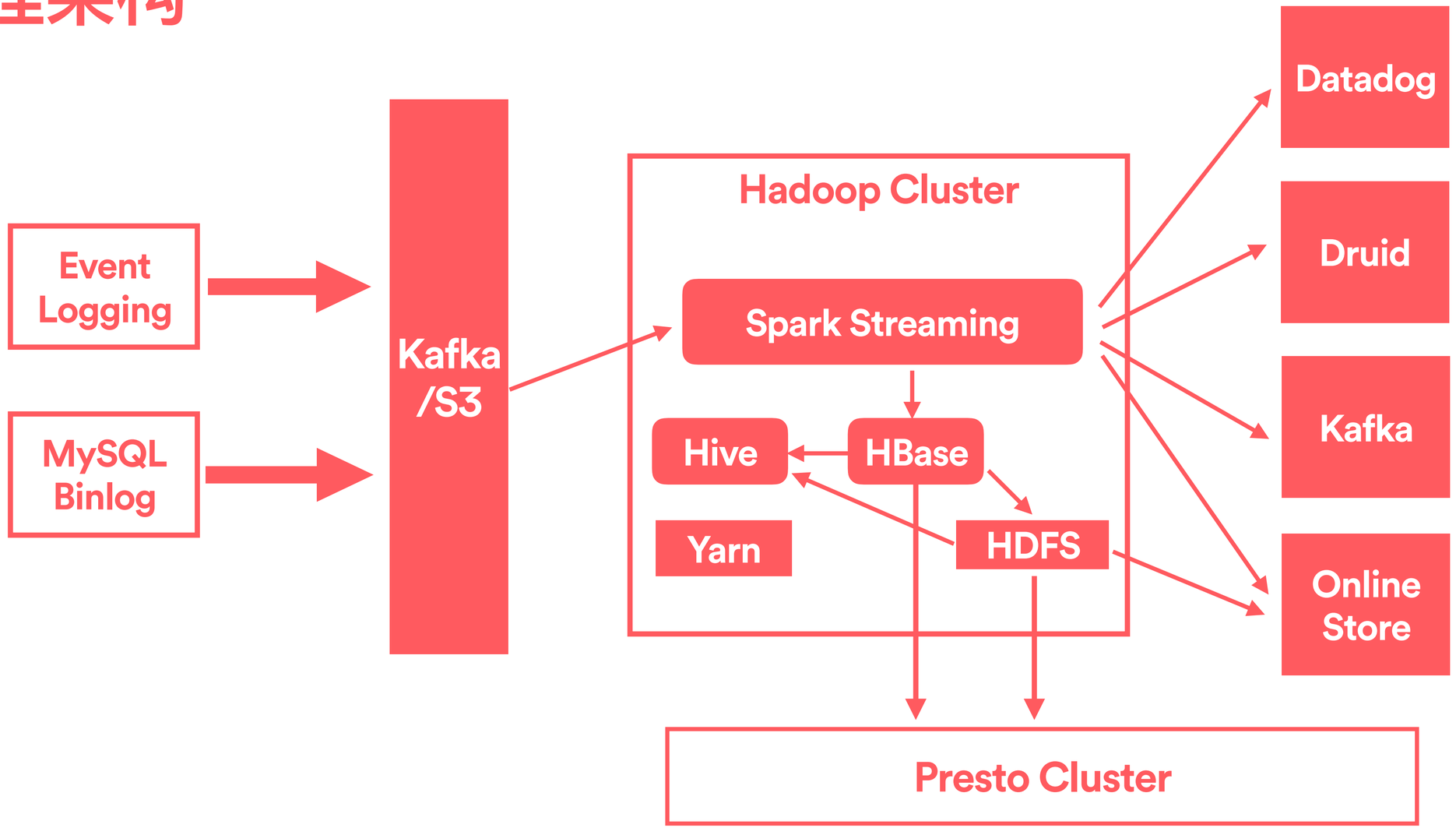
全球房源
3,000,000+

Airbnb 的数据基础架构

批处理架构


Airflow
Event Logging
Kafka
MySQL Data
Sqoop
Gold Cluster
Hive
Yarn
HDFS
ReAir
Silver Cluster
Hive
Yarn
HDFS
Spark Clusters
Spark
S3
Presto Cluster
SQL Lab
Superset
Tableau

流处理架构


Event Logging
MySQL Binlog
Kafka /S3
Hadoop Cluster
Spark Streaming
Hive
HBase
Yarn
HDFS
Presto Cluster
Datadog
Druid
Kafka
Online Store

大数据 @ Airbnb

Airbnb大数据应用模式的演进

从MySQL数据库中导出商务数据，主要用于报表和结果验证

使用事件日志进行实时验证，监控和用户行为分析

海量数据离线分析，提供正确的产品决策

在线数据产品，提高用户体验

案例1 房东报表

高度定制的访问流量和房客统计报表

案例2 房源实时统计

在产品页面显示房源实时统计数据

为房客提供最新房源热度，质量等一系列有用信息

机器学习 @ Airbnb

应用领域

- 市场推广
- 产品搜索
- 风险控制
- 用户体验
- 客户服务

大规模数据产品开发中的问题

通用数据平台的必要性

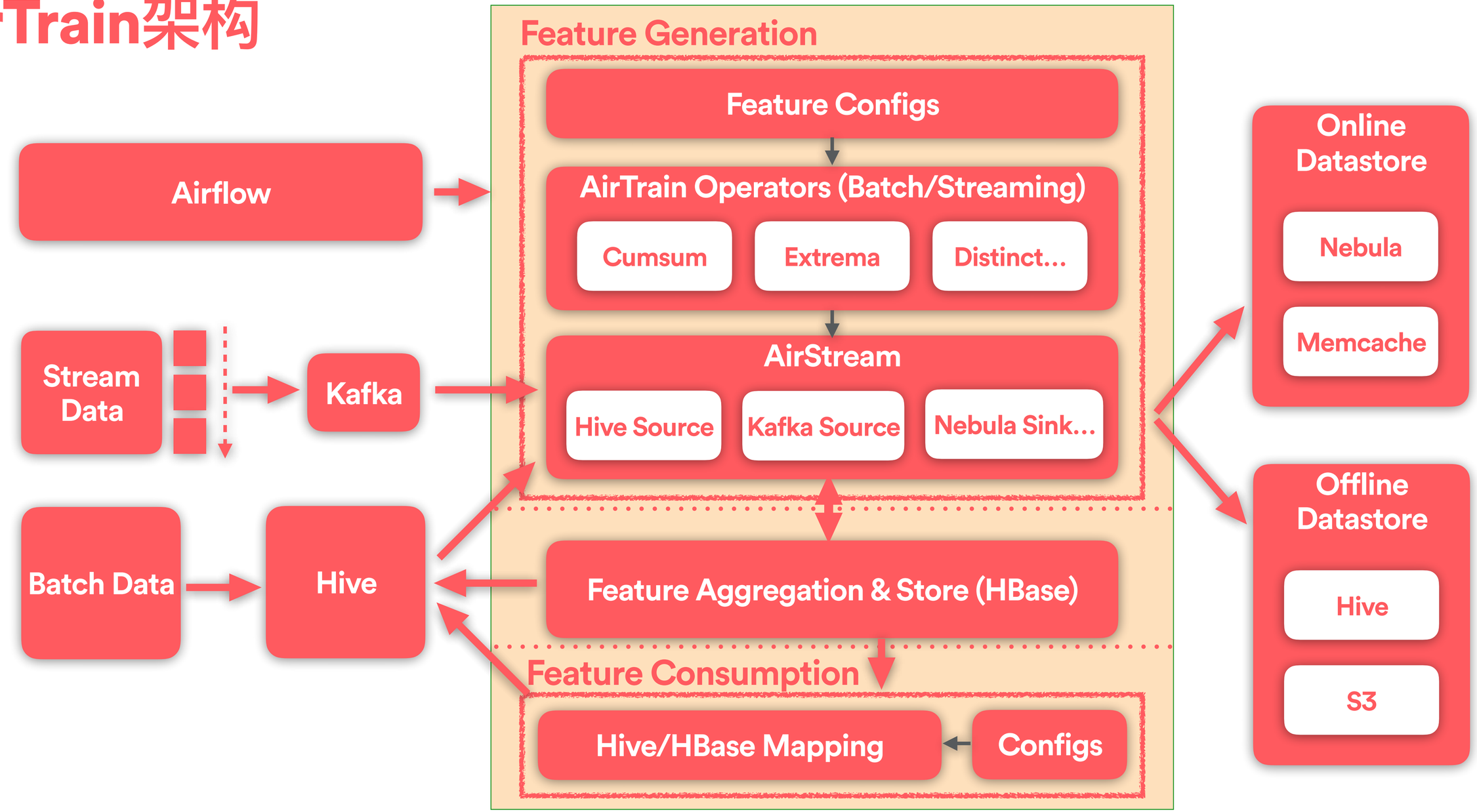
- **资源浪费:** 不同产品部门重复开发维护类似应用，大量数据重复计算，存储
- **信息孤岛:** 不同应用间信息不能共享
- **数据权威性:** 缺乏权威数据定义和来源
- **数据一致性:** 大量数据在准确性和实时性方面存在差异，包括离线 / 在线，训练 / 评分
- **产品开发难度和质量:** 产品部门缺乏基础架构开发经验

AirTrain简介

AirTrain

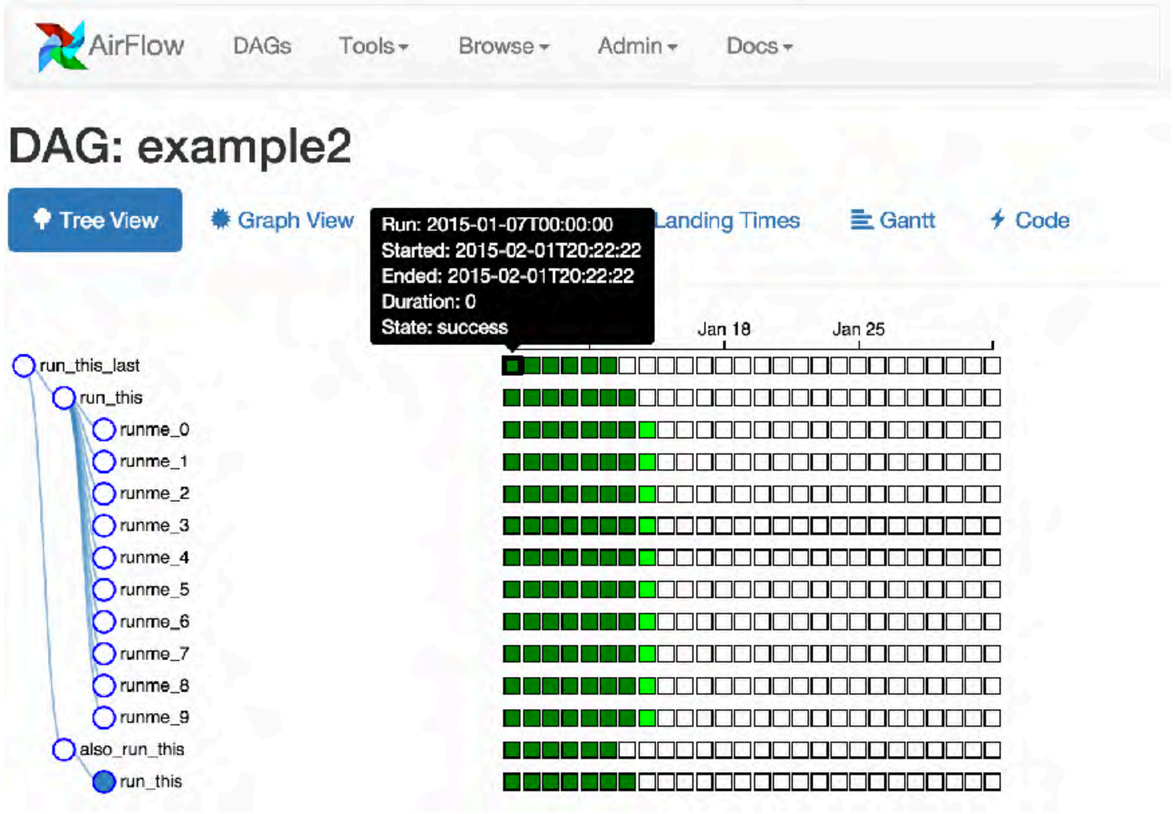
- 设计目标: 通用数据平台
- 提供数据生产（derivation），聚集（aggregation）和存储（storage）解决方案
- 标准化数据计算逻辑和管理，降低数据产品开发入门门槛
- 商务逻辑通过配置文件表达，集中分级管理
- 权威数据源，支持离线和在线应用
- 配置文件驱动的机器学习训练集生成

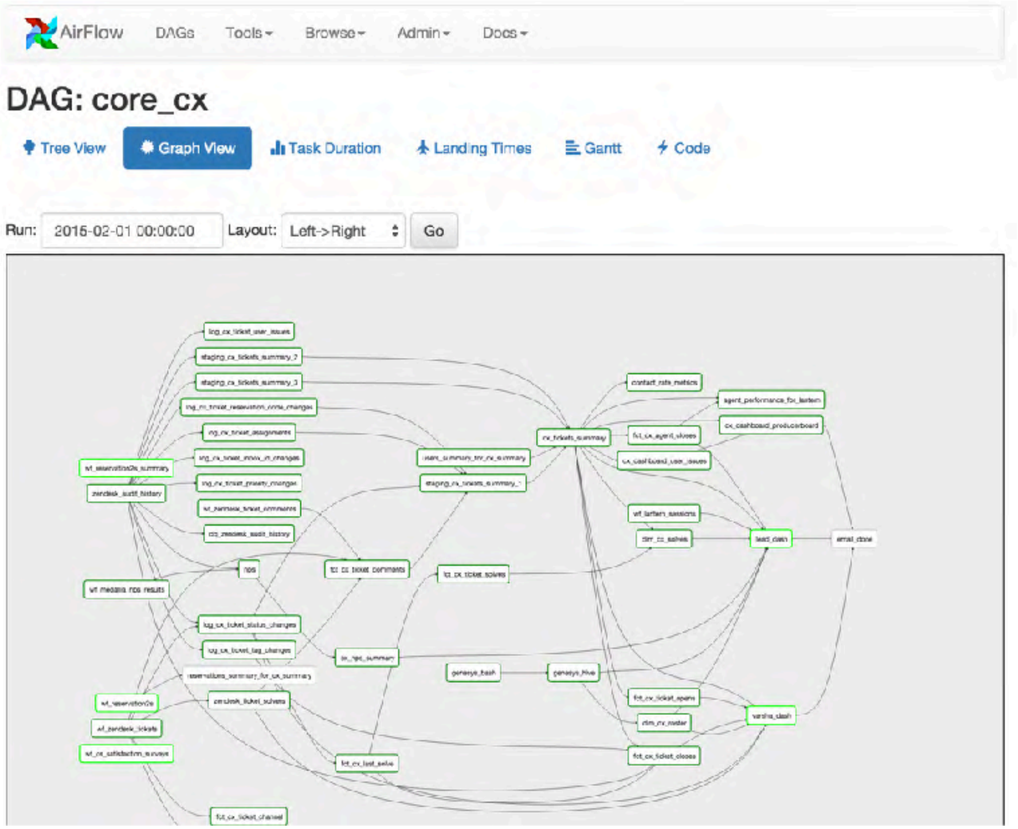
AirTrain架构


Feature Generation
Feature Configs
AirTrain Operators (Batch/Streaming)
Cumsum
Extrema
Distinct...
AirStream
Hive Source
Kafka Source
Nebula Sink...
Airflow
Stream Data
Kafka
Batch Data
Hive
Feature Aggregation & Store (HBase)
Feature Consumption
Hive/HBase Mapping
Configs
Online Datastore
Nebula
Memcache
Offline Datastore
Hive
S3

工作调度：基于Airflow的调度引擎

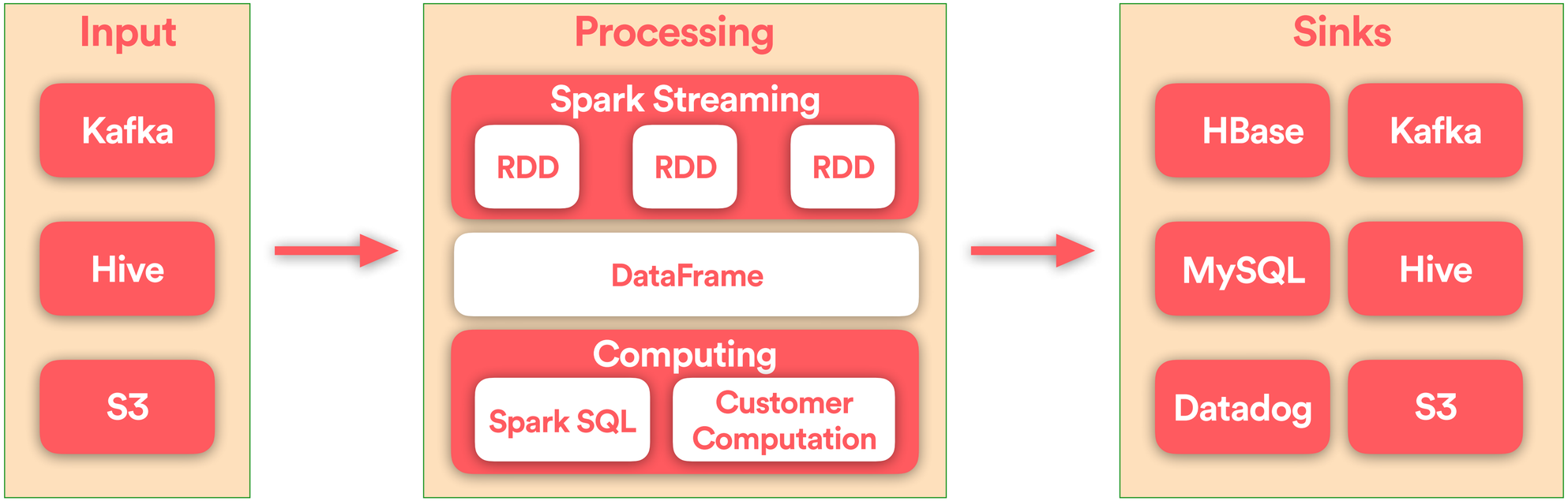
- Airbnb自主研发并open source的工作流管理引擎 https://airflow.incubator.apache.org/





数据生产：基于AirStream的计算引擎

- Airbnb自主研发的用户Spark 应用开发，部署，监控平台



数据生产：模块化的计算逻辑

- 使用数据计算模块提供统一计算逻辑和接口
  - 通过简单配置实现数据计算，无需编程
  - 保障批处理和流处理数据计算的一致性
  - 保障训练和评分时数据处理的一致性
- 模块可通过用户配置文件进行参数定制
- 提供常用模块库以满足大部分开发团队需要
  - Spark SQL
  - 积累计数器
  - 极限值
- 模块可按plugin模式扩展

```
streaming_update = {
    source = {
        type: kafka
        config: {
            topic: mytopic,
            broker: "kafka-main.synapse:3229"
        }
    }
    type = update_first
    pkey_column = [user_id, ip]
    message_timestamp_field = created_at
    column_projection = """
        min(created_at) AS ts_first_obs_v1
    """
    filter = """
        ip is not NULL AND user_id is not NULL AND user_id != 0
    """
}
batch_update = {
    type = update_first
    ...
}
```

数据聚集与存储：基于HBase的特征库

单一数据源（Single source of truth）+ 合并批处理与流处理数据

所有特征自带时间戳和版本控制

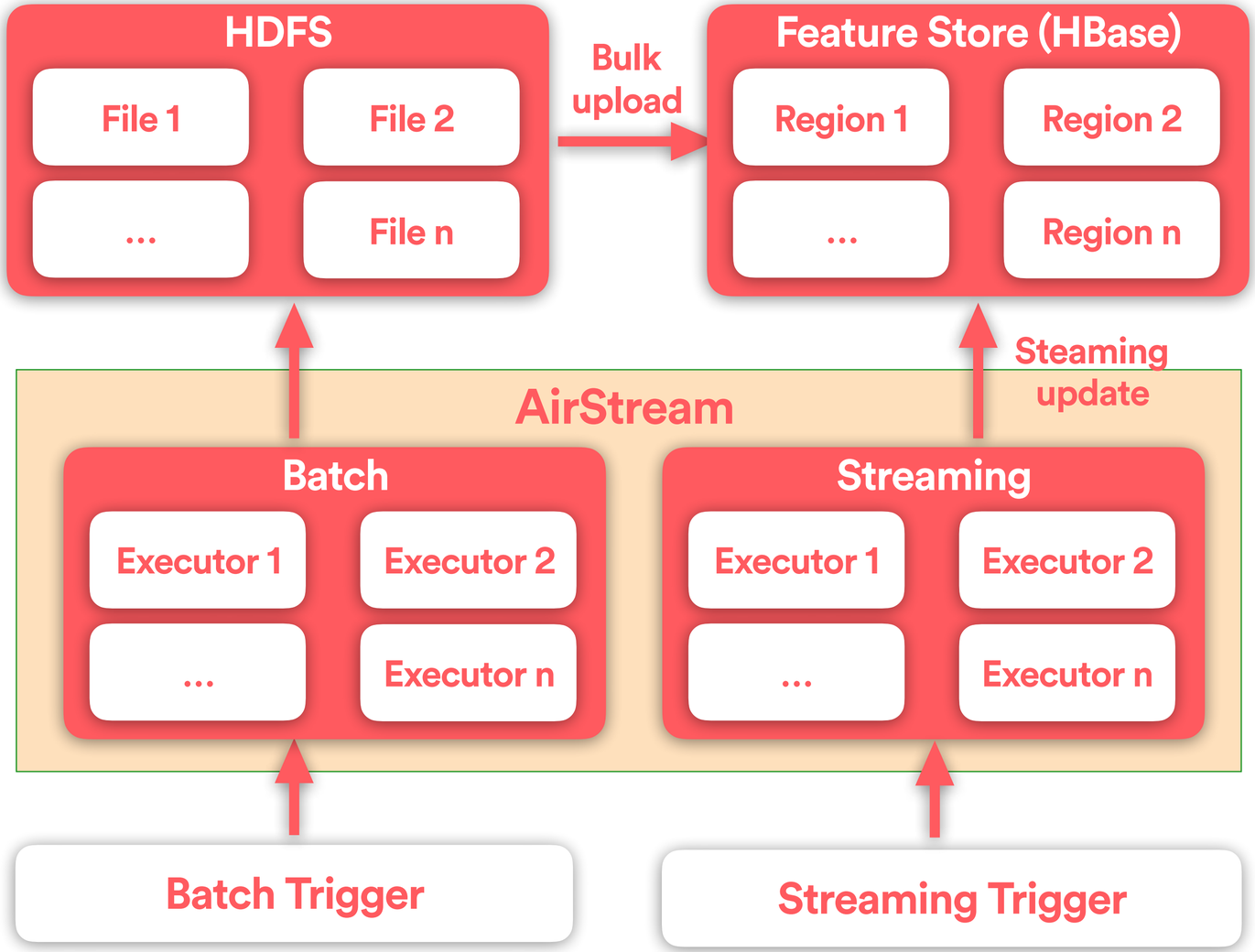
保留历史 + 支持PIT（Point In Time）查找

支持批量导入（Bulk upload） + 流导入（Streaming update）

支持大吞吐数据导出+ 扫描 + 在线随机读取

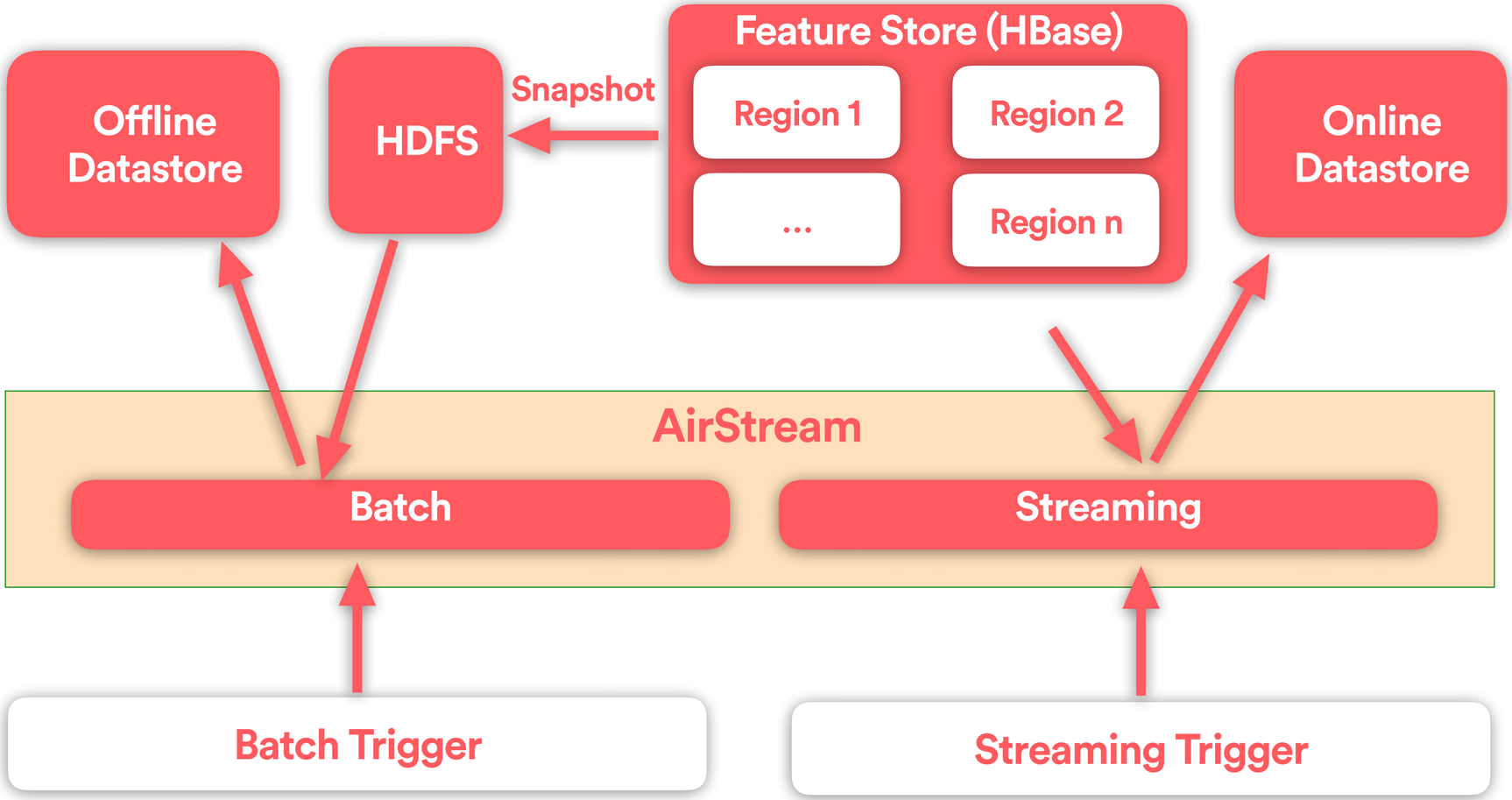
数据导入

- 数据导入前按 Key分区
- 使用 bulk upload 对大批量数据进行导入

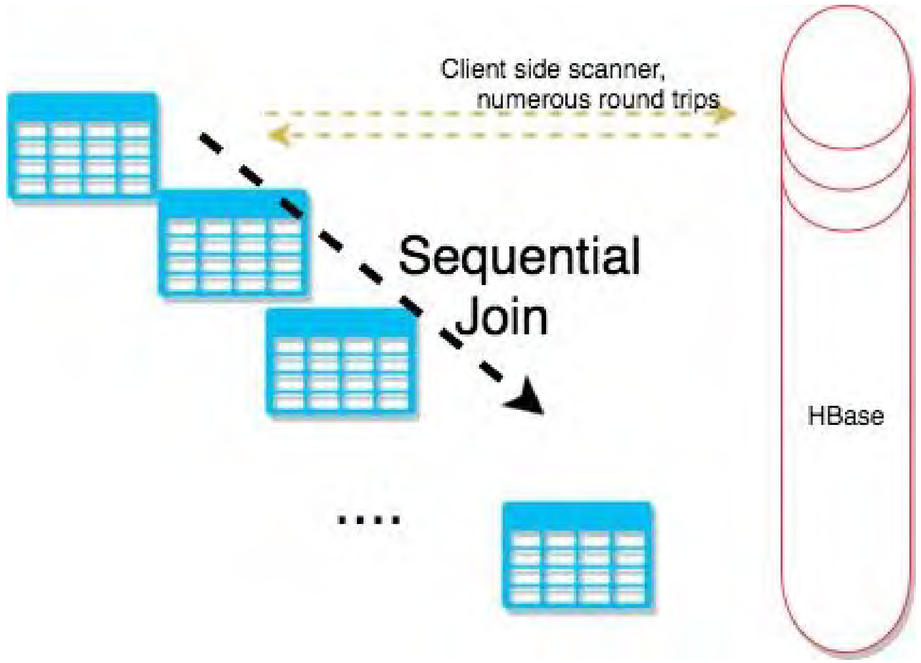


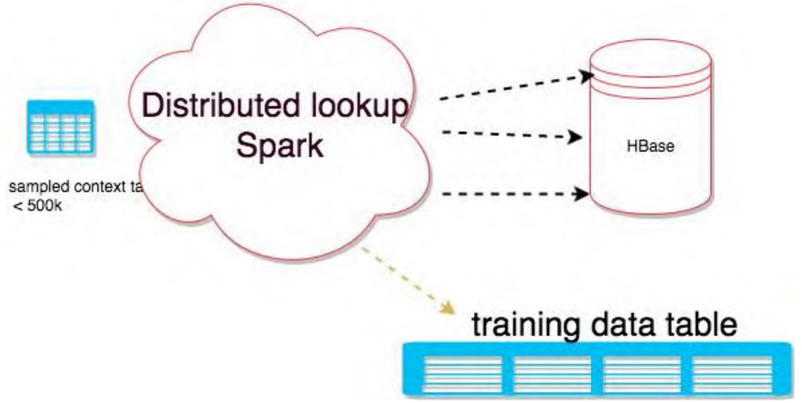
数据导出

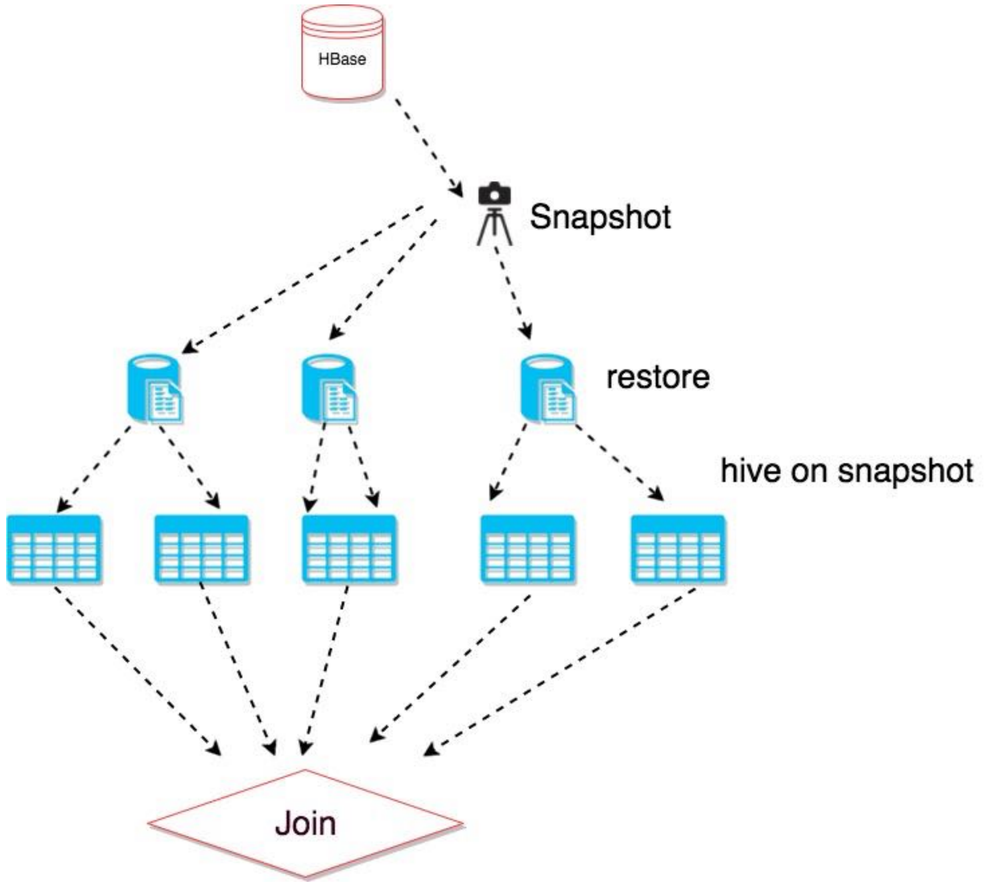
- Snapshot大吞吐量导出offline数据
- 流处理小批量实时计算并导出online数据



机器学习训练集的生成


Client side scanner,
numerous round trips
Sequential
Join
HBase


Distributed lookup
Spark
HBase
training data table

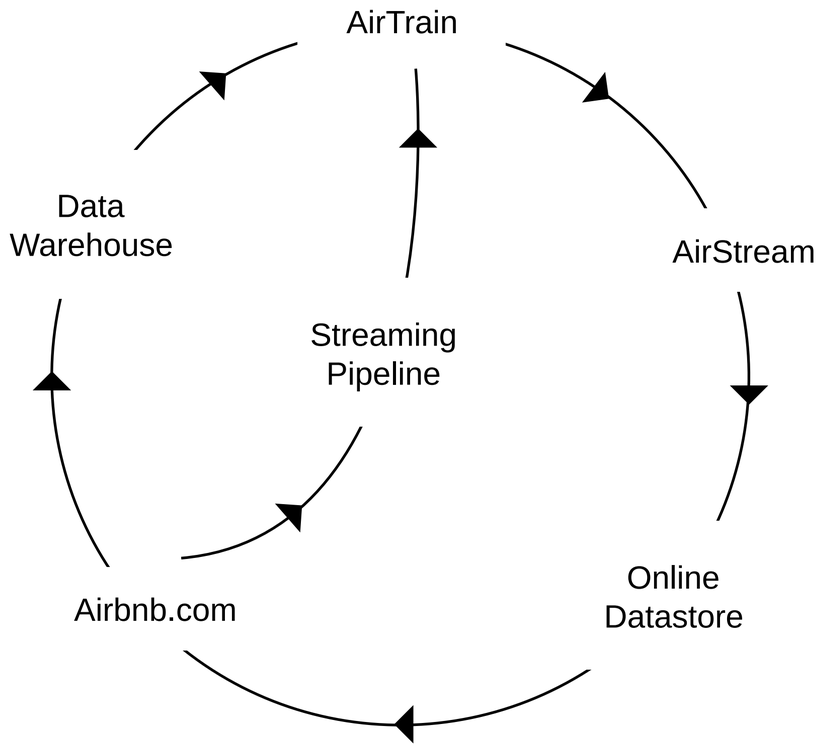

HBase
Snapshot
restore
hive on snapshot
Join

用户配置集中分级管理

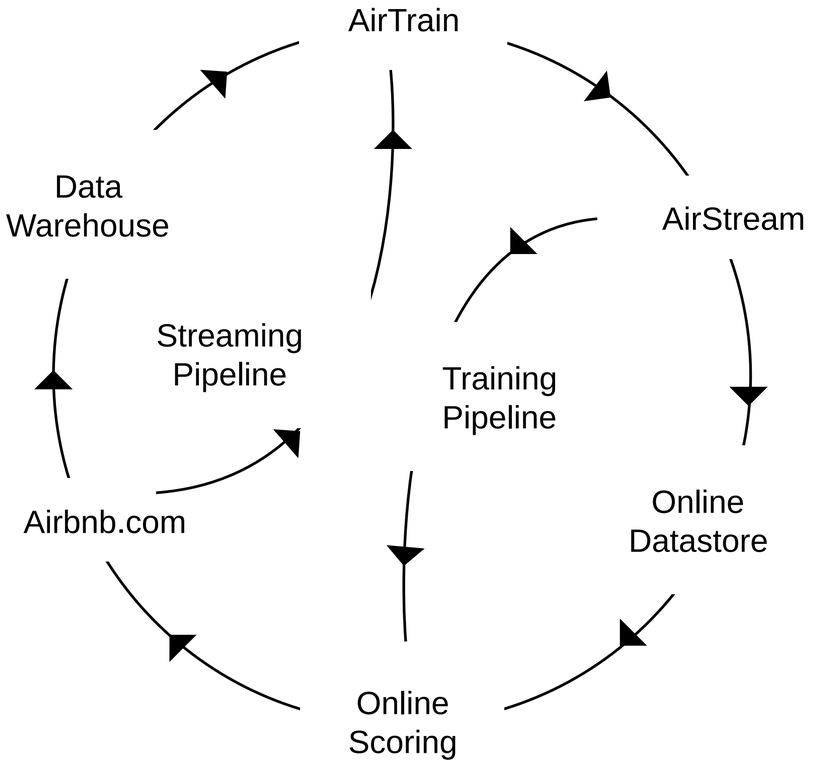
- **数据按重要性分级**
  - core: 核心商业数据，专门团队维护
  - core-ml: 核心机器学习数据，专门团队维护
  - ad-hoc: 用户自定义数据，对全公司开放
- **配置文件按目录分类**
  - 不同级别数据分类
  - 不同特征名分类
- **特征名通过配置文件路径自动生成以保障唯一性**

AirTrain与Airbnb数据产品

General Data Product



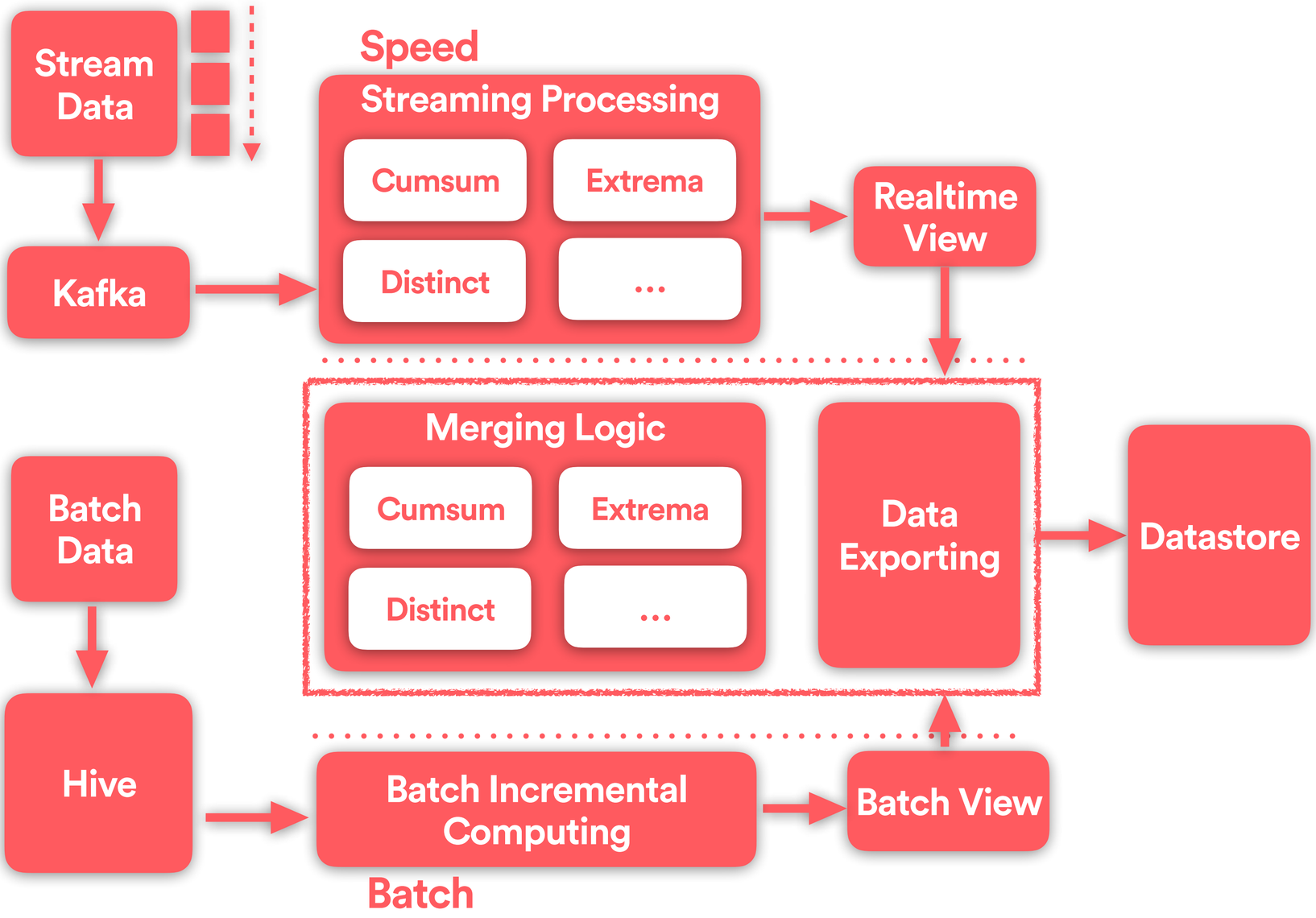
ML Data Product



批处理与流处理数据的归并

数据归并流程

- **lambda 架构:** 批处理 source of truth + 流处理实事更新
- **数据计算模块:** 数据计算模块包含批处理和流处理公共逻辑
- **共用配置文件:** 同一配置文件中使用数据计算模块来保障逻辑的一致性
- **共享计算平台:** 共用底层计算平台（AirStream）
- **数据聚集:** 数据按Key聚集，使用时间戳进行管理
- **数据归并:** 数据导出时运行归并逻辑



案例：Streaming Cumulative Sum

Key	Streaming sum	Streaming sum snapshot	Batch sum	Time	Final sum
0147ae14listing.listing_id=3808475	100	100	1000	t	1000
0147ae14listing.listing_id=3808475	101			t+1	1001
0147ae14listing.listing_id=3808475	102			t+2	1002
0147ae14listing.listing_id=3808475	…	…	…	…	…
0147ae14listing.listing_id=3808475	150	150	1040	t+n	1040
0147ae14listing.listing_id=3808475	151			t+n+1	1041
0147ae14listing.listing_id=3808475	152			t+n+2	1042
0147ae14listing.listing_id=3808475	153			t+n+3	1043
0147ae14listing.listing_id=3808475	154			t+n+4	1044

Final sum = Batch sum + Streaming sum - Streaming sum snapshot

总结

- 提供完整数据生产（derivation），聚集（aggregation）和存储（storage）解决方案的通用数据平台
- 标准化数据计算逻辑和配置管理
- 支持离线和在线应用
- 支持普通数据产品以及机器学习应用

谢谢!

Join Airbnb

Learn More

https://www.airbnb.com/careers/locations/beijing-china