

# Flink Application in Alibaba and Ali-cloud

公司：阿里巴巴

职位：高级开发工程师

演讲者：张冬/元奇



01

什么是实时计算

02

实时计算应用场景

03

Flink在阿里巴巴

04

Flink在云上

# 01

## 什么是实时计算

01

## 什么是实时计算-定义

**A real-time system has been described as one which "controls an environment by receiving data, processing them, and returning the results sufficiently quickly to affect the environment at that time"**

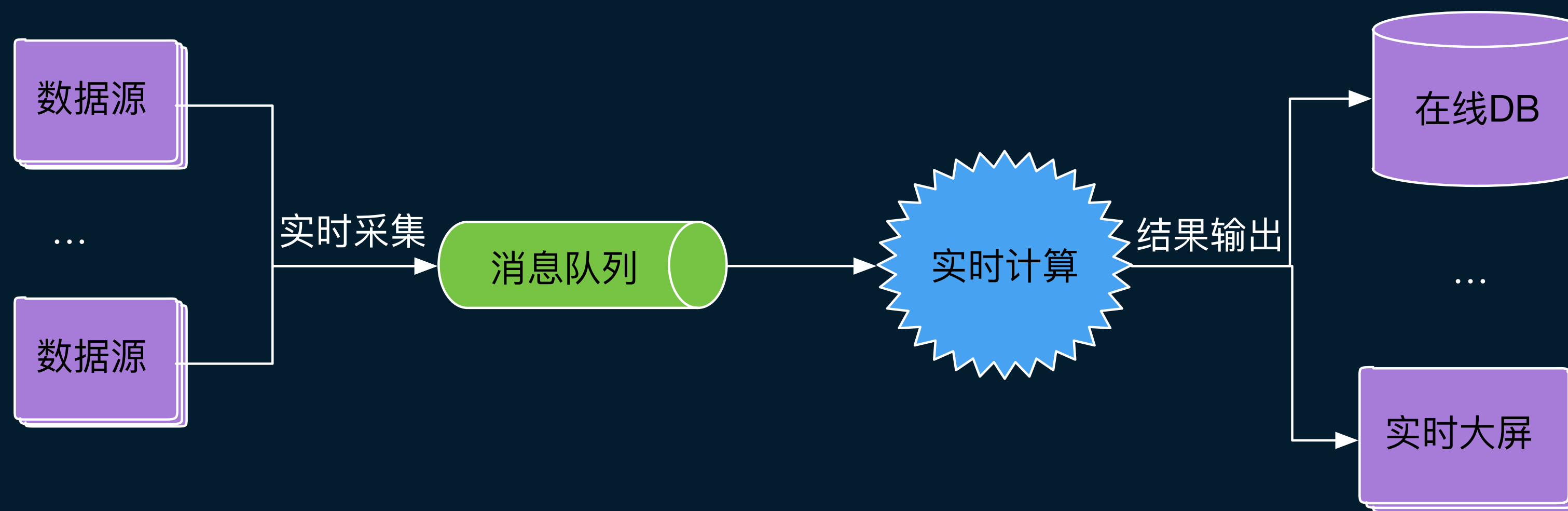
**--wikipedia**

# 什么是实时计算-特征

对比指标	批量计算	实时计算
数据集特征	预先准备好 通常计算过程中数据集不会变化	数据边生成边处理
数据范围	全部数据集或者大部分	最近时间窗口内
数据大小	大批量数据	单条或者一小批
时延	分钟、小时甚至天	秒级、亚秒级
使用场景	H+1、T+1	实时或近实时场景

01

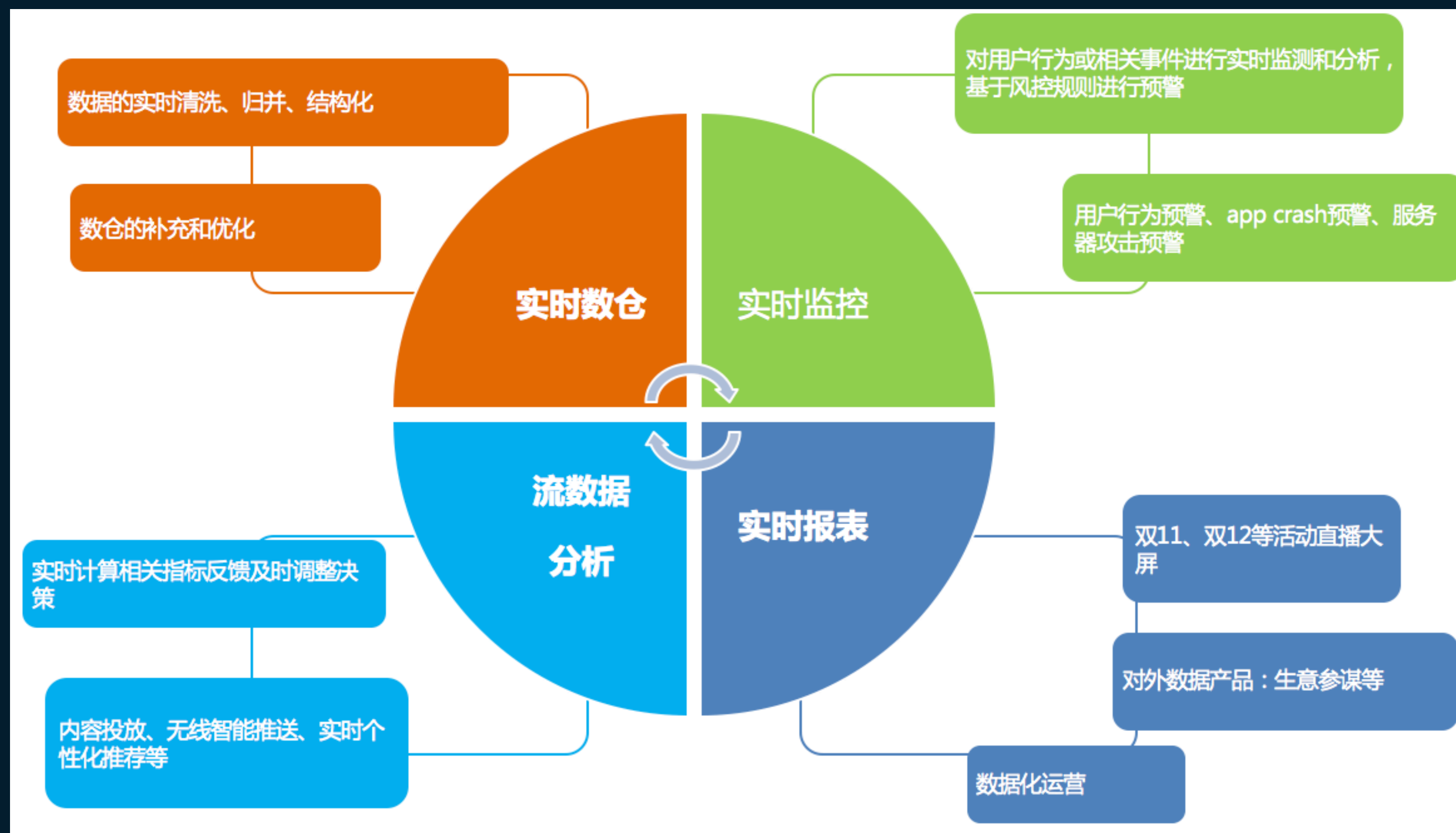
# 什么是实时计算-数据处理流程



# 02

## 实时计算应用场景

## 实时计算应用场景





## 实时计算应用场景

### 实时ETL：

集成流计算现有的诸多数据通道和SQL灵活的加工能力，对流式数据进行实时清洗、归并、结构化处理。同时，为离线数仓进行有效的补充和优化，为数据实时传输的提供可计算通道。

### 实时报表：

实时化采集、加工流式数据存储。实时监控和展现业务、客户各类指标，让数据化运营实时化。

### 监控预警：

对系统和用户行为进行实时检测和分析。实时监测和发现危险行为。

### 在线系统：

实时计算各类数据指标，并利用实时结果及时调整在线系统相关策略。在各类内容投放、无线智能推送领域有大量的应用。

# 03

## Flink在阿里巴巴



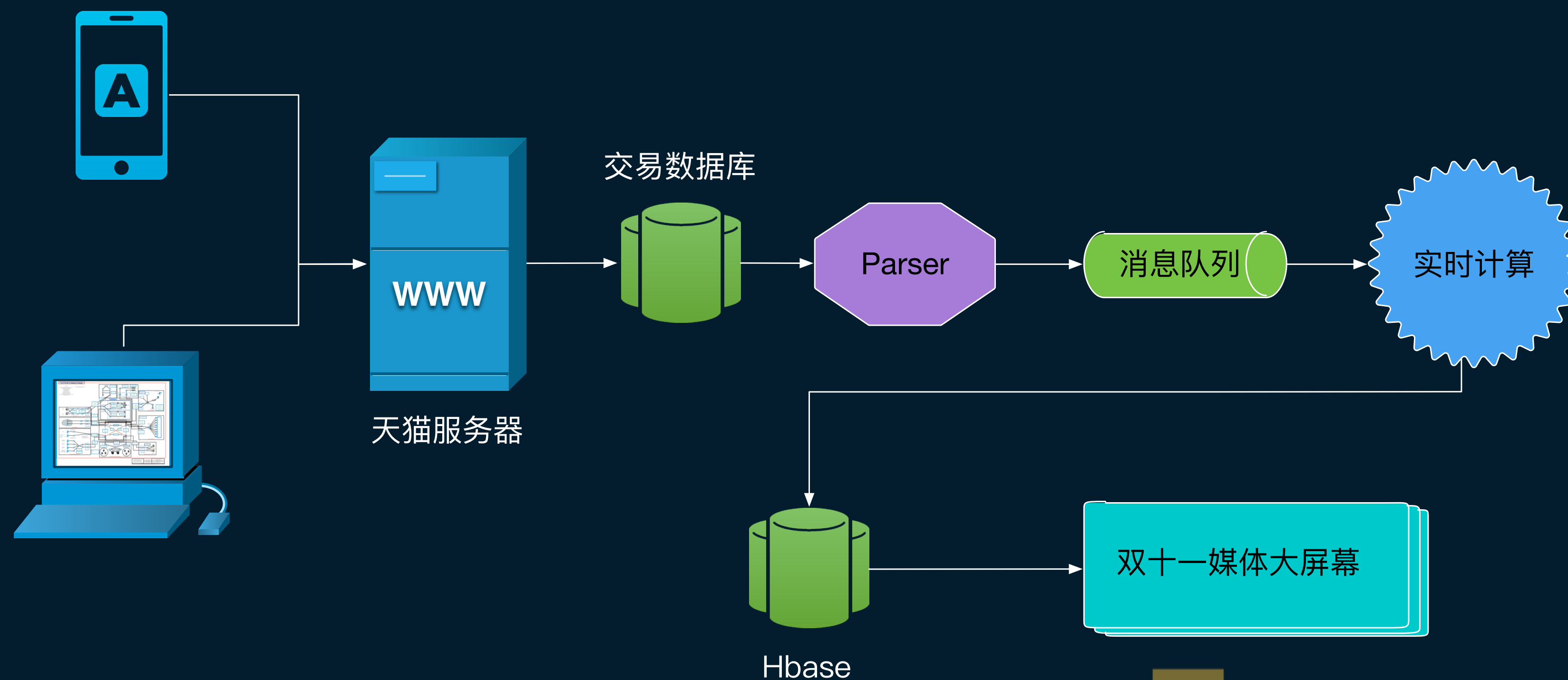
## Flink在阿里巴巴





# Flink在阿里巴巴

媒体大屏数据链路，全链路延迟<3s



## Flink在阿里巴巴

- 实时日志处理峰值每秒17亿条（约合每秒1.7TB），相当于1秒钟内读完120万本2018年新版新华字典的数据量。
- Flink承载Alibaba 绝大部分部门的实时计算，作业数万级别，机器数万台规模。

## Flink在阿里巴巴

- 菜鸟智能物流。
- 盒马交易实时数仓。
- 蚂蚁内容推荐系统。
- 实时广告投放。
- 支付宝实时风控。

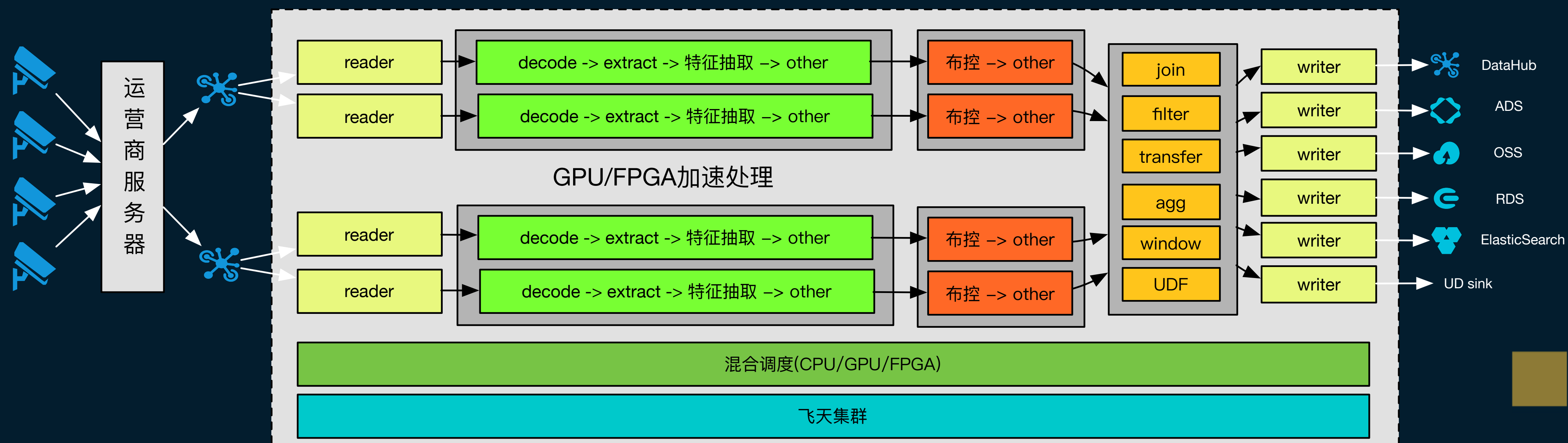


# 04

## Flink在云上

# Flink在云上

## 城市大脑



视频接入

实时计算(Flink)集群

数据输出

## Flink在云上

### 城市大脑

- 视频接入：大华、海康等标准对接。
- 非结构化处理：图像识别，转结构化。
- 硬件加速：FPGA/GPU。
- 安防布控，罪犯识别，交通智能调度



**THANKS**

