

# Apache Kafka的过去, 现在, 和未来

饶军

Confluent联合创始人



基于实践经验总结和提炼的品牌专栏  
尽在【极客时间】



重拾极客时间，提升技术认知



全球技术领导力峰会

通往**年薪百万**的CTO的路上，  
如何打造自己的技术**领导力**？

扫描二维码了解详情



# 2010的领英

人才 ⇔ 机会



200M+

Members Worldwide

2 new

Members Per Second

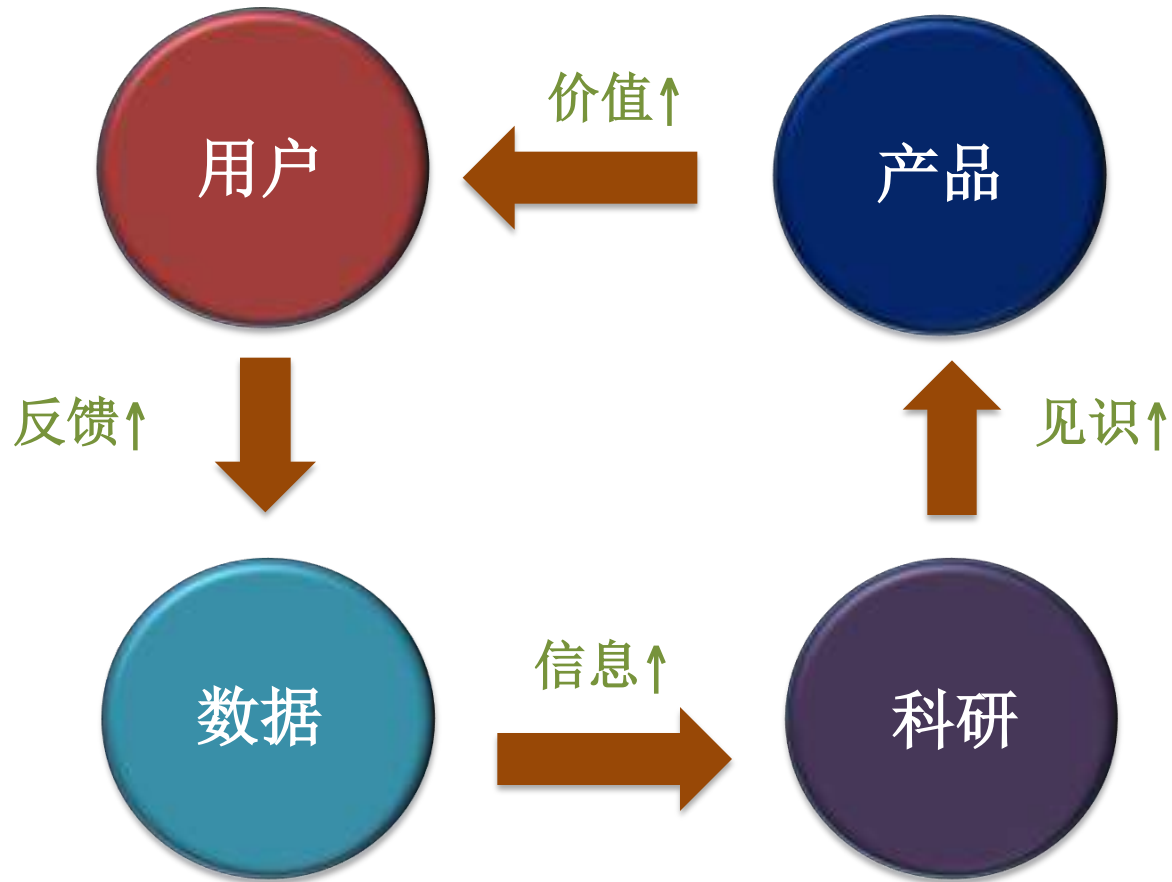
100M+

Monthly Unique Visitors

2M+

Company Pages

# 数据驱动



















# 数据源的多样性

- 交易数据
- 用户行为
- 运营服务监测指标
- 应用程序日志
- IOT
- ...

# People You May Know

## People You May Know <sup>*beta*</sup> See people from different parts of your professional life



- |   |   |
|---|---|
|  <p><b>Peter Carlin</b> <sup>(2nd)</sup><br/>at Microsoft<br/>Greater Seattle Area</p> <p><a href="#">Connect</a>  5 shared connections</p>                               |  <p><b>Dmitri Smirnov</b> <sup>(2nd)</sup><br/>at Microsoft<br/>San Francisco Bay Area</p> <p><a href="#">Connect</a>  7 shared connections</p>                                       |
|  <p><b>Josh Walker</b> <sup>(2nd)</sup><br/>Principal Software Engineer at LinkedIn<br/>San Francisco Bay Area</p> <p><a href="#">Connect</a>  18 shared connections</p>  |  <p><b>Subrahmanyam Sanka (Subbu)</b><br/>DW/BI at LinkedIn<br/>San Francisco Bay Area</p> <p><a href="#">Connect</a>  8 shared connections</p>                                       |
|  <p><b>Walt Roloson</b> <sup>(2nd)</sup><br/>Business Operations Manager, Investor<br/>San Francisco Bay Area</p> <p><a href="#">Connect</a>  2 shared connections</p> |  <p><b>Vinoth Chandar</b> <sup>(2nd)</sup><br/>Engineer, Distributed Data Systems at LinkedIn<br/>San Francisco Bay Area</p> <p><a href="#">Connect</a>  17 shared connections</p> |
|  <p><b>Alison King</b> <sup>(2nd)</sup><br/>HR assistant at LinkedIn<br/>San Francisco Bay Area</p> <p><a href="#">Connect</a>  32 shared connections</p>             |  <p><b>Ricky Ho</b> <sup>(2nd)</sup><br/>Principal Architect<br/>San Francisco Bay Area</p> <p><a href="#">Connect</a>  5 shared connections</p>                                  |



**3X POINTS ON AIRFARE**

**2X POINTS ON GAS & GROCERIES**

**1X POINT ON EVERYTHING ELSE**

EARN 25,000 MEMBERSHIP REWARDS® POINTS

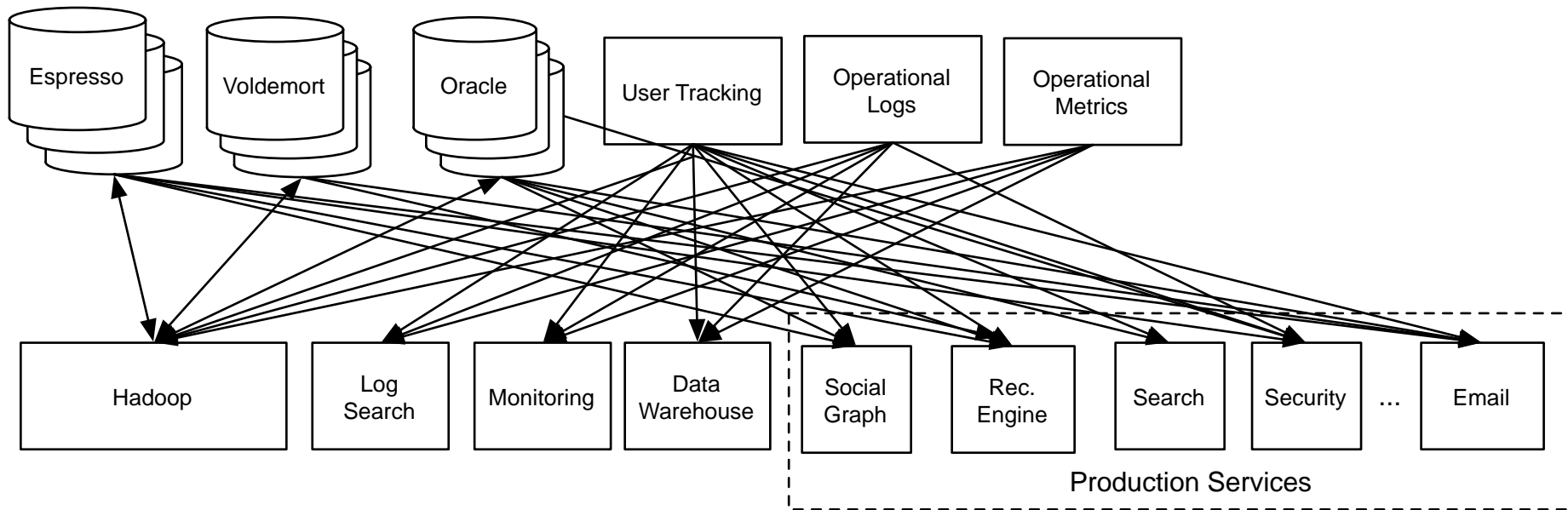
\$0 INTRO ANNUAL FEE FOR THE FIRST YEAR



[APPLY NOW](#)

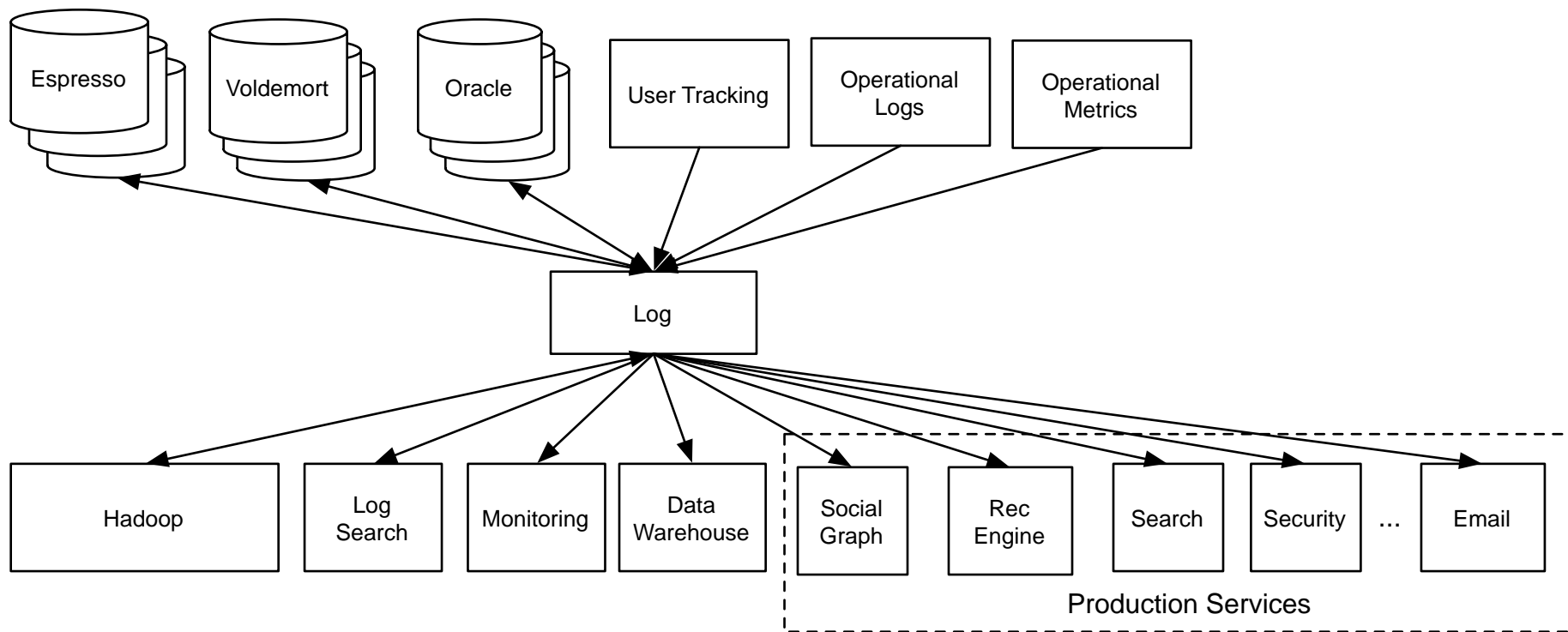
Terms, Conditions, and Restrictions apply.

# 点到点数据集成

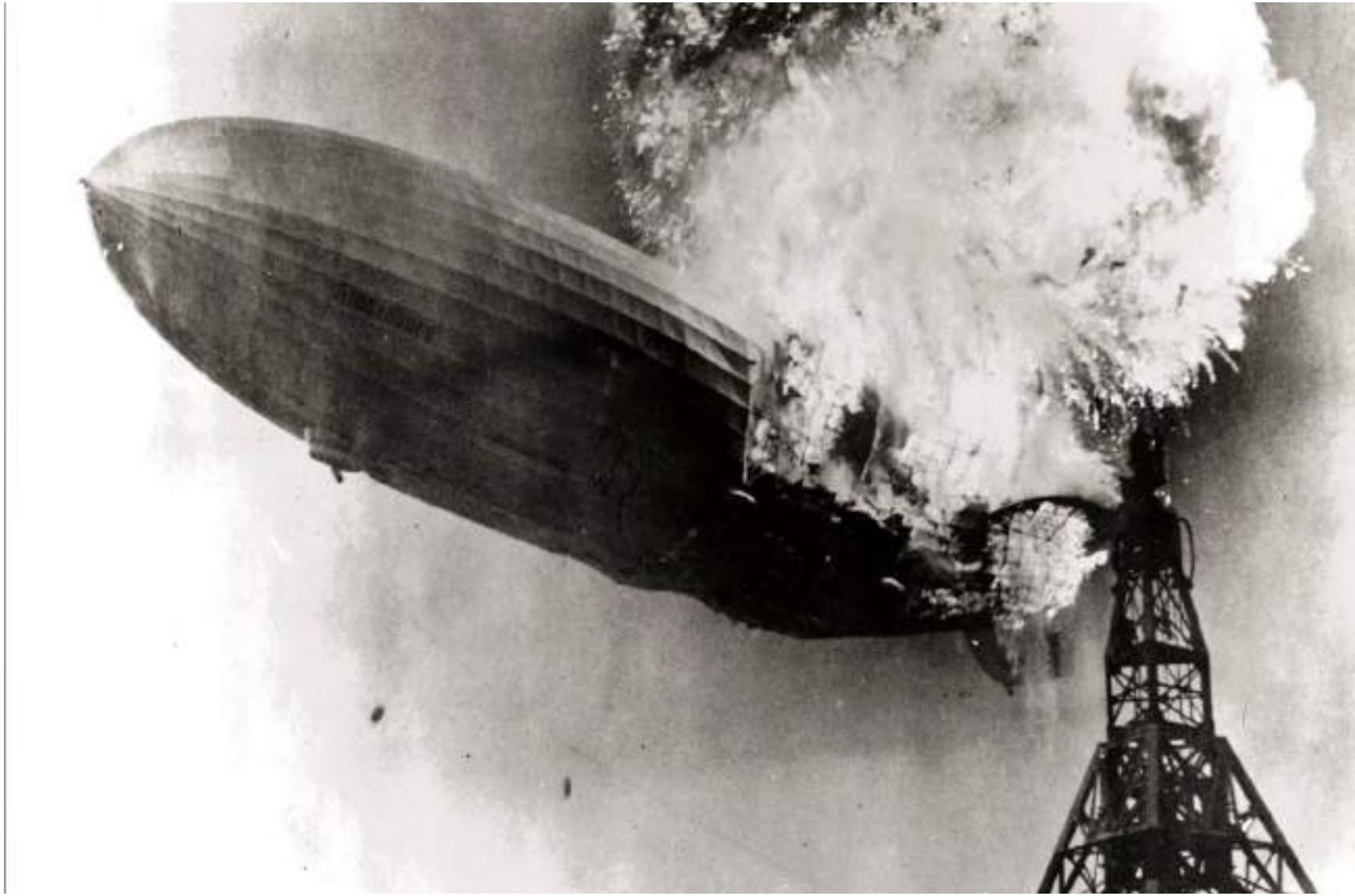




# 理想构架



首次尝试：不想重造轮子

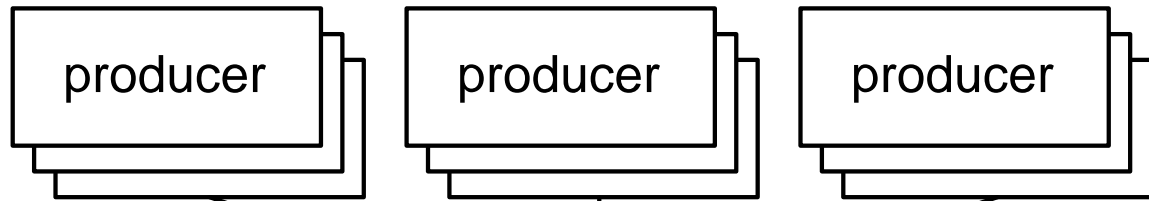


# Kafka第一版

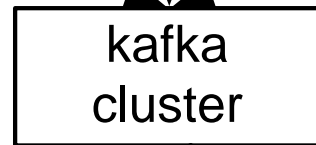
- 高吞吐发布订阅消息系统

# 分布式构架

生产



服务

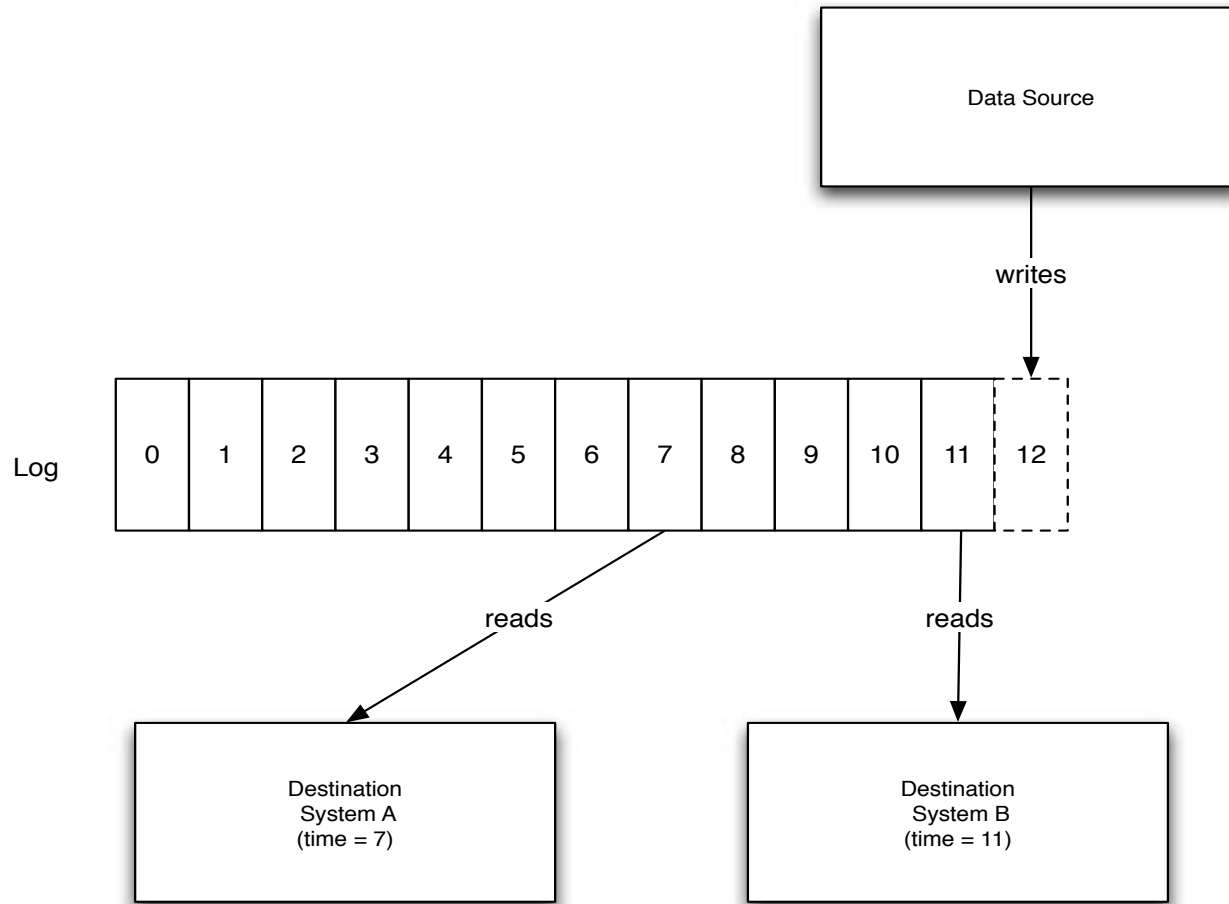


消费



消息源(Topics)可以分区

# 简单实用的日志存储



# 两个优化

- 批处理
  - 生产, 服务, 消费
- 压缩
  - Gzip, snappy, lz4
  - 端到端

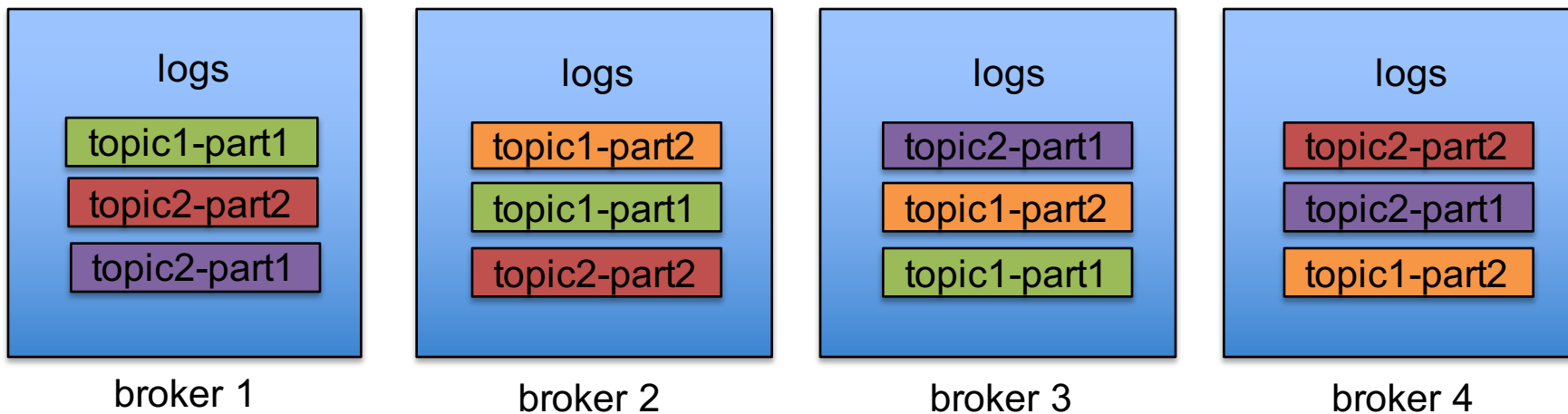
# Kafka在领英2011

- 生产：46万消息 / 秒
- 消费：2百万消息 / 秒
- 几万个生产微服务线程
  
- 数据民主化！



# Kafka第二版

- 高可靠性



# Kafka于2011年加入Apache



10大旅游企业中的6个



10大银行中的7个



10大保险中的8个



10大电信中的9个

# Confluent于2014年成立

- 打造基于Kafka的数据流平台

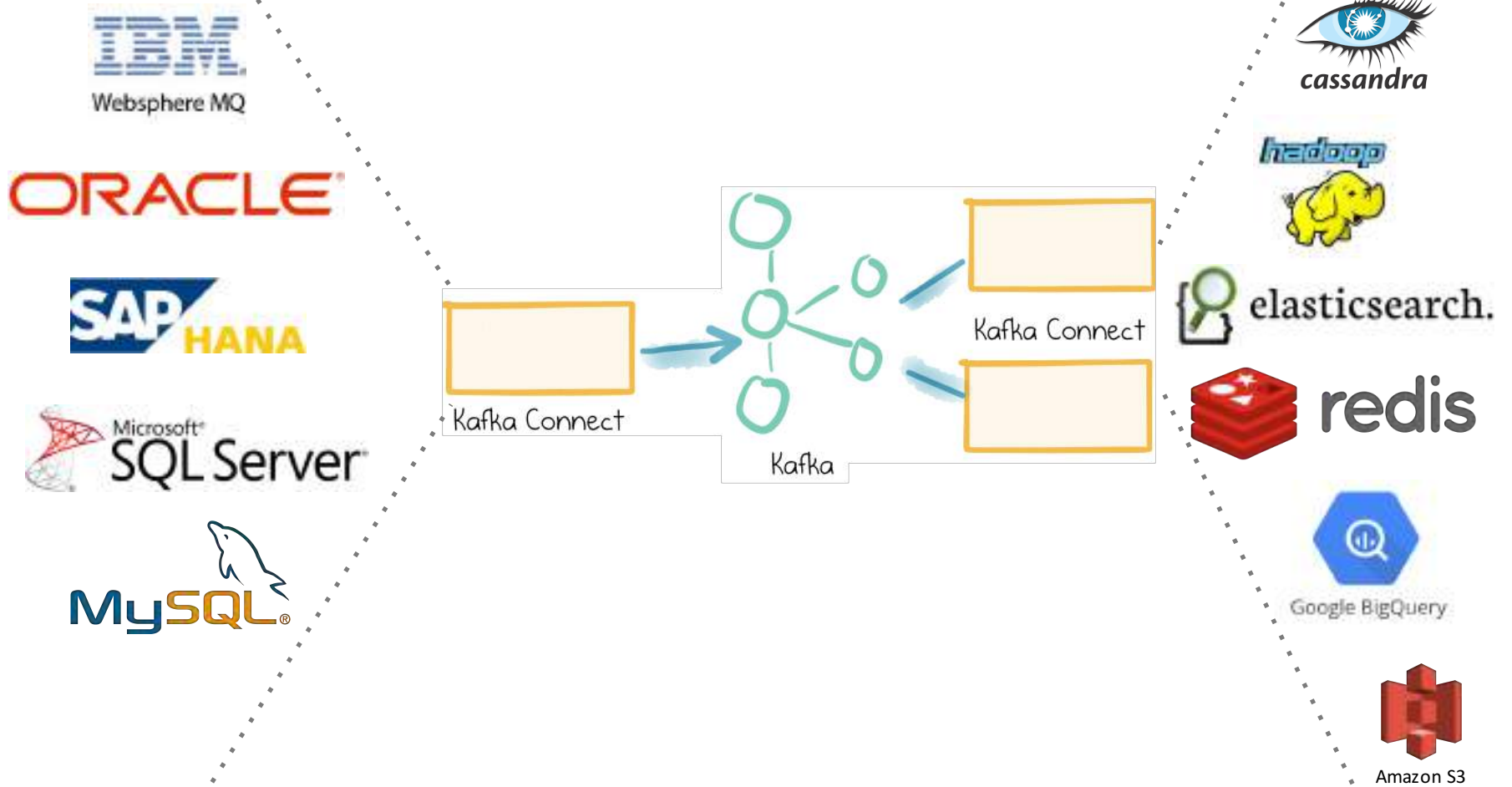
# Kafka的发展

- 企业层功能
- 数据流处理

# 数据安全

- 客户端认证
  - SSL
  - SASL
- 授权

# Kafka Connect



# 数据流处理

- 转换
- 充实
- 聚合

# Kafka Streams

```
KStream<Integer, Integer> input = builder.stream("numbers-topic");

// Stateless computation
KStream<Integer, Integer> doubled = input.mapValues(v -> v * 2);

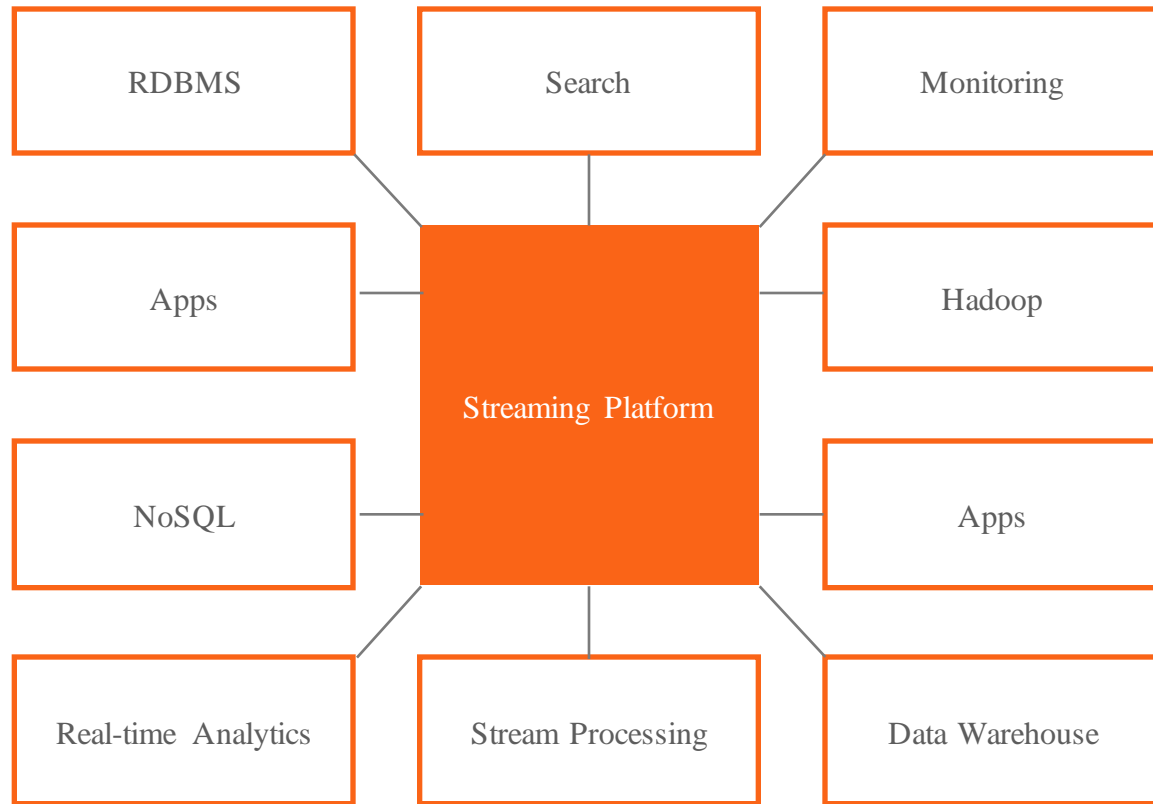
// Stateful computation
KStream<Integer, Integer> sumOfOdds = input
    .filter((k,v) -> v % 2 != 0)
    .selectKey((k, v) -> 1)
    .reduceByKey((v1, v2) -> v1 + v2, "sum-of-odds")
    .toStream();
```



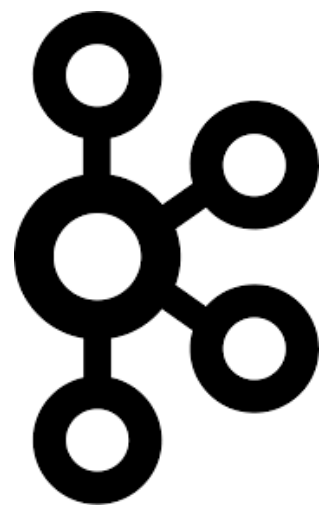
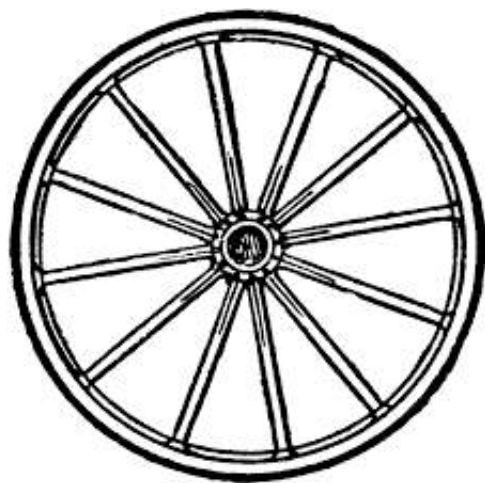
# KSQL (from Confluent)

```
CREATE STREAM vip_actions AS
  SELECT userid, page, action
  FROM clickstream c
  LEFT JOIN users u
    ON c.userid = u.user_id
  WHERE u.level = 'Platinum';
```

# Kafka的未来：实时数据流平台



# 结束语



# GMTC 2018

## 全球大前端技术大会

—— 大前端的下一站 ——



<<扫码了解更多详情>>



关注 ArchSummit 公众号  
获取国内外一线架构设计  
了解上千名知名架构师的实践动向



Apple • Google • Microsoft • Facebook • Amazon 腾讯 • 阿里 • 百度 • 京东 • 小米 • 网易 • 微博

深圳站：2018年7月6-9日    北京站：2018年12月7-10日

主办方 **Geekbang** **InfoQ**  
极客邦科技

# QCon

全球软件开发大会【2018】

# 上海站

2018年10月18-20日

**7折** 预售中, 现在报名立减2040元  
团购享更多优惠, 截至2018年7月1日



极客邦科技  
企业培训与咨询

Geekbang

扫码关注  
获取更多培训信息

