

## 推酷

- [文章](#)
- [站点](#)
- [主题](#)
- [活动](#)
- [公开课](#)
- [APP](#) [差](#)
- [周刊](#)
  - [编程狂人](#)
  - [设计匠艺](#)
  - [创业周刊](#)
  - [科技周刊](#)
  - [Guru Weekly](#)
  - [一周拾遗](#)

# Kafka是个奇葩! ——Linkin [登录](#)

## 论文学习笔记

时间 2015-10-13 16:09:39 [机会永远留给有准备的人](#)

原文 <http://thinkinginjavablog.sinaapp.com/?p=649>

主题 [Apache Kafka](#)

## Kafka是啥?

是个消息中间件吗? 那和市面上其他一堆堆的中间件例如ActiveMQ, RabbitMQ有什么区别?

答案只有一个:

Kafka是个集群的消息中间件+存储, 一个节点可以存储几T的数据!

为啥一个中间件需要存储数据呢?

## 慢慢道来……

原来, 对于Linkin这样的互联网企业来说, 用户和网站上产生的 数据有三种 :

1. 需要实时响应的交易数据, 用户提交一个表单, 输入一段内容, 这种数据最后是存放在关系数据库(Oracle, MySQL)中的, 有些需要事务支持。
2. 活动流数据, 准实时的, 例如页面访问量、用户行为、搜索情况, 这些数据可以产生啥? 广播、排序、个性化推荐、运营监控等。这种数据一般是前端服务器先写文件, 然后通过批量的方式把文件倒到Hadoop这种大数据分析器里面慢慢整。
3. 各个层面程序产生的日志, 例如httpd的日志、tomcat的日志、其他各种程序产生的日志。码农专用, 这种数据一个是用来说监控报警, 还有就是用来做分析。

Linkin 的牛逼之处, 就在于他们发现了原先2,3的数据处理方式有问题, 对于2而言, 原来动辄一两个钟头批处理一次的方式已经不行了, 用户在一次购买完之后最好马上就能看到相关的推荐。而对于3而言, 传统的syslog模式等也不好, 而且很多情况下2和3用的是同一批数据, 只是数据消费者不一样。

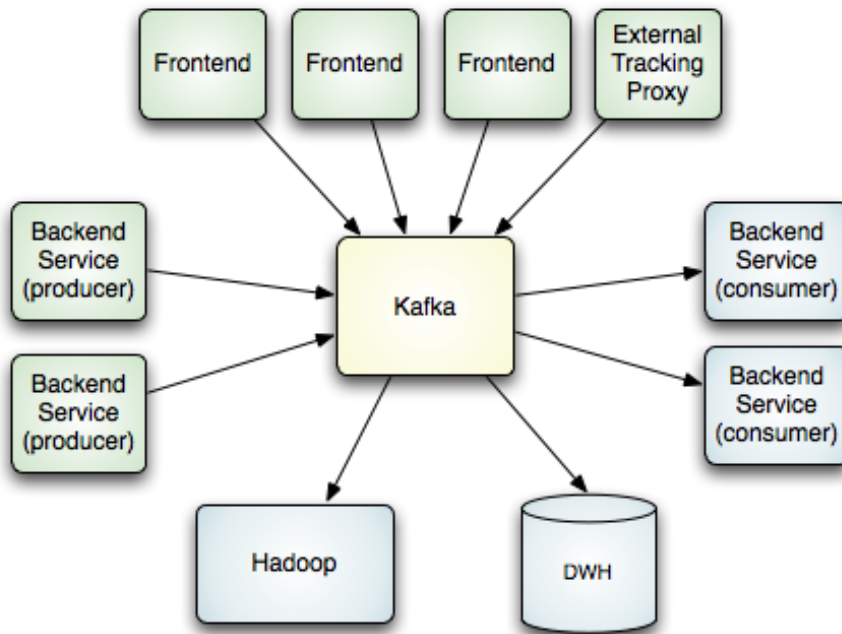
这2种数据的特点是:

1. 准实时，不需要秒级响应，分钟级别即可。
2. 数据量巨大，是交易数据的10倍以上。
3. 数据消费者众多，例如评级、投票、排序、个性化推荐、安全、运营监控、程序监控、后期报表等

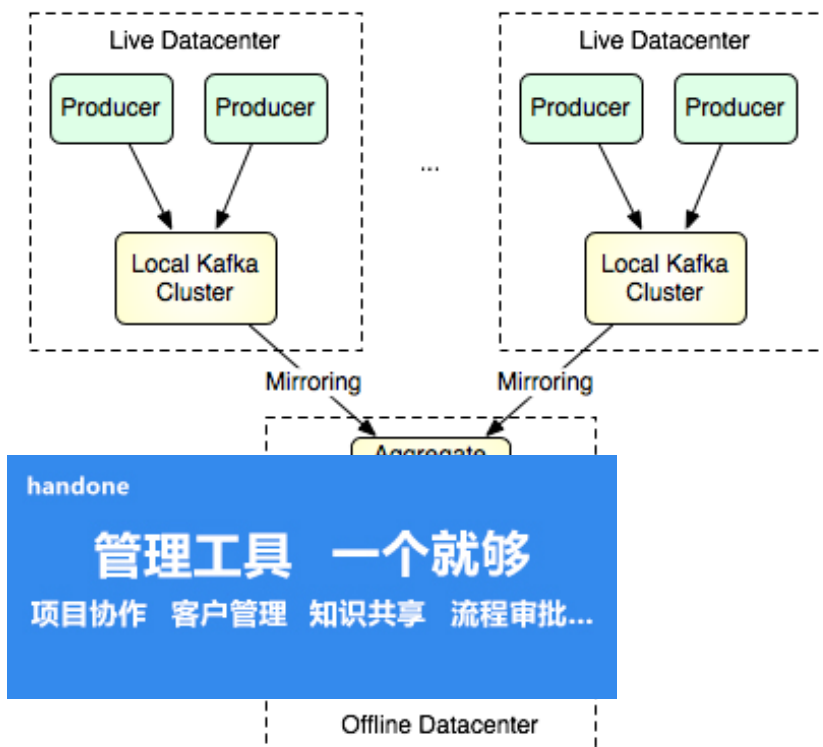
于是，Linkin就自己开发了一套系统，专门用来处理这种性质的数据，这就是Kafka

那么，在整个实践过程中Linkin做了什么样的设计， 解决了什么问题？

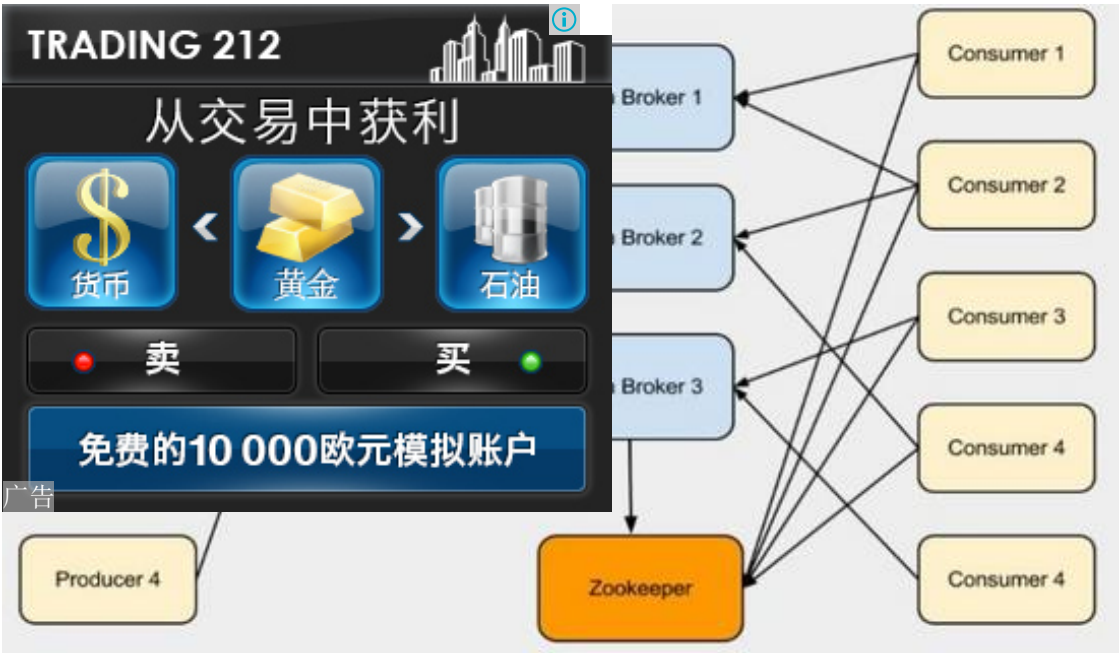
首先看下数据流动图：



多数据中心怎么管理数据：



集群本身的架构图



Kafka内部架构图，分为数据产生者(Producer)，数据中间者(Broker)，数据消费者(Consumer)

显然，这是一个 集群的发布/订阅系统 ，有如下几个特点

1. 生产者是推数据(Push)，消费者是拉数据(Pull)。存在数据复用，在Linkin平均生产1条消息会被消费5.5次。
2. 数据生产者和数据消费者的速度不对等，所以要把数据沉淀在Kafka内慢慢处理，Linkin一般在集群内放7天的数据。
3. 性能上追求高吞吐，保证一定的延时性之内。这方面做了大量优化，包括没有全局hash，批量发送，跨数据中心压缩等等。
4. 容错性上使用的“至少传输一次”的语义。不保证强一次，但避免最多传一次的情况。
5. 集群中数据分区，保证单个数据消费者可以读到某话题(topic)的某子话题(例如某用户的数据)的所有数据，避免全局读数据
6. 数据规范性，所有数据分为数百个话题，然后在数据的源头——生产者(Producer)这边就用Schema来规范数据，这种理念使得后期的数据传输、序列化、压缩、消费都有了统一的规范，同时也解决了这个领域非常麻烦的数据版本不兼容问题——生产者一改代码，消费者就抓瞎。
7. 用于监控，这个系统的威力在于，前面所有生产系统的数据流向，通过这个系统都能关联起来，用于日常的运营也好，用于数据审计，用于运维级别的监控也好都是神器啊！

所以，Kafka的设计基本上目前这个领域的唯一选择。我也看了很多 其他实现 ，包括：

scribe(Facebook)	2	C++	已停止更新，不建议使用
flume(Apache, Cloudera)	1	Java	配置较重
chukwa(Hadoop)	12	Java	2012发布最后一版，不建议使用
fluentd	1	Ruby	比较活跃，看起来不错
logstash	12345	JRuby	功能全，据说有不少小bug
splunk	12345	C/Python	商业闭源，功能强大，可做参考
timetunnel(Alibaba)	2	Java	基于thrift，10年左右成熟
kafka(Linkin)	2 4	Scala	性能强劲，设计巧妙，可以作为基础设施
Samza(Linkin)	12345		=Kafka+YARN+Hadoop

rabbitmq/activemq/qpid	2	Java	传统消息中间件
Storm(twitter)	3	Clojure	实时计算系统
Jstorm(Alibaba)	3	Java	storm的Java版, 据说更稳定
S4(Yahoo)	3	Java	2013年已停止维护
Streambase(IBM)	3	Java	商业产品, 作为参考
HStreaming	3	Java	商业产品, 作为参考
spark	3	Scala	基于Hadoop
mongodb	4	C++	比较浪费硬盘
mysql	4	C++	无需多说
hdfs/hbase	4	Java	无需多说

=

1. 数据采集组件
2. 数据传输组件
3. 数据实时计算/索引/搜索组件
4. 数据存储/持久化组件
5. 数据展示/查询/报警界面组件

从数据传输这块的设计理念来说, Kafka是最为先进的,

在目前的各种实现中, 我猜测可以和Kafka一战的也就只有Splunk了

后面我会分析一下这2个软件的设计和实现

欲知后事如何, 且听下回分解 ~~

## 主要参考文章

[日志: 每个软件工程师都应该知道的有关实时数据的统一概念](#) —— 这篇比较抽象, 高屋建瓴, 理论先行

[Building LinkedIn' s Real-time Activity Data Pipeline](#) —— 实践层的论文, 把做事情的前因后果都写明白了

[分布式发布订阅消息系统 Kafka 架构设计](#) —— 落地设计

## 次要参考文章

《Building LinkedIn' s Real-time Activity Data Pipeline》

《分布式发布订阅消息系统 Kafka 架构设计》

《StreamBase简介》

《Yahoo! s4和Twitter storm的粗略比较》

《最火爆的开源流式系统Storm vs 新星Samza》

《架构之淘宝实时数据传输平台: TimeTunnel介绍》

《Graylog2 简介》

《logstash 还是不行》

《日志收集以及分析: Splunk 》

《LogStash日志分析系统》

《LogStash, 使日志管理更简单》

《logstash VS splunk》

《个性化离线实时分析系统pora》

《日志: 每个软件工程师都应该知道的有关实时数据的统一概念》

《基于Flume的美团日志收集系统(二)改进和优化》

《基于Flume的美团日志收集系统(一)架构和设计》

《对互联网海量数据实时计算的理解》

《流式日志系统启示录》

《flume-ng+Kafka+Storm+HDFS 实时系统搭建》



分享



收藏

纠错

 简道云

数据搜集  管理  分析  协作 

不用Excel, 简道云全部能搞定

#### 推荐文章

- 1. [距离全方位掌握Apache spark 2.0你就差最后两步](#)
- 2. [构建一套成功大数据基础设施需要遵循的七项要诀](#)
- 3. [最伟大的程序员列表](#)
- 4. [你还需要知道的9个大数据技术](#)
- 5. [Hive和HBase整合](#)
- 6. [如何在Hive构造Dual表](#)

#### 相关推刊



by [xunlei1221](#)

《默认推刊》 6

我来评几句

请输入评论内容...

登录后评论

已发表评论数 (0)

相关站点

机会永远留给有准备的人

[机会永远留给有准备的人](#)

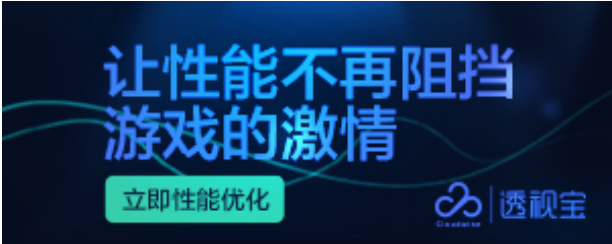
+ 订阅

热门文章

- 1. [【知识库专访】蒋守壮：Hive性能优化实战分享](#)
- 2. [距离全方位掌握Apache spark 2.0你就差最后两步](#)
- 3. [构建一套成功大数据基础设施需要遵循的七项要诀](#)
- 4. [HBase集群管理](#)
- 5. [最伟大的程序员列表](#)







收藏到推刊

[创建推刊](#)

收藏

取消

推刊名(必填)

请填写推刊名

推刊描述

描述不能大于100个字符!

权限设置:

☒ 公开

☐ 仅自己可见

创建

取消

×

文章纠错

邮箱地址

错误类型

正文不准确 ▼

补充信息

网站相关

- [关于我们](#)
- [移动应用](#)
- [建议反馈](#)

提交

关注我们



[推酷网](#)



tuicool2012

友情链接

[人人都是产品经理](#) [PM256](#) [移动信息化](#) [行晓网](#) [智城外包网](#) [虎嗅](#) [IT耳朵](#) [创媒工场](#) [经理人分享](#) [市场部网](#) [砍柴网](#) [CocoaChina](#) [北风网](#) [云智慧](#) [我赢职场](#) [大数据时代](#) [奇笛网](#) [咕噜网](#) [红联linux](#) [Win10之家](#) [鸟哥笔记](#) [爱游戏](#) [投资潮](#) [31会议网](#) [极光推送](#) [Teambition](#) [硅谷网](#) [leangoo](#) [ZEALER中国](#) [OpenSNS](#) [小牛学堂](#) [handone](#) [Scrum中文网](#) [比戈大牛](#) [又拍云](#) [更多链接](#)  
[>>](#)