

03 | Kafka只是消息引擎系统吗？

2019-06-08 胡夕



你好，我是胡夕。今天我们来聊一个老生常谈的话题：**Kafka**只是消息引擎系统吗？

要搞清楚这个问题，我们不可避免地要了解一下**Apache Kafka**的发展历程。有的时候我们会觉得说了解一个系统或框架的前世今生似乎没什么必要，直接开始学具体的技术不是更快更好吗？其实，不论是学习哪种技术，直接扎到具体的细节中，亦或是从一个很小的点开始学习，你很快就会感到厌烦。为什么呢？因为你虽然快速地搞定了某个技术细节，但无法建立全局的认知观，这会导致你只是在单个的点上有所进展，却没法将其串联成一条线进而扩展成一个面，从而实现系统地学习。

我这么说是有依据的，因为这就是我当初学习**Kafka**的方式。你可能不会相信，我阅读**Kafka**源码就是从**utils**包开始的。显然，我们不用看源码也知道这玩意是干什么用的，对吧？就是个工具类包嘛，而且这种阅读源码的方式是极其低效的。就像我说的，我是在一个点一个点地学习，但全部学完之后压根没有任何感觉，依然不了解**Kafka**，因为不知道这些包中的代码组合在一起能达成什么效果。所以我说它是很低效的学习方法。

后来我修改了学习的方法，转而从自上而下的角度去理解**Kafka**，竟然发现了很多之前学习过程中忽略掉的东西。更特别的是，我发现这种学习方法能够帮助我维持较长时间的学习兴趣，不会阶段性地产生厌烦情绪。特别是在了解**Apache Kafka**整个发展历史的过程中我愉快地学到了很多运营大型开源软件社区的知识和经验，可谓是技术之外的一大收获。

纵观**Kafka**的发展脉络，它的确是从消息引擎起家的，但正如文章标题所问，**Apache Kafka**真

的只是消息引擎吗？通常，在回答这个问题之前很多文章可能就要这样展开了：那我们先来讨论下什么是消息引擎以及消息引擎能做什么事情。算了，我还是直给吧，就不从“唐尧虞舜”说起了。这个问题的答案是，**Apache Kafka是消息引擎系统，也是一个分布式流处理平台（Distributed Streaming Platform）**。如果你通读全文但只能记住一句话，我希望你记住的就是这句。再强调一遍，**Kafka是消息引擎系统，也是分布式流处理平台**。

众所周知，**Kafka**是**LinkedIn**公司内部孵化的项目。根据我和**Kafka**创始团队成员的交流以及查阅到的公开信息显示，**LinkedIn**最开始有强烈的数据强实时处理方面的需求，其内部的诸多子系统要执行多种类型的数据处理与分析，主要包括业务系统和应用程序性能监控，以及用户行为数据处理等。

当时他们碰到的主要问题包括：

- 数据正确性不足。因为数据的收集主要采用轮询（**Polling**）的方式，如何确定轮询的间隔时间就变成了一个高度经验化的事情。虽然可以采用一些类似于启发式算法（**Heuristic**）来帮助评估间隔时间值，但一旦指定不当，必然会造成较大的数据偏差。
- 系统高度定制化，维护成本高。各个业务子系统都需要对接数据收集模块，引入了大量的定制开销和人工成本。

为了解决这些问题，**LinkedIn**工程师尝试过使用**ActiveMQ**来解决这些问题，但效果并不理想。显然需要有一个“大一统”的系统来取代现有的工作方式，而这个系统就是**Kafka**。

Kafka自诞生伊始是以消息引擎系统的面目出现在大众视野中的。如果翻看**0.10.0.0**之前的官网说明，你会发现**Kafka**社区将其清晰地定位为一个分布式、分区化且带备份功能的提交日志（**Commit Log**）服务。

这里引出一个题外话，你可能好奇**Kafka**这个名字的由来，实际上**Kafka**作者之一**Jay Kreps**曾经谈及过命名的原因。

因为**Kafka**系统的写性能很强，所以找了个作家的名字来命名似乎是一个好主意。大学期间我上了很多文学课，非常喜欢**Franz Kafka**这个作家，另外为开源软件起这个名字听上去很酷。

言归正传，**Kafka**在设计之初就旨在提供三个方面的特性：

- 提供一套**API**实现生产者和消费者；
- 降低网络传输和磁盘存储开销；
- 实现高伸缩性架构。

在专栏后面的课程中，我们将陆续探讨**Kafka**是如何做到以上三点的。总之随着**Kafka**的不断完善，**Jay**等大神们终于意识到将其开源惠及更多的人是一个非常棒的主意，因此在**2011年Kafka**正式进入到**Apache**基金会孵化并于次年**10月**顺利毕业成为**Apache**顶级项目。

开源之后的Kafka被越来越多的公司应用到它们企业内部的数据管道中，特别是在大数据工程领域，Kafka在承接上下游、串联数据流管道方面发挥了重要的作用：所有的数据几乎都要从一个系统流入Kafka然后再流向下流的另一个系统中。这样的使用方式屡见不鲜以至于引发了Kafka社区的思考：与其我把数据从一个系统传递到下一个系统中做处理，我为何不自己实现一套流处理框架呢？基于这个考量，Kafka社区于0.10.0.0版本正式推出了流处理组件Kafka Streams，也正是从这个版本开始，Kafka正式“变身”为分布式的流处理平台，而不仅仅是消息引擎系统了。今天Apache Kafka是和Apache Storm、Apache Spark和Apache Flink同等级的实时流处理平台。

诚然，目前国内对Kafka是流处理平台的认知还尚不普及，其核心的流处理组件Kafka Streams更是少有大厂在使用。但我们也欣喜地看到，随着在Kafka峰会上各路大神们的鼎力宣传，如今利用Kafka构建流处理平台的案例层出不穷，而了解并有意愿使用Kafka Streams的厂商也是越来越多，因此我个人对于Kafka流处理平台的前景也是非常乐观的。

你可能会有这样的疑问：作为流处理平台，Kafka与其他主流大数据流式计算框架相比，优势在哪里呢？我能想到的有两点。

第一点是更容易实现端到端的正确性（Correctness）。Google大神Tyler曾经说过，流处理要最终替代它的“兄弟”批处理需要具备两点核心优势：**要实现正确性和提供能够推导时间的工具。实现正确性是流处理能够匹敌批处理的基石。**正确性一直是批处理的强项，而实现正确性的基石则是要求框架能提供精确一次处理语义，即处理一条消息有且只有一次机会能够影响系统状态。目前主流的大数据流处理框架都宣称实现了精确一次处理语义，但这是有限定条件的，即它们只能实现框架内的精确一次处理语义，无法实现端到端的。

这是为什么呢？因为当这些框架与外部消息引擎系统结合使用时，它们无法影响到外部系统的处理语义，所以如果你搭建了一套环境使得Spark或Flink从Kafka读取消息之后进行有状态的数据计算，最后再写回Kafka，那么你能只能保证在Spark或Flink内部，这条消息对于状态的影响只有一次。但是计算结果有可能多次写入到Kafka，因为它们不能控制Kafka的语义处理。相反地，Kafka则不是这样，因为所有的数据流转和计算都在Kafka内部完成，故Kafka可以实现端到端的精确一次处理语义。

可能助力Kafka胜出的第二点是它自己对于流式计算的定位。官网上明确标识Kafka Streams是一个用于搭建实时流处理的客户端库而非是一个完整的功能系统。这就是说，你不能期望着Kafka提供类似于集群调度、弹性部署等开箱即用的运维特性，你需要自己选择适合的工具或系统来帮助Kafka流处理应用实现这些功能。

读到这你可能会说这怎么算是优点呢？坦率来说，这的确是一个“双刃剑”的设计，也是Kafka社区“剑走偏锋”不正面PK其他流计算框架的特意考量。大型公司的流处理平台一定是大规模部署的，因此具备集群调度功能以及灵活的部署方案是不可或缺的要素。但毕竟这世界上还存在着很多中小企业，它们的流处理数据量并不巨大，逻辑也并不复杂，部署几台或十几台机器足以应

付。在这样的需求之下，搭建重量级的完整性平台实在是“杀鸡焉用牛刀”，而这正是Kafka流处理组件的用武之地。因此从这个角度来说，未来在流处理框架中，Kafka应该是有一席之地的。

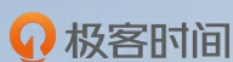
除了消息引擎和流处理平台，Kafka还有别的用途吗？当然有！你能想象吗，Kafka能够被用作分布式存储系统。Kafka作者之一Jay Kreps曾经专门写过一篇文章阐述为什么能把[Kafka用作分布式存储](#)。不过我觉得你姑且了解下就好了，我从没有见过在实际生产环境中，有人把Kafka当作持久化存储来用。

说了这么多，我只想阐述这样的观点：Apache Kafka从一个优秀的消息引擎系统起家，逐渐演变成现在分布式的流处理平台。你不仅要熟练掌握它作为消息引擎系统的非凡特性及使用技巧，最好还要多了解下其流处理组件的设计与案例应用。

开放讨论

你觉得Kafka未来的演进路线是怎麼样的？如果你是Kafka社区的“掌舵人”，你准备带领整个社区奔向什麼方向呢？（提示下，你可以把自己想象成Linus再去思考）

欢迎写下你的思考和答案，我们一起讨论。如果你觉得有所收获，也欢迎把文章分享给你的朋友。



Kafka 核心技术与实战

全面提升你的 Kafka 实战能力

胡夕

人人贷计算平台部总监

Apache Kafka Contributor



新版升级：点击「 请朋友读」，20位好友免费读，邀请订阅更有**现金**奖励。

精选留言



Michael IYZY

13



学到了。刚接触，对一次性处理语义的概念和背后的含义不太明确，能否结合实例讲解比较一下...

2019-06-08

作者回复

举个例子，如果我们使用**Kafka**计算某网页的**PV**——我们将每次网页访问都作为一个消息发送的**Kafka**。**PV**的计算就是我们统计**Kafka**总共接收了多少条这样的消息即可。精确一次处理语义表示每次网页访问都会产生且只会产生一条消息，否则有可能产生多条消息或压根不产生消息。

2019-06-10



梁亮

7

推荐大家去搜索一个**Confluence**的演讲，题目是**ETL is dead**，其中讲到了**Kafka**在流处理平台的来龙去脉

2019-06-08



清晨吼于林

2

老师您好~~~

我了解的：一个**partition**在一个**group**内，只能被一个消息者进程消费（一个**jvm**，启动了一个**java**进程）。

问题前提：经过分区算法的匹配，**A partition**被**B**消费者消费。

我的问题：在这个**B**的消费者里面，假如我用多线程消费（多个线程，每个线程维护了一个**KafkaConsumer**实例。而不是一个**KafkaConsumer**然后多个**worker**线程消费的模式），那这多个线程都能从这个**A partition**里面取到消息嘛？

2019-06-12

作者回复

同一个组下有多少消费者实例不是看进程数或线程数，而是看创建的**KafkaConsumer**实例数。

所以在你的场景中，**B**消费者不是一个，而是多个，因为**B**进程启动了多个线程，而每个线程都维护了单独的**KafkaConsumer**实例。

2019-06-12



燕子上

1

还是那句话：**Apache Kafka**是消息引擎系统，也是一个分布式流处理平台！主：消息引擎，辅：流处理

2019-06-09



南辕北辙

1

目前团队还是在用**flink**做实时流计算，希望好好跟着老师学一下**kafka**的流处理，再来好好对比一下

2019-06-09



Shane

0

老师，能举个例子说明下流出来和批处理的区别吗？

目前我的理解就是批处理是一次请求中包含多条消息？然后消费者取出这一整个请求内容进行处理消费。流处理就是每个请求每次只发送一条消息，所以消费者也只能每次消费一条？

感觉自己理解的应该不怎么正确呢？网络上的解释也是非常虚，想看看老师有啥指导的吗？

2019-06-15



demmm

👍 0

消息引擎系统，也是一个分布式流处理平台

想问下这两个概念到底有什么区别呢

2019-06-13

作者回复

严格来说这两个是完全不同领域内的东西。各自都有响当当的理论、框架。

2019-06-13



Java农民

👍 0

基本的运维kafka-manager

2019-06-13



趙衍

👍 0

关于【但是计算结果有可能多次写入到 Kafka，因为它们不能控制 Kafka的语义处理】。我想问老师，Kafka不是在0.11版本实现了exactly once，保证一条消息只会被消费一次吗，为什么说计算结果还有可能会被多次写入到Kafka呢？

2019-06-12

作者回复

嗯嗯，这说的就是0.11之前的故事。事实上，Apache Flink从1.4开始推出了支持E2E Exactly-Once语义的两阶段SinkFunction。它用的就是Kafka 0.11的事务

2019-06-13



霄嵩

👍 0

老师写的很用心，加油！

2019-06-11



风中花

👍 0

其实要学会是什么，最终得选择交给自己，掌握得知识广度和厚度来，做出更好的选择。百花齐放啊，，，，

2019-06-11



pepezzzz

👍 0

题外话“唐尧虞舜”

2019-06-11



文竹

👍 0

目前kafka相关组件太分散以及通过命令行运维，希望有一个平台产品集成所有组件（基础运维，ksql，AI等）

2019-06-11

作者回复

Confluent公司的Confluent Kafka做到了这一点。社区版的Kafka没有这样的功能~

2019-06-11



永光

0

老师好，

对于流处理不是很了解，然后你说的Kafka流处理的两个优点都没有get到，呜呜呜。

1、第一点是更容易实现端到端的正确性。

目前主流的大数据流处理框架都宣称实现了精确一次处理语义，但这是有限定条件的，即它们只能实现框架内的精确一次处理语义，无法实现端到端的。

主流的大数据流处理框架为什么不能实现端到端的精确一次处理语义，什么是端到端的精确一次处理语义呀？

2、它自己对于流式计算的定位。这是在说Kafka流处理相比于其他的更轻量级？

2019-06-10

作者回复

1. E2E你可以近似地认为是从生产者端到消费者端或者是跨上下游系统。这种场景下如何确保消息或事件对最终状态的影响有且只有一次是极有挑战的课题。

2. 可以这么理解。Kafka Streams目前定位是流处理库（library），而不是平台（platform）。这也说明了它和其他流处理框架的差异化。

2019-06-10



陈华应

0

为中小型公司提供开箱即用的流处理平台，低成本，低学习曲线，支持与其他框架集成~~

2019-06-09



言希

0

胡大神，最近生产环境上遇到问题，数据在kafka中积压了一段时间，重新启动消费者后显示加入消费组成功但是一直不消费数据，current-offset一直没变。能帮忙提供排查思路嘛|

2019-06-09

作者回复

看看是否存在消息corrupted导致的hang住——可能的原因是网络传输中的某些瞬时故障，不过应该极为罕见。可以启动一个Console consumer去消费下试试

2019-06-10



Francis

0

我觉得kafka还有很多点需要优化，exactly once 的传输保障，消息的延迟问题，不确定问题，不稳定问题，都是金融级应用所不承受的问题，在高可用和性能之间的取舍和平衡

2019-06-09



DNT

0



虽然Kafka的优势是吞吐量，但是我想请教一下latency的问题。对于生产级的硬件（ssd，Linux系统优化，最新cpu，0.1以下的ping），一个Kafka的消息从发出到被收到，合理的latency应该是多少范围内？

2019-06-09

作者回复

这个没有通用的范围值，但我像毫秒到几十毫秒级的消息传输应该是能够达到的吧。

2019-06-10



明翼

0

希望胡大神多讲点kafka的流计算的方面内容，我还是比较敢兴趣，像这种只有API实时流客户端的系统，整个处理流程如何，虽然以前看了些不完备，谢谢。

2019-06-08

不瘦二十斤
不改头像

jeffery

0

整章介绍Kafka 是消息引擎系统，也是一个分布式流处理平台 是不是有点太奢华了！阿里的收购flink 必hot

2019-06-08