



下载APP



## 结束语 | 源码学习，我们才刚上路呢

2020-07-18 胡夕

Kafka核心源码解读

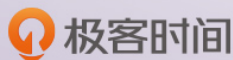
[进入课程 >](#)**胡夕**

Apache Kafka Committer, 老虎证券技术总监

你好,我是胡夕。

作为 IT 从业者,我们不光要写代码,更要读代码。阅读优秀框架的源码,已经从“BEST DO”变成了“MUST DO”,不要担心坚持不下来,这条路上充满了同行者,我们结伴向前。

# Source Code

**讲述：胡夕**

时长 08:16 大小 7.58M



你好，我是胡夕。《Kafka 核心源码解读》课到今天就告一段落了，又到了和你说再见的时候。

当初接下这个任务时，我记得小编和我说，咱们这个课程将会是极客时间的第一个源码类的专栏课，我顿时觉得既自豪又忐忑。

自豪的是，我能够作为首个源码专栏课作者，给你分享 Kafka 源码的学习心得；忐忑的是，源码并不好讲，如果讲解的粒度太粗，就会流于表面，不够深入，而粒度太细的话就很容易陷入流水账似的代码走读，这其中的度，需要仔细拿捏才行。



不过，我依然还是愿意试一试。

因为我始终觉得，现在的编程教育对于阅读源码的重视程度远远不足。一直以来，学校或各类培训机构只看重编程技能的灌输，而忽略了带领大家一同学习业界流行框架的源码编写方法，但这恰恰是非常关键的。如果你没有阅读过大量质量上乘的代码，那么，你就不可能写出优秀的代码。

就像美国作家斯蒂芬·金所说：“If you don't have time to read, you don't have the time (or the tools) to write. Simple as that.” 虽然这句话是关于写作的，但是对于我们写代码来说，同样也是适用的。学习比自己更有经验的人写出的源码，一定会直接或间接地影响你自己的编码水平。

我曾在 [开篇词](#) 中提到过阅读源码的 3 个好处，但实际上，对于一个每天都要写代码的程序员来说，**我们阅读源码的最大目的，就是能够写出更好的代码**。这么多年阅读 Kafka 源码的经历，让我对此深有体会，因此，我很想把我的阅读经验整理出来，分享给你，尽可能帮你更加科学高效地阅读源码，从而写出更好的代码。

从 2019 年 12 月开始筹备这门课，到今天交付完成，共经历了近八个月的时间，我精心挑选并讲解了 Kafka Broker 端主要组件的核心源码。我个人觉得，总体上还是成功的。有很多模块的源码分析，我自己在写的时候都收获良多。因此，我在想，如果你认真地学完了前面的内容，应该也是有些收获的。

当然，无论你眼下掌握了多少，这些在过去更新的知识就停留在过去了，等着你时时去翻阅，去消化，去借助它们解决现在以及未来的问题。此刻，我们就要暂时跟它们说再见了，我想再和你聊一聊未来的事情，未来学习源码的事情。

课程结束后，你应该如何继续你的 Kafka 源码阅读计划呢？我根据自己的经验，再给你分享一些建议。

**首先，我建议你研读一下 Kafka Clients 端的代码。**

你可能会问，难道不是应该先阅读 Broker 端其他组件的代码吗，为什么要读 Clients 端的代码？实际上，Clients 工程包含的已经不仅仅是客户端的源码了，它还包含了各类支撑服务器端组件的基础代码，比如通信通道、消息集合，等等。了解它们的实现，可以补齐你对 Kafka 代码框架的整体认知。**通读并搞懂 Clients 和 Core 工程下的代码，是精通 Kafka 实现原理的必要步骤。**

**其次，把课程中涉及到组件的测试用例代码学习一遍，最好是打上断点，边运行边**

**Debug。**很多时候，光看组件实现源码，还是会很难理解它真正运行起来是什么样的，而测试用例就能够很好地带领我们一探究竟。

我举个真实的例子。当初，我自己学习 Log Cleaner 源码时，就非常吃力。我把 LogCleaner 类看了很多遍，却始终不得要领。后来，我突然想到，不是有个 LogCleanerTest 吗？于是，我就翻开里面的各种 test×××方法，挨个儿琢磨，终于明白了这个组件实际的工作原理。因此，我想特别提醒你一句，千万不要忽视测试用例代码的威力。

**最后，你一定要读一读 Log Cleaner 组件的代码。**

它主要用于完成带 Key 消息的日志清理，在其他场合的出境机会不多，所以我并没有专门在课程里讲它，但是，它其实也是 Broker 端非常重要的组件。

我们熟知的位移主题，就是依靠这个机制，定期删除过期位移消息和过期注册消息的。事实上，清除过期数据或重复数据，是任何一个数据存储系统都要解决的关键课题。这个组件的代码量不多，你可以在很短的时间内掌握它。

Okay，现在阅读的方向有了，但你依然可能会遇到一个难题，那就是，看不懂怎么办？

虽然我给出了一些源码阅读方法，但是，毕竟，理解别人写的代码是一件很有挑战的事情，因此，我的答案是：“别着急！”

刚开始时，你可能压根不知道某个方法或类的代码是做什么的，这是非常正常的现象。这个时候，你要做的，就是初步了解一下代码结构和轮廓，知道有哪些类，类中定义了哪些 API。这就像学习写作一样，努力搞懂一些知识范畴之外的代码，就可以扩大你的代码功力的外延。如果一直徜徉在简单、枯燥、易理解的代码海洋中，就很难真正地有所进步。

同时，你也不要幻想，代码只读一遍就可以了！阅读源码这件事，一定要重复很多次之后，才会有“质”的变化。

不知道你是否听说过“洋葱源码阅读法”，它指的是，我们读源码，就要像剥洋葱一样，第一遍的阅读仅仅是获取整体项目结构和概要，接下来的多遍阅读，才会逐步揭开代码的

本质。切记：**每次只剥一层**。

除此之外，如果说我还有什么私藏的方法，那就是亲自动手去修改源码。

实际上，只有你对阅读的代码进行了一些改进或变更，才能真正地体会到阅读源码的快乐。因此，在阅读 Kafka 或其他框架的源码时，你不妨亲自动手改一改代码，做一些实验 Debug 一下，去体会一下你所做的改动给 Kafka 带来的变化。相信我，你会爱上这个过程。

**最后，我还想说的一点是，不要恐惧。**阻碍很多人阅读源码的第一大原因就是恐惧。一想到有那么多行代码要读，就本能地竖起了放弃的大旗。其实，再宏大的框架，也是一行行代码堆砌起来的。我一直秉承着一句七字箴言：**你敢写我就敢读**！现在我把它送给你，希望在你想要放弃的时候，把它拿出来，鼓励一下自己。

好了，这一季的课程就到这里结束吧。今天没有太多华丽的表达，全部都是些朴实的建议。其实，所有的一切，我其实都只是想表达，源码学习，我们才刚刚上路。

阅读源码仿佛一场漫长的登山之旅，半山腰很挤，我们山顶上见。

我给你准备了一份 [📄 结课问卷](#)，希望你能在问卷里聊一聊你对这门课的看法。欢迎你点击下面的图片，用 1~2 分钟时间填写问卷，期待你畅所欲言。

**胡夕**

Apache Kafka Committer, 老虎证券技术总监

感谢一起走过的这段时间，非常想听听你对我和这个课程的反馈和建议。在 2020 年 8 月 2 日前提交，将有机会获得



原创 | 手绘护腕垫

价值 **¥69**

或



极客时间课程阅码

价值 **¥99****填写问卷** [提建议](#)



更多课程推荐

# 设计模式之美

前 Google 工程师手把手教你写高质量代码

王争

前 Google 工程师

《数据结构与算法之美》专栏作者



涨价倒计时 🕒

限时秒杀 **¥149**，7月31日涨价至 **¥299**

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 期末测试 | 一套习题，测试你的掌握程度

## 精选留言 (4)

写留言



陪你去砍刘星宇

2020-07-21

胡老师您好，感谢您的课程，我到了很多。请问是否有机会再开一门课程讲讲connect，谢谢老师。



1



Roger宇

2020-07-18

胡老师。有个问题我实在是没有办法了，但探索欲折磨我想知道答案，还希望老师能帮帮我。关于kafka消息的有序性，我们都知道同一个分区内可以保证消息的有序性。但我百度谷歌遍了网上也找不到任何人可以解释为什么kafka能做到这一点。我从LogSegment找到Log，再到MemoryRecords到MemoryRecordsBuilder然后到RecordBatch再到Record sAccumulator。尽管log.append和MemoryRecordsBuilder.append等方法确实加锁...  
展开 ∨

作者回复: 网上搜不到因为它压根就不是由Kafka保证的，而是由操作系统的页缓存机制来保存的。

3

1



伯安知心  
2020-07-19

胡老师，以后还会有课程吗？期待～  
展开

作者回复: 嗯，暂时没有了～



伯安知心  
2020-07-18

感谢～  
展开