

开篇词 | 深度拆解核心原理，轻松掌握所有消息队列

许文强 · 深入拆解消息队列47讲



你好，我是文强。一名在消息队列领域一线摸爬滚打很多年的开发者，很高兴在这里遇见你！

跟大多数人一样，在职业生涯前期，我是一个纯粹的消息队列使用者，在日常开发过程中大量使用 Kafka 和 RabbitMQ。后来我到了腾讯云消息队列的团队，主要负责多款消息队列产品的商业化、内核优化、架构升级、运维运营体系建设、成本优化、大规模集群稳定性等工作。这段经历让我对消息队列这个领域有了非常深刻的理解，同时也打磨了我的技术能力和解决问题的能力，积累了大量的实战经验。

从而我也理解作为消息队列的使用者、研发人员、运维人员、架构师，他们希望获得什么，解决什么问题。

作为使用者，我们会希望针对我们的业务场景来更好地完成技术选型，从而选出符合我们业务需求、系统架构的最优组件。我们一般会从功能特性、性能、稳定性出发，查阅资料，横向对比多款消息队列，从而做出选择。

作为研发人员，我们希望了解消息队列的底层原理、设计思考、实现方案，从而来提升我们自己的技术能力、技术视野，提升竞争力。

作为运维人员，我们希望能够理解用到的消息队列的系统架构、设计原理等，从而判断这款产品是否足够稳定，是否有隐藏的风险，长期的运维成本是否合理，监控体系是否完善等等。从而评估系统的风险点，以及时规避。

作为架构师，我们希望自己有足够的技术视野，能够了解到每款消息队列的功能清单、系统架构的优劣势、潜在的风险、成本结构、社区活跃度等等信息。让我们在设计架构的过程中可以有足够的理论依据辅助我们做出合理的决策。

但不管什么角色，都会遇到一个关键问题，那就是：**消息队列那么多，看起来那么复杂，我们是否做出了最优选择？**

而要解决这个问题，我们是不是要把业界那么多的主流消息队列都学一遍，从原理到源码，这样的成本会不会太高了呢？

我的答案是不用，我们有更简单的方法能达成这个目的。

消息队列那么多，我该怎么学习？

我们知道，中间件作为三大基础软件之一，消息队列是其中重要的组成部分。

早年业界消息队列演进的主要推动力在于功能（如延迟消息、事务消息、顺序消息等）、场景（实时场景、大数据场景等）、分布式集群的支持等等。近几年，随着云原生架构和 Serverless 的普及，业界 MQ 主要向实时消息和流消息的融合架构、Serverless、Event、协议兼容等方面演进。从而实现计算、存储的弹性，实现集群的 Serverless 化。

从架构设计角度来看，消息队列在设计思想上基本是一致的，在架构演进过程中存在相互借鉴。这也给我们学习消息队列提供了一个便捷的路径，即只要我们从**需求出发，理解设计原理、主流技术方案、方案之间的优劣、选型过程主要的思考点**，那么我们再往下学习具体某一款消息队列就会变得非常简单。

举个简单的例子，我们在选型时经常会问，哪款消息队列性能最高？

此时如果我们知道了通信协议、网络模型、存储结构、生产消费原理、集群部署、分片副本等因素是跟性能密切相关的，如果我们还进一步知道了哪种网络模型性能最高，哪种存储结构在哪种场景下性能最高，然后你再结合具体某个消息队列的实现机制，回答这个问题就非常简单了。

那我们具体该怎么做呢？掌握的关键是什么？

掌握消息队列的关键路径是什么？

在我看来，掌握的核心关键点是成体系、系统、全面的知识结构。只要掌握了核心和顶层设计原理，不管有多少消息队列都能轻松驾驭。

以最近消息队列领域最火的 Apache Pulsar 举例。从它的设计思想中，你会看到 Kafka、RocketMQ、RabbitMQ 的影子。从架构的角度，Pulsar Broker 和 Kafka 的设计几乎是一模一样的。所以你只要理解了 Kafka、RocketMQ、RabbitMQ，再来理解 Pulsar，就是事半功倍的。

另外，为了深入理解某款消息队列，我们经常会去研究源码。不过这又会碰到一个老大难问题，源码太多，甚至看不懂。其实你的源码能不能看懂是依赖于对原理的理解程度，而原理又依赖于体系化学习和梳理，简单来说就是“知其然才知所以然”。总之会绕回来。

那么在源码学习上，其实是可以变被动为主动的。这里我非常建议你，掌握了原理后尝试去开源社区成为 Contributor，贡献代码、文档等等，我自己是参与了 Apache Kafka、Apache RocketMQ、Apache Pulsar 社区的一些工作，对我的帮助很大。包括我后面在腾讯云升级改造很多适合公有云场景的产品特性时，我也从开源社区中沉淀了很多去完成我司自己的 Kafka、RabbitMQ、Pulsar 的从 0 到 1。这种帮助它很难去量化，但会潜移默化影响你解决问题的思维模式。

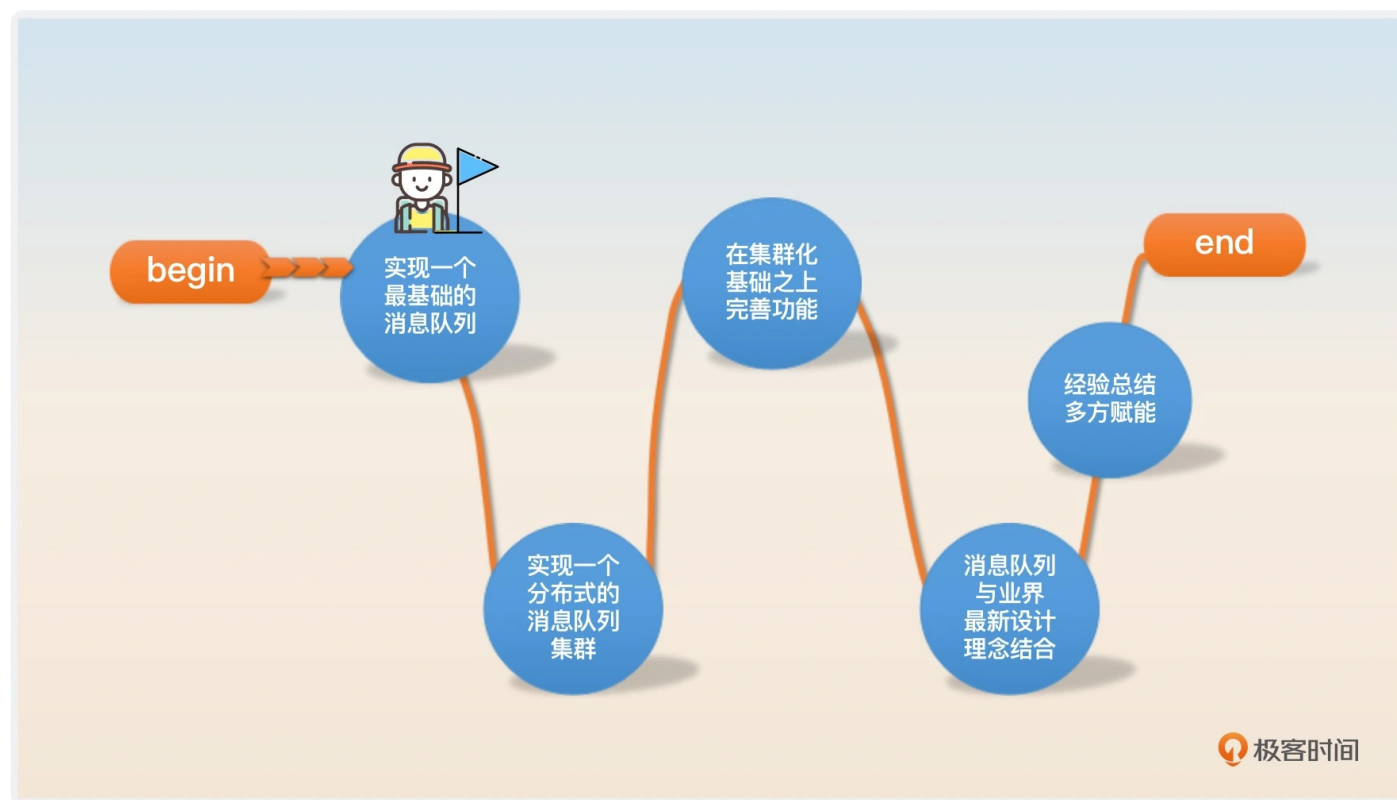
这时你还会收获一个惊喜！因为在实际的软件开发中，消息队列作为基础软件，它可以串联起分布式系统设计、网络编程、操作系统、存储原理、计算机硬件、数据结构、数据一致性等等

多个技术点。在学习的过程中，你会不知不觉地复习、串联起这些知识点。这不仅会让你对消息队列的使用更加得心应手，还会让你在日常面试中有不错的表现。

所以基于我的经验和视野出发，我们这门课就有了一个清晰的学习路径。

这门课能为你带来什么？

我会带你深入拆解消息队列的各个模块的主流技术方案、选型思考、技术实现，从而让你对消息队列的技术架构设计有一个清晰、整体、全面的认识。



这门课我们设置了六个模块，分为预习篇、基础篇、进阶篇、功能篇、架构升级篇、经验总结篇。

预习篇：我将结合自己的实践经历，带你了解主流 MQ 的发展脉络，从中明确未来发展方向，并就后续课程频繁提及的基础概念做一个对齐。

基础篇：从功能上来看一个最基础的消息队列应该具备生产、存储、消费的能力。我将带你从通信协议、网络模块、存储模块、生产者、消费者五个部分，来分析一个最基础的消息队列应

该考虑什么，如何选型以及如何设计实现。最后会围绕着这五个部分分析 4 款主流消息队列的设计实现。

进阶篇：我们将在最基础的消息队列之上，从集群瓶颈和可靠性风险分析、如何构建集群、如何确保数据一致性、集群的安全控制、集群的可观测性、编码技巧、集群优化等七个方面，来分析实现一个分布式的消息队列集群应该考虑什么、要怎么做，有哪些方案可以选择以及各种方案的优劣。最后还是会围绕着这七个方面分析 4 款主流消息队列的设计实现。

功能篇：在集群化的消息队列的基础上，我们将探讨在集群上如何实现顺序消息、幂等消息、延时消息、事务消息、死信队列、优先级队列、消息查询、支持 Schema、支持 WebSocket 等功能。我们将详细分析实现这些功能要考虑什么，如何做方案设计、技术选型，以及最后是怎样实现这些功能的。

架构升级篇：在前面四个模块，我们已经实现了一个集群化、功能丰富的消息队列。接下来我们将探讨在云原生架构演进、降本增效诉求增大、Serverless/Event 概念兴起等背景下，消息队列是如何跟进业界最新设计理念，做架构升级以满足系统弹性和降本诉求的。我们将从存算分离架构、分层存储、Serverless/Event 架构、集群容灾、数据连接、消息中台等六个方面来展开讲解技术上如何实现、如何考虑、能达到什么效果、有哪些风险、能满足哪些场景等等。

经验总结篇：这个模块比较特别，是基于我自己多年的观察所产出的分享，从业务视角输出，赋能业务和产品，也为自己赋能，期待能帮到你。如果有更多希望讨论的话题，我们也可以在这个模块详聊。



尽最大的努力，做最好的自己

消息队列是一个复杂的基础软件，但我希望从庖丁解牛的角度出发，跟你一起从设计一个最简单的消息队列开始，将其扩展为一个集群，再往集群添加各种功能，最后我们会结合云原生、Serverless 的架构理念，对架构做一个全面升级。

知识最怕孤岛，无法串联就容易忘记。在学习期间，你还会发现我们几乎会用到计算机领域的所有基础知识点，包括但不限于分布式、操作系统、网络、存储等等。所以这门课程并不只是一门消息队列的原理课，也可以是一门分布式系统的原理课。不仅能在工作中帮到你，也能在面试中帮到你。

最后我想说，在互联网降本增效的浪潮下，我希望我和你都能做到：尽最大的努力，做最好的自己。我经常给自己说的一句话就是：加油，lobo。此刻与你共勉，一起开启这段学习之旅！

© 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。

精选留言 (12)



张申傲

2023-06-26 来自北京

消息队列是分布式中间件的集大成者，其中涉及到协议设计、网络通信、数据存储、分布式系统架构等多个方面的技术，掌握好消息队列可以使得技术能力有一个整体的提升。看到课程内容感觉老师真的是用心了，希望跟着老师的专栏学下来，并梳理内化成自己的知识体系~

作者回复: 老哥加油，一起学习。之前跟我朋友讲过，我看我自己写的东西，也觉得能学到东西，——。

主要人脑的记忆力实在有限，有系统的梳理才能记得住，才能回想起来。

共 3 条评论 >



9



LargeOrange

2023-06-29 来自上海

老师看到您是腾讯的，问一个问题，之前面试腾讯云子公司，面试官和我聊kafka，问了我一个问题你觉得kafka有什么致命的缺点。我实在想不出来，能帮忙解释下么。我觉得kafka作为业界非常主流的消息队列中间件，如果有什么致命弱点，估计早就翻天了吧？

作者回复: 哈哈，这个问题肯定不会是我问的。

在我看来kafka 有痛点，但是没有致命缺点。从面试的角度，我觉得可以这么回答：

当前版本的 kafka 没有致命缺点，但是下面几个痛点：

1. 弹性扩容的能力，在Broker负载高的时候，没办法快速扩容，需要先迁移数据，扩容速度会比较慢，极端情况下，可能无法扩容，影响业务。
2. 消费分组Rebalance速度在一些极端情况(如分区和消费者比较多)会比较慢，而Rebalance会导致消费暂停，从而影响消费性能。
3. 基于zookeeper的架构，可能会出现Zookeeper、Controller、Broker之间的元数据不一致。可能会导致集群异常。另外zookeeper的存在，增加了kafka的运维复杂度，所以社区在支持去zookeeper的版本。
4. 功能层面，kafka支持的功能比较简单，目前主要支持生产、消费、事务、幂等功能。大家希望kafka 支持更多的功能，比如延时消息、死信队列、消息轨迹等，但是社区当前不支持。

共 2 条评论 >



5



木几、

2023-06-26 来自福建

期待已久，果断上车

编辑回复: 欢迎，一起开启一段奇妙旅程！



👍 2



XXG

2023-06-26 来自山东

尽最大的努力，做最好的自己。

编辑回复: 没毛病！



👍 1



humor

2023-07-12 来自浙江

冒昧的问一句，这门课跟 消息队列高手课的区别是啥

编辑回复: 哈喽，同学你好，这是个好问题！小编来详细回复一下。

1. 这两门课的学习人群是接近的，整体看上手难度更低。
2. 内容定位为《消息队列高手课》进阶篇内容的系统化。以四款业界经典消息队列和最新技术架构为学习重点，详解技术点，深入关键功能设计，挖掘设计方案，分析方案之间的差异，带你掌握消息队列的架构全貌。不啃源码，内容设计上会更加轻松易懂，接近当下工作场景。
3. 交付核心重在探索实现集群化、功能丰富的消息队列，以及消息队列如何跟进近几年业界最新的设计理念，包含架构演进过程中比较新的值得学习的内容。



Z.

2023-07-06 来自北京

上海滩大佬出品应该不会差吧

作者回复: 谢谢老哥，一起学习~





码小呆

2023-07-01 来自广东

感觉, 消息队列,mysql,redis 是必学的,也是面试必问的....

作者回复: 是的 深入了解一个后 再去了解其他的 效果会更好

共 2 条评论 >



春

2023-06-29 来自广东

跟大佬一起学习。共勉

作者回复: 加油, 老哥

我也是 加油 lobo



长脖子树

2023-06-27 来自浙江

上海滩许文强 哈哈哈

作者回复: hhhh



水瓶

2023-06-27 来自广东

看了一整个目录页才决定入手, 对小新很友好, 版块清晰, 简明实用, 一直想找一个老师实实在在的带着走, 哈哈, 等到了, 不负期待。



Mr.J

2023-06-26 来自北京

正好消息队列是我薄弱的地方, 跟着老师学习

编辑回复: 期待能帮助你巩固好这部分知识。



aoe

2023-06-26 来自浙江

看到了有一个可爱的小娃娃的学习路线图，就激动的加入了课程

编辑回复: 欢迎，期待闯关打卡！

