

# Kafka权威指南

O'REILLY®

TURING 图灵程序设计丛书



# Kafka权威指南

Kafka: The Definitive Guide

Kafka核心作者和业界一流一线人员共同执笔  
全面介绍Kafka设计原理和架构细节

[美] Neha Narkhede Gwen Shapira Todd Palino 著  
薛命灯 译

中国工信出版集团 人民邮电出版社

[Kafka权威指南\\_下载链接1](#)

著者:Neha Narkhede

出版者:人民邮电出版社

出版时间:2017-12-26

装帧:平装

isbn:9787115473271

每个应用程序都会产生数据，包括日志消息、度量指标、用户活动记录、响应消息等。如何移动数据，几乎变得与数据本身一样重要。如果你是架构师、开发者或者产品工程师，同时也是Apache

Kafka新手，那么这本实践指南将会帮助你成为流式平台上处理实时数据的专家。

本书由出身于LinkedIn的Kafka核心作者和一线技术人员共同执笔，详细介绍了如何部署Kafka集群、开发可靠的基于事件驱动的微服务，以及基于Kafka平台构建可伸缩的流式应用程序。通过详尽示例，你将会了解到Kafka的设计原则、可靠性保证、关键API，以及复制协议、控制器和存储层等架构细节。

- 了解发布和订阅消息模型以及该模型如何被应用在大数据生态系统中
- 学习使用Kafka生产者和消费者来生成消息和读取消息
- 了解Kafka保证可靠性数据传递的模式和场景需求
- 使用Kafka构建数据管道和应用程序的最佳实践
- 在生产环境中管理Kafka，包括监控、调优和维护
- 了解Kafka的关键度量指标
- 探索Kafka如何成为流式处理利器

作者介绍：

Neha Narkhede，Confluent联合创始人、CTO，曾在LinkedIn主导基于Kafka和Apache Samza构建流式基础设施，是Kafka作者之一。

Gwen Shapira，Confluent系统架构师，帮助客户构建基于Kafka的系统，在可伸缩数据架构方面拥有十余年经验；曾任Cloudera公司解决方案架构师。另著有《Hadoop应用架构》。

Todd Palino，LinkedIn主任级SRE，负责部署管理大型的Kafka、Zookeeper和Samza集群。

### 【译者简介】

薛命灯，毕业于厦门大学软件学院，十余年软件开发和架构经验，InfoQ高级社区编辑。译有《硅谷革命》《生产微服务》等书。微信公众号CodeDeep。

目录: 序 xiii

前言 xv

第1章 初识Kafka 1

1.1 发布与订阅消息系统 1

1.1.1 如何开始 2

1.1.2 独立的队列系统 3

1.2 Kafka登场 4

1.2.1 消息和批次 4

1.2.2 模式 4

1.2.3 主题和分区 5

1.2.4 生产者和消费者 5

1.2.5 broker和集群 6

1.2.6 多集群 7

1.3 为什么选择Kafka 8

|                            |    |
|----------------------------|----|
| 1.3.1 多个生产者                | 8  |
| 1.3.2 多个消费者                | 8  |
| 1.3.3 基于磁盘的数据存储            | 9  |
| 1.3.4 伸缩性                  | 9  |
| 1.3.5 高性能                  | 9  |
| 1.4 数据生态系统                 | 9  |
| 1.5 起源故事                   | 11 |
| 1.5.1 LinkedIn的问题          | 11 |
| 1.5.2 Kafka的诞生             | 12 |
| 1.5.3 走向开源                 | 12 |
| 1.5.4 命名                   | 13 |
| 1.6 开始Kafka之旅              | 13 |
| 第 2 章 安装Kafka              | 14 |
| 2.1 要事先行                   | 14 |
| 2.1.1 选择操作系统               | 14 |
| 2.1.2 安装Java               | 14 |
| 2.1.3 安装Zookeeper          | 15 |
| 2.2 安装Kafka Broker         | 17 |
| 2.3 broker配置               | 18 |
| 2.3.1 常规配置                 | 18 |
| 2.3.2 主题的默认配置              | 19 |
| 2.4 硬件的选择                  | 23 |
| 2.4.1 磁盘吞吐量                | 23 |
| 2.4.2 磁盘容量                 | 23 |
| 2.4.3 内存                   | 23 |
| 2.4.4 网络                   | 24 |
| 2.4.5 CPU                  | 24 |
| 2.5 云端的Kafka               | 24 |
| 2.6 Kafka集群                | 24 |
| 2.6.1 需要多少个broker          | 25 |
| 2.6.2 broker 配置            | 25 |
| 2.6.3 操作系统调优               | 26 |
| 2.7 生产环境的注意事项              | 28 |
| 2.7.1 垃圾回收器选项              | 28 |
| 2.7.2 数据中心布局               | 29 |
| 2.7.3 共享Zookeeper          | 29 |
| 2.8 总结                     | 30 |
| 第 3 章 Kafka生产者——向Kafka写入数据 | 31 |
| 3.1 生产者概览                  | 32 |
| 3.2 创建Kafka生产者             | 33 |
| 3.3 发送消息到Kafka             | 34 |
| 3.3.1 同步发送消息               | 35 |
| 3.3.2 异步发送消息               | 35 |
| 3.4 生产者的配置                 | 36 |
| 3.5 序列化器                   | 39 |
| 3.5.1 自定义序列化器              | 39 |
| 3.5.2 使用Avro序列化            | 41 |
| 3.5.3 在Kafka里使用Avro        | 42 |
| 3.6 分区                     | 45 |
| 3.7 旧版的生产者API              | 46 |
| 3.8 总结                     | 47 |
| 第 4 章 Kafka消费者——从Kafka读取数据 | 48 |
| 4.1 KafkaConsumer概念        | 48 |
| 4.1.1 消费者和消费者群组            | 48 |

|                               |    |
|-------------------------------|----|
| 4.1.2 消费者群组和分区再均衡             | 51 |
| 4.2 创建Kafka消费者                | 52 |
| 4.3 订阅主题                      | 53 |
| 4.4 轮询                        | 53 |
| 4.5 消费者的配置                    | 55 |
| 4.6 提交和偏移量                    | 57 |
| 4.6.1 自动提交                    | 58 |
| 4.6.2 提交当前偏移量                 | 59 |
| 4.6.3 异步提交                    | 59 |
| 4.6.4 同步和异步组合提交               | 61 |
| 4.6.5 提交特定的偏移量                | 61 |
| 4.7 再均衡监听器                    | 62 |
| 4.8 从特定偏移量处开始处理记录             | 64 |
| 4.9 如何退出                      | 66 |
| 4.10 反序列化器                    | 67 |
| 4.11 独立消费者——为什么以及怎样使用没有群组的消费者 | 71 |
| 4.12 旧版的消费者API                | 71 |
| 4.13 总结                       | 72 |
| 第5章 深入Kafka                   | 73 |
| 5.1 集群成员关系                    | 73 |
| 5.2 控制器                       | 74 |
| 5.3 复制                        | 74 |
| 5.4 处理请求                      | 76 |
| 5.4.1 生产请求                    | 78 |
| 5.4.2 获取请求                    | 78 |
| 5.4.3 其他请求                    | 80 |
| 5.5 物理存储                      | 81 |
| 5.5.1 分区分配                    | 81 |
| 5.5.2 文件管理                    | 82 |
| 5.5.3 文件格式                    | 83 |
| 5.5.4 索引                      | 84 |
| 5.5.5 清理                      | 84 |
| 5.5.6 清理的工作原理                 | 84 |
| 5.5.7 被删除的事件                  | 86 |
| 5.5.8 何时会清理主题                 | 86 |
| 5.9 总结                        | 86 |
| 第6章 可靠的数据传递                   | 87 |
| 6.1 可靠性保证                     | 87 |
| 6.2 复制                        | 88 |
| 6.3 broker配置                  | 89 |
| 6.3.1 复制系数                    | 89 |
| 6.3.2 不完全的首领选举                | 90 |
| 6.3.3 最少同步副本                  | 91 |
| 6.4 在可靠的系统里使用生产者              | 92 |
| 6.4.1 发送确认                    | 92 |
| 6.4.2 配置生产者的重试参数              | 93 |
| 6.4.3 额外的错误处理                 | 94 |
| 6.5 在可靠的系统里使用消费者              | 94 |
| 6.5.1 消费者的可靠性配置               | 95 |
| 6.5.2 显式提交偏移量                 | 95 |
| 6.6 验证系统可靠性                   | 97 |
| 6.6.1 配置验证                    | 98 |
| 6.6.2 应用程序验证                  | 98 |
| 6.6.3 在生产环境监控可靠性              | 99 |

6.7 总结 100

第7章 构建数据管道 101

7.1 构建数据管道时需要考虑的问题 102

7.1.1 及时性 102

7.1.2 可靠性 102

7.1.3 高吞吐量和动态吞吐量 103

7.1.4 数据格式 103

7.1.5 转换 104

7.1.6 安全性 104

7.1.7 故障处理能力 104

7.1.8 耦合性和灵活性 105

7.2 如何在Connect API和客户端API之间作出选择 105

7.3 Kafka Connect 106

7.3.1 运行Connect 106

7.3.2 连接器示例——文件数据源和文件数据池 107

7.3.3 连接器示例——从MySQL到ElasticSearch 109

7.3.4 深入理解Connect 114

7.4 Connect之外的选择 116

7.4.1 用于其他数据存储的摄入框架 116

7.4.2 基于图形界面的ETL工具 117

7.4.3 流式处理框架 117

7.5 总结 117

第8章 跨集群数据镜像 118

8.1 跨集群镜像的使用场景 118

8.2 多集群架构 119

8.2.1 跨数据中心通信的一些现实情况 119

8.2.2 Hub和Spoke架构 120

8.2.3 双活架构 121

8.2.4 主备架构 123

8.2.5 延展集群 127

8.3 Kafka的MirrorMaker 128

8.3.1 如何配置 129

8.3.2 在生产环境部署MirrorMaker 130

8.3.3 MirrorMaker调优 132

8.4 其他跨集群镜像方案 134

8.4.1 优步的uReplicator 134

8.4.2 Confluent的Replicator 135

8.5 总结 135

第9章 管理Kafka 136

9.1 主题操作 136

9.1.1 创建主题 137

9.1.2 增加分区 138

9.1.3 删主题 138

9.1.4 列出集群里的所有主题 139

9.1.5 列出主题详细信息 139

9.2 消费者群组 140

9.2.1 列出并描述群组 140

9.2.2 删除群组 142

9.2.3 偏移量管理 142

9.3 动态配置变更 143

9.3.1 覆盖主题的默认配置 143

9.3.2 覆盖客户端的默认配置 145

9.3.3 列出被覆盖的配置 145

9.3.4 移除被覆盖的配置 146

|                       |     |
|-----------------------|-----|
| 9.4 分区管理              | 146 |
| 9.4.1 首选的首领选举         | 146 |
| 9.4.2 修改分区副本          | 147 |
| 9.4.3 修改复制系数          | 150 |
| 9.4.4 转储日志片段          | 151 |
| 9.4.5 副本验证            | 152 |
| 9.5 消费和生产             | 153 |
| 9.5.1 控制台消费者          | 153 |
| 9.5.2 控制台生产者          | 155 |
| 9.6 客户端ACL            | 157 |
| 9.7 不安全的操作            | 157 |
| 9.7.1 移动集群控制器         | 157 |
| 9.7.2 取消分区重分配         | 157 |
| 9.7.3 移除待删除的主题        | 158 |
| 9.7.4 手动删除主题          | 158 |
| 9.8 总结                | 159 |
| 第 10 章 监控Kafka        | 160 |
| 10.1 度量指标基础           | 160 |
| 10.1.1 度量指标在哪里        | 160 |
| 10.1.2 内部或外部度量        | 161 |
| 10.1.3 应用程序健康检测       | 161 |
| 10.1.4 度量指标的覆盖面       | 161 |
| 10.2 broker的度量指标      | 162 |
| 10.2.1 非同步分区          | 162 |
| 10.2.2 broker度量指标     | 166 |
| 10.2.3 主题和分区的度量指标     | 173 |
| 10.2.4 Java虚拟机监控      | 174 |
| 10.2.5 操作系统监控         | 175 |
| 10.2.6 日志             | 176 |
| 10.3 客户端监控            | 177 |
| 10.3.1 生产者度量指标        | 177 |
| 10.3.2 消费者度量指标        | 179 |
| 10.3.3 配额             | 181 |
| 10.4 延时监控             | 182 |
| 10.5 端到端监控            | 183 |
| 10.6 总结               | 183 |
| 第 11 章 流式处理           | 184 |
| 11.1 什么是流式处理          | 185 |
| 11.2 流式处理的一些概念        | 186 |
| 11.2.1 时间             | 187 |
| 11.2.2 状态             | 188 |
| 11.2.3 流和表的二元性        | 188 |
| 11.2.4 时间窗口           | 189 |
| 11.3 流式处理的设计模式        | 190 |
| 11.3.1 单个事件处理         | 191 |
| 11.3.2 使用本地状态         | 191 |
| 11.3.3 多阶段处理和重分区      | 193 |
| 11.3.4 使用外部查找——流和表的连接 | 193 |
| 11.3.5 流与流的连接         | 195 |
| 11.3.6 乱序的事件          | 195 |
| 11.3.7 重新处理           | 196 |
| 11.4 Streams示例        | 197 |
| 11.4.1 字数统计           | 197 |
| 11.4.2 股票市场统计         | 199 |

|                         |     |
|-------------------------|-----|
| 11.4.3 填充点击事件流          | 201 |
| 11.5 Kafka Streams的架构概览 | 202 |
| 11.5.1 构建拓扑             | 202 |
| 11.5.2 对拓扑进行伸缩          | 203 |
| 11.5.3 从故障中存活下来         | 205 |
| 11.6 流式处理使用场景           | 205 |
| 11.7 如何选择流式处理框架         | 206 |
| 11.8 总结                 | 208 |
| 附录A 在其他操作系统上安装Kafka     | 209 |
| 作者介绍                    | 214 |
| 封面介绍                    | 214 |
| • • • • • (收起)          |     |

[Kafka权威指南](#) [\\_下载链接1](#)

## 标签

Kafka

消息队列

分布式

大数据

计算机

架构

大数据流

计算机科学

## 评论

入门的书籍，看技术数学有时候不能看的太细致，只要记住其中的技术点，等到用的时

候，能找到相应的书籍去参考，首先要保证技术的全面性，这就要求涉猎大量书籍，所以读所有技术书籍的节奏应该是不一致的，这本书读的太慢了，下次有些内容没必要细读

走马观花的看了一圈，书比较浅显，只能当作入门了解用吧

翻译没什么大问题。== p79 “跟随者副本也不行”弄反了。原文：follower replica, even though they are consumers, are exempt from this.

Kafka 不同于 amqp 的消息队列，本身设计相当白盒所以适用场景更加广泛，这几年已经成为构建数据管道的事实标准了。作者水平很高，翻译也不错，真的是一句废话也没有。

书很薄，简单清晰有实例，入门了解看这本很好

Day 82 入门款小书，Kafka还是有许多细节值得推敲 #百日早起学习挑战

#内容循序渐进，翻译行云流水，大赞

译者叫薛命灯，看过他翻译的也是图灵系列的《生产微服务》，真的叫一个行云流水。本书2天读毕释卷，已然对kafka有了整体清晰的认识，值得一读的好书。

基本覆盖所有关键点。

密度比较高的一本书，系统地介绍了Kafka的方方面面，也有作者的很多经验之谈，挺好。

-----  
作为开发人员，需要好好的理解消费者、生产者、kafka高可用原理这几个章节，关于kafka调优、MirrorMaker、监控等几个章节可以适当的放一放。

-----  
字儿真多

-----  
书比较薄，内容比较简单，但是可以了解到kafka新版本的一些特性。英文比较好的童鞋，可以直接去kafka官网免费下载pdf即可

-----  
姑且也算读过吧

-----  
看完了，忘记了

-----  
实用，深入

\*-----  
写的挺好的，虽然由于篇幅没有说具体的实现细节，不过给出的整体的框架的设计，以及思考问题的方式都是很有借鉴意义的。<sup>\*</sup>阿里开源的rocketmq，支持事物处理，看过一些文章，大概的原理知道，找机会看看实现的机制<sup>\*</sup>最近刚获得InfoWorld最佳开源数据平台的Apache Pulsar，也很有意思，分了一层将存储交给BookKeeper，这个跟google chubby的log是基于某个大学搞得容错db类似的想法。<sup>\*</sup>基于现有的东西来build一些新的东西。联想到了技术的本质和演进，哈哈哈

-----  
4.5/5。挺基础的书，如果之前用过kafka，就会看得比较快，但也能学到东西，算是全方面入门吧。翻译还行，但中间章节部分拗口，需要对比一下原文【有史以来最快的  
一本技术书，用了大概48个小时读完

一本偏原理的书，看一次不能完全消化，有需要可以多翻特定章节，感觉会有不少收获。  
。

---

## 偏向于入门

---

[Kafka权威指南 下载链接1](#)

## 书评

毕竟是kafka作者写的，在操作系统层级讲的挺好的，入门也容易。作者写的挺精要的，kafka最重要的点在于多节点的选举策略，恢复系统到指定状态，客户端不误读，这些上面都有讲的，剩下的就是配置配置配置。入门可以看看，想真的上线使用，最好还是读文档和翻源代码。我在实际使...

---

与Apache Pulsar、AWS Kinesis、Aliyun LogHub等一些新实时数据存储系统相比，Kafka胜在：1. 它是log模型的引领一个时代，参考Jay Kreps: 《The Log: What every software engineer should know about real-time data's unifying abstraction》。2. 在2010年代，优秀的分布式...

---

这本书是为使用Kafka API 开发应用程序的工程师和在生产环节安装、配置、调优、监控 kafka 的运维工程师而写的。  
这本书已经明确指出了，这不适合初学者。如果对消息中间件、流处理之类没有必要的理论类储备和实际应用经验，读取来恐怕有很多陌生的名词，往往不能领会作者在某些...

---

本书由出身于LinkedIn的Kafka核心作者和一线技术人员共同执笔，详细介绍了如何部署Kafka集群、开发可靠的基于事件驱动的微服务，以及基于Kafka平台构建可伸缩的流式应用程序。通过详尽示例，你将会了解到Kafka的设计原则、可靠性保证、关键API，以及复制协议、控制器和存储层等...

Learn how to take full advantage of Apache Kafka, the distributed, publish-subscribe queue for handling real-time data feeds. With this comprehensive book, you'll understand how Kafka works and how it's designed. Authors Neha Narkhede, Gwen Shapira, and Tod...

---

[Kafka权威指南\\_下载链接1](#)