

生态体-创始人

码龄3年

暂无认证

76

13万+

147万+

4005

原创

周排名

总排名

访问

等级

4

0

2

0

6

积分

粉丝

获赞

评论

收藏

私信

关注

搜博主文章

- 热门文章
- bios更新工具_5分钟教会您升级bios主板,华硕主板BIOS教程

992
- cut最后几位 shell_shell中字符串截取命令: cut, printf, awk, sed

283
- 双代号网络图节点时间参数_双代号网络图的绘制与6个时间参数的计算,一篇全学会! ...

214
- v割与邮票孔_pcb板V割和邮票孔的区别

208
- 提取过程_[i]论文阅读石榴皮提取物对生猪肉饼品质的影响

169

- 最新文章
- php图片缩放比例缩放,php图片等比例放大与缩小的方法举例
- windows php redis扩展安装,怎么在Windows下安装PHP的Redis扩展服务
- 车牌识别_php,ecLinker-API_Shop_OCR-车牌识别_API接口_PHP调用示例代码
- 2021

03月7篇

02月30篇

01月30篇
- 2020年

15篇

目录

Consumer

分区分配的演变（Rebalance）

投递模式

创建一个Kafka消费者

kafka 消费者组 消费者订阅不同主题_深入了解kafka系列-消费者

原创生态体-创始人2021-01-22 02:07:3691收藏

文章标签：

kafka 消费者组 消费者订阅不同主题

工程师小C的小店

我也想开通小店

Python编程三剑客: Python编程从入门到实践第2版+快速上手第2版+极客编程...

作者: [美] 埃里克·马瑟斯 (Eric Matthes) 出版社: 人民邮电出版社

好评: 100% 销售量: 0

¥172.9

更多

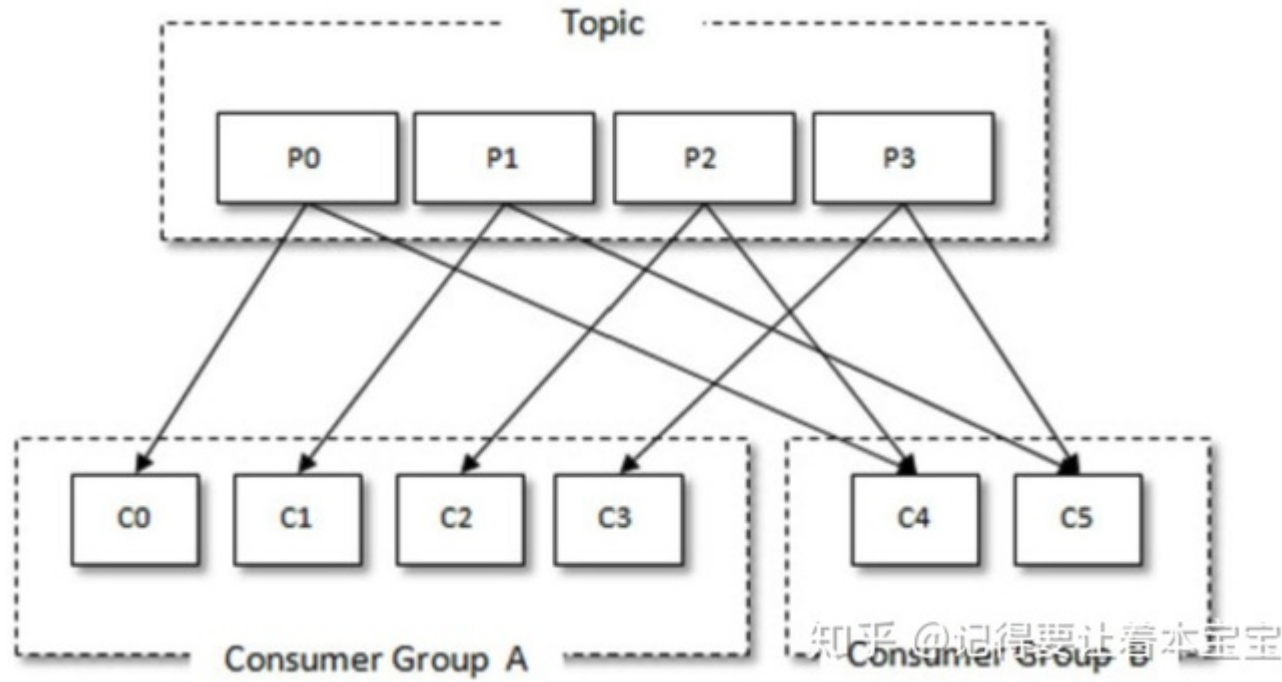
前言

与生产者对应的是消费者，应用程序可以通过KafkaConsumer来订阅主题，并从订阅的主题中拉取消息。不过在使用KafkaConsumer消费消息之前需要先了解消费者和消费组的概念，否则无法理解如何使用KafkaConsumer。

Consumer

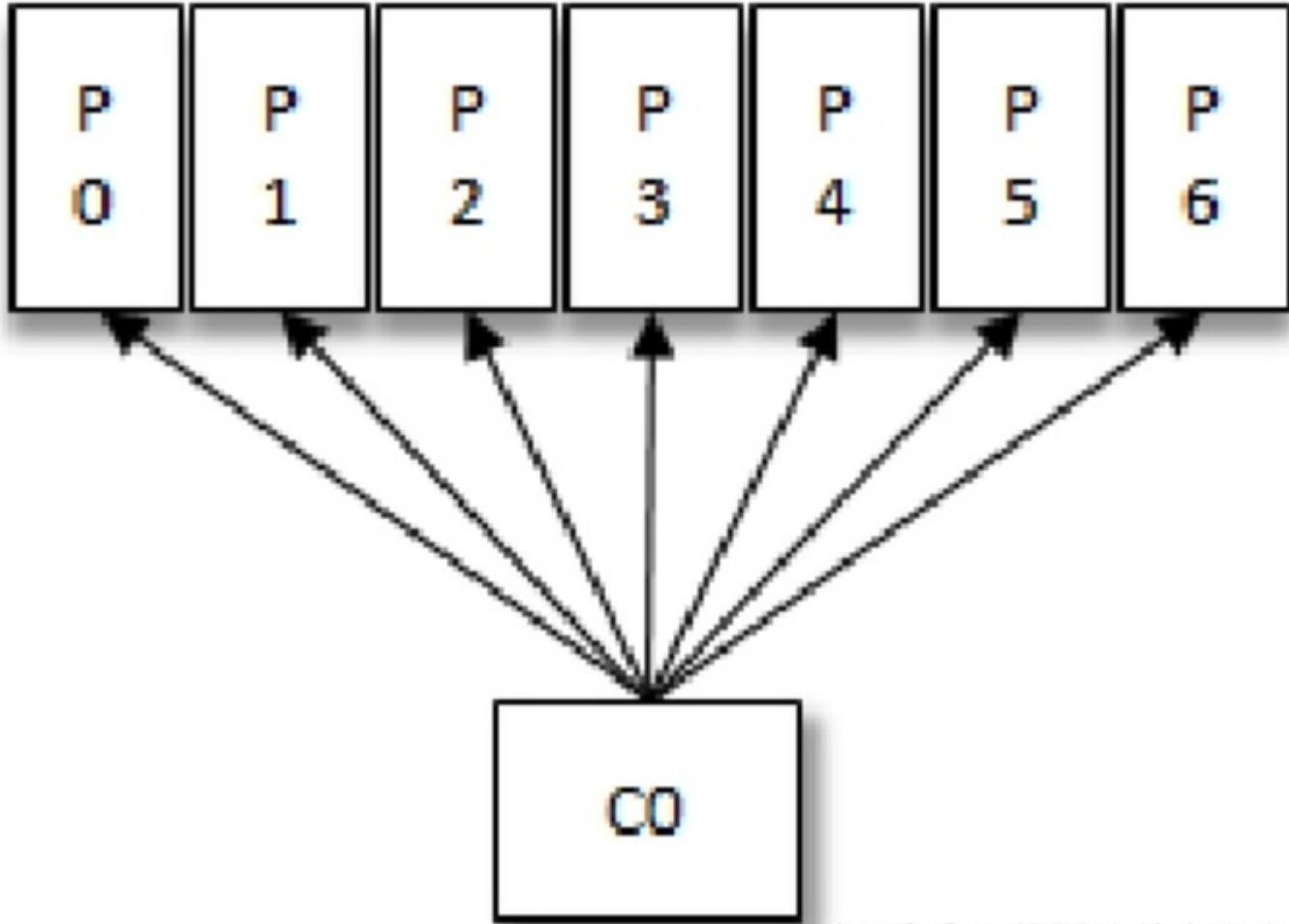
- 消费者（Consumer）负责订阅Kafka中的主题（Topic），并且从订阅的主题上拉取消息。与其他一些消息中间件不同的是：在Kafka的消费理念中还有一层消费组（Consumer Group）的概念，每个消费者都有一个对应的消费组。

当消息发布到主题后，只会被投递给订阅它的每个消费组中的一个消费者。如图所示，某个主题中共有4个分区（Partition）：P0、P1、P2、P3。有两个消费组A和B都订阅了这个主题，消费组A中有4个消费者（C0、C1、C2和C3），消费组B中有2个消费者（C4和C5）。按照Kafka默认的规则，最后的分配结果是消费组A中的每一个消费者分配到1个分区，消费组B中的每一个消费者分配到2个分区，两个消费组之间互不影响。每个消费者只能消费所分配到的分区中的消息。换言之，每一个分区只能被一个消费组中的一个消费者所消费。

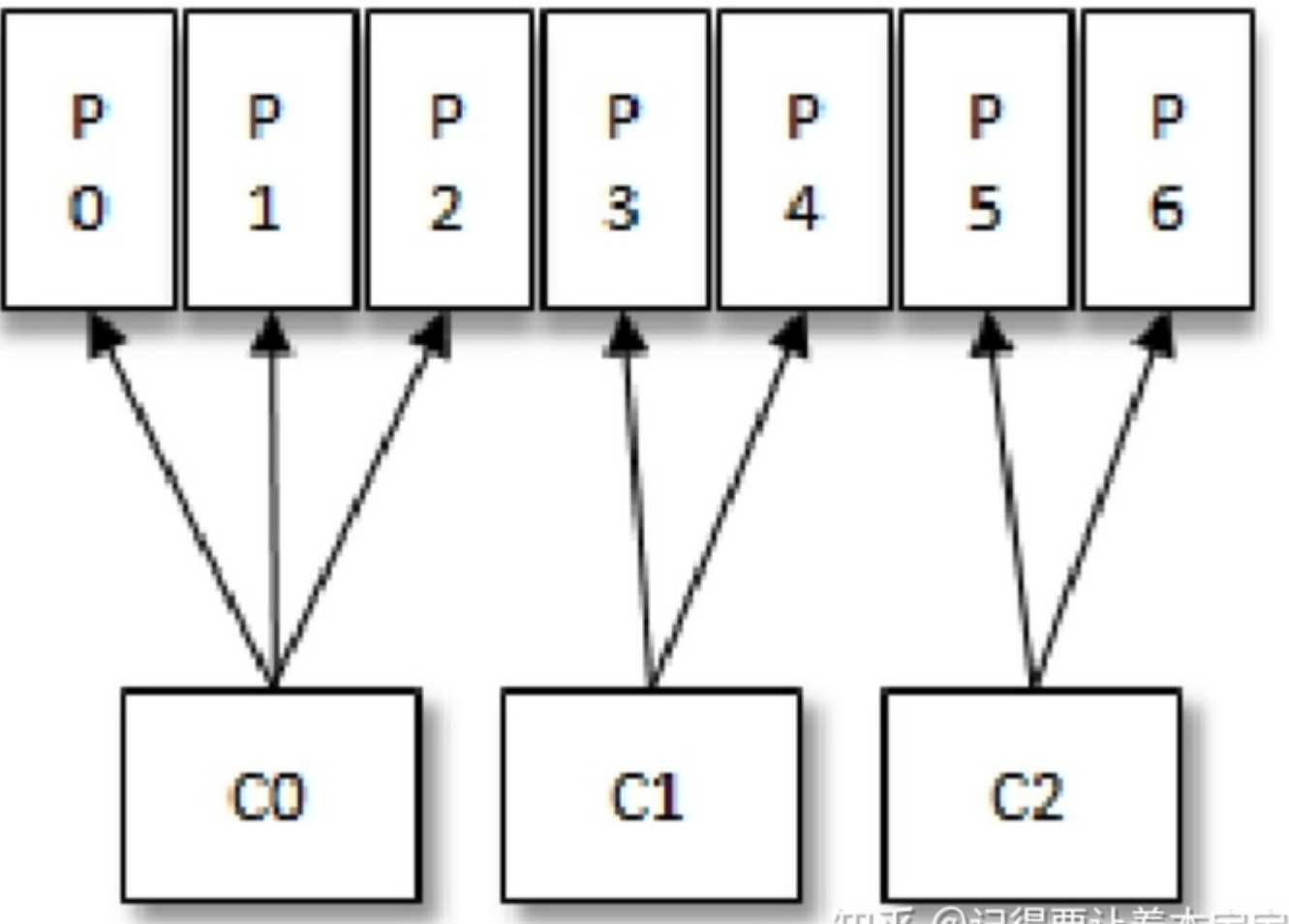


分区分配的演变（Rebalance）

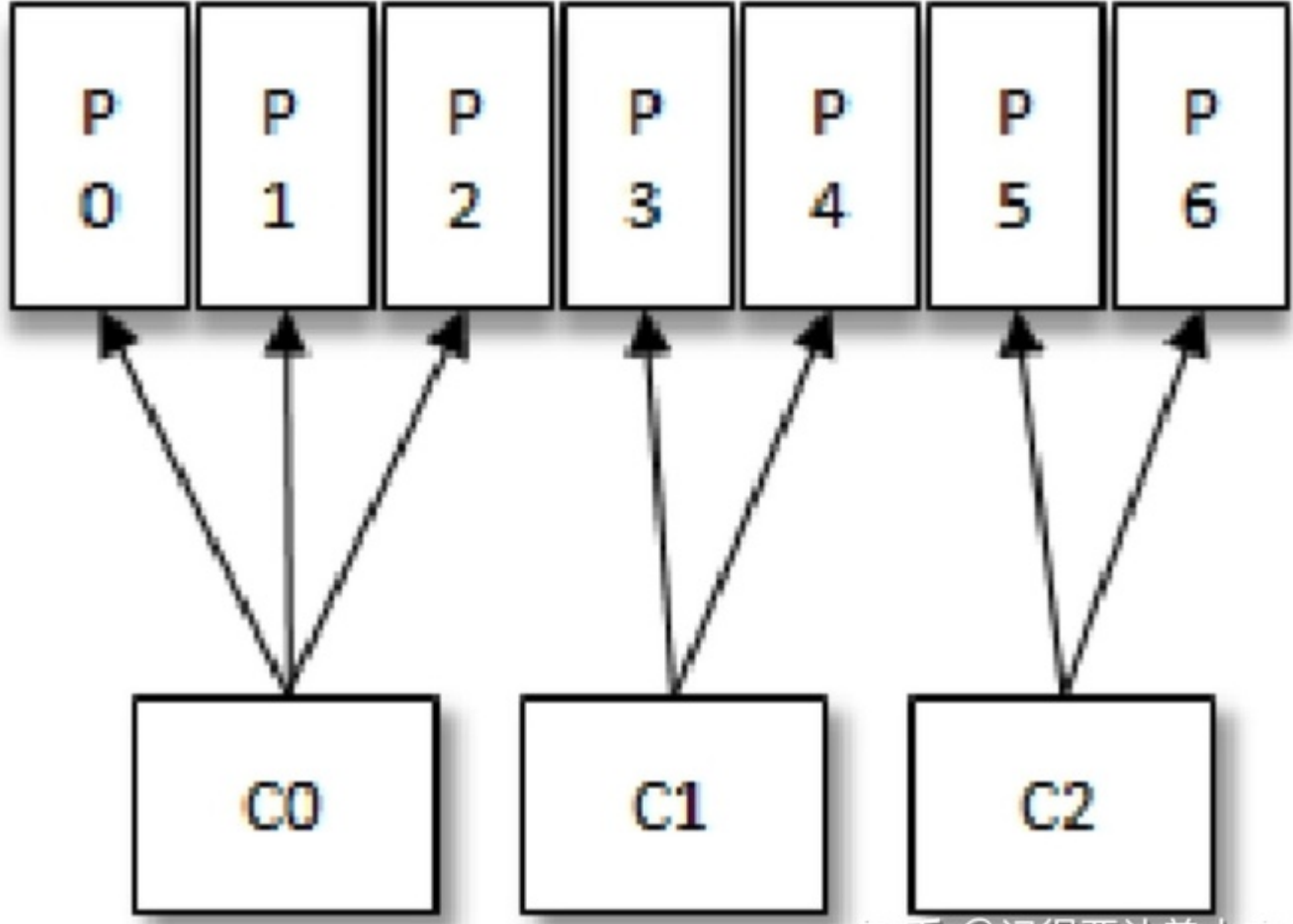
我们再来看一下消费组内的消费者个数变化时所对应的分区分配的演变。假设目前某消费组内只有一个消费者C0，订阅了一个主题，这个主题包含7个分区：P0、P1、P2、P3、P4、P5、P6。也就是说，这个消费者C0订阅了7个分区，具体分配情形如图。



消费者与消费组此时消费组内又加入了一个新的消费者C1，按照既定的逻辑，需要将原来消费者C0的部分分区分配给消费者C1消费，如下图所示。消费者C0和C1各自负责消费所分配到的分区，彼此之间并无逻辑上的干扰。

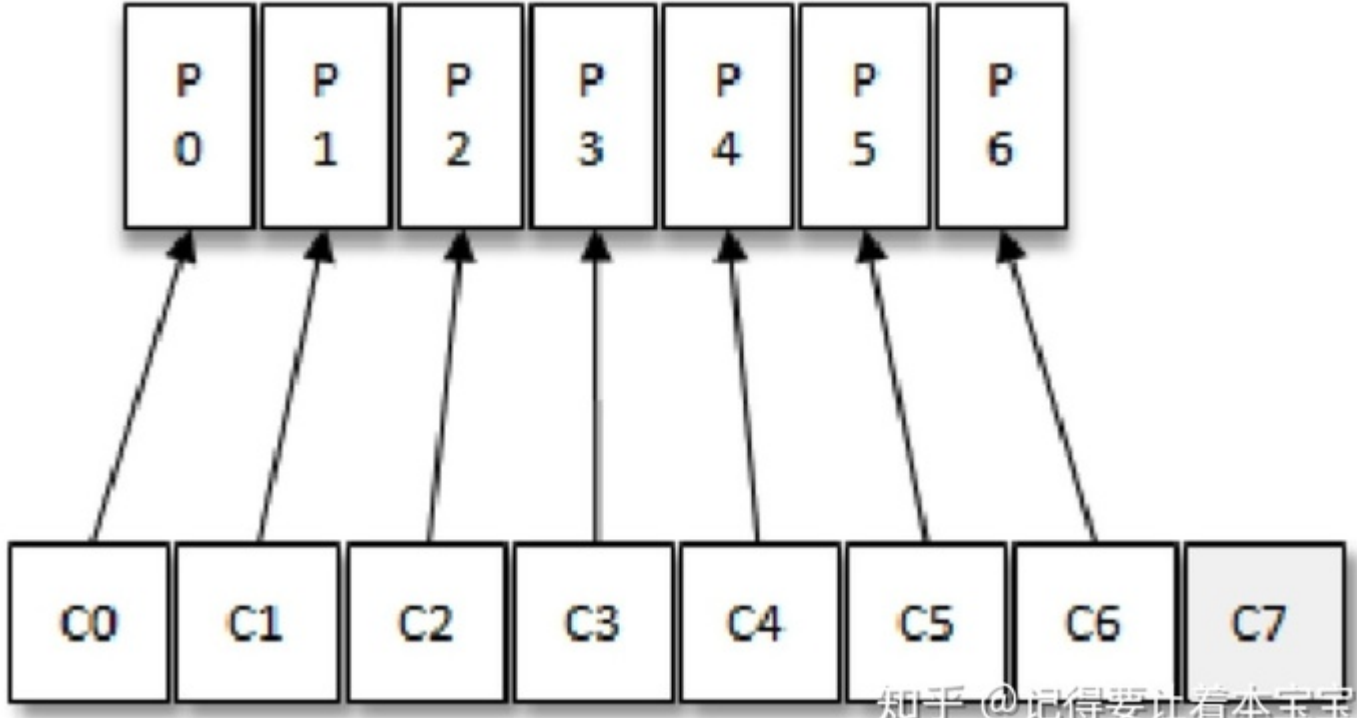


紧接着消费组内又加入了一个新的消费者C2，消费者C0、C1和C2按照下图方式各自负责消费所分配到的分区。



知乎 @记得要让本宝宝

消费者与消费组这种模型可以让整体的消费能力具备横向伸缩性，我们可以增加（或减少）消费者的个数来提高（或降低）整体的消费能力。对于分区数固定的情况，一味地增加消费者并不会让消费能力一直得到提升，如果消费者过多，出现了消费者的个数大于分区个数的情况，就会有消费者分配不到任何分区。参考图如下，一共有8个消费者，7个分区，那么最后的消费者C7由于分配不到任何分区而无法消费任何消息。



知乎 @记得要让本宝宝

投递模式

以上分配逻辑都是基于默认的分区分配策略进行分析的，可以通过消费者客户端参数 `partition.assignment.strategy` 来设置消费者与订阅主题之间的分区分配策略，有关分区分配的更多细节可以再接下来的系列继续聊。

对于消息中间件而言，一般有两种消息投递模式：

点对点（P2P，Point-to-Point）模式: 点对点模式是基于队列的，消息生产者发送消息到队列，消息消费者从队列中接收消息。

发布/订阅（Pub/Sub）模式: 发布订阅模式定义了如何向一个内容节点发布和订阅消息，这个内容节点称为主题（Topic），主题可以认为是消息传递的中介，消息发布者将消息发布到某个主题，而消息订阅者从主题中订阅消息。主题使得消息的订阅者和发布者互相保持独立，不需要进行接触即可保证消息的传递，发布/订阅模式在消息的一对多广播时采用。

Kafka 同时支持两种消息投递模式，而这正是得益于消费者与消费组模型的契合：

- 如果所有的消费者都隶属于同一个消费组，那么所有的消息都会被均衡地投递给每一个消费者，即每条消息只会被一个消费者处理，这就相当于点对点模式的应用。
- 如果所有的消费者都隶属于不同的消费组，那么所有的消息都会被广播给所有的消费者，即每条消息会被所有的消费者处理，这就相当于发布/订阅模式的应用

消费组是一个逻辑上的概念，它将旗下的消费者归为一类，每一个消费者只隶属于一个消费组。每一个消费组都会有一个固定的名称，消费者在进行消费前需要指定其所属消费组的名称，这个可以通过消费者客户端参数 `group.id` 来配置，默认值为空字符串。消费者并非逻辑上的概念，它是实际的应用实例，它可以是一个线程，也可以是一个进程。同一个消费组内的消费者既可以部署在同一台机器上，也可以部署在不同的机器上。

创建一个Kafka消费者

- 以下代码段显示了如何创建KafkaConsumer：

```
``$xslt Properties props = new Properties(); props.put("bootstrap.servers", "broker1:9092,broker2:9092"); props.put("group.id", "CountryCounter"); props.put("key.deserializer", "org.apache.kafka.common.serialization.StringDeserializer"); props.put("value.deserializer", "org.apache.kafka.common.serialization.StringDeserializer");

KafkaConsumer consumer = new KafkaConsumer(props);
```

```
1 * 订阅主题
2
3
4 ...
5 consumer.subscribe(Collections.singletonList("customerCountries"));
```

- 要订阅所有test主题，我们可以：

```
consumer.subscribe(Pattern.compile("test.*"));
```

- 轮询循环

消费者API的核心是一个简单的循环，用于轮询服务器以获取更多数据。一旦用户订阅了主题，轮询循环便会处理协调，分区重新平衡，心跳和数据获取的所有详细信息，从而为开发人员提供了一个干净的API，该API仅从分配的分区中返回可用数据。消费者的主体如下所示

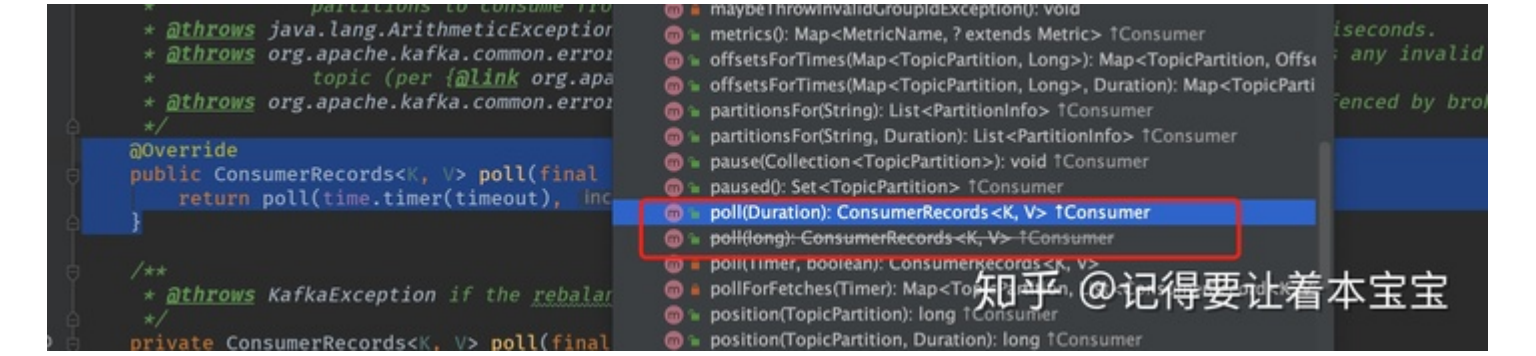
```
1 try {
2     while (true) { 1
3         ConsumerRecords<String, String> records = consumer.poll(100); 2
4         for (ConsumerRecord<String, String> record : records) 3
5             {
6                 log.debug("topic = %s, partition = %d, offset = %d,"
7                     customer = %s, country = %sn",
8                     record.topic(), record.partition(), record.offset(),
9                     record.key(), record.value());
10
11                 int updatedCount = 1;
12                 if (custCountryMap.containsKey(record.value())) {
13                     updatedCount = custCountryMap.get(record.value()) + 1;
14                 }
15                 custCountryMap.put(record.value(), updatedCount)
16
17                 JSONObject json = new JSONObject(custCountryMap);
18                 System.out.println(json.toString(4)) 4
19             }
20     }
21 } finally {
22     consumer.close(); 5
23 }
24
25
```




```
2
3
4
5 public class StringDeserializer implements Deserializer<String> {
6     private String encoding = "UTF8";
7
8     @Override
9     public void configure(Map<String, ?> configs, boolean isKey) {
10         String propertyName = isKey ? "key.deserializer.encoding" : "value.deserializ
11         Object encodingValue = configs.get(propertyName);
12         if (encodingValue == null)
13             encodingValue = configs.get("deserializer.encoding");
14         if (encodingValue instanceof String)
15             encoding = (String) encodingValue;
16     }
17
18     @Override
19     public String deserialize(String topic, byte[] data) {
20         try {
21             if (data == null)
22                 return null;
23             else
24                 return new String(data, encoding);
25         } catch (UnsupportedEncodingException e) {
26             throw new SerializationException("Error when deserializing byte[] to stri
27         }
28     }
29 }
```

- 消息消费

Kafka中的消费是基于拉模式的。消息的消费一般有两种模式：推模式和拉模式。推模式是服务端主动将消息推送给消费者，而拉模式是消费者主动向服务端发起请求来拉取消息。从轮询循环代码清单中可以看出，Kafka中的消息消费是一个不断轮询的过程，消费者所要做的就是重复地调用poll（）方法，而poll（）方法返回的是所订阅的主题（分区）上的一组消息。对于poll（）方法而言，如果某些分区中没有可供消费的消息，那么此分区对应的消息拉取的结果就为空；如果订阅的所有分区中都没有可供消费的消息，那么poll（）方法返回为空的集合。



poll（long）方法中timeou的时间单位固定为毫秒，而poll（Duration）方法可以根据Duration中的ofMillis（）、ofSeconds（）、ofMinutes（）、ofHours（）等多种不同的方法指定不同的时间单位，灵活性更强。并且poll（long）方法也已经被标注为@Deprecated，虽然目前还可以使用，如果条件允许的话，还是推荐使用poll（Duration）的方式。

我们在消费消息的时候可以直接对 ConsumerRecord 中感兴趣的字段进行具体的业务逻辑处理。

poll（）方法的返回值类型是 ConsumerRecords，它用来表示一次拉取操作所获得的消息集，内部包含了若干 ConsumerRecord，它提供了一个iterator（）方法来循环遍历消息集内部的消息，iterator（）方法的定义如下：

```
1 @Override
2 public Iterator<ConsumerRecord<K, V>> iterator() {
3     return new ConcatenatedIterable<>(records.values()).iterator();
4 }
```

在 ConsumerRecords 类中还提供了几个方法来方便开发人员对消息集进行处理：count（）方法用来计算出消息集中的消息个数，返回类型是int；isEmpty（）方法用来判断消息集是否为空，返回类型是boolean；empty（）方法用来获取一个空的集合，返回类型是ConsumerRecord < K, V >。

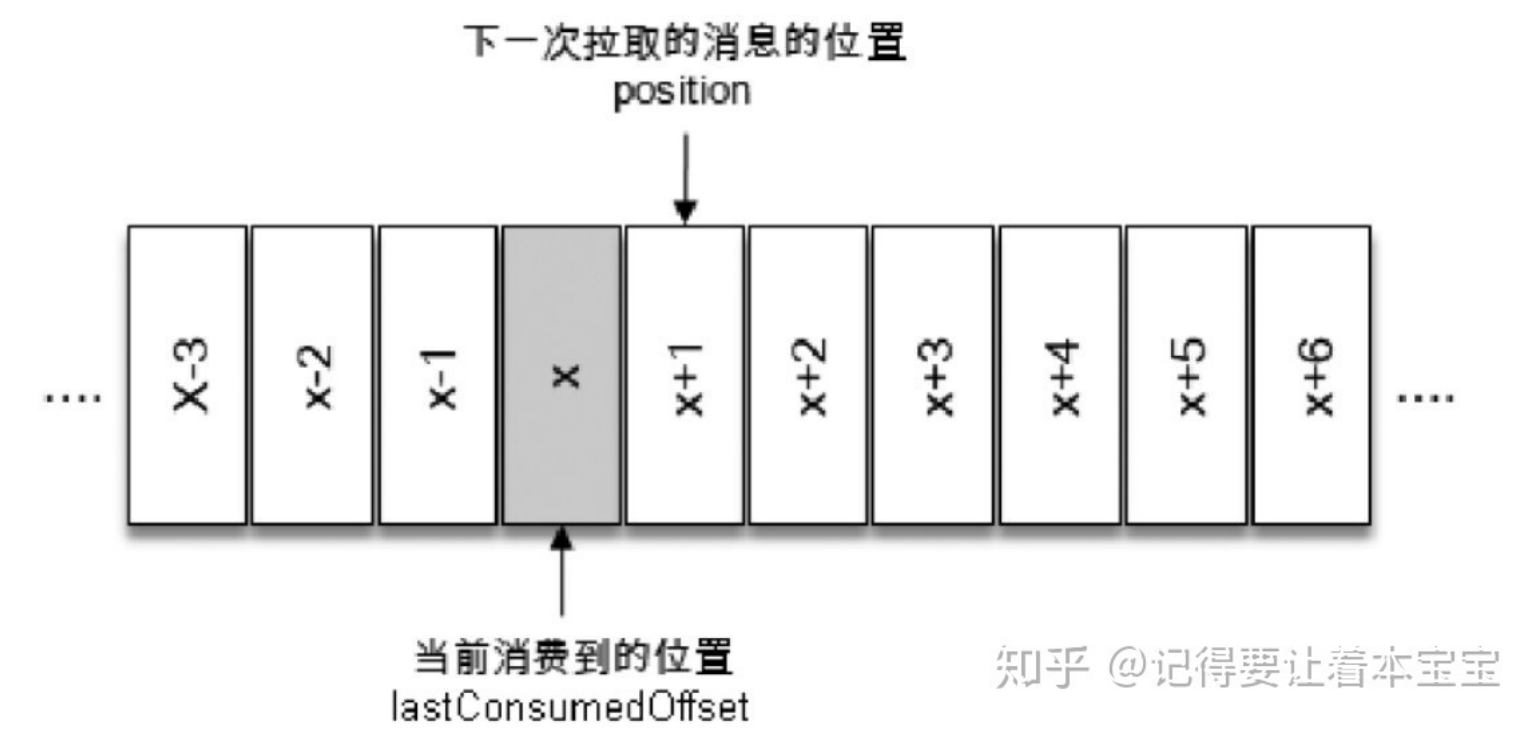
到目前为止，可以简单地认为poll（）方法只是拉取一下消息而已，但就其内部逻辑而言并不简单，它涉及消费位移、消费者协调器、组协调器、消费者的选举、分区分配的分发、再均衡的逻辑、心跳等内容

- 位移提交

对于Kafka中的分区而言，它的每条消息都有唯一的offset，用来表示消息在分区中对应的位置。对于消费者而言，它也有一个offset的概念，消费者使用offset来表示消费到分区中某个消息所在的位置。单词“offset”可以翻译为“偏移量”，也可以翻译为“位移”，很多同学可能并没有过多地在意这一点：在很多中文资料中都会交叉使用“偏移量”和“位移”这两个词，并没有很严谨地进行区分。

我对offset做了一些区分：对于消息在分区中的位置，我们将offset称为“偏移量”；对于消费者消费到的位置，将offset称为“位移”，有时候也会更明确地称之为“消费位移”。做这一区分的目的是让读者在遇到offset的时候可以很容易甄别出是在讲分区存储层面的内容，还是在讲消费层面的内容

在每次调用poll（）方法时，它返回的是还没有被消费过的消息集（当然这个前提是消息已经存储在Kafka中了，并且暂不考虑异常情况的发生），在旧消费者客户端中，消费位移是存储在ZooKeeper中的。而在新消费者客户端中，消费位移存储在Kafka内部的主题__consumer_offsets中。这里把将消费位移存储起来（持久化）的动作称为“提交”，消费者在消费完消息之后需要执行消费位移的提交。



- 指定位移消费

正是有了消费位移的持久化，才使消费者在关闭、崩溃或者在遇到再均衡的时候，可以让接替的消费者能够根据存储的消费位移继续进行消费，可是有一个问题则是__consumer_offsets位移信息过期而被删除后，它也没有可以查找的消费位移，这个时候就会根据消费者客户端参数auto.offset.reset的配置来决定从何处开始进行消费

除了查找不到消费位移，位移越界也会触发 auto.offset.reset 参数的执行，然而有些时候，我们需要一种更细粒度的掌控，让我们从特定的位移处开始拉取消息，哎！这个时候 KafkaConsumer 中的 seek（）方法正好提供了这个功能，让我们得以追前消费或回溯消费。seek（）方法的具体定义如下：

public void seek(TopicPartition partition, long offset) {} seek（）方法为我们提供了从特定位置读取消息的能力，我们可以通过这个方法来向前跳过若干消息，也可以通过这个方法来向后回溯若干消息，这样为消息的消费提供了很大的灵活性

原创不易，如果觉得有点用的话，请毫不留情点个赞，转发一下，这将是持续输出优质文章的最强动力。

工程师小C的小店

我也想开通小店

缓存中间件Redis技术入门与应用场景实战（SpringBoot2.x + 抢红包系统设计...
讲师：修罗debug
好评：100% 销售量：1


¥ 99

更多

亿速云服务器免备案CN2高速直连
亿速云，CN2高速稳定独享带宽目前还有优惠活动低至29元每月，香港服务器CN2高速直连低延时

广告 亿速云


Kafka系列（四）Kafka消费者：从Kafka中读取数据 u012501054的博客 3万+
本系列文章为对《Kafka：The Definitive Guide》的学习整理，希望能够帮助到大家应用从Kafka中读取数据需要使用Kafka...



优质评论可以帮助作者获得更高权重

抢沙发

评论



生态体-创始人|**博主**：这篇文章对你有帮助吗？作为一名程序工程师，在评论区留下你的困惑或你的见解，大家一起来交流吧！

相关推荐

- kafka 消费者组 消费者订阅不同主题_kafka消费者组与重...

3-25

消费组组(Consumer group)可以说是kafka很有亮点的一个设计。传统的消息引擎处理模型主要有两种,队列模型,和发布-订...
- kafka 消费者组 消费者订阅不同主题_Kafka 为什么使用...

3-30

如果有了消费者组,组内的成员就可以分担这10个分区的压力,提高消费性能。 2.2 消费模式灵活 假设有4个消费者订阅一个...
- kafka 重新分配节点_详细解析kafka之_kafka消费者组与重平衡机制

weixin_42350506的博客

65

消费组组(Consumer group)可以说是kafka很有亮点的一个设计。传统的消息引擎处理模型主要有两种，队列模型，和发布...
- kafka消费者多个topic配置

澹若水的专栏

1万+

我们的项目是springboot集成kafka的，配置文件在application.yml里如下： #kafka配置 kafka: consumer: bootstrap: maxPol...
- kafka 消费者组 消费者订阅不同主题_详细解析kafka之 k...

4-6

消费组组(Consumer group)可以说是kafka很有亮点的一个设计。传统的消息引擎处理模型主要有两种,队列模型,和发布-订...
- 深入了解kafka系列-消费者_小技术君

4-3

深入了解kafka系列-消费者 前言 与生产者对应的是消费者,应用程序可以通过KafkaConsumer来订阅主题,并从订阅的主题中...
- Kafka消费者——从 Kafka读取数据

weixin_33693070的博客

647

应用程序使用 KafkaConsumer向 Kafka 订阅主题，并从订阅的主题上接收消息。从 Kafka 读取数据不同于从其他消息系...
- Kafka消费组(consumer group)

weixin_33713503的博客

2174

一直以来都想写一点关于kafka consumer的东西，特别是关于新版consumer的中文资料很少。最近Kafka社区邮件组已经在...
- kafka多个消费者消费一个topic_你真的了解 Kafka 消费...

3-27

消费者和消费组 我们先来说什么是消费者和消费组。消费者是从订阅的主题中消费消息,一群消费者组成的一个消费组。...
- @kafkalistener消费不到消息_深入了解kafka系列-消费者

1-19

前言与生产者对应的是消费者,应用程序可以通过KafkaConsumer来订阅主题,并从订阅的主题中拉取消息。不过在使用Kafk...
- kafka不同组消费同一主题topic生产者配置

LMC技术人生专栏

2453

因数涉及到数据的原子性，必须多个分组消费同一消费主题，写入数据库后，可以自行回滚数据，重新消费，不影响其...
- 关于kafka更改消费者对应分组下的offset值

阿正的博客

1574

kafka的offset保存位置分为两种情况 0.9.0.0版本之前默认保存在zookeeper当中 0.9.0.0版本之后保存在broker对应的topic当...
- kafka 消费者消息确认_超详细讲解Kafka消费者,让你看完...

4-2

深入学习Kafka数据消费大致流程 如何创建并使用Kafka消费者 Kafka消费者常用配置 1.1 概念入门 1.1.1 消费者和消费组 K...
- Kafka消费者群组消费不到消息解决办法

Wengzhengcun的博客

2417

测试环境发Kafka消息，不能消费，我让测试重启一下brokekr，再发消息，发现能正常消费了。
- 聊聊Kafka(三)Kafka消费者与消费组

403

Kafka消费者与消费组简介消费者概念入门消费者、消费组心跳机制消息接收必要参数配置订阅反序列化位移提交消费者位...
- Kafka下的生产消费者模式与订阅发布模式

茅庐

4万+

生产消费者模式与订阅发布模式是使用消息中间件时常用的两种模式，用于功能解耦和分布式系统间的消息通信。本文在总...
- Kafka消费者 之 如何订阅主题或分区

CREATE_17的博客

242

每一个成功人士的背后，必定曾经做出过勇敢而又孤独的决定。放弃不难，但坚持很酷~一、消费者配置在创建真正消费者...
- 用Kafka自带脚本实现Kafka消费情况实时监控

ChengYanan的博客

357

查看consumer组内消费的offset 在\${KAFKA_HOME}/bin下 ./kafka-run-class.sh kafka.tools.ConsumerOffsetChecker --zook...
- kafka 发布订阅_在Kafka中发布订阅模型

最佳 Java 编程

782

kafka 发布订阅 这是第四个柱中的一系列关于同步客户端集成与异步系统（1，2，3）。在这里，我们将尝试了解Kafka...
- kafka 查看消费者组

鸭梨的博客

1590

用kafka-consumer-groups.sh 命令可以查看消费者组,其中console-consumer-89657是消费的时候如果不设置消费者组自动...
- Kafka学习笔记: Kafka 的消费者组

张伯毅的专栏

478

Consumer Group 是 Kafka 提供的可扩展且具有容错性的消费者机制。既然是一个组，那么组内必然可以有多个消费者或...
- Kafka中的消费者组（Consumer Group）

象在舞的技术专栏

3万+

1. 消费者组（Consumer Group） 消费者组是Kafka实现单播和广播两种消息模型的手段。同一个topic，每个消费者组...
- kafka | 消费者组到底是什么？

sinat_27143551的博客

863

消费者组，即 Consumer Group，应该算是 Kafka 比较有亮点的设计了。那么何谓 Consumer Group 呢？用一句话概括就...
- kafka为什么使用消费者组

疯一样的女子

487

1. 消费者组的特点 这是 kafka 集群的典型部署模式。消费组保证了：一个分区只可以被消费组中的一个消费者所消费 一...

©2020 CSDN 皮肤主题: 数字20 设计师:CSDN官方博客 返回首页

关于我们 招贤纳士 广告服务 开发助手 400-660-0108 kefu@csdn.net 在线客服 工作时间 8:30-22:00

公安备案号11010502030143 京ICP备19004658号 京网文〔2020〕1039-165号 经营性网站备案信息 北京互联网违法和不良信息举报中心 网络110报警服务 中国互联网举报中心 家长监护 Chrome商店下载 ©1999-2021北京创新乐知网络技术有限公司 版权与免责声明 版权申诉

