

## 19 | CommitFailedException异常怎么处理？

2019-07-16 胡夕



你好，我是胡夕。今天我来跟你聊聊CommitFailedException异常的处理。

说起这个异常，我相信用过Kafka Java Consumer客户端API的你一定不会感到陌生。所谓CommitFailedException，顾名思义就是Consumer客户端在提交位移时出现了错误或异常，而且还是那种不可恢复的严重异常。如果异常是可恢复的瞬时错误，提交位移的API自己就能规避它们了，因为很多提交位移的API方法是支持自动错误重试的，比如我们在上一期中提到的commitSync方法。

每次和CommitFailedException一起出现的，还有一段非常著名的注释。为什么说它很“著名”呢？第一，我想不出在近50万行的Kafka源代码中，还有哪个异常类能有这种待遇，可以享有这么大段的注释，来阐述其异常的含义；第二，纵然有这么长的文字解释，却依然有很多人对该异常想表达的含义感到困惑。

现在，我们一起领略下这段文字的风采，看看社区对这个异常的最新解释：

Commit cannot be completed since the group has already rebalanced and assigned the partitions to another member. This means that the time between subsequent calls to poll() was longer than the configured max.poll.interval.ms, which typically implies that the poll loop is spending too much time message processing. You can address this either by increasing max.poll.interval.ms or by reducing the maximum size of batches returned in poll() with max.poll.records.

这段话前半部分的意思是，本次提交位移失败了，原因是消费者组已经开启了Rebalance过程，并且将要提交位移的分区分配给了另一个消费者实例。出现这个情况的原因是，你的消费者实例连续两次调用poll方法的时间间隔超过了期望的max.poll.interval.ms参数值。这通常表明，你的消费者实例花费了太长的时间进行消息处理，耽误了调用poll方法。

在后半部分，社区给出了两个相应的解决办法（即橙色字部分）：

1. 增加期望的时间间隔max.poll.interval.ms参数值。
2. 减少poll方法一次性返回的消息数量，即减少max.poll.records参数值。

在详细讨论这段文字之前，我还想提一句，实际上这段文字总共有3个版本，除了上面的这个最新版本，还有2个版本，它们分别是：

Commit cannot be completed since the group has already rebalanced and assigned the partitions to another member. This means that the time between subsequent calls to poll() was longer than the configured session.timeout.ms, which typically implies that the poll loop is spending too much time message processing. You can address this either by increasing the session timeout or by reducing the maximum size of batches returned in poll() with max.poll.records.

Commit cannot be completed since the group has already rebalanced and assigned the partitions to another member. This means that the time between subsequent calls to poll() was longer than the configured max.poll.interval.ms, which typically implies that the poll loop is spending too much time message processing. You can address this either by increasing the session timeout or by reducing the maximum size of batches returned in poll() with max.poll.records.

这两个较早的版本和最新版相差不大，我就不详细解释了，具体的差异我用橙色标注了。我之所以列出这些版本，就是想让你在日后看到它们时能做到心中有数，知道它们说的是一个事情。

其实不论是哪段文字，它们都表征位移提交出现了异常。下面我们就来讨论下该异常是什么时候被抛出的。从源代码方面来说，CommitFailedException异常通常发生在手动提交位移时，即用户显式调用KafkaConsumer.commitSync()方法时。从使用场景来说，有两种典型的场景可能遭遇该异常。

## 场景一

我们先说说最常见的场景。当消息处理的总时间超过预设的max.poll.interval.ms参数值时，Kafka Consumer端会抛出CommitFailedException异常。这是该异常最“正宗”的登场方式。你只需要写一个Consumer程序，使用KafkaConsumer.subscribe方法随意订阅一个主题，之后设置Consumer端参数max.poll.interval.ms=5秒，最后在循环调用KafkaConsumer.poll方法之间，插入

`Thread.sleep(6000)`和手动提交位移，就可以成功复现这个异常了。在这里，我展示一下主要的代码逻辑。

```
...
Properties props = new Properties();
...
props.put("max.poll.interval.ms", 5000);
consumer.subscribe(Arrays.asList("test-topic"));

while (true) {
    ConsumerRecords<String, String> records =
    consumer.poll(Duration.ofSeconds(1));

    // 使用Thread.sleep模拟真实的消息处理逻辑
    Thread.sleep(6000L);

    consumer.commitSync();
}
```

如果要防止这种场景下抛出异常，你需要简化你的消息处理逻辑。具体来说有4种方法。

1. **缩短单条消息处理的时间**。比如，之前下游系统消费一条消息的时间是100毫秒，优化之后成功地下降到50毫秒，那么此时Consumer端的TPS就提升了一倍。
2. **增加Consumer端允许下游系统消费一批消息的最大时长**。这取决于Consumer端参数`max.poll.interval.ms`的值。在最新版的Kafka中，该参数的默认值是5分钟。如果你的消费逻辑不能简化，那么提高该参数值是一个不错的办法。值得一提的是，Kafka 0.10.1.0之前的版本是没有这个参数的，因此如果你依然在使用0.10.1.0之前的客户端API，那么你需要增加`session.timeout.ms`参数的值。不幸的是，`session.timeout.ms`参数还有其他的含义，因此增加该参数的值可能会有其他方面的“不良影响”，这也是社区在0.10.1.0版本引入`max.poll.interval.ms`参数，将这部分含义从`session.timeout.ms`中剥离出来的原因之一。
3. **减少下游系统一次性消费的消息总数**。这取决于Consumer端参数`max.poll.records`的值。当前该参数的默认值是500条，表明调用一次KafkaConsumer.poll方法，最多返回500条消息。可以说，该参数规定了单次poll方法能够返回的消息总数的上限。如果前两种方法对你都不适用的话，降低此参数值是避免CommitFailedException异常最简单的手段。
4. **下游系统使用多线程来加速消费**。这应该算是“最高级”同时也是最难实现的解决办法了。具体的思路就是，让下游系统手动创建多个消费线程处理poll方法返回的一批消息。之前你使用Kafka Consumer消费数据更多是单线程的，所以当消费速度无法匹及Kafka Consumer消息返回的速度时，它就会抛出CommitFailedException异常。如果是多线程，你就可以灵

活地控制线程数量，随时调整消费承载能力，再配以目前多核的硬件条件，该方法可谓是防止CommitFailedException最高档的解决之道。事实上，很多主流的大数据流处理框架使用的都是这个方法，比如Apache Flink在集成Kafka时，就是创建了多个KafkaConsumerThread线程，自行处理多线程间的数据消费。不过，凡事有利就有弊，这个方法实现起来并不容易，特别是在多个线程间如何处理位移提交这个问题上，更是极容易出错。在专栏后面的内容中，我将着重和你讨论一下多线程消费的实现方案。

综合以上这4个处理方法，我个人推荐你首先尝试采用方法1来预防此异常的发生。优化下游系统的消费逻辑是百利而无一害的法子，不像方法2、3那样涉及到Kafka Consumer端TPS与消费延时（Latency）的权衡。如果方法1实现起来有难度，那么你可以按照下面的法则来实践方法2、3。

首先，你需要弄清楚你的下游系统消费每条消息的平均延时是多少。比如你的消费逻辑是从Kafka获取到消息后写入到下游的MongoDB中，假设访问MongoDB的平均延时不超过2秒，那么你可以认为消息处理需要花费2秒的时间。如果按照max.poll.records等于500来计算，一批消息的总消费时长大约是1000秒，因此你的Consumer端的max.poll.interval.ms参数值就不能低于1000秒。如果你使用默认配置，那默认值5分钟显然是不够的，你将有很大概率遭遇CommitFailedException异常。将max.poll.interval.ms增加到1000秒以上的做法就属于上面的第2种方法。

除了调整max.poll.interval.ms之外，你还可以选择调整max.poll.records值，减少每次poll方法返回的消息数。还拿刚才的例子来说，你可以设置max.poll.records值为150，甚至更少，这样每批消息的总消费时长不会超过300秒（ $150 \times 2 = 300$ ），即max.poll.interval.ms的默认值5分钟。这种减少max.poll.records值的做法就属于上面提到的方法3。

## 场景二

Okay，现在我们已经说完了关于CommitFailedException异常的经典发生场景以及应对办法。从理论上讲，关于该异常你了解到这个程度，已经足以帮助你应对应用开发过程中由该异常带来的“坑”了。但其实，该异常还有一个不太为人所知的出现场景。了解这个冷门场景，可以帮助你拓宽Kafka Consumer的知识面，也能提前预防一些古怪的问题。下面我们就来说说这个场景。

之前我们花了很多时间学习Kafka的消费者，不过大都集中在消费者组上，即所谓的Consumer Group。其实，Kafka Java Consumer端还提供了一个名为Standalone Consumer的独立消费者。它没有消费者组的概念，每个消费者实例都是独立工作的，彼此之间毫无联系。不过，你需要注意的是，独立消费者的位移提交机制和消费者组是一样的，因此独立消费者的位移提交也必须遵守之前说的那些规定，比如独立消费者也要指定group.id参数才能提交位移。你可能会觉得奇怪，既然是独立消费者，为什么还要指定group.id呢？没办法，谁让社区就是这么设计的呢？总之，消费者组和独立消费者在使用之前都要指定group.id。

现在问题来了，如果你的应用中同时出现了设置相同`group.id`值的消费者组程序和独立消费者程序，那么当独立消费者程序手动提交位移时，**Kafka**就会立即抛出**CommitFailedException**异常，因为**Kafka**无法识别这个具有相同`group.id`的消费者实例，于是就向它返回一个错误，表明它不是消费者组内合法的成员。

虽然说这个场景很冷门，但也并非完全不会遇到。在一个大型公司中，特别是那些将**Kafka**作为全公司级消息引擎系统的公司中，每个部门或团队都可能有自己的消费者应用，谁能保证各自的**Consumer**程序配置的`group.id`没有重复呢？一旦出现不凑巧的重复，发生了上面提到的这种场景，你使用之前提到的哪种方法都不能规避该异常。令人沮丧的是，无论是刚才哪个版本的异常说明，都完全没有提及这个场景，因此，如果是这个原因引发的**CommitFailedException**异常，前面的4种方法全部都是无效的。

更为尴尬的是，无论是社区官网，还是网上的文章，都没有提到过这种使用场景。我个人认为，这应该算是**Kafka**的一个bug。比起返回**CommitFailedException**异常只是表明提交位移失败，更好的做法应该是，在**Consumer**端应用程序的某个地方，能够以日志或其他方式友善地提示你错误的原因，这样你才能正确处理甚至是预防该异常。

## 小结

总结一下，今天我们详细讨论了**Kafka Consumer**端经常碰到的**CommitFailedException**异常。我们从它的含义说起，再到它出现的时机和场景，以及每种场景下的应对之道。当然，我也留了个悬念，在专栏后面的内容中，我会详细说说多线程消费的实现方式。希望通过今天的分享，你能清晰地掌握**CommitFailedException**异常发生的方方面面，从而能在今后更有效地应对此异常。

## 预防CommitFailedException异常的4种方法

- 缩短单条消息处理的时间。
- 增加Consumer端允许下游系统消费一批消息的最大时长。
- 减少下游系统一次性消费的消息总数。
- 下游系统使用多线程来加速消费。



### 开放讨论

请比较一下今天我们提到的预防该异常的4种方法，并说说你对它们的理解。

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。





小生向北

4

`max.poll.interval.ms`是指两次`poll()`的最大间隔时间，`kafka`消费者以轮询的方式来拉取消息，并且一次拉取批量的消息（默认500条），而批量的大小是通过`max.poll.records`来控制的。两次`poll()`的实际时间取决于 单条消息的处理时间\*一次拉取的消息量（500），当超过`max.poll.interval.ms`配置的时间`Kafka server`认为`kafka consumer`掉线了，于是就执行分区再均衡将这个`consumer`踢出消费者组。但是`consumer`又不知道服务端把自己给踢出了，下次在执行`poll()`拉取消息的时候（在`poll()`拉取消息之前有个自动提交`offset`的操作），就会触发该问题。可见第2,3种方案是通过调整`Kafka consumer`的配置参数来缩短业务总的处理时间或者增加服务端判断时长，比较容易实现；第1种就跟业务有关了，比较难搞，有些业务可能就是要这么长的时间，很难再缩短；第4种方案就更复杂了，要把同步消息转换成异步，交给其它线程来处理，这时需要把`auto.commit.enable=false`，手动提交`offset`，并且`consumer`是线程不安全的，异步线程何时处理完，何时该提交，在哪提交，也是应用需要考虑的问题！希望胡老师针对第4种方案重点探讨一下！

2019-07-17



ban

2

老师，1、请问`Standalone Consumer` 的独立消费者一般什么情况会用到  
2、`Standalone Consumer` 的独立消费者 使用跟普通消费者组有什么区别的。

2019-07-16

作者回复

1. 很多流处理框架的`Kafka connector`都没有使用`consumer group`，而是直接使用`standalone consumer`，因为`group`机制不好把控
2. `standalone consumer`没有`rebalance`，也没有`group`提供的负载均衡，你需要自己实现。其他方面（比如位移提交）和`group`没有太大的不同

2019-07-17



ban

1

老师，我想问下`max.poll.interval.ms`两者`session.timeout.ms`有什么联系，可以说0.10.1.0 之前的客户端 API，相当于`session.timeout.ms`代替了`max.poll.interval.ms`吗？  
比如说`session.timeout.ms`是5秒，如果消息处理超过5秒，也算是超时吗？

2019-07-16

作者回复

嗯，我更愿意说是`max.poll.interval.ms`承担了`session.timeout.ms`的部分功能。在没有`max.poll.interval.ms`和单独的心跳线程之前，如果`session.timeout.ms = 5s`，消息处理超过了5s，那么`consumer`就算是超时

2019-07-17



windcaller

0



To use this mode, instead of subscribing to the topic using `subscribe`, you just call `assign(Collection)` with the full list of partitions that you want to consume.

```
String topic = "foo";
TopicPartition partition0 = new TopicPartition(topic, 0);
TopicPartition partition1 = new TopicPartition(topic, 1);
consumer.assign(Arrays.asList(partition0, partition1));
```

Once assigned, you can call `poll` in a loop, just as in the preceding examples to consume records. The group that the consumer specifies is still used for committing offsets, but now the set of partitions will only change with another call to `assign`. Manual partition assignment does not use group coordination, so consumer failures will not cause assigned partitions to be rebalanced. Each consumer acts independently even if it shares a groupid with another consumer. To avoid offset commit conflicts, you should usually ensure that the groupid is unique for each consumer instance.

老师 standalone mode 是上面这段内容吗？

2019-07-31

作者回复

是的。使用`assign`的consumer就是standalone consumer

2019-07-31



有时也，命也，运也，如之奈何？

0

老师kafka死信该怎么去实现的？

2.0之后增加了如下配置：

`errors.tolerance = all`

`errors.deadletterqueue.topic.name = ""`？

2019-07-30

作者回复

还不够，你需要使用Kafka Connect组件才能实现。见：<https://www.confluent.io/blog/kafka-connect-deep-dive-error-handling-dead-letter-queues>

2019-07-30



windcaller

0

我未在kafka官网、stackoverflow、google 找到任何关于 standalone kafka consumer的例子，还望老师给个链接学习学习

2019-07-30

作者回复

Standalone consumer的提法并未出现在官方文档中，你可以在javadoc中看到一些：<https://kafka.apache.org/23/javadoc/org/apache/kafka/clients/consumer/KafkaConsumer.html#manualassignment>

2019-07-30



老角

0





7 23

0

胡老师，手动提交时，当前后两次poll时间超过期望的max.poll.interval.ms时，会触发Rebalance。那么假如是自动提交时，会触发Rebalance吗？

假如，手动提交场景，consumer消费端处理业务时间过长（特殊case导致的），发生了Rebalance，那么该consumer实例被踢出了，那么它永远‘死掉’了吗，还是会再次通过heartbeat检测让它复活？

2019-07-26

作者回复

不会

2019-07-27



曹伟雄

0

老师，请教2个问题，谢谢！

我想问下0.10.1.0之后的session.timeout.ms还有什么作用呢？

standalone consumer和group consumer在配置上如何区分？

2019-07-19

作者回复

用于侦测会话超时。standalone consumer和group的区分体现在API上

2019-07-19



郭刚

0

[2019-07-17 18:53:36,230] ERROR [ReplicaManager broker=2] Error processing append operation on partition \_\_consumer\_offsets-49 (kafka.server.ReplicaManager)

org.apache.kafka.common.errors.NotEnoughReplicasException: The size of the current ISR Set(2) is insufficient to satisfy the min.isr requirement of 2 for partition \_\_consumer\_offsets-49

把min.insync.replicas = 2改成1，消费者就可以运行了，这种flower失效的情况怎么处理呢？

2019-07-17



郭刚

0

老师，kafka报错，哪个论坛比较火？

2019-07-17



juan

0

请问如何计算单条消息处理的时间，比如收到一条消息之后要调用A,B,C,D,E 五个方法处理，其中处理完B,E之后将结果发给别的kafka，处理时间是处理完A,B时间还是处理完A,B,C,D,E 的总时间？

2019-07-17

作者回复

这取决于你对如何才算处理完一条消息的定义。另外从Kafka中拿到消息后剩下的事情就完全由你负责了，因此如何计算处理时间应该是你说了算的：)

2019-07-17



电光火石

0

我在线上遇到这个问题，想问一下老师：在新的consumer加入，发生repartition的时候，是

否也会抱这个错，谢谢了！

2019-07-16



曾轼麟

👍 0

老师我这边遇到了一个奇怪的情况，kafka生产者发送消息能创建topic,但是消息怎么都发不上去broker。并且在kafka-logs底下有一个和刚刚那条消息key值一样的文件夹，

并且打印出如下的日志：

```
[2019-07-16 22:34:16,380] INFO [Log partition=23bb7ffd-4aa5-42e2-9d84-c90f4566c15b-2, dir=/tmp/kafka-logs] Loading producer state till offset 0 with message format version 2 (kafka.log.Log)
```

```
[2019-07-16 22:34:16,381] INFO [Log partition=23bb7ffd-4aa5-42e2-9d84-c90f4566c15b-2, dir=/tmp/kafka-logs] Completed load of log with 1 segments, log start offset 0 and log end offset 0 in 3 ms (kafka.log.Log)
```

```
[2019-07-16 22:34:16,383] INFO Created log for partition 23bb7ffd-4aa5-42e2-9d84-c90f4566c15b-2 in /tmp/kafka-logs with properties {compression.type -> producer, message.format.version -> 2.0-IV1, file.delete.delay.ms -> 60000, max.message.bytes -> 1000012, min.compaction.lag.ms -> 0, message.timestamp.type -> CreateTime, message.downconversion.enable -> true, min.insync.replicas -> 1, segment.jitter.ms -> 0, preallocate -> false, min.cleanable.dirty.ratio -> 0.5, index.interval.bytes -> 40}
```

2019-07-16

作者回复

这都是info级别的log，没有看出有什么问题。最好确认去leader副本所在的broker去看日志

2019-07-17



永恒记忆

👍 0

老师好，我上一个问题的具体描述是这样的，今天讲CommitFailedException的例子是调用consumer.commitSync();手动提交offset，确实当消息处理的总时间超过预设的max.poll.interval.ms时会报这个异常，但是如果是自动提交offset的情况下，也就是把enable.auto.commit=true，然后删除consumer.commitSync();代码，其它代码不变，也是max.poll.interval.ms=5s，然后循环中sleep(6s)，发现不会报异常并且会一直重复消费，想问下这是什么原因呢？

2019-07-16

作者回复

嗯，是的。这个异常只是在手动提交时抛出的。

2019-07-17



Li Shunduo

👍 0

假如broker集群整个挂掉了，过段时间集群恢复后，consumer group会自动恢复消费吗？还是需要手动重启consumer机器？

2019-07-16

作者回复

consumer有重连机制

2019-07-16



永恒记忆

0

老师好，今天讲CommitFailedException的例子是调用consumer.commitSync();手动提交offset，确实当消息处理的总时间超过预设的max.poll.interval.ms时会报这个异常，但是如果是自动提交offset的情况下，不会报异常并且会一直重复消费，想问下这是为什么呢？

2019-07-16

作者回复

不知道你的代码是怎么写的。你是说循环调用poll然后返回相同的方法，是吗

2019-07-16



cricket1981

0

想进一步了解学习standalone consumer有什么资料推荐吗？

2019-07-16

作者回复

没有。你可以自己试试。其实用法与consumer group差不多

2019-07-16



cricket1981

0

"不幸的是，session.timeout.ms 参数还有其他的含义，因此增加该参数的值可能会有其他方面的“不良影响”，这也是社区在 0.10.1.0 版本引入 max.poll.interval.ms 参数，将这部分含义从 session.timeout.ms 中剥离出来的原因之一。"--->能细述一下不良影响吗？

2019-07-16

作者回复

之前版本中session.timeout.ms有多重含义，session过期时间、消息处理逻辑最大时间等

2019-07-16



永恒记忆

0

老师好，看你解释注释那段的时候提到消费者组已经开启了Rebalance过程，但是看后面的介绍，出现异常好像并不一定有Rebalance，是这样吗？

2019-07-16

作者回复

所以这段注释是有歧义或不准确的地方啊。

2019-07-16



德惠先生

0

希望老师可以更加具体的说说，rebalance的细节，比如某个consumer发生full gc的场景，它的partition是怎么被分配走的，重连之后提交会发生什么

2019-07-16

作者回复

假设full gc导致所有线程STW，从而心跳中断，导致被踢出group，Coordinator向其他存活consumer发送心跳response，通知它们开启新一轮rebalance。

2019-07-16