

## 05 | 分布式事务：如何保证多个系统间的数据是一致的？

李玥 · 后端存储实战课



你好，我是李玥。

上节课，我和你一起通过账户系统学习了数据库事务，事务很好地解决了交易类系统的数据一致性问题。

事务的原子性和持久性可以确保在一个事务内，更新多条数据，要么都成功，要么都失败。在一个系统内部，我们可以使用数据库事务来保证数据一致性。那如果一笔交易，涉及到跨多个系统、多个数据库的时候，用单一的数据库事务就没办法解决了。

在之前大系统的时代，普遍的做法是，在设计时尽量避免这种跨系统跨数据库的交易。

但是，现在的技术趋势是云原生和微服务，微服务它的理念是什么？大系统被打散成多个小的微服务，每个微服务独立部署并且拥有自己的数据库，大数据库也被打散成多个小的数据库。跨越微服务和数据库的交易就成为一种越来越普遍的情况。我们的业务系统微服务化之后，不可避免地要面对跨系统的数据一致性问题。

如何解决这种跨系统、跨数据库的数据一致性问题呢？你可能会脱口而出：分布式事务。但是，分布式事务可不像数据库事务那样，在开始和结尾分别加上 `begin` 和 `commit`，剩下的问题数据库都可以帮我们搞定。在分布式环境下，没有这么简单的事儿，为什么？

因为在分布式环境中，一个交易将会被分布到不同的系统中，多个微服务进程内执行计算，在多个数据库中执行数据更新操作，这个场景比数据库事务支持的单进程单数据库场景复杂太多了。所以，并没有什么分布式事务服务或者组件能在分布式环境下，提供接近数据库事务的数据一致性保证。

今天这节课我们就来说一下，如何用分布式事务的方法，来解决微服务系统中，我们实际面临的分布式数据一致性问题。

## 到底什么是分布式事务？

在学习分布式事务这个概念之前，我先跟你说一下为什么一定要搞懂概念。我们这门课程是一门实战课，一般来说，我们更关注的是如何解决实际问题，而不是理论和概念，所以你看，我们在讲解数据库事务的时候，讲的内容是如何用事务解决交易的问题，而没讲 MySQL 是如何实现 ACID 的。因为数据库已经把事务封装的非常好了，我们只需要掌握如何使用就可以很好地解决问题。

但分布式事务不是这样的，我刚刚说了，并没有一种分布式事务的服务或者组件，能帮我们很简单地就解决分布式系统下的数据一致性问题。我们在使用分布式事务时，更多的情况是，用分布式事务的理论来指导设计和开发，自行来解决数据一致性问题。也就是说，要解决分布式一致性问题，你必须掌握几种分布式事务的实现原理。

我们在讲解数据库事务时讲到了事务的 ACID 四个特性，我们知道即使是数据库事务，它考虑到性能的因素，大部分情况下不能也不需要百分之百地实现 ACID，所以才有了事务四种隔离级别。

理论上，分布式事务也是事务，也需要遵从 ACID 四个特性，但实际情况是，在分布式系统中，因为必须兼顾性能和高可用，所以是不可能完全满足 ACID 的。我们常用的几种分布式事务的实现方法，都是“残血版”的事务，而且相比数据库事务，更加的“残血”。

分布式事务的解决方案有很多，比如：2PC、3PC、TCC、Saga 和本地消息表等等。这些方法，它的强项和弱项都不一样，适用的场景也不一样，所以最好这些分布式事务你都能够掌握，这样才能在面临实际问题的时候选择合适的方法。这里面，2PC 和本地消息表这两种分布式事务的解决方案，比较贴近于我们日常开发的业务系统。

## 2PC：订单与优惠券的数据一致性问题

2PC 也叫二阶段提交，是一种常用的分布式事务实现方法。我们用订单和优惠券的例子来说明一下，如何用 2PC 来解决订单系统和促销系统的数据一致性问题。在我们购物下单时，如果使用了优惠券，订单系统和优惠券系统都要更新自己的数据，才能完成“在订单中使用优惠券”这个操作。

订单系统需要：

1. 在“订单优惠券表”中写入订单关联的优惠券数据；
2. 在“订单表”中写入订单数据。

订单系统内两个操作的一致性问题的直接使用数据库事务来解决。促销系统需要操作就比较简单，把刚刚使用的那张优惠券的状态更新成“已使用”就可以了。我们需要这两个系统的数据更新操作保持一致，要么都更新成功，要么都更新失败。

接下来我们来看 2PC 是怎么解决这个问题的。2PC 引入了一个事务协调者的角色，来协调订单系统和促销系统，协调者对客户端提供一个完整的“使用优惠券下单”的服务，在这个服务的内部，协调者再分别调用订单和促销的相应服务。

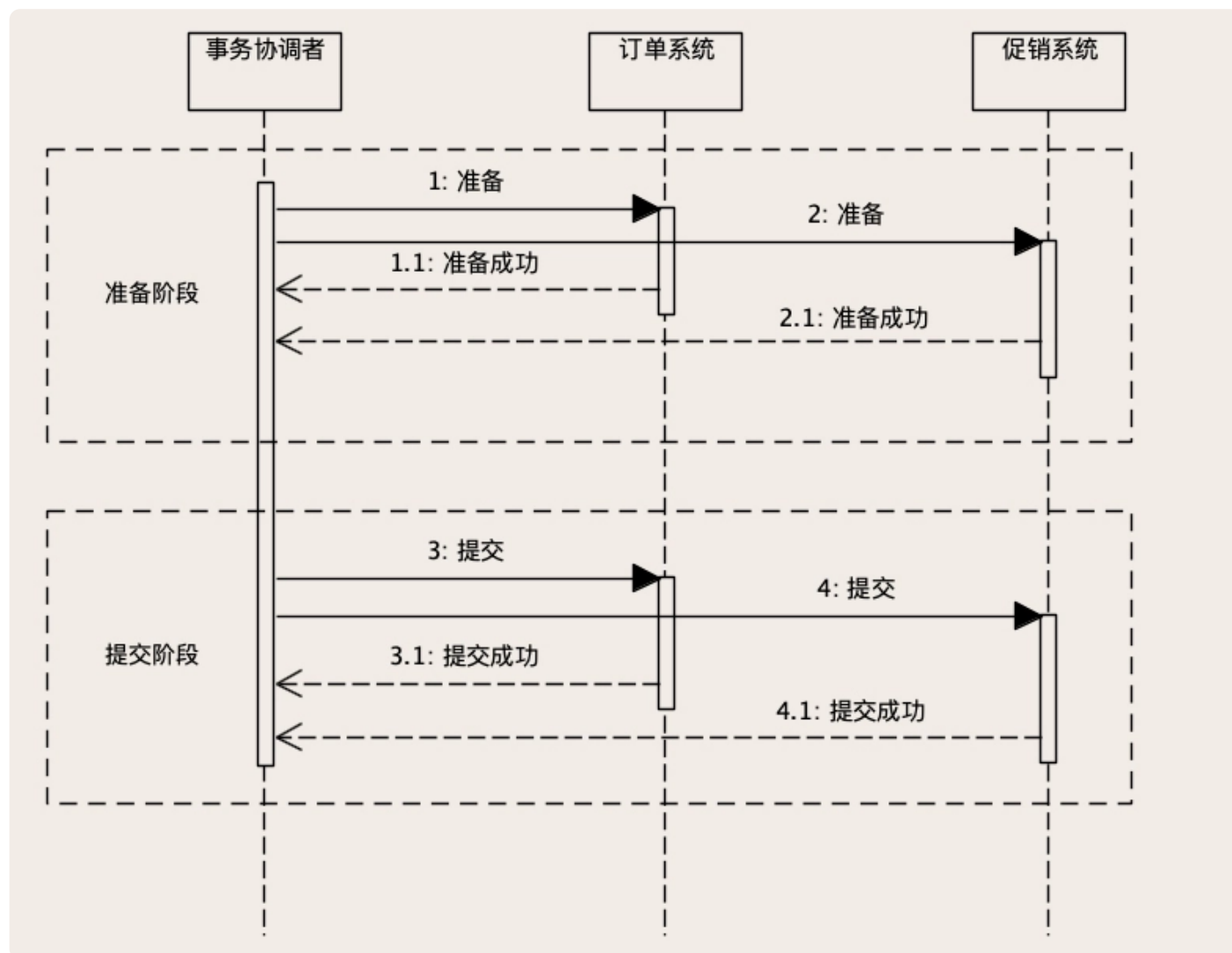
所谓的二阶段指的是准备阶段和提交阶段。在准备阶段，协调者分别给订单系统和促销系统发送“准备”命令，订单和促销系统收到准备命令之后，开始执行准备操作。准备阶段都需要做哪些事儿呢？你可以理解为，除了提交数据库事务以外的所有工作，都要在准备阶段完成。比如说订单系统在准备阶段需要完成：

1. 在订单库开启一个数据库事务；
2. 在“订单优惠券表”写入这条订单的优惠券记录；

### 3. 在“订单表”中写入订单数据。

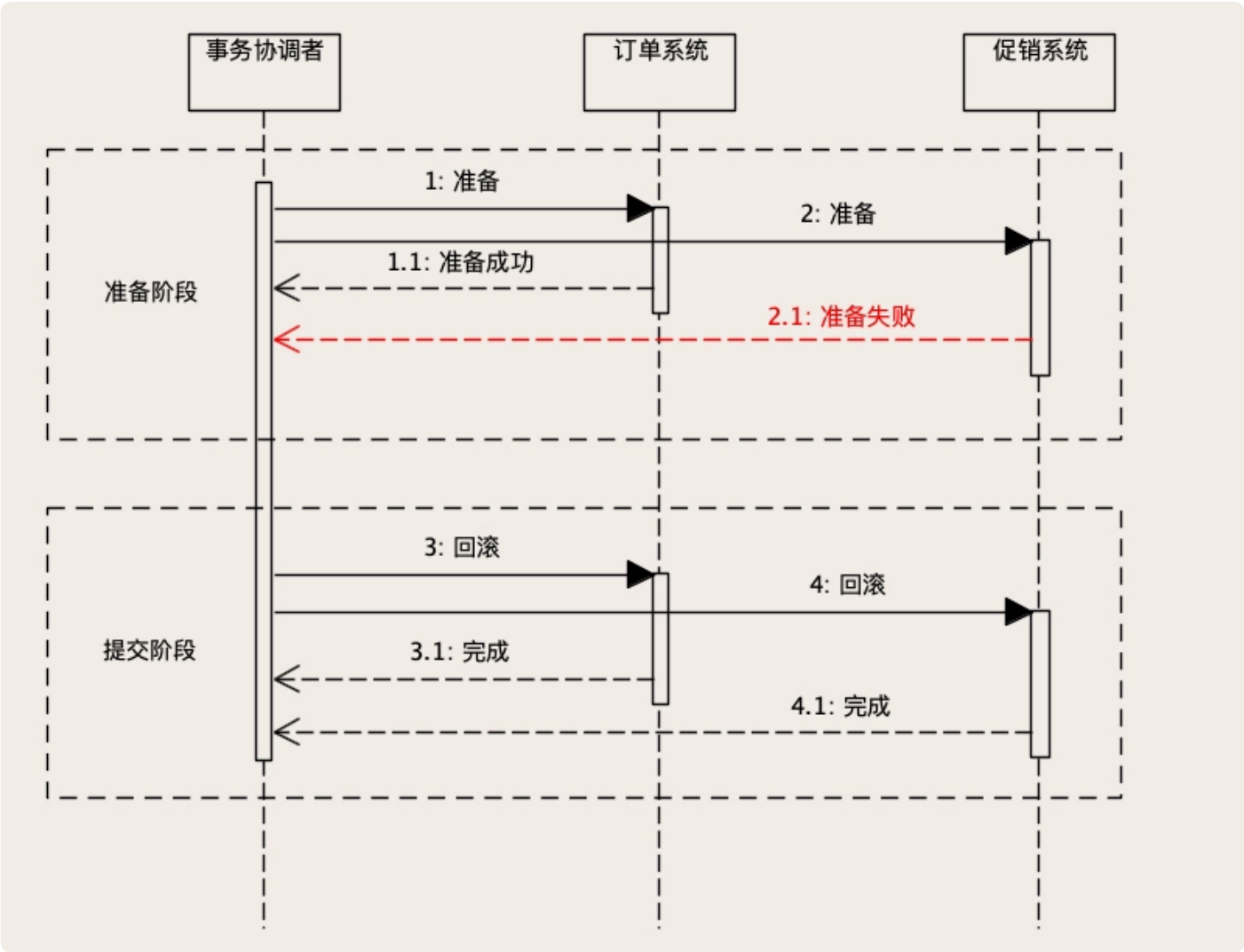
注意，到这里我们没有提交订单数据库的事务，最后给事务协调者返回“准备成功”。类似的，促销服务在准备阶段，需要在促销库开启一个数据库事务，更新优惠券状态，但是暂时不要提交这个数据库事务，给协调者返回“准备成功”。协调者在收到两个系统“准备成功”的响应之后，开始进入第二阶段。

等两个系统都准备好了之后，进入提交阶段。提交阶段就比较简单了，协调者再给这两个系统发送“提交”命令，每个系统提交自己的数据库事务，然后给协调者返回“提交成功”响应，协调者收到所有响应之后，给客户端返回成功响应，整个分布式事务就结束了。以下是这个过程的时序图：



这是正常情况，接下来才是重点：异常情况下怎么办？

我们还是分两个阶段来说明。在准备阶段，如果任何一步出现错误或者是超时，协调者就会给两个系统发送“回滚事务”请求。每个系统在收到请求之后，回滚自己的数据库事务，分布式事务执行失败，两个系统的数据库事务都回滚了，相关的所有数据回滚到分布式事务执行之前的状态，就像这个分布式事务没有执行一样。以下是异常情况的时序图：



如果准备阶段成功，进入提交阶段，这个时候就“只有华山一条路”，整个分布式事务**只能成功，不能失败**。

如果发生网络传输失败的情况，需要反复重试，直到提交成功为止。如果这个阶段发生宕机，包括两个数据库宕机或者订单服务、促销服务所在的节点宕机，还是有可能出现订单库完成了提交，但促销库因为宕机自动回滚，导致数据不一致的情况。但是，因为提交的过程非常简单，执行很快，出现这种情况的概率非常小，所以，从实用的角度来说，2PC 这种分布式事务的方法，实际的数据一致性还是非常好的。

在实现 2PC 的时候，没必要单独启动一个事务协调服务，这个协调服务的工作最好和订单服务或者优惠券服务放在同一个进程里面，这样做有两个好处：

参与分布式事务的进程更少，故障点也就更少，稳定性更好；

减少了一些远程调用，性能也更好一些。

2PC 是一种强一致的设计，它可以保证原子性和隔离性。只要 2PC 事务完成，订单库和促销库中的数据一定是一致的状态，也就是我们总说的，要么都成功，要么都失败。

所以 2PC 比较适合那些对数据一致性要求比较高的场景，比如我们这个订单优惠券的场景，如果一致性保证不好，有可能会被黑产利用，一张优惠券反复使用，那样我们的损失就大了。

2PC 也有很明显的缺陷，整个事务的执行过程需要阻塞服务端的线程和数据库的会话，所以，2PC 在并发场景下的性能不会很高。并且，协调者是一个单点，一旦过程中协调者宕机，就会导致订单库或者促销库的事务会话一直卡在等待提交阶段，直到事务超时自动回滚。

卡住的这段时间内，数据库有可能会锁住一些数据，服务中会卡住一个数据库连接和线程，这些都会造成系统性能严重下降，甚至整个服务被卡住。

所以，只有在需要强一致、并且并发量不大的场景下，才考虑使用 2PC。

## 本地消息表：订单与购物车的数据一致性问题

2PC 它的适用场景其实是很窄的，更多的情况下，只要保证数据最终一致就可以了。比如说，在购物流程中，用户在购物车界面选好商品后，点击“去结算”按钮进入订单页面创建一个新订单。这个过程我们的系统其实做了两件事儿。

第一，订单系统需要创建一个新订单，订单关联的商品就是购物车中选择的那些商品。

第二，创建订单成功后，购物车系统需要把订单中的这些商品从购物车里删掉。

这也是一个分布式事务问题，创建订单和清空购物车这两个数据更新操作需要保证，要么都成功，要么都失败。但是，清空购物车这个操作，它对一致性要求就没有扣减优惠券那么高，订



单创建成功后，晚几秒钟再清空购物车，完全是可以接受的。只要保证经过一个小的延迟时间后，最终订单数据和购物车数据保持一致就可以了。

本地消息表非常适合解决这种分布式最终一致性的问题。我们一起来看一下，如何使用本地消息表来解决订单与购物车的数据一致性问题。

本地消息表的实现思路是这样的，订单服务在收到下单请求后，正常使用订单库的事务去更新订单的数据，并且，在执行这个数据库事务过程中，在本地记录一条消息。这个消息就是一个日志，内容就是“清空购物车”这个操作。因为这个日志是记录在本地的，这里面没有分布式的问题，那这就是一个普通的单机事务，那我们就可以让订单库的事务，来保证记录本地消息和订单库的一致性。完成这一步之后，就可以给客户端返回成功响应了。

然后，我们再用一个异步的服务，去读取刚刚记录的清空购物车的本地消息，调用购物车系统的服务清空购物车。购物车清空之后，把本地消息的状态更新成已完成就可以了。异步清空购物车这个过程中，如果操作失败了，可以通过重试来解决。最终，可以保证订单系统和购物车系统它们的数据是一致的。

这里面，本地消息表，你可以选择存在订单库中，也可以用文件的形式，保存在订单服务所在服务器的本地磁盘中，两种方式都是可以的，相对来说，放在订单库中更简单一些。

消息队列 RocketMQ 提供一种事务消息的功能，其实就是本地消息表思想的一个实现。使用事务消息可以达到和本地消息表一样的最终一致性，相比我们自己来实现本地消息表，使用起来更加简单，你也可以考虑使用。（我在 [🔗《消息队列高手课》](#) 的专栏中的“[🔗如何利用事务消息实现分布式事务？](#)”这节课中有详细的讲解。）

如果看事务的 ACID 四个特性，本地消息表这种方法，它只能满足 D（持久性），A（原子性）C（一致性）、I（隔离性）都比较差，但是，它的优点非常突出。

首先，实现简单，在单机事务的基础上稍加改造就可以实现分布式事务，另外，本地消息表的性能非常好，和单机事务的性能几乎没有差别。在这个基础上，还提供了大部分情况下都能接受的“数据最终一致性”的保证，所以，本地消息表是更加实用的分布式事务实现方法。

当然，即使能接受数据最终一致，本地消息表也不是什么场景都可以适用的。它有一个前提条件就是，异步执行的那部分操作，不能有依赖的资源。比如说，我们下单的时候，除了要清空购物车以外，还需要锁定库存。

库存系统锁定库存这个操作，虽然可以接受数据最终一致，但是，锁定库存这个操作是有一个前提的，这个前提是：库存中得有货。这种情况就不适合使用本地消息表，不然就会出现用户下单成功后，系统的异步任务去锁定库存的时候，因为缺货导致锁定失败。这样的情况就很难处理了。

## 小结

这节课我们讲解了，如何用分布式事务的几种方法来解决分布式系统中的数据一致性问题。对于订单和优惠券这种需要强一致的分布式事务场景，可以采用 2PC 的方法来解决问题。

2PC 它的优点是强一致，但是性能和可用性上都有些缺陷。本地消息表适用性更加广泛，虽然在数据一致性上有所牺牲，只能满足最终一致性，但是有更好的性能，实现简单，系统的稳定性也很好，是一种非常实用的分布式事务的解决方案。

无论是哪种分布式事务方法，其实都是把一个分布式事务，拆分成多个本地事务。**本地事务可以用数据库事务来解决，那分布式事务就专注于解决如何让这些本地事务保持一致的问题。**我们在遇到分布式一致性问题的時候，也要基于这个思想来考虑问题，再结合实际的情况选择分布式事务的方法。

## 思考题

2PC 也有一些改进版本，比如 3PC、TCC 这些，它们大体的思想和 2PC 是差不多的，解决了 2PC 的一些问题，但是也会带来新的问题，实现起来也更复杂，限于篇幅我们没法每个都详细地去讲解。在理解了 2PC 的基础上，课后请你自行去学习一下 3PC 和 TCC，然后对比一下，2PC、3PC 和 TCC 分别适用于什么样的业务场景？

欢迎你在留言区与我讨论，如果你觉得今天的内容对你有帮助，也欢迎把它分享给你的朋友。

© 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 (46)



李玥 置顶

2020-03-07

Hi, 我是李玥。

这里说一下上节课的思考题：

课后希望你能动手执行一下我们今天这节课中给出的例子，看一下多个事务并发更新同一个账户时，RC 和 RR 两种不同的隔离级别，在行为上有什么不同？

这个思考题的主要目的还是希望你不要光是听和看，还要能真正动手去试一下，以便加深理解。RC和RR在并发更新数据的时候，都需要对数据加锁（一般是行锁）在二个事务同时更新一条记录的时候，先更新的那个事务会抢占到锁，在它结束事务之前，其它需要更新这条记录的事务都会卡住等待这个锁。这一点二种隔离级别是一样的。



👍 45



一剑

2020-03-07

可以这么考虑：作为电商，可以容忍异常导致的多锁定库存，不能容忍少锁定库存（超卖）。本地消息表异步调用会导致超卖；2PC/3PC同步性能调用性能太差。所以可以由调用方先后执行锁定库存及创建订单的接口：如果锁定库存成功，创建订单成功，返回成功；如果锁定库存成功，创建订单失败，返回失败，由调用方重试，可能会导致多锁定库存；如果锁定库存失败，则不再创建订单，返回失败，由调用方重试，可能会导致多锁定库存；那些因为异常导致的异常锁定，可以通过一些手段事后补偿，比如：找出长时间未释放的异常库存锁定，然后进行释放。

共 2 条评论 >

👍 32



小袁

2020-03-18

老师为啥不说下2pc情况下，一个db宕机重启的问题呢？是不是由一个脚本定期扫描发现有哪些订单出现异常的呢？

作者回复: 由于事务执行器有超时机制，数据库本身也有事务超时机制，如果是第一阶段失败或者超时的话，会自动回滚。

对于第二阶段的问题，如果db宕机，确实有可能出现数据不一致的问题，对于这种情况，可以用脚本根据业务做一些补偿处理。

共 3 条评论 >

👍 19



stanley

2020-03-29

在本地消息表方案的例子中，后续需要去清空购物车，同时要把之前记录在本地的那一条记录删掉，这是不是另一个分布式事务的问题？

作者回复: 这个过程就不需要再用事务保证了，否则就死循环了。

可以先清空购物车，再删除本地消息（一般都是更新状态）。即使出现不一致，购物车清了，本地消息表没更新成功，下次再执行一次清空购物车，也问题不大。

共 5 条评论 >

👍 13



几字凉了秋、

2020-04-23

老师，库存系统，锁定库存的一般操作是什么样的，可以说一说吗

作者回复: 一般是直接减掉，但需要记录锁定库存的商品、时间戳，后续超时未支付的情况下，再把库存释放出来。

共 2 条评论 >

👍 8



Dovelol

2020-03-15

老师好，想问下2pc性能不好，该如何优化呢？我听说蚂蚁金服有方案是在大促的高峰将2阶段的第二阶段延后到低峰在唤起执行，因为如果第一阶段资源已经预留，基本上最终状态也已经确定了，业界有这种方案吗？

作者回复: 你说的这种降级方案感觉像是降级成了本地消息表。

共 9 条评论 >

👍 7



1

2020-03-10

实现2pc是自己写吗？还是有第三方可以用？

作者回复：还是要自己实现的。

共 5 条评论 >



7



乖, 摸摸头

2020-05-11

锁定库存锁定，和用户余额 锁定 这种具体是怎么做的了？是直接减掉吗？比如我在下单的时候 库存减了，用户余额也减了，但是我没支付？这时我怎么处理了

作者回复：一般是先锁定库存，然后定时扫描未支付的订单，自动取消超时未支付订单，取消订单的流程中自然就会释放库存。

你会发现，很多秒杀活动要求支付时间都很短，就是这个原因。

共 2 条评论 >



6



旭东(Frank)

2020-03-20

老师好，订单和优惠券的两阶段提交，先提交优惠券还是先提交订单有没细微区别？

作者回复：理论上这两个RM是并行提交的，无所谓谁先谁后。要说细微区别，就是如果某一个数据库宕机后，数据不一致的情况会有区别。



6



ray

2020-03-18

老师您好，

我对2pc有个疑问，如果是走http协定（发出response后连结就关闭了），实作上一般的程序要如何在参与者完成准备阶段后，持续将事务保持在未提交的状态，等待协调者告知可以提交事务？

谢谢老师的解答^^

作者回复: 如果是HTTP协议的话, 可以在RM (事务参与方) Session中, 或者内存中维护进行中的事务, 通过事务ID区分。

第二阶段提交命令从协调者发到RM上之后, 根据事务ID查找对应的数据库会话, 进行提交。

共 3 条评论 >

👍 4



ple

2020-03-07

有一个问题, 如果在订单生成的时候锁库存可以解决文章中提到的, 本地消息表的缺陷么? 消息表的确是不灵活

共 3 条评论 >

👍 4



划过天空阿忠

2020-03-27

就拿下单需要锁库存这个业务来说, 下单接口调锁库存接口, 锁定库存成功之后下单接口提交事务, 锁定库存失败下单接口回滚事务。

其中最关键的一步是锁库存之后要重新调一下库存查询接口, 检查是否真正锁定, 若没有真正锁定, 订单服务手动抛异常, 回滚订单事务。

小弟浅薄理解, 各位指点一下 😊

作者回复: 🍌🍌🍌

共 6 条评论 >

👍 3



etdick

2020-03-12

老师, 我想问问, JD有没有把MySQL的纳入k8S的管理。对数据库这样有状态的节点, k8S是否有问题。

作者回复: 京东有自己的弹性数据库, 但不是简单的把MySQL跑在K8s中, 存储类系统如果要容器化, 一般需要先实现存储计算分离, 把自己变成无状态的节点。这里面有很多问题需要解决, 我们也在持续探索和尝试。



👍 2



流年

2020-11-09

老师，为什么不发送个广播消息，然后其他服务来订阅这个消息来执行后续的业务操作呢？执行失败了重试不就好了

共 3 条评论 >



1



张三丰

2020-06-07

如果出现下边的情况该怎么解决呢？

还是有可能出现订单库完成了提交，但促销库因为宕机自动回滚，导致数据不一致的情况。但是，因为提交的过程非常简单，执行很快，出现这种情况的概率非常小。

作者回复：一般是事后审计和补偿来解决。



2



啸歌

2020-05-05

老师只说了2PC的缺点：

- 1、单点问题
- 2、数据不一致
- 3、性能问题

问题1：那如果我在高并发场景下就是需要强一致性，有什么解决方案呢？

问题2：那2PC具体是怎么落地的呢？协调者有哪些选择？

问题3：最后说了下单扣库存，这一步怎么保证一致性？放在redis减库存还是放在数据库里减库存呢？

老师讲的很精彩，但是意犹未尽呀！！！！

共 1 条评论 >



1



被过去推开

2020-04-27

通读了两遍，收益良多。特别是本地消息表，不仅效率高，而且满足很多的场景。缺点就是不是强一致性



1



winzheng

2020-03-16

不论哪种方案服务要根据实际情况实践，去感受方案，直接想找一个组件来解决业务问题，只能是白日梦。



👍 2



hello

2020-03-07

老师，您好，现在微服务越来越流行，导致系统动不动就微服务化，然而一般的开发人员常年都是CRUD根本不想动脑子，服务化后引入的事务问题，就想着能像操作本地事务那样一两个注解就搞定（无脑最好），这种想法嘛大家都认同，但真正到落地上面，限于自身小公司技术实力，也限于自身认知，一直没有找到特别满意的处理方式。我一直关注开源的方案（例如：阿里的seata，还有一个是TiDB），一个是在应用层解决事务，一个是在数据库底层解决事务，数据库底层解决事务，我比较认同，但限于自身对分布式共识算法的认知，对TiDB底层的原理知之甚少，而且目前业务体量还没到那个级别。老师能否在分布式存储事务方面给点建议或是可供参考的资料，多谢！

共 1 条评论 >

👍 1



约书亚

2020-03-07

其实我感觉本地事务表的设计思想 对比 rocketmq那种事务消息的设计思想还是有挺大差异的，不过他们的适用场景确实一致.....



👍 1