

## 08 | 答疑解惑：缓存失效的策略应该怎么定？

2018-11-14 许令波

如何设计一个秒杀系统

[进入课程 >](#)



十一黄金周的时候，极客时间团队邀请到了前阿里巴巴高级技术专家许令波专门撰写了《如何设计一个秒杀系统》专栏，希望带你透彻理解秒杀系统的各个关键技术点，并借助“秒杀”这个互联网高并发场景中的典型代表，带你了解如何打造一个超大流量并发读写、高性能，以及高可用的系统架构。

专栏虽然只有短短 7 篇，但却持续获得大量用户的支持和赞誉。留言区，我们更是可以看到大量从学习角度或业务角度出发提出的各种问题。为此，我们也特别邀请专栏作者许令波就一些关键或普遍的问题进一步“加餐”解答，希望能够给你更好的帮助。

**1. “06 | 秒杀系统‘减库存’设计的核心逻辑”一文中，很多用户比较关注应用层排队的问题，大家主要的疑问就是应用层用队列接受请求，然后结果怎么返回的问题。**

其实我这里所说的排队，更多地是说在服务端的服务调用之间采用排队的策略。例如，秒杀需要调用商品服务、调用价格优惠服务或者是创建订单服务，由于调用这些服务出现性能瓶

再比如说，正常整个网站上每秒只有几万个请求，这几万个请求可能是非常分散的，那么假如现在有一个秒杀商品，这个秒杀商品带来的瞬间请求一下子就打满了我们的服务器资源，这样就会导致那些正常的几万个请求得不到正常的服务，这个情况对系统来说是绝对不合理的，也是应该避免的。

所以我们设计了一些策略，把秒杀系统独立出来，部署单独的一些服务器，也隔离了一些热点的数据库，等等。但是实际上不能把整个秒杀系统涉及的所有系统都独立部署一套，不然这样代价太大。

既然不能所有系统都独立部署一套，势必就会存在一部分系统不能区分秒杀请求和正常请求，那么要如何防止前面所说的问题出现呢？通常的解决方案就是在部分服务调用的地方对请求进行 Hash 分组，来限制一部分热点请求过多地占用服务器资源，分组的策略就可以根据商品 ID 来进行 Hash，热点商品的请求始终会进入一个分组中，这样就解决了前面的问题。

我看问的问题很多是说对秒杀的请求进行排队如何把结果通知给用户，我并不是说在用户 HTTP 请求时采用排队的策略（也就是把用户的所有秒杀请求都放到一个队列进行排队，然后在队列里按照进入队列的顺序进行选择，先到先得），虽然这看起来还是一个挺合理的设计，但是实际上并没有必要这么做！

为什么？因为我们服务端接受请求本身就是按照请求顺序处理的，而且这个处理在 Web 层是实时同步的，处理的结果也会立马就返回给用户。但是我前面也说了，整个请求的处理涉及很多服务调用也涉及很多其他的系统，也会有部分的处理需要排队，所以可能有部分先到的请求由于后面的一些排队的服务拖慢，导致最终整个请求处理完成的时间反而比较后面的请求慢的情况。

这种情况理论上的确存在，你可能会说这样可能会不公平，但是这的确没有办法，这种所谓的“不公平”，并不是由于人为设置的因素导致的。

你可能会问（如果你一定要问），采用请求队列的方式能不能做？我会说“能”，但是有两点问题：

不也就没有惊喜了？

至于大家在纠结异步请求如何返回结果的问题，其实有多种方案。

一是页面中采用轮询的方式定时主动去服务端查询结果，例如每秒请求一次服务端看看有没有处理结果（现在很多支付页面都采用了这种策略），这种方式的缺点是服务端的请求数会增加不少。

二是采用主动 push 的方式，这种就要求服务端和客户端保持连接了，服务端处理完请求主动 push 给客户端，这种方式的缺点是服务端的连接数会比较多。

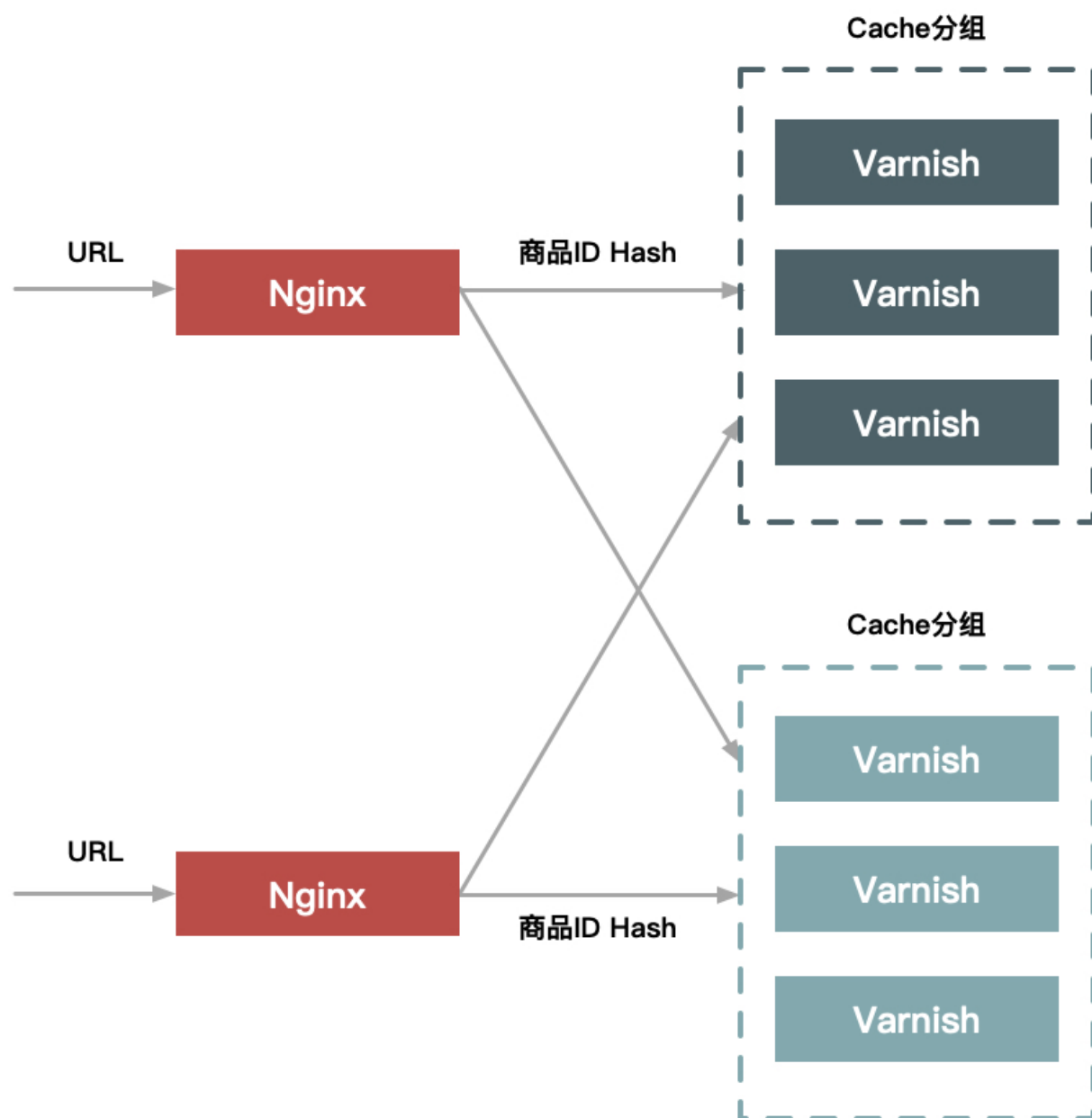
还有一个问题，就是如果异步的请求失败了，怎么办？对秒杀来说，我觉得如果失败了直接丢弃就好了，最坏的结果就是这个人没有抢到而已。但是你非要纠结的话，就要做异步消息的持久化以及重试机制了，要保证异步请求的最终正确处理一般都要借助消息系统，即消息的最终可达，例如阿里的消息中间件是能承诺只要客户端消息发送成功，那么消息系统一定会保证消息最终被送到目的地，即消息不会丢。因为客户端只要成功发送一条消息，下游消费方就一定会消费这条消息，所以也就不存在消息发送失败的问题了。

**2. 在“02 | 如何才能做好动静分离？有哪些方案可选？”一文中，有介绍静态化的方案中关于 Hash 分组的问题。**

大家可能通常理解 Hash 分组，像 Cache 这种可能一个 key 对应的数据只存在于一个实例中，这样做其实是为了保证缓存命中率，因为所有请求都被路由到一个缓存实例中，除了第一次没有命中外，后面的都会命中。

但是这样也存在一个问题，就是如果热点商品过于集中，Cache 就会成为瓶颈，这时单个实例也支撑不了。像秒杀这个场景中，单个商品对 Cache 的访问会超过 20w 次，一般单 Cache 实例都扛不住这么大的请求量。所以需要采用一个分组中有多个实例缓存相同的数据（冗余）的办法来支撑更大的访问量。

你可能会问：一个商品数据存储多个 Cache 实例中，如何保证数据一致性呢？（关于失效问题大家问得也比较多，后面再回答。）这个专栏中提的 Hash 分组都是基于 Nginx+Varnish 实现的，Nginx 把请求的 URL 中的商品 ID 进行 Hash 并路由到一个



另外一个问题，关于 Hash 分组大家关注比较多的是命中率的问题，就是 Cache 机器越多命中率会越低。

这个其实很好理解，Cache 实例越多，那么这些 Cache 缓存数据需要访问的次数也就越多。例如我有 3 个 Redis 实例，需要 3 个 Redis 实例都缓存商品 A，那么至少需要访问 3 次才行，而且是这 3 次访问刚好落到不同的 Redis 实例中。那么从第 4 次访问开始才会被命中，如果仅仅是一个 Redis 实例，那么第二次访问时其实就能命中了。所以理论上 Cache 实例多会影响命中率。



多情况下，还没等到所有 Cache 实例填满，该商品就已经失效了。所以，我们要根据商品的重复访问量来合理地设置 Cache 分组。

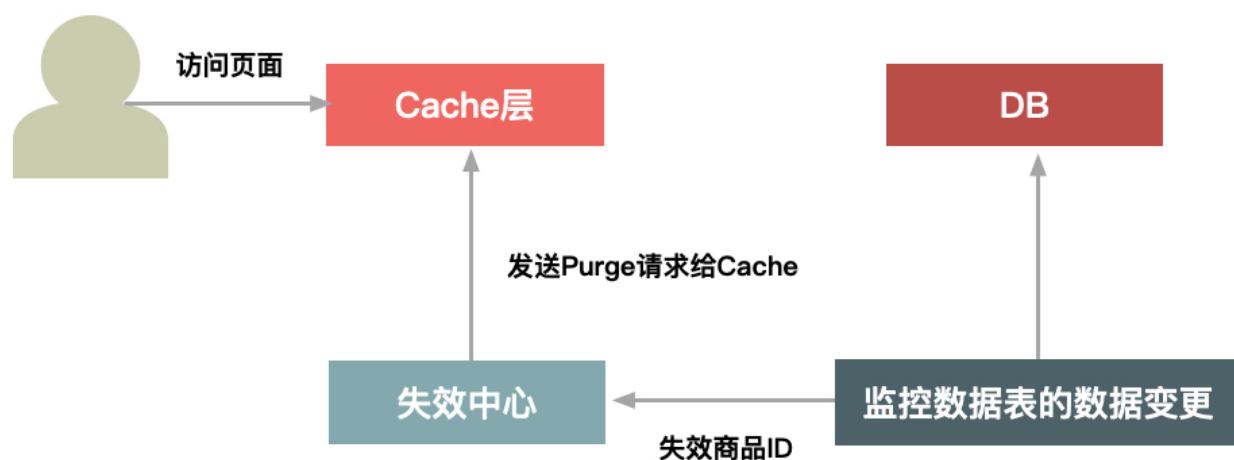
### 3. 在“02 | 如何才能做好动静分离？有哪些方案可选？”和“04 | 流量削峰这事应该怎么做？”两篇文章中，关于 Cache 失效的问题。

首先，咱们要有个共识，有 Cache 的地方就必然存在失效问题。为啥要失效？因为要保证数据的一致性。所以要用到 Cache 必然会问如何保证 Cache 和 DB 的数据一致性，如果 Cache 有分组的话，还要保证一个分组中多个实例之间数据的一致性，就像保证 MySQL 的主从一致一样。

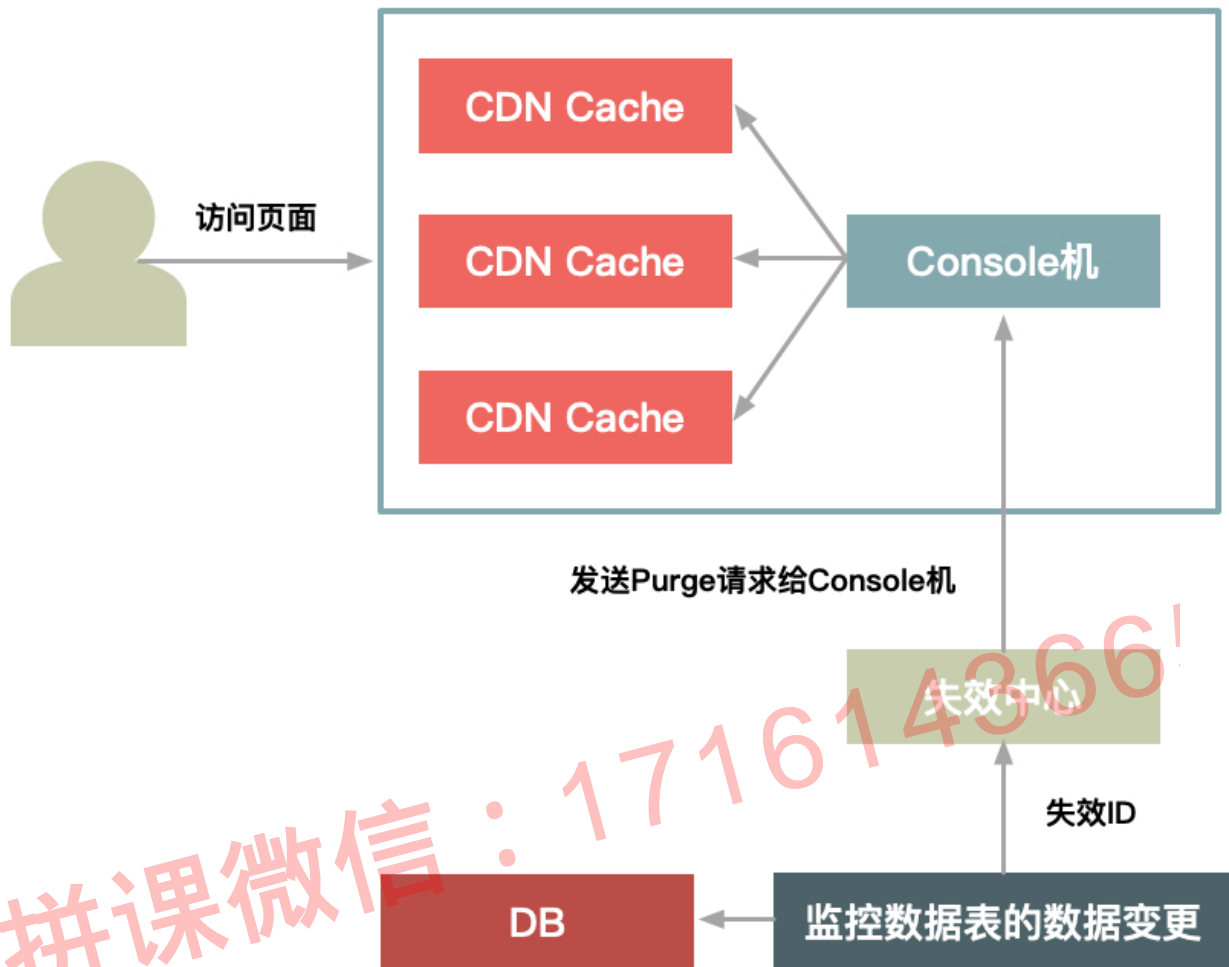
其实，失效有主动失效和被动失效两种方式。

被动失效，主要处理如模板变更和一些对时效性不太敏感数据的失效，采用设置一定时间长度（如只缓存 3 秒钟）这种自动失效的方式。当然，你也要开发一个后台管理界面，以便能够在紧急情况下手工失效某些 Cache。

主动失效，一般有 Cache 失效中心监控数据库表变化发送失效请求、系统发布也需要清空 Cache 数据等几种场景。其中失效中心承担了主要的失效功能，这个失效中心的逻辑图如下：



失效中心会监控关键数据表的变更（有个中间件来解析 MySQL 的 binlog，然后发现有 Insert、Update、Delete 等操作时，会把变更前的数据以及要变更的数据转成一个消息发送给订阅方），通过这种方式来发送失效请求给 Cache，从而清除 Cache 数据。如果



这种失效有失效中心将失效请求发送给每个 CDN 节点上的 Console 机，然后 Console 机来发送失效请求给每台 Cache 机器。

# 如何设计一个秒杀系统

大并发高可用秒杀系统的设计之道

许令波 前阿里巴巴 高级技术专家



新版升级：点击「👤请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

© 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。

上一篇 07 | 准备Plan B：如何设计兜底方案？

## 精选留言 (21)

💬 写留言



张雨桐老豆

2018-12-06

👍 27

这方面的东西看了很多，但是，大多数程序员是接触不到或者这方面的知识让你去搭建设计一个这样的系统，一方面公司没有这么大的流量，除了一些bat公司.另一方面，就算有这个留了，一般的系统都是现成的.一个100人的团队，也许就那么一两个人负责，大多数还是在硬编码，所以说大多数的程序员都停留在一个理论阶段，很少有实践的地方

展开 ∨



我来也

2018-11-14

👍 5

还有加餐，厉害了👍

展开 ∨

**古夜**

2019-01-07

👍 4

第一个问题:同一个商品比如苹果，多个这种苹果的ID是否相同？

二:这个其实很好理解，Cache 实例越多，那么这些 Cache 缓存数据需要访问的次数也就越多。例如我有 3 个 Redis 实例，需要 3 个 Redis 实例都缓存商品 A，那么至少需要访问 3 次才行，而且是这 3 次访问刚好落到不同的 Redis 实例中。那么从第 4 次访问开始才...  
展开 ▾

作者回复: 第一个问题没明白你要问什么？

第二个问题：我已经解释的这么详细了，不知道还怎么解释😂！！所谓命中就是在没有提前填充缓存的情况下，必须要访问一次cache这个商品才会被缓存起来，这样第二次再访问时cache才会被命中。

**Do**

2019-02-03

👍 2

你好，分组是为避免热点，这是ok的，我所说的空间浪费3倍不是指分组的浪费，我指的是组内的机器。如你的例子，一个组里面有3个varson实例，那么假设这3个varson实例的内存是1GB，那么其实这3个varson实例也就只能提供1GB的的缓存。假如这个组命名为A组，分别给组内varson实例分别命名为A1,A2A3,假设每组的热点数据是2GB,那么是否有以下几个缺陷...

展开 ▾

作者回复: 你没理解我的意思，建议你仔细看看我说的话，一个组内的三台实例，同一份商品数据只会存一份。

一个商品是随机路由到3个分组的，但是一个商品始终只会命中一个分组中的一个实例

**小胡子**

2019-01-08

👍 1



**耶愿**

2018-12-20

1

许老师，问个问题，关于并发写的问题，目前我知道MySQL和redis并发写即使做过优化，tps也就是上千。而淘宝的写入tps至少上万，想知道淘宝是怎么做到上万写入tps的，而用户没有等待的感觉，非常感谢！虽然其它同学有留言，但感觉他们的方法都是过家家的方法，我想知道淘宝这是怎么实现的。

展开 ∨

作者回复: 不知道你说的tps先·并发写上千是什么场景下得出的，就我了解即使是MySQL没经过单机写肯定也不止上千QPS。阿里在数据库的并发写肯定是做了很多的优化，我建议可以订阅一下隔壁的MySQL课程。另外我们说的并发写是有纬度概念的，比如单机还是单库还是单表，还是指一个业务在同一时刻的并发，都不一样。比如阿里的双十一并发下单支持10w的QPS，虽然是10w但是落到实际的数据库层多个库的多台机器上，因为我们可以根据用户请求的商品ID进行分库分表，这样可以大大减少并发度。

**maomaosty...**

2019-04-11

1

失效中心可以直接调用cdn的接口服务吧

展开 ∨

作者回复: 嗯

**北极的大企...**

2019-04-11

1

比较想问的是学完框架后，是先学设计模式还是先学JVM原理，并发与线程安全，然后中间件和架构设计，数据库设计，Linux学习，跨语言学习顺序，这些都是按照什么样的顺序学习的，还有源码阅读顺序

作者回复: 这个没有统一的答案 😊

量。

.....

展开 ▾

作者回复: 😊

◀

▶



晓洁

2019-02-20



什么时候上新呢？

展开 ▾

作者回复: 你还想看哪些方面的？😊

◀

▶



Do

2019-01-28



你好，在2. 在“02 | 如何才能做好动静分离？有哪些方案可选？”一文中，有介绍静态化的方案中关于 Hash 分组的问题。”中varnish是做了分组，分为2组，每组3台机器。我的问题是，分支并且组内多个varnish实例虽然能避免单点和热点问题，但是组内的机器其实存储的数据很大部分是一样的，varnish是单实例部署，实例间没有共享内存，这样会不会导致内存浪费。例如存储1GB的HTTP缓存，那么这个架构的话实际上就是消耗3GB存储，...

展开 ▾

作者回复: 首先分为2组，是浪费了一倍的空间，不是3倍。

其次，设置分组是会浪费空间的，浪费空间是为了保证热点问题。

当然也有一个简单的热点方案就是在每个Nginx上增加一个很小的热点模块，这个模块只会缓存少量的热点商品，例如top1000个热点商品。这样也可以起到一定的保护作用，这样就不用通过分组的实现了，当然采用哪种方式还是要看情况而定

◀

▶

在cache中，不设置过期呢，等秒杀结束再删掉？

展开 ∨

作者回复: 秒杀商品中的部分信息一致不变的话（例如商品描述信息），是可以一直缓存在cache中的，等结束了再删除



oTo123

2019-01-18



老师，请问一下，一个其他问题。商品详情页这种页面，有商品数据，商家数据，评价数据，商品推荐等。这些是放到一个接口返回呢（这样接口那耦合了多个模块的数据），还是每个部分都开一个接口给前端调用（这样减少了网络开销次数）？在性能需求下怎么权衡呢？

展开 ∨

作者回复: 如你所说这是个平衡的问题，没有标准答案，是合并还是分开要具体根据测试结果来判断。就根据经验来判断，像一些评价和推荐数据一般都是单独请求比较好，一个是这些数据比较多，而且服务端也比较耗时所以单独请求会比较合理一些



外星人

2018-12-26



如果是秒杀1个商品10个库存，服务端请求排队1000个(10倍库存量或更多)，超出1000个直接返回“秒杀结束”，这种方式有啥不好呢？

展开 ∨

作者回复: 排队本身没啥问题，就是觉得没啥必要：)



放

2018-12-11



感谢老师的分享，即使不做架构师，对以后写代码的思路也有很大的提升！

3.2.0 的，我现在要用spring-session实现session共享，我看网上都是要求spring版本4.x以上，我在本地试了一下，发现session是能存到redis里的，但是我把项目部署到服务器上时却存不到redis里，不知道是什么原因，还望许神有空帮忙分析一下哈

展开 ∨

作者回复: 找你身边的同学帮你看看吧>o<



jack

2018-12-01



针对Cache和DB的数据一致性问题，如果DB是A团队业务使用的，Cache是B团队业务使用。DB对于B团队来说不可控，可以通过什么方式对A提出较少的需求（比如是要在DB所在机器安装监控binlog的中间件吗？比如要求A团队打开DB的binlog？），实现读取Cache和A业务的DB数据一致性。

展开 ∨

作者回复: 你说的场景是一种实现思路，当然也可以直接通过应用程序直接操作db和cache保持一致性



Adam

2018-11-28



解析binlog的中间件是什么？有开源的产品么？不然自己实现的话代价就比较大了

作者回复: Canal可以去了解一下



func

2018-11-21



赞一个

展开 ∨



下载APP



mysql的日志分析方式有哪些？

oracle有类似binlog的日志分析方式吗？

作者回复: 阿里有个来源的mysql工具分析工具canal

