

## 开篇词 | 秒杀系统架构设计都有哪些关键点？

2018-09-25 许令波

如何设计一个秒杀系统

[进入课程 >](#)



讲述：秭明

时长 05:59 大小 2.74M



你好，我是许令波，花名“君山”。说起来我的职业生涯算是比较简单，2009 年大学毕业后就进入了淘宝，一直工作了七年多。这七年多的时间里，我有幸看到了淘宝业务的快速增长，并且以开发者的身份参与其中。

说实话，作为一名程序员，我的技术能力也在公司业务的快速增长过程中得到了历练，并积累了一些大流量高并发网站架构设计和优化的经验，尤其是针对“秒杀”这个场景。因为我确信，那个时候我们肯定是对系统做了足够多的极致优化，才能扛住当时洪峰般的流量请求。

记得早期的时候，淘宝商品详情系统的 PV 还差不多是 1 亿的样子，但是到 2016 年差不多已经升至 50 亿了。尤其是 2012 年到 2014 年那个时间段，“秒杀”活动特别流行，用户的参与热情一浪高过一浪，系统要面对流量也是成倍增长。

而每一次的秒杀活动对技术团队来说都是一次考验。现在想起来，那个时候我们整个团队，无所畏惧，逐步迭代创新，然后解决一个个难题的过程，也是极具挑战性和成就感的事情。

记得有一年，为了应对“双十一”，我们整个商品详情团队对系统做了很多优化，我们自认为已经是整个公司最牛的系统了，性能也已经是“业界之巅”。

但是那年“双十一”的晚上，我们的系统还是遇到了瓶颈。当时老大就跑过来盯着我们，问我们什么时候能够恢复，我们整个团队都承担着巨大的心理压力。

事后我们复盘宕机的原因，发现当时的秒杀流量远远超过了我们的预想，我们根本没想到大家的参与热情能有那么高。于是我们按照这个增长率去预估下一年的流量和服务器，粗算下来，我记得差不多要增加 2000 台服务器，简直不可思议。

怎么可能真正增加这么多机器，所以这也就倒逼我们必须找出一些特殊的手段来优化系统。后面，经过一段时间的调研和分析，我们想到了把整个系统进行动静分离改造的解决方案。

秒杀系统也差不多那个时候才从商品详情系统独立出来成为一个独立产品的。因为我见证了秒杀系统的建设过程，所以也有颇多感慨。秒杀系统的迭代又是一个升级打怪的过程，我们也都是遇到问题解决问题，逐一优化。

那么，如何才能更好地理解秒杀系统呢？我觉得作为一个程序员，你首先需要从高维度出发，从整体上思考问题。在我看来，**秒杀其实主要解决两个问题，一个是并发读，一个是并发写**。并发读的核心优化理念是尽量减少用户到服务端来“读”数据，或者让他们读更少的数据；并发写的处理原则也一样，它要求我们在数据库层面独立出来一个库，做特殊的处理。另外，我们还要针对秒杀系统做一些保护，针对意料之外的情况设计兜底方案，以防止最坏的情况发生。

而从一个架构师的角度来看，要想打造并维护一个超大流量并发读写、高性能、高可用的系统，在整个用户请求路径上从浏览器到服务端我们要遵循几个原则，就是要保证用户请求的数据尽量少、请求数尽量少、路径尽量短、依赖尽量少，并且不要有单点。这些关键点我会在后面的文章里重点讲解。

**其实，秒杀的整体架构可以概括为“稳、准、快”几个关键字。**

所谓“稳”，就是整个系统架构要满足高可用，流量符合预期时肯定要稳定，就是超出预期时也同样不能掉链子，你要保证秒杀活动顺利完成，即秒杀商品顺利地卖出去，这个是最基本的前提。

然后就是“准”，就是秒杀 10 台 iPhone，那就只能成交 10 台，多一台少一台都不行。一旦库存不对，那平台就要承担损失，所以“准”就是要求保证数据的一致性。

最后再看“快”，“快”其实很好理解，它就是说系统的性能要足够高，否则你怎么支撑这么大的流量呢？不光是服务端要做极致的性能优化，而且在整个请求链路上都要做协同的优化，每个地方快一点，整个系统就完美了。

所以从技术角度上看“稳、准、快”，就对应了我们架构上的高可用、一致性和高性能的要求，我们的专栏也将主要围绕这几个方面来展开，具体如下。

**高性能。** 秒杀涉及大量的并发读和并发写，因此支持高并发访问这点非常关键。本专栏将从设计数据的动静分离方案、热点的发现与隔离、请求的削峰与分层过滤、服务端的极致优化这 4 个方面重点介绍。

**一致性。** 秒杀中商品减库存的实现方式同样关键。可想而知，有限数量的商品在同一时刻被很多倍的请求同时来减库存，减库存又分为“拍下减库存”“付款减库存”以及预扣等几种，在大并发更新的过程中都要保证数据的准确性，其难度可想而知。因此，我将用一篇文章来专门讲解如何设计秒杀减库存方案。

**高可用。** 虽然我介绍了很多极致的优化思路，但现实中总难免出现一些我们考虑不到的情况，所以要保证系统的高可用和正确性，我们还要设计一个 PlanB 来兜底，以便在最坏情况发生时仍然能够从容应对。专栏的最后，我将带你思考可以从哪些环节来设计兜底方案。

最后，很幸运能在极客时间遇到你，希望这节课能让你彻底理解大并发、高性能、高可用秒杀系统的设计之道，并能够在思考解决类似问题时有更准确的思考和判断。

# 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。

下一篇 01 | 设计秒杀系统时应该注意的5个架构原则

## 精选留言 (67)

 写留言



郭蕾 置顶

2018-09-26

 28

一共7篇，10月1日正式更新，每天一篇。详细的介绍大家可以在专栏简介页面了解。



杰

2018-09-27

 82

知识无价，你可能小赚，但我永远不亏。

展开 



null

2018-09-27

 40

之前开发过一个秒杀系统：

1. 前端限流，5 秒内只提交一个请求，静态资源存放于 CDN。
2. 后端 redis 对 uid 限流，同样 5 秒内提交一个请求。
3. 请求保存队列，队列长度为库存 2 倍。为什么是 2 倍呢？拍脑门决定的，防止前面预订失败，后面补上。...

展开 ∨



**chon**

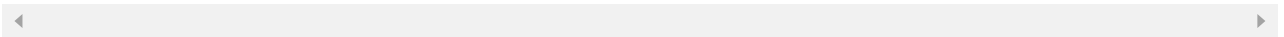
2018-09-28

👍 21

建议增加如何用啥样的工具发现某个地方存在瓶颈。很多人不知道如何定位，即使学了，也不知道如何发现，定位和量化自己项目中问题

展开 ∨

作者回复: 会有介绍的



**龙猫**

2018-09-26

👍 10

一顿兰州拉面的价钱，划算！！！！

展开 ∨



**aoe**

2018-09-26

👍 8

向年轻人学习

展开 ∨



**long.mr**

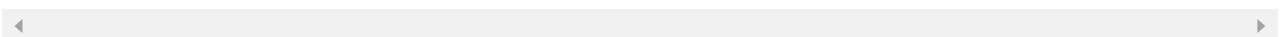
2018-09-25

👍 7

问一哈，秒杀系统，判断先后顺序的依据是请求到达的时间戳吗？会不会有一个服务器存在串行的请求的情况，这样的话会在一定程度上影响公平性吗？

展开 ∨

作者回复: 为啥要时间戳，谁先到谁先执行啊







dragonwsh

2018-09-25

👍 7

看到是许神的视频果断买了，相信许神的质量。

展开 ▾



云学

2018-09-29

👍 6

如何发现瓶颈点，怎么判断某个地方是否有优化的空间，有哪些推荐的工具，请作者分享一下

作者回复: 后面文章有介绍



青梅煮酒

2018-09-25

👍 4

非常期待“君山”老师后续的课程

展开 ▾



彭正聪

2018-10-23

👍 3

问一个问题，存不存在2个请求完全同时的情况。这样先后顺序怎么区分？

作者回复: 你说的问题其实也是并发问题，对秒杀来说，谁先减库存成功就算谁秒杀成功，假如有两个完全同时的同求减库存，在数据库层可能要数据库来区分谁先完成了😁



Geek\_d6c37...

2018-09-28

👍 3

有开源项目吗

展开 ▾



ds.Yang™...

2018-09-26

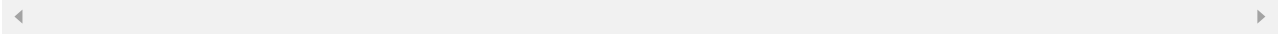
👍 3

拼课微信：171614366!

相见恨晚，最近也在读许老师的web技术内幕，很不错

展开 ▾

作者回复: 谢谢支持，这个小专栏我们就是想交付一个具体的场景解决方案，希望能帮到你



**张斌**

2018-09-25

👍 3

之前看过许先生的书，实战经历很丰富

展开 ▾



**智心悦**

2018-09-29

👍 2

作为不回家的菜鸟，给自己一个国庆学习的动力！

展开 ▾



**过河卒子**

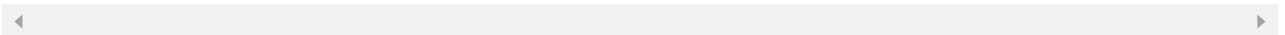
2018-09-27

👍 2

怎么没有看到具体课程介绍呀！

展开 ▾

作者回复: APP右上角有个有图标啊



**Exception**

2018-09-26

👍 2

之前买过许神的书籍，受益颇多，一看标题，直接秒杀

展开 ▾



**JackGe**

2018-09-25

👍 2

支持令波 😊

展开 ▾



**小柯基**

2018-09-25

👍 2

跟着javaweb技术内幕进来的

展开 ▾

---



**刘涛**

2018-09-30

👍 1

会有代码吗，还是只有一些文字描述？

展开 ▾