



下载APP



## 期中测试题答案 | 这些问题，你都答对了吗？

2020-10-07 蒋德钧

Redis核心技术与实战

[进入课程 >](#)



讲述：蒋德钧

时长 06:33 大小 6.01M



你好，我是蒋德钧。今天，我来公布一下主观题的答案。

### 第一题

Redis 在接收多个网络客户端发送的请求操作时，如果有一个客户端和 Redis 的网络连接断开了，Redis 会一直等待该客户端恢复连接吗？为什么？

答案：

Redis 不会等待客户端恢复连接。



原因是，Redis 的网络连接是由操作系统进行处理的，操作系统内核负责监听网络连接套接字上的连接请求或数据请求，而 Redis 采用了 IO 多路复用机制 epoll，不会阻塞在某一个特定的套接字上。epoll 机制监测到套接字上有请求到达时，就会触发相应的事件，并把事件放到一个队列中，Redis 就会对这个事件队列中的事件进行处理。这样一来，Redis 只用查看和处理事件队列，就可以了。当客户端网络连接断开或恢复时，操作系统会进行处理，并且在客户端能再次发送请求时，把接收到的请求以事件形式通知 Redis。

## 第二题

Redis 的主从集群可以提升数据可靠性，主节点在和从节点进行数据同步时，会使用两个缓冲区：复制缓冲区和复制积压缓冲区。这两个缓冲区的作用各是什么？会对 Redis 主从同步产生什么影响吗？

答案：

首先来说一下复制缓冲区。

**作用：**主节点开始和一个从节点进行全量同步时，会为从节点创建一个输出缓冲区，这个缓冲区就是复制缓冲区。当主节点向从节点发送 RDB 文件时，如果又接收到了写命令操作，就会把它们暂存在复制缓冲区中。等 RDB 文件传输完成，并且在从节点加载完成后，主节点再把复制缓冲区中的写命令发给从节点，进行同步。

**对主从同步的影响：**如果主库传输 RDB 文件以及从库加载 RDB 文件耗时长，同时主库接收的写命令操作较多，就会导致复制缓冲区被写满而溢出。一旦溢出，主库就会关闭和从库的网络连接，重新开始全量同步。所以，我们可以通过调整 client-output-buffer-limit slave 这个配置项，来增加复制缓冲区的大小，以免复制缓冲区溢出。

再来看看复制积压缓冲区。

**作用：**主节点和从节点进行常规同步时，会把写命令也暂存在复制积压缓冲区中。如果从节点和主节点间发生了网络断连，等从节点再次连接后，可以从复制积压缓冲区中同步尚未复制的命令操作。

**对主从同步的影响：**如果从节点和主节点间的网络断连时间过长，复制积压缓冲区可能被新写入的命令覆盖。此时，从节点就没有办法和主节点进行增量复制了，而是只能进行全

量复制。针对这个问题，应对的方法是调大复制积压缓冲区的大小（可以参考 [第 6 讲](#) 中对 `repl_backlog_size` 的设置）。

### 第三题

假设在业务场景中，我们有 20GB 的短视频属性信息（包括短视频 ID、短视频基本信息，例如短视频作者、创建时间等）要持久化保存，并且线上负载以读为主，需要能快速查询到这些短视频信息。

现在，针对这个需求，我们想使用 Redis 来解决，请你来设计一个解决方案。我来提几个问题，你可以思考下。

首先，你会用 Redis 的什么数据类型来保存数据？如果我们只用单个实例来运行的话，你会采用什么样的持久化方案来保证数据的可靠性？

另外，如果不使用单实例运行，我们有两个备选方案：一个是用两台 32GB 内存的云主机来运行主从两个 Redis 实例；另一个是用 10 台 8GB 的云主机来运行 Redis Cluster，每两台云主机分别运行一个 Redis 实例主库和从库，分别保存 4GB 数据，你会用哪种方案呢？请聊一聊你的想法。

答案：

Redis 的 Hash 类型属于典型的集合类型，可以保存 key-value 形式的数据。而且，当 Hash 类型中保存较多数据时，它的底层是由哈希表实现的。哈希表的存取复杂度是  $O(1)$ ，所以可以实现快速访问。在这道题中，短视频属性信息属于典型 key-value 形式，所以，我们可以使用 Hash 类型保存短视频信息。具体来说，就是将一个短视频 ID 作为 Hash 集合的 key，将短视频的其他属性信息作为 Hash 集合内部的键值对，例如“作者”：“实际姓名”，“创建时间”：“实际时间”。这样既满足了保存数据的需求，也可以利用 Hash 快速查询的特点，快速查到相应的信息。

Redis 的 AOF 日志会记录客户端发送给实例的每一次写操作命令，在 Redis 实例恢复时，可以通过重新运行 AOF 文件中的命令，实现恢复数据的目的。在这道题的业务场景中，负载以读为主，因此，写命令不会太多，AOF 日志文件的体量不会太大，即使实例故障了，也可以快速完成恢复。所以，当使用单实例运行时，我们可以使用 AOF 日志来做持久化方案。

关于使用多实例的运行方案：两种方案各有优势，我们分析一下。

## 方案一

优势：可以节省云主机数量和成本。虽然主从节点进行第一次全量同步时，RDB 文件较大，耗时会长些，但是因为写请求少，所以复制缓冲区的压力不大。

不足：如果网络环境不好，需要频繁地进行全量同步的话，这种方案的优势就小了，每次全量同步时的 RDB 生成和传输压力都很大。

## 方案二

优势：每个实例只用保存 4GB 数据，和从库同步时的压力较小。而且，这种方案的可扩展性更好，如果有新增数据，可以更好地应对。

不足：需要较多的云主机，运维和资源成本较高。

好了，这节课就到这里。假期很快就要结束了，希望你抓住最后的几天时间，好好地巩固一下所学的内容。我们下节课见。

提建议

## 更多课程推荐

# 数据结构与算法之美

为工程师量身打造的数据结构与算法私教课

王争

前 Google 工程师



立省 ¥40

破 90000 订阅特惠，到手价 ¥89

© 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 期中测试题 | 一套习题，测出你的掌握程度

## 精选留言 (4)

写留言



徐小熊

2020-10-09

老师，我想问一下，如果我们使用主从库模式，读写分离，会出现高并发情况下主库的写命令还未同步到从库的情况，这个时候又有读命令发送到从库，是不是就会读不到本来应该写入的数据呢？

1



yeek

2020-10-09

客户端连接断开的补充猜测：

1. epoll只是负责帮我们维护连接，当客户端断连之后，epoll不会自己帮我们删除无效的连接，redis服务端有个空闲链接检测机制，需手动开启，用于定期检查释放无效的连接，删除epoll中的fd...

展开 ▾



漫步oo0云端

2020-10-09

请问第三题的视频为什么首选不是保存在string上？string也是保存在全局哈希表中，也很快啊？



王聪

2020-10-07

哈哈，国庆终于把进度补上来了

展开 ▾

