

10个常见的Redis面试"刁难"问题

导读：在程序员面试过程中Redis相关的知识是常被问到的话题。作为一名在互联网技术行业打击过成百上千名的资深技术面试官，本文作者总结了面试过程中经常问到的问题。十分值得一读。

作者简介：钱文品（老钱），互联网分布式高并发技术十年老兵，目前任掌阅科技资深后端工程师。熟练使用 Java、Python、Golang 等多种计算机语言，开发过游戏，制作过网站，写过消息推送系统和MySQL 中间件，实现过开源的 ORM 框架、Web 框架、RPC 框架等

Redis在互联网技术存储方面使用如此广泛，几乎所有的后端技术面试官都要在Redis的使用和原理方面对小伙伴们进行各种刁难。作为一名在互联网技术行业打击过成百上千名【请允许我夸张一下】的资深技术面试官，看过了无数落寞的身影失望的离开，略感愧疚，故献上此文，希望各位读者以后面试势如破竹，永无失败！

Redis有哪些数据结构？

字符串String、字典Hash、列表List、集合Set、有序集合SortedSet。

如果你是Redis中高级用户，还需要加上下面几种数据结构
HyperLogLog、Geo、Pub/Sub。

如果你说还玩过Redis Module，像BloomFilter，RedisSearch，Redis-ML，面试官得眼睛就开始发亮了。

使用过Redis分布式锁么，它是怎么回事？

先拿setnx来争抢锁，抢到之后，再用expire给锁加一个过期时间防止锁忘记了释放。

这时候对方会告诉你说你回答得不错，然后接着问如果在setnx之后执行expire之前进程意外crash或者要重启维护了，那会怎么样？

这时候你要给予惊讶的反馈：唉，是喔，这个锁就永远得不到释放了。紧接着你需要抓一抓自己得脑袋，故作思考片刻，好像接下来的结果是你主

动思考出来的，然后回答：我记得set指令有非常复杂的参数，这个应该是可以同时把setnx和expire合成一条指令来用的！对方这时会显露笑容，心里开始默念：嗯，这小子还不错。

假如Redis里面有1亿个key，其中有10w个key是以某个固定的已知的前缀开头的，如果将它们全部找出来？

使用keys指令可以扫出指定模式的key列表。

对方接着追问：如果这个redis正在给线上的业务提供服务，那使用keys指令会有什么问题？

这个时候你要回答redis关键的一个特性：redis的单线程的。keys指令会导致线程阻塞一段时间，线上服务会停顿，直到指令执行完毕，服务才能恢复。这个时候可以使用scan指令，scan指令可以无阻塞的提取出指定模式的key列表，但是会有一定的重复概率，在客户端做一次去重就可以了，但是整体所花费的时间会比直接用keys指令长。

使用过Redis做异步队列么，你是怎么用的？

一般使用list结构作为队列，rpush生产消息，lpop消费消息。当lpop没有消息的时候，要适当sleep一会再重试。

如果对方追问可不可以不用sleep呢？list还有个指令叫blpop，在没有消息的时候，它会阻塞住直到消息到来。

如果对方追问能不能生产一次消费多次呢？使用pub/sub主题订阅者模式，可以实现1:N的消息队列。

如果对方追问pub/sub有什么缺点？在消费者下线的情况下，生产的消息会丢失，得使用专业的消息队列如rabbitmq等。

如果对方追问redis如何实现延时队列？我估计现在你很想把面试官一棒打死如果你手上有一根棒球棍的话，怎么问的这么详细。但是你很克制，然后神态自若的回答道：使用sortedset，拿时间戳作为score，消息内容作为key调用zadd来生产消息，消费者用zrangebyscore指令获取N秒之前的数据轮询进行处理。

到这里，面试官暗地里已经对你竖起了大拇指。但是他不知道的是此刻你却竖起了中指，在椅子背后。

如果有大量的key需要设置同一时间过期，一般需要注意什么？

如果大量的key过期时间设置的过于集中，到过期的那个时间点，redis可能会出现短暂的卡顿现象。一般需要在时间上加一个随机值，使得过期时间分散一些。

Redis如何做持久化的？

bgsave做镜像全量持久化，aof做增量持久化。因为bgsave会耗费较长时间，不够实时，在停机的时候会导致大量丢失数据，所以需要aof来配合使用。在redis实例重启时，优先使用aof来恢复内存的状态，如果没有aof日志，就会使用rdb文件来恢复。

如果再问aof文件过大恢复时间过长怎么办？你告诉面试官，Redis会定期做aof重写，压缩aof文件日志大小。如果面试官不够满意，再拿出杀手锏答案，Redis4.0之后有了混合持久化的功能，将bgsave的全量和aof的增量做了融合处理，这样既保证了恢复的效率又兼顾了数据的安全性。这个功能甚至很多面试官都不知道，他们肯定会对你刮目相看。

如果对方追问那如果突然机器掉电会怎样？取决于aof日志sync属性的配置，如果不要求性能，在每条写指令时都sync一下磁盘，就不会丢失数据。但是在高性能的要求下每次都sync是不现实的，一般都使用定时sync，比如1s1次，这个时候最多就会丢失1s的数据。

Pipeline有什么好处，为什么要用pipeline？

可以将多次IO往返的时间缩减为一次，前提是pipeline执行的指令之间没有因果相关性。使用redis-benchmark进行压测的时候可以发现影响redis的QPS峰值的一个重要因素是pipeline批次指令的数目。

Redis的同步机制了解么？

Redis可以使用主从同步，从从同步。第一次同步时，主节点做一次bgsave，并同时后续修改操作记录到内存buffer，待完成后将rdb文件全量同步到复制节点，复制节点接受完成后将rdb镜像加载到内存。加载

完成后，再通知主节点将期间修改的操作记录同步到复制节点进行重放就完成了同步过程。

是否使用过Redis集群，集群的原理是什么？

Redis Sentinel着眼于高可用，在master宕机时会自动将slave提升为master，继续提供服务。

Redis Cluster着眼于扩展性，在单个redis内存不足时，使用Cluster进行分片存储。

相关阅读：

- [用最少的机器支撑万亿级访问，微博6年Redis优化历程](#)
- [首发 | 360开源的类Redis存储系统:Pika](#)
- [Redis实战：如何构建类微博的亿级社交平台](#)
- [Codis作者黄东旭细说分布式Redis架构设计和踩过的那些坑](#)
- [Redis架构之防雪崩设计：网站不宕机背后的兵法](#)
- [同程旅游缓存系统设计:如何打造Redis时代的完美体系\(含PPT\)](#)
- [挑战Kafka！Redis5.0重量级特性Stream尝鲜](#)



高可用架构

改变互联网的构建方式



长按二维码 关注「高可用架构」公众号