# Redis

## 为什么 Redis 单线程却能支撑高并发？

redis中使用I/O多路复用技术来防止I/O阻塞

在 I/O 多路复用模型中，最重要的函数调用就是 select，该方法的能够同时监控多个文件描述符的可读可写情况，当其中的某些文件描述符可读或者可写时， select方法就会返回可读以及可写的文件描述符个数。与此同时也有其它的 I/O 多路复用函数 epoll/kqueue/evport

## Redis为什么设计成单线程？

执行两个任务，在不涉及I/O操作的时候，单线程依次执行两个任务比双线程分别执行两个任务要快。Redis不涉及I/O操作，因此设计为单线程是效率最高的！

## Redis的性能瓶颈在哪？

一般在两个地方：

其一是机器内存大小，内存大小关系到Redis存储的数据量

其二是网络带宽，Redis客户端执行一条命令分为四个过程：发送命令、命令排队、命令执行、返回结果。而其中发送命令+返回结果这一过程被称为Round Trip Time（RTT，往返时间）

## 如何保证缓存与数据库的双写一致性？

一般来说，如果允许缓存可以稍微的跟数据库偶尔有不一致的情况，也就是说如果你的系统不是严格要求 “缓存+数据库” 必须保持一致性的话，最好不要做这个方案，即：读请求和写请求串行化，串到一个内存队列里去。

串行化可以保证一定不会出现不一致的情况，但是它也会导致系统的吞吐量大幅度降低，用比正常情况下多几倍的机器去支撑线上的一个请求。

**最经典的缓存+数据库读写的模式，就是 Cache Aside Pattern**。

* 读的时候，先读缓存，缓存没有的话，就读数据库，然后取出数据后放入缓存，同时返回响应。
* 更新的时候，**先更新数据库，然后再删除缓存**。

为什么是删除缓存，而不是更新缓存？因为很多数据更新频繁，但查询不一定频繁。

**最初级的缓存不一致问题及解决方案**

问题：先更新数据库，再删除缓存。如果删除缓存失败了，那么会导致数据库中是新数据，缓存中是旧数据，数据就出现了不一致。

解决思路：先删除缓存，再更新数据库。如果数据库更新失败了，那么数据库中是旧数据，缓存中是空的，那么数据不会不一致。因为读的时候缓存没有，所以去读了数据库中的旧数据，然后更新到缓存中。

**比较复杂的数据不一致问题分析**

数据发生了变更，先删除了缓存，然后要去修改数据库，此时还没修改。一个请求过来，去读缓存，发现缓存空了，去查询数据库，查到了修改前的旧数据，放到了缓存中。随后数据变更的程序完成了数据库的修改。完了，数据库和缓存中的数据不一样了...

解决思路：将操作都发送到串行化队列中执行。但是这个队列就是性能的瓶颈，这时就需要以某个字段为key，散列到分布式队列中进行处理。然而还有一个隐藏的问题：就是有些词访问频繁，有些不平凡，这样也可能会造成某个队列负载过大。

# 缓存问题

## 缓存穿透

正常情况下，我们去查询数据都是存在。那么请求去查询一条压根数据库中根本就不存在的数据，也就是缓存和数据库都查询不到这条数据，但是请求每次都会打到数据库上面去。这种查询不存在数据的现象我们称为缓存穿透。

解决办法：1.缓存空值；2.BloomFilter（BloomFilter 类似于一个hase set 用来判断某个元素（key）是否存在于某个集合中）

如何选择：针对这种key异常多，请求重复率比较低的数据，我们就没有必要进行缓存，使用第二种方案直接过滤掉。对于空数据的key有限的，重复率比较高的，我们则可以采用第一种方式进行缓存。

## 缓存击穿

在平常高并发的系统中，大量的请求同时查询一个 key 时，此时 这个key 正好失效了，就会导致大量的请求都打到数据库上面去。这种现象我们成为击穿。

如何解决：上面的现象是多个线程同时去查询数据库的这条数据，那么我们可以在第一个查询数据的请求上使用一个互斥锁来锁住它。其他的线程走到这一步拿不到锁就等着，等第一个线程查询到了数据，然后做缓存。后面的线程进来发现已经有缓存了，就直接走缓存。

## 缓存雪崩

缓存的情况是说，当某一时刻发生大规模的缓存失效的情况。比如你的缓存服务宕机了，会有大量的请求进来直接打到DB上面。结果就是DB 称不住，挂掉。

解决办法：事前：使用集群缓存，保证缓存服务的高可用（如果是使用 Redis，可以使用 主从+哨兵 ，Redis Cluster 来避免 Redis 全盘崩溃的情况）事中：使用 ehcache 本地缓存 + Hystrix 限流&降级 ,避免 MySQL 被打死的情况发生。事后：开启 Redis 持久化机制，尽快恢复缓存集群

## 热点数据集中失效问题怎么解决

我们在设置缓存的时候，一般会给缓存设置一个失效时间，过了这个时间，缓存就失效了。对于一些热点的数据来说，当缓存失效以后会存在大量的请求过来，然后打到数据库去，从而可能导致数据库崩溃的情况。

解决办法：设置不同的失效时间；互斥锁（在第一个请求去查询数据库的时候对他加一个互斥锁，其余的查询请求都会被阻塞住，直到锁被释放，从而保护数据库。）

# redis的热key问题

热key，就是瞬间有几十万的请求去访问redis上某个固定的key，从而压垮缓存服务的情情况。

hot key和big key问题，大家一定要有所了解。

## 怎么发现热key

方法一:凭借业务经验，进行预估哪些是热key

方法二:在客户端进行收集

方法三:在Proxy层做收集

方法四:用redis自带命令

(1)monitor命令，该命令可以实时抓取出redis服务器接收到的命令，然后写代码统计出热key是啥。当然，也有现成的分析工具可以给你使用，比如redis-faina。但是该命令在高并发的条件下，有内存增暴增的隐患，还会降低redis的性能。

(2)hotkeys参数，redis 4.0.3提供了redis-cli的热点key发现功能，执行redis-cli时加上–hotkeys选项即可。但是该参数在执行的时候，如果key比较多，执行起来比较慢。

方法五:自己抓包评估

## 如何解决

(1)利用二级缓存:比如利用ehcache，或者一个HashMap都可以。在你发现热key以后，把热key加载到系统的JVM中。

(2)备份热key:不要让key走到同一台redis上不就行了。我们把这个key，在多个redis上都存一份不就好了。接下来，有热key请求进来的时候，我们就在有备份的redis上随机选取一台，进行访问取值，返回数据。

## 业内方案

在客户端进行收集热key：可以使用TMC方案

TMC 对原生jedis包的JedisPool和Jedis类做了改造，在JedisPool初始化过程中集成TMC“热点发现”+“本地缓存”功能Hermes-SDK包的初始化逻辑。也就说人家改写了jedis原生的jar包，加入了Hermes-SDK包

那Hermes-SDK包用来干嘛？

OK，就是做热点发现和本地缓存。从监控的角度看，该包对于Jedis-Client的每次key值访问请求，Hermes-SDK 都会通过其通信模块将key访问事件异步上报给Hermes服务端集群，以便其根据上报数据进行“热点探测”。

# Redis妙用

redis是键值对的数据库，常用的五种数据类型为字符串类型（string），散列类型（hash），列表类型（list），集合类型（set），有序集合类型（zset）

Redis用作缓存，主要两个用途：高性能，高并发，因为内存天然支持高并发

## 应用场景

**分布式锁（string）**

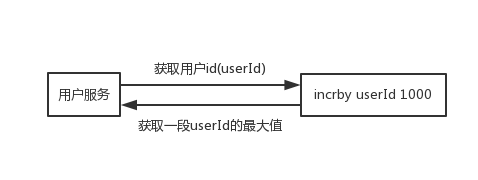
setnx key value，当key不存在时，将 key 的值设为 value ，返回1。若给定的 key 已经存在，则setnx不做任何动作，返回0。

当setnx返回1时，表示获取锁，做完操作以后del key，表示释放锁，如果setnx返回0表示获取锁失败，整体思路大概就是这样，细节还是比较多的，有时间单开一篇来讲解

**计数器（string）**

如知乎每个问题的被浏览器次数

**分布式全局唯一id（string）**



直接获取一段userId的最大值，缓存到本地慢慢累加，快到了userId的最大值时，再去获取一段，一个用户服务宕机了，也顶多一小段userId没有用到

**消息队列（list）**

在list里面一边进，一边出即可

**抽奖活动（set）**

# 参加抽奖活动:

sadd key {userId}

# 获取所有抽奖用户，大轮盘转起来:

smembers key

# 抽取count名中奖者，并从抽奖活动中移除:

spop key count

# 抽取count名中奖者，不从抽奖活动中移除:

srandmember key count

**实现点赞，签到，like等功能(set)**

# 1001用户给8001帖子点赞:

sadd like::8001 1001

# 取消点赞:

srem like::8001 1001

# 检查用户是否点过赞:

sismember like::8001 1001

# 获取点赞的用户列表:

smembers like::8001

# 获取点赞用户数:

scard like::8001

**实现关注模型，可能认识的人（set）**

seven关注的人

sevenSub -> {qing, mic, james}

青山关注的人

qingSub->{seven,jack,mic,james}

Mic关注的人

MicSub->{seven,james,qing,jack,tom}

# 返回sevenSub和qingSub的交集，即seven和青山的共同关注

sinter sevenSub qingSub -> {mic,james}

# 我关注的人也关注他,下面例子中我是seven

# qing在micSub中返回1，否则返回0

sismember micSub qing

sismember jamesSub qing

# 我可能认识的人,下面例子中我是seven

# 求qingSub和sevenSub的差集，并存在sevenMayKnow集合中

sdiffstore sevenMayKnow qingSub sevenSub -> {seven,jack}

**电商商品筛选（set）**

每个商品入库的时候即会建立他的静态标签列表如，品牌，尺寸，处理器，内存

# 将拯救者y700P-001和ThinkPad-T480这两个元素放到集合brand::lenovo

sadd brand::lenovo 拯救者y700P-001 ThinkPad-T480

sadd screenSize::15.6 拯救者y700P-001 机械革命Z2AIR

sadd processor::i7 拯救者y700P-001 机械革命X8TIPlus

# 获取品牌为联想，屏幕尺寸为15.6，并且处理器为i7的电脑品牌(sinter为获取集合的交集)

sinter brand::lenovo screenSize::15.6 processor::i7 -> 拯救者y700P-001

**排行版（zset）**

redis的zset天生是用来做排行榜的、好友列表, 去重, 历史记录等业务需求

# user1的用户分数为 10

zadd ranking 10 user1

zadd ranking 20 user2

取分数最高的3个用户

zrevrange ranking 0 2 withscores

## 过期策略

生产中一般采用定期删除+惰性删除两种策略

定期删除:redis 会将每个设置了过期时间的 key 放入到一个独立的字典中，以后会定期遍历这个字典来删除到期的 key。

定期删除策略

Redis 默认会每秒进行十次过期扫描（100ms一次），过期扫描不会遍历过期字典中所有的 key，而是采用了一种简单的贪心策略。从过期字典中随机 20 个 key；删除这 20 个 key 中已经过期的 key；如果过期的 key 比率超过 1/4，那就重复步骤 1；

惰性删除

除了定期遍历之外，它还会使用惰性策略来删除过期的 key，所谓惰性策略就是在客户端访问这个 key 的时候，redis 对 key 的过期时间进行检查，如果过期了就立即删除，不会给你返回任何东西。

为什么要采用定期删除+惰性删除2种策略呢？

如果过期就删除。假设redis里放了10万个key，都设置了过期时间，你每隔几百毫秒，就检查10万个key，那redis基本上就死了，cpu负载会很高的，消耗在你的检查过期key上了。

所以说用了上述2种策略后，下面这种现象就不难解释了：数据明明都过期了，但是还占有着内存

## 内存淘汰策略

这个问题可能有小伙伴们遇到过，放到Redis中的数据怎么没了？

因为Redis将数据放到内存中，内存是有限的，比如redis就只能用10个G，你要是往里面写了20个G的数据，会咋办？当然会干掉10个G的数据，然后就保留10个G的数据了。那干掉哪些数据？保留哪些数据？当然是干掉不常用的数据，保留常用的数据了

**Redis提供的内存淘汰策略有如下几种：**

1.noeviction 不会继续服务写请求 (DEL 请求可以继续服务)，读请求可以继续进行。这样可以保证不会丢失数据，但是会让线上的业务不能持续进行。这是默认的淘汰策略。

2.volatile-lru 尝试淘汰设置了过期时间的 key，最少使用的 key 优先被淘汰。没有设置过期时间的 key 不会被淘汰，这样可以保证需要持久化的数据不会突然丢失。（这个是使用最多的）

3.volatile-ttl 跟上面一样，除了淘汰的策略不是 LRU，而是 key 的剩余寿命 ttl 的值，ttl 越小越优先被淘汰。

4.volatile-random 跟上面一样，不过淘汰的 key 是过期 key 集合中随机的 key。

5.allkeys-lru 区别于 volatile-lru，这个策略要淘汰的 key 对象是全体的 key 集合，而不只是过期的 key 集合。这意味着没有设置过期时间的 key 也会被淘汰。

6.allkeys-random 跟上面一样，不过淘汰的策略是随机的 key。allkeys-random 跟上面一样，不过淘汰的策略是随机的 key。

## 持久化策略

Redis的数据是存在内存中的，如果Redis发生宕机，那么数据会全部丢失，因此必须提供持久化机制。

Redis 的持久化机制有两种，第一种是快照（RDB），第二种是 AOF 日志。快照是一次全量备份，AOF 日志是连续的增量备份。快照是内存数据的二进制序列化形式，在存储上非常紧凑，而 AOF 日志记录的是内存数据修改的指令记录文本。AOF 日志在长期的运行过程中会变的无比庞大，数据库重启时需要加载 AOF 日志进行指令重放，这个时间就会无比漫长。所以需要定期进行 AOF 重写，给 AOF 日志进行瘦身。

RDB是通过Redis主进程fork子进程，让子进程执行磁盘 IO 操作来进行 RDB 持久化，AOF 日志存储的是 Redis 服务器的顺序指令序列，AOF 日志只记录对内存进行修改的指令记录。即RDB记录的是数据，AOF记录的是指令

**RDB和AOF到底该如何选择？**

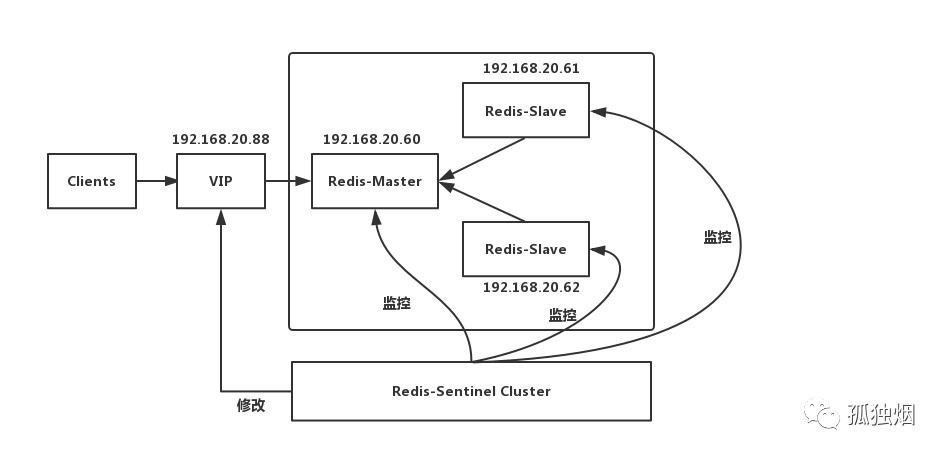
不要仅仅使用 RDB，因为那样会导致你丢失很多数据，因为RDB是隔一段时间来备份数据

也不要仅仅使用 AOF，因为那样有两个问题，第一，通过 AOF 做冷备没有RDB恢复速度快; 第二，RDB 每次简单粗暴生成数据快照，更加健壮，可以避免 AOF 这种复杂的备份和恢复机制的 bug

用RDB恢复内存状态会丢失很多数据，重放AOP日志又很慢。Redis4.0推出了混合持久化来解决这个问题。将 rdb 文件的内容和增量的 AOF 日志文件存在一起。这里的 AOF 日志不再是全量的日志，而是自持久化开始到持久化结束的这段时间发生的增量 AOF 日志，通常这部分 AOF 日志很小。于是在 Redis 重启的时候，可以先加载 rdb 的内容，然后再重放增量 AOF 日志就可以完全替代之前的 AOF 全量文件重放，重启效率因此大幅得到提升。

# 那些年用过的Redis集群架构

## Replication+Sentinel



这里Sentinel的作用有三个:

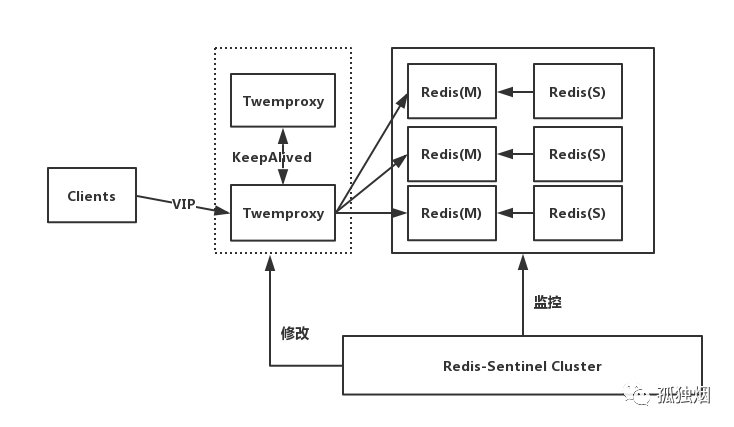
监控:Sentinel 会不断的检查主服务器和从服务器是否正常运行。

通知:当被监控的某个Redis服务器出现问题，Sentinel通过API脚本向管理员或者其他的应用程序发送通知。

自动故障转移:当主节点不能正常工作时，Sentinel会开始一次自动的故障转移操作，它会将与失效主节点是主从关系的其中一个从节点升级为新的主节点，并且将其他的从节点指向新的主节点。

工作原理如下，当Master宕机的时候，Sentinel会选举出新的Master，并根据Sentinel中client-reconfig-script脚本配置的内容，去动态修改VIP(虚拟IP)，将VIP(虚拟IP)指向新的Master。我们的客户端就连向指定的VIP即可！

## Proxy+Replication+Sentinel



工作原理如下

前端使用Twemproxy+KeepAlived做代理，将其后端的多台Redis实例分片进行统一管理与分配

每一个分片节点的Slave都是Master的副本且只读

Sentinel持续不断的监控每个分片节点的Master，当Master出现故障且不可用状态时，Sentinel会通知/启动自动故障转移等动作

Sentinel 可以在发生故障转移动作后触发相应脚本（通过 client-reconfig-script 参数配置 ），脚本获取到最新的Master来修改Twemproxy配置

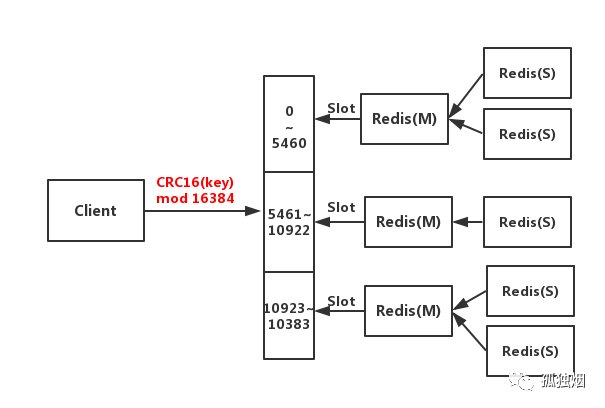
缺陷:

(1)部署结构超级复杂

(2)可扩展性差，进行扩缩容需要手动干预

(3)运维不方便

## Redis Cluster



工作原理如下

客户端与Redis节点直连,不需要中间Proxy层，直接连接任意一个Master节点

根据公式HASH\_SLOT=CRC16(key) mod 16384，计算出映射到哪个分片上，然后Redis会去相应的节点进行操作

具有如下优点:

(1)无需Sentinel哨兵监控，如果Master挂了，Redis Cluster内部自动将Slave切换Master

(2)可以进行水平扩容

(3)支持自动化迁移，当出现某个Slave宕机了，那么就只有Master了，这时候的高可用性就无法很好的保证了，万一Master也宕机了，咋办呢？ 针对这种情况，如果说其他Master有多余的Slave ，集群自动把多余的Slave迁移到没有Slave的Master 中。

缺点:

(1)批量操作是个坑

(2)资源隔离性较差，容易出现相互影响的情况。

## Redis Cluster面试问题

**问题1:懂Redis事务么？**

正常版：Redis事务是一些列redis命令的集合,blabla…

高调版: 我们在生产上采用的是Redis Cluster集群架构，不同的key是有可能分配在不同的Redis节点上的，在这种情况下Redis的事务机制是不生效的。其次，Redis事务不支持回滚操作，简直是鸡肋！所以基本不用！

**问题2:Redis的多数据库机制，了解多少？**

正常版：Redis支持多个数据库，并且每个数据库的数据是隔离的不能共享，单机下的redis可以支持16个数据库（db0 ~ db15）

高调版: 在Redis Cluster集群架构下只有一个数据库空间，即db0。因此，我们没有使用Redis的多数据库功能！

**问题3:Redis集群机制中，你觉得有什么不足的地方吗？**

正常版: 不知道

高调版: 假设我有一个key，对应的value是Hash类型的。如果Hash对象非常大，是不支持映射到不同节点的！只能映射到集群中的一个节点上！还有就是做批量操作比较麻烦！

**问题4:懂Redis的批量操作么？**

正常版: 懂一点。比如mset、mget操作等，blabla

高调版: 我们在生产上采用的是Redis Cluster集群架构，不同的key会划分到不同的slot中，因此直接使用mset或者mget等操作是行不通的。

**问题5:那在Redis集群模式下，如何进行批量操作？**

正常版:不知道

高调版:这个问题其实可以写一篇文章了，改天写。这里说一种有一个很简单的答法，足够面试用。即:

如果执行的key数量比较少，就不用mget了，就用串行get操作。如果真的需要执行的key很多，就使用Hashtag保证这些key映射到同一台Redis节点上。简单来说语法如下

对于key为{foo}.student1、{foo}.student2，{foo}student3，这类key一定是在同一个redis节点上。因为key中“{}”之间的字符串就是当前key的hash tags， 只有key中{ }中的部分才被用来做hash，因此计算出来的redis节点一定是同一个!

ps:如果你用的是Proxy分片集群架构，例如Codis这种，会将mget/mset的多个key拆分成多个命令发往不同得Redis实例，这里不多说。我推荐答的还是Redis Cluster。

**问题6:你们有对Redis做读写分离么？**

正常版:没有做，至于原因额。。。额。。。额。。没办法了，硬着头皮扯~

高调版:不做读写分离。我们用的是Redis Cluster的架构，是属于分片集群的架构。而Redis本身在内存上操作，不会涉及IO吞吐，即使读写分离也不会提升太多性能，Redis在生产上的主要问题是考虑容量，单机最多10-20G，key太多降低Redis性能.因此采用分片集群结构，已经能保证了我们的性能。其次，用上了读写分离后，还要考虑主从一致性，主从延迟等问题，徒增业务复杂度。