



下载APP



## 02 | 数据结构：快速的Redis有哪些慢操作？

2020-08-03 蒋德钧

Redis核心技术与实战

[进入课程 >](#)**讲述：蒋德钧**

时长 19:06 大小 17.50M



你好，我是蒋德钧。

一提到 Redis，我们的脑子里马上就会出现一个词：“快。”但是你有没有想过，Redis 的快，到底是快在哪里呢？实际上，这里有一个重要的表现：它接收到一个键值对操作后，能以**微秒级别**的速度找到数据，并快速完成操作。

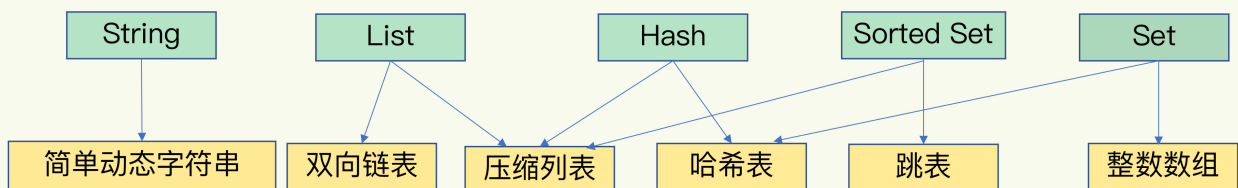
数据库这么多，为啥 Redis 能有这么突出的表现呢？一方面，这是因为它是内存数据库，所有操作都在内存上完成，内存的访问速度本身就很快。另一方面，这要归功于它的**数据结构**。这是因为，键值对是按一定的数据结构来组织的，操作键值对最终就是对数据结构进行增删改查操作，所以高效的数据结构是 Redis 快速处理数据的基础。这节课，我就来和你聊聊数据结构。



说到这儿，你肯定会说：“这个我知道，不就是 String（字符串）、List（列表）、Hash（哈希）、Set（集合）和 Sorted Set（有序集合）吗？”其实，这些只是 Redis 键值对中值的数据类型，也就是数据的保存形式。而这里，我们说的数据结构，是要去看看它们的底层实现。

简单来说，底层数据结构一共有 6 种，分别是简单动态字符串、双向链表、压缩列表、哈希表、跳表和整数数组。它们和数据类型的对应关系如下图所示：

### Redis数据类型和底层数据结构的对应关系



可以看到，String 类型的底层实现只有一种数据结构，也就是简单动态字符串。而 List、Hash、Set 和 Sorted Set 这四种数据类型，都有两种底层实现结构。通常情况下，我们会把这四种类型称为集合类型，它们的特点是**一个键对应了一个集合的数据**。

看到这里，其实有些问题已经值得我们去考虑了：

这些数据结构都是值的底层实现，键和值本身之间用什么结构组织？

为什么集合类型有那么多底层结构，它们都是怎么组织数据的，都很快吗？

什么是简单动态字符串，和常用的字符串是一回事吗？

接下来，我就和你聊聊前两个问题。这样，你不仅可以知道 Redis “快” 的基本原理，还可以借此理解 Redis 中有哪些潜在的“慢操作”，最大化 Redis 的性能优势。而关于简单动态字符串，我会在后面的课程中再和你讨论。

我们先来看看键和值之间是用什么结构组织的。

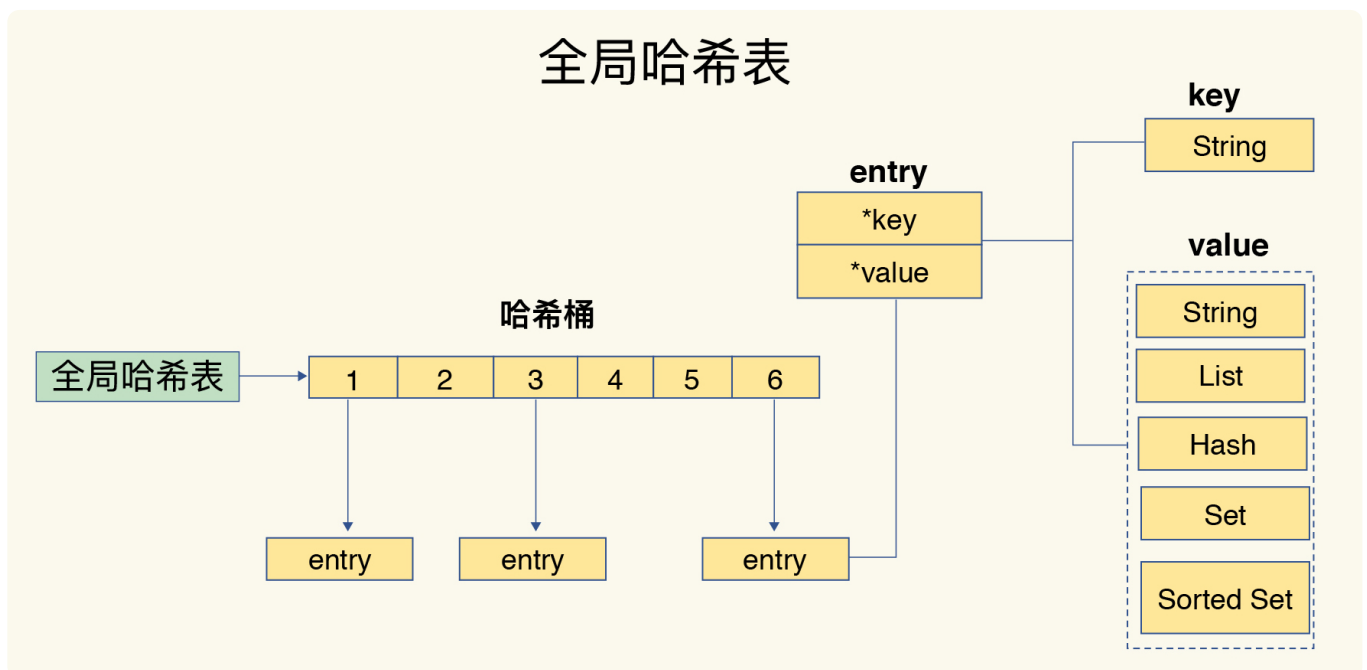
### 键和值用什么结构组织？

为了实现从键到值的快速访问，Redis 使用了一个哈希表来保存所有键值对。

一个哈希表，其实就是一个数组，数组的每个元素称为一个哈希桶。所以，我们常说，一个哈希表是由多个哈希桶组成的，每个哈希桶中保存了键值对数据。

看到这里，你可能会问了：“如果值是集合类型的话，作为数组元素的哈希桶怎么来保存呢？”其实，哈希桶中的元素保存的并不是值本身，而是指向具体值的指针。这也就是说，不管值是 String，还是集合类型，哈希桶中的元素都是指向它们的指针。

在下图中，可以看到，哈希桶中的 entry 元素中保存了 \*key 和 \*value 指针，分别指向了实际的键和值，这样一来，即使值是一个集合，也可以通过 \*value 指针被查找到。



因为这个哈希表保存了所有的键值对，所以，我也把它称为**全局哈希表**。哈希表的最大好处很明显，就是让我们可以用  $O(1)$  的时间复杂度来快速查找到键值对——我们只需要计算键的哈希值，就可以知道它所对应的哈希桶位置，然后就可以访问相应的 entry 元素。

你看，这个查找过程主要依赖于哈希计算，和数据量的多少并没有直接关系。也就是说，不管哈希表里有 10 万个键还是 100 万个键，我们只需要一次计算就能找到相应的键。

但是，如果你只是了解了哈希表的  $O(1)$  复杂度和快速查找特性，那么，当你往 Redis 中写入大量数据后，就可能发现操作有时候会突然变慢了。这其实是因为你忽略了一个潜在的风险点，那就是**哈希表的冲突问题和 rehash 可能带来的操作阻塞**。

## 为什么哈希表操作变慢了？

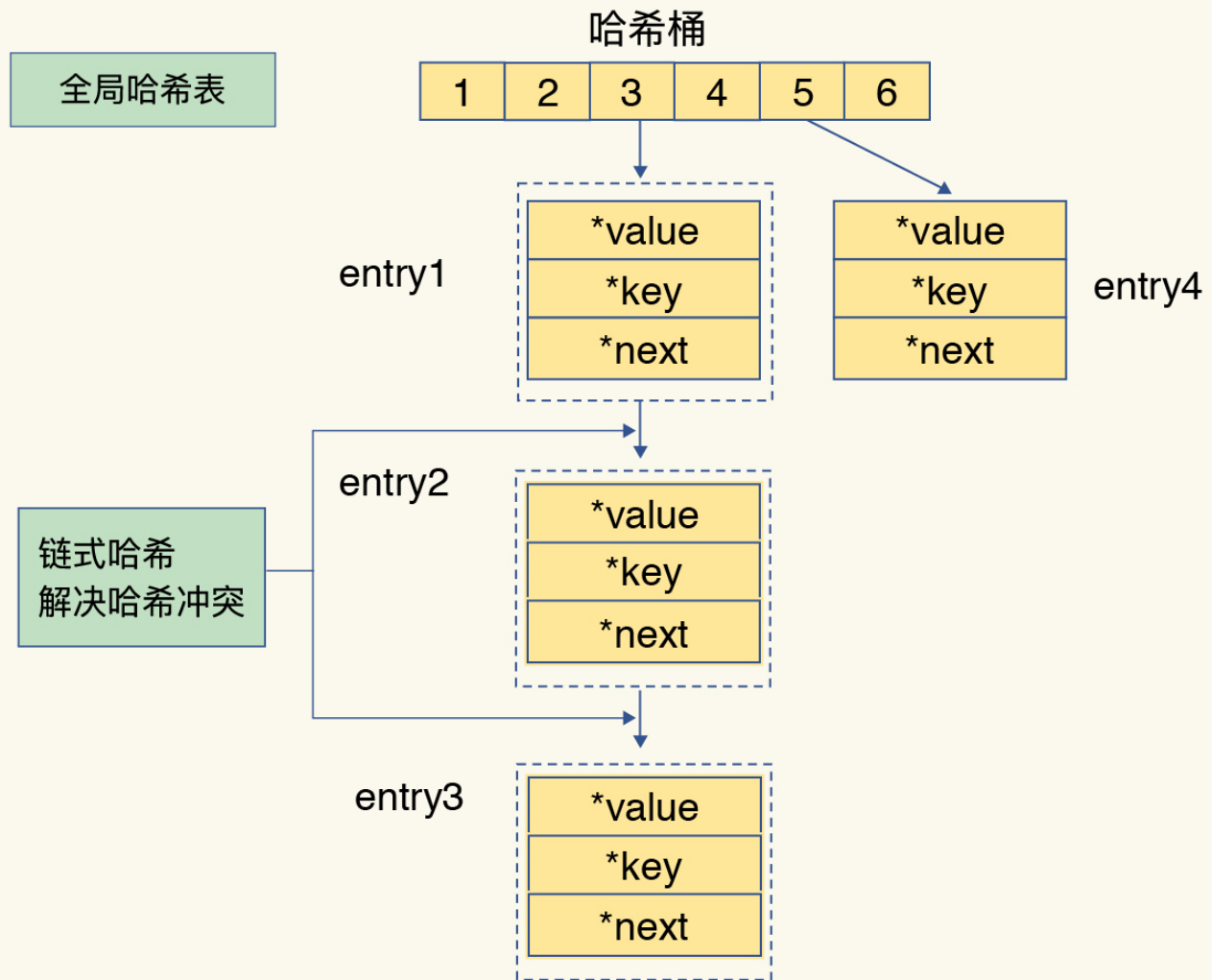
当你往哈希表中写入更多数据时，哈希冲突是不可避免的问题。这里的哈希冲突，也就是指，两个 key 的哈希值和哈希桶计算对应关系时，正好落在了同一个哈希桶中。

毕竟，哈希桶的个数通常要少于 key 的数量，这也就是说，难免会有一些 key 的哈希值对应到了同一个哈希桶中。

Redis 解决哈希冲突的方式，就是链式哈希。链式哈希也很容易理解，就是指**同一个哈希桶中的多个元素用一个链表来保存，它们之间依次用指针连接。**

如下图所示：entry1、entry2 和 entry3 都需要保存在哈希桶 3 中，导致了哈希冲突。此时，entry1 元素会通过一个\*next指针指向 entry2，同样，entry2 也会通过\*next指针指向 entry3。这样一来，即使哈希桶 3 中的元素有 100 个，我们也可以通过 entry 元素中的指针，把它们连起来。这就形成了一个链表，也叫作哈希冲突链。

# 哈希表的哈希冲突



但是，这里依然存在一个问题，哈希冲突链上的元素只能通过指针逐一查找再操作。如果哈希表里写入的数据越来越多，哈希冲突可能也会越来越多，这就会导致某些哈希冲突链过长，进而导致这个链上的元素查找耗时长，效率降低。对于追求“快”的 Redis 来说，这是不太能接受的。

所以，Redis 会对哈希表做 rehash 操作。rehash 也就是增加现有的哈希桶数量，让逐渐增多的 entry 元素能在更多的桶之间分散保存，减少单个桶中的元素数量，从而减少单个桶中的冲突。那具体怎么做呢？

其实，为了使 rehash 操作更高效，Redis 默认使用了两个全局哈希表：哈希表 1 和哈希表 2。一开始，当你刚插入数据时，默认使用哈希表 1，此时的哈希表 2 并没有被分配空间。随着数据逐步增多，Redis 开始执行 rehash，这个过程分为三步：



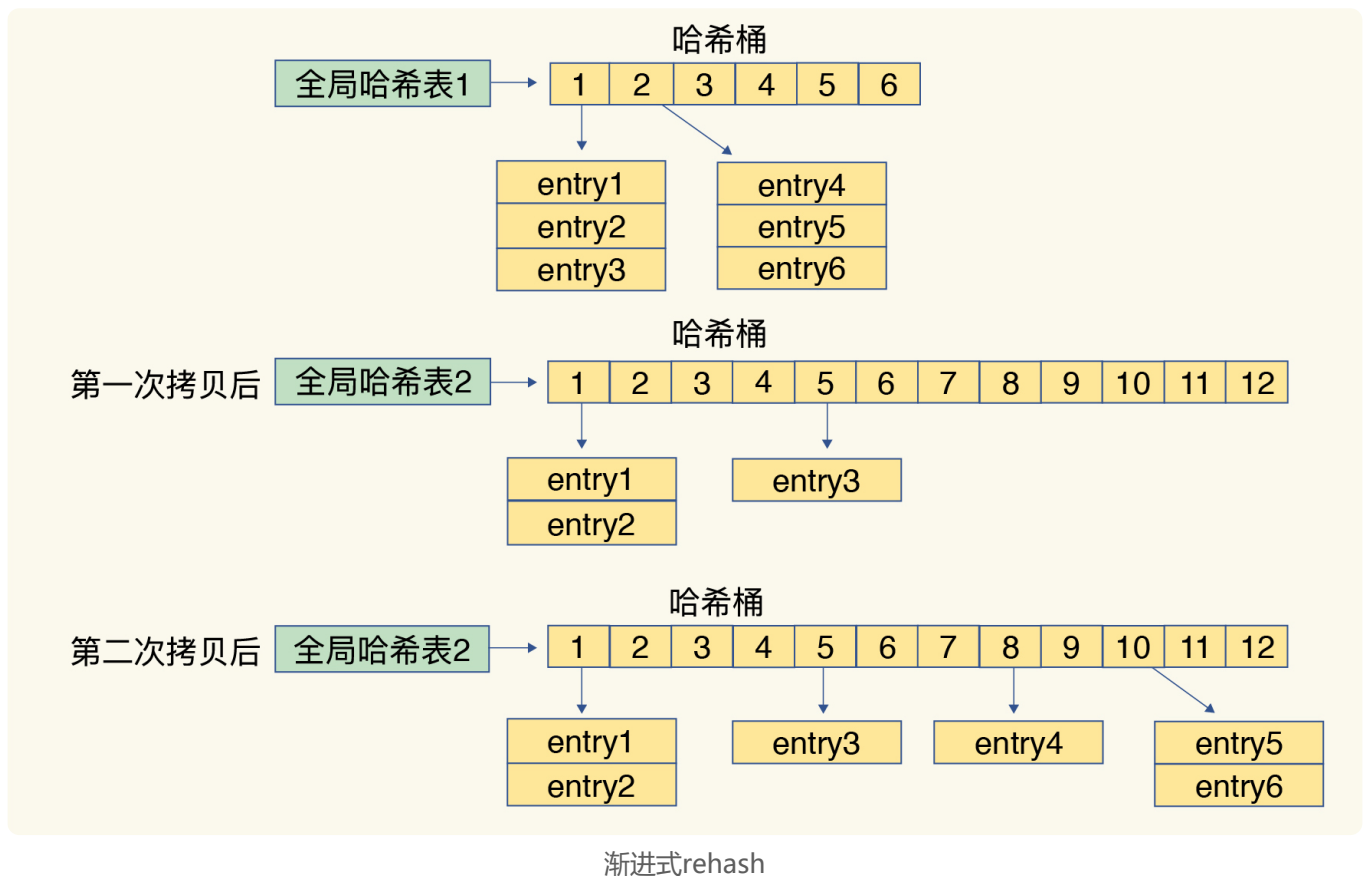
1. 给哈希表 2 分配更大的空间，例如是当前哈希表 1 大小的两倍；
2. 把哈希表 1 中的数据重新映射并拷贝到哈希表 2 中；
3. 释放哈希表 1 的空间。

到此，我们就可以从哈希表 1 切换到哈希表 2，用增大的哈希表 2 保存更多数据，而原来的哈希表 1 留作下一次 rehash 扩容备用。

这个过程看似简单，但是第二步涉及大量的数据拷贝，如果一次性把哈希表 1 中的数据都迁移完，会造成 Redis 线程阻塞，无法服务其他请求。此时，Redis 就无法快速访问数据了。

为了避免这个问题，Redis 采用了**渐进式 rehash**。

简单来说就是在第二步拷贝数据时，Redis 仍然正常处理客户端请求，每处理一个请求时，从哈希表 1 中的第一个索引位置开始，顺带着将这个索引位置上的所有 entries 拷贝到哈希表 2 中；等处理下一个请求时，再顺带拷贝哈希表 1 中的下一个索引位置的 entries。如下图所示：



这样就巧妙地把一次性大量拷贝的开销，分摊到了多次处理请求的过程中，避免了耗时操作，保证了数据的快速访问。

好了，到这里，你应该就能理解，Redis 的键和值是怎么通过哈希表组织的了。对于 String 类型来说，找到哈希桶就能直接增删改查了，所以，哈希表的  $O(1)$  操作复杂度也就是它的复杂度了。

但是，对于集合类型来说，即使找到哈希桶了，还要在集合中再进一步操作。接下来，我们来看集合类型的操作效率又是怎样的。

## 集合数据操作效率

和 String 类型不同，一个集合类型的值，第一步是通过全局哈希表找到对应的哈希桶位置，第二步是在集合中再增删改查。那么，集合的操作效率和哪些因素相关呢？

首先，与集合的底层数据结构有关。例如，使用哈希表实现的集合，要比使用链表实现的集合访问效率更高。其次，操作效率和这些操作本身的执行特点有关，比如读写一个元素的操作要比读写所有元素的效率高。

接下来，我们就分别聊聊集合类型的底层数据结构和操作复杂度。

## 有哪些底层数据结构？

刚才，我也和你介绍过，集合类型的底层数据结构主要有 5 种：整数数组、双向链表、哈希表、压缩列表和跳表。

其中，哈希表的操作特点我们刚刚已经学过了；整数数组和双向链表也很常见，它们的操作特征都是顺序读写，也就是通过数组下标或者链表的指针逐个元素访问，操作复杂度基本是  $O(N)$ ，操作效率比较低；压缩列表和跳表我们平时接触得可能不多，但它们也是 Redis 重要的数据结构，所以我来重点解释一下。

压缩列表实际上类似于一个数组，数组中的每一个元素都对应保存一个数据。和数组不同的是，压缩列表在表头有三个字段 `zlbytes`、`zltail` 和 `zllen`，分别表示列表长度、列表尾的偏移量和列表中的 `entry` 个数；压缩列表在表尾还有一个 `zlend`，表示列表结束。

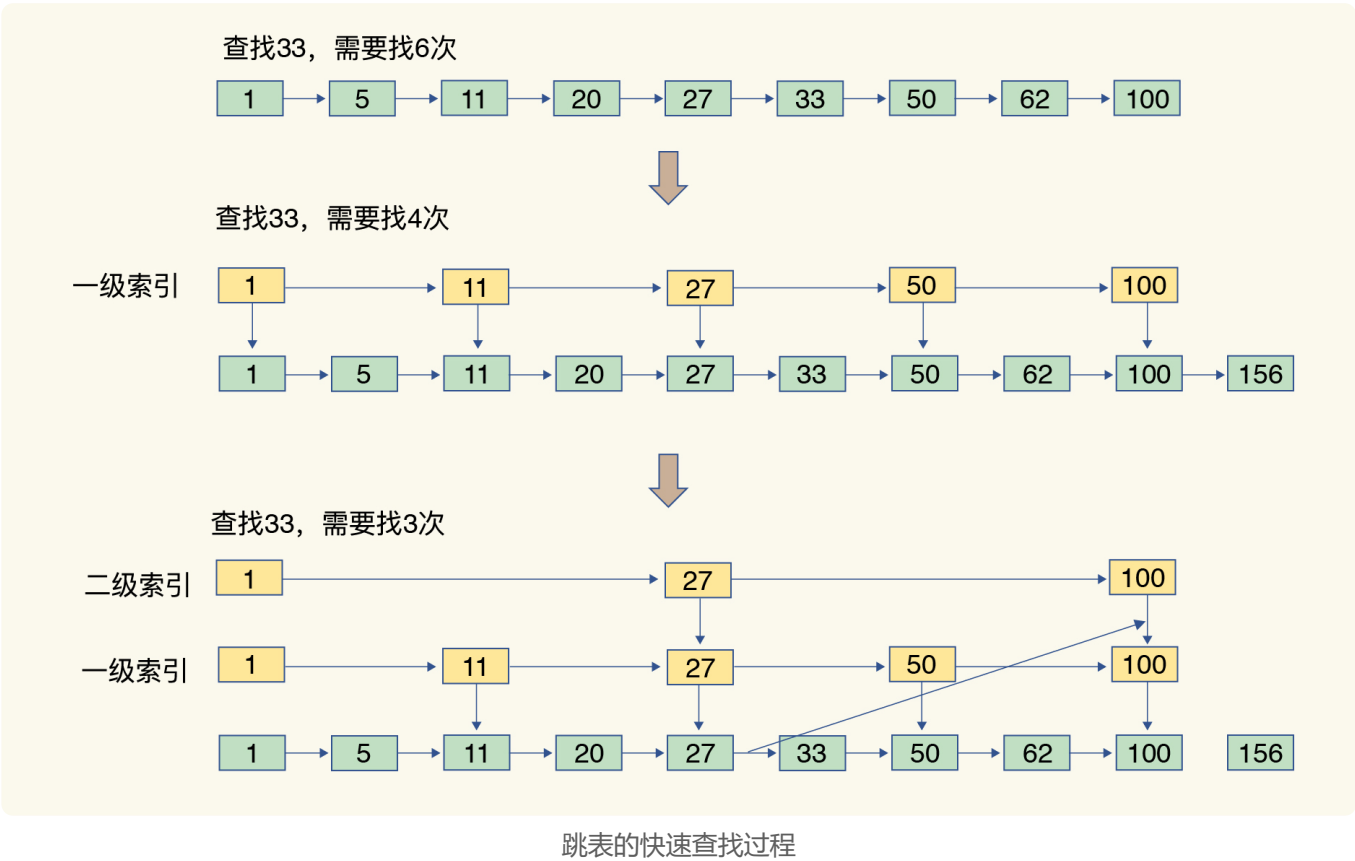
# 压缩列表的查找

zlbytes	zltail	zllen	entry1	entry2	...	entryN	zlend
---------	--------	-------	--------	--------	-----	--------	-------

在压缩列表中，如果我们要查找定位第一个元素和最后一个元素，可以通过表头三个字段的长度直接定位，复杂度是  $O(1)$ 。而查找其他元素时，就没有这么高效了，只能逐个查找，此时的复杂度就是  $O(N)$  了。

我们再来看下跳表。

有序链表只能逐一查找元素，导致操作起来非常缓慢，于是就出现了跳表。具体来说，跳表在链表的基础上，**增加了多级索引，通过索引位置的几个跳转，实现数据的快速定位**，如下图所示：



如果我们要在链表中查找 33 这个元素，只能从头开始遍历链表，查找 6 次，直到找到 33 为止。此时，复杂度是  $O(N)$ ，查找效率很低。



为了提高查找速度，我们来增加一级索引：从第一个元素开始，每两个元素选一个出来作为索引。这些索引再通过指针指向原始的链表。例如，从前两个元素中抽取元素 1 作为一级索引，从第三、四个元素中抽取元素 11 作为一级索引。此时，我们只需要 4 次查找就能定位到元素 33 了。

如果我们还想再快，可以再增加二级索引：从一级索引中，再抽取部分元素作为二级索引。例如，从一级索引中抽取 1、27、100 作为二级索引，二级索引指向一级索引。这样，我们只需要 3 次查找，就能定位到元素 33 了。

可以看到，这个查找过程就是在多级索引上跳来跳去，最后定位到元素。这也正好符合“跳”表的叫法。当数据量很大时，跳表的查找复杂度就是  $O(\log N)$ 。

好了，我们现在可以按照查找的时间复杂度给这些数据结构分下类了：

## 数据结构的时间复杂度

名称	时间复杂度
哈希表	$O(1)$
跳表	$O(\log N)$
双向链表	$O(N)$
压缩列表	$O(N)$
整数数组	$O(N)$

### 不同操作的复杂度

集合类型的操作类型很多，有读写单个集合元素的，例如 HGET、HSET，也有操作多个元素的，例如 SADD，还有对整个集合进行遍历操作的，例如 SMEMBERS。这么多操作，它们的复杂度也各不相同。而复杂度的高低又是我们选择集合类型的重要依据。

我总结了一个“四句口诀”，希望能帮助你快速记住集合常见操作的复杂度。这样你在使用过程中，就可以提前规避高复杂度操作了。

单元素操作是基础；

范围操作非常耗时；

统计操作通常高效；

例外情况只有几个。

**第一，单元素操作，是指每一种集合类型对单个数据实现的增删改查操作。**例如，Hash 类型的 HGET、HSET 和 HDEL，Set 类型的 SADD、SREM、SRANDMEMBER 等。这些操作的复杂度由集合采用的数据结构决定，例如，HGET、HSET 和 HDEL 是对哈希表做操作，所以它们的复杂度都是  $O(1)$ ；Set 类型用哈希表作为底层数据结构时，它的 SADD、SREM、SRANDMEMBER 复杂度也是  $O(1)$ 。

这里，有个地方你需要注意一下，集合类型支持同时对多个元素进行增删改查，例如 Hash 类型的 HMGET 和 HMSET，Set 类型的 SADD 也支持同时增加多个元素。此时，这些操作的复杂度，就是由单个元素操作复杂度和元素个数决定的。例如，HMSET 增加  $M$  个元素时，复杂度就从  $O(1)$  变成  $O(M)$  了。

**第二，范围操作，是指集合类型中的遍历操作，可以返回集合中的所有数据，**比如 Hash 类型的 HGETALL 和 Set 类型的 SMEMBERS，或者返回一个范围内的部分数据，比如 List 类型的 LRANGE 和 ZSet 类型的 ZRANGE。**这类操作的复杂度一般是  $O(N)$ ，比较耗时，我们应该尽量避免。**

不过，Redis 从 2.8 版本开始提供了 SCAN 系列操作（包括 HSCAN，SSCAN 和 ZSCAN），这类操作实现了渐进式遍历，每次只返回有限数量的数据。这样一来，相比于 HGETALL、SMEMBERS 这类操作来说，就避免了一次性返回所有元素而导致的 Redis 阻塞。

**第三，统计操作，是指集合类型对集合中所有元素个数的记录，**例如 LLEN 和 SCARD。这类操作复杂度只有  $O(1)$ ，这是因为当集合类型采用压缩列表、双向链表、整数数组这些数据结构时，这些结构中专门记录了元素的个数统计，因此可以高效地完成相关操作。

第四，例外情况，是指某些数据结构的特殊记录，例如**压缩列表和双向链表都会记录表头和表尾的偏移量**。这样一来，对于 List 类型的 LPOP、RPOP、LPUSH、RPUSH 这四个操作来说，它们是在列表的头尾增删元素，这就可以通过偏移量直接定位，所以它们的复杂度也只有  $O(1)$ ，可以实现快速操作。

## 小结

这节课，我们学习了 Redis 的底层数据结构，这既包括了 Redis 中用来保存每个键和值的全局哈希表结构，也包括了支持集合类型实现的双向链表、压缩列表、整数数组、哈希表和跳表这五大底层结构。

Redis 之所以能快速操作键值对，一方面是因为  $O(1)$  复杂度的哈希表被广泛使用，包括 String、Hash 和 Set，它们的操作复杂度基本由哈希表决定，另一方面，Sorted Set 也采用了  $O(\log N)$  复杂度的跳表。不过，集合类型的范围操作，因为要遍历底层数据结构，复杂度通常是  $O(N)$ 。这里，我的建议是：**用其他命令来替代**，例如可以用 SCAN 来代替，避免在 Redis 内部产生费时的全集合遍历操作。

当然，我们不能忘了复杂度较高的 List 类型，它的两种底层实现结构：双向链表和压缩列表的操作复杂度都是  $O(N)$ 。因此，我的建议是：**因地制宜地使用 List 类型**。例如，既然它的 POP/PUSH 效率很高，那么就将它主要用于 FIFO 队列场景，而不是作为一个可以随机读写的集合。

Redis 数据类型丰富，每个类型的操作繁多，我们通常无法一下子记住所有操作的复杂度。所以，最好的办法就是**掌握原理，以不变应万变**。这里，你可以看到，一旦掌握了数据结构基本原理，你可以从原理上推断不同操作的复杂度，即使这个操作你不一定熟悉。这样一来，你不用死记硬背，也能快速合理地做出选择了。

## 每课一问

整数数组和压缩列表在查找时间复杂度方面并没有很大的优势，那为什么 Redis 还会把它们作为底层数据结构呢？

数据结构是了解 Redis 性能的必修课，如果你身边还有不太清楚数据结构的朋友，欢迎你把今天的内容分享给他 / 她，期待你在留言区和我交流讨论。

42 人觉得很赞 | 提建议

# Redis 核心技术与实战

## 从原理到实战，彻底吃透 Redis

蒋德钧  
中科院计算所副研究员



新版升级：点击「👤 请朋友读」，20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 01 | 基本架构：一个键值数据库包含什么？

下一篇 03 | 高性能IO模型：为什么单线程Redis能那么快？

### 精选留言 (63)

[写留言](#)

Kaito

2020-08-04

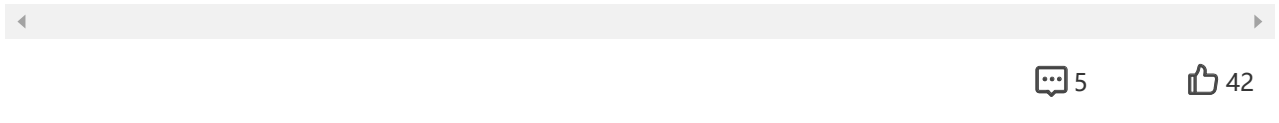
两方面原因：

1、内存利用率，数组和压缩列表都是非常紧凑的数据结构，它比链表占用的内存要更少。Redis是内存数据库，大量数据存到内存中，此时需要做尽可能的优化，提高内存的利用率。...

展开 ∨

作者回复: 回答的很好！

我再提个小问题，如果在数组上是随机访问，对CPU高速缓存还友好不？：)

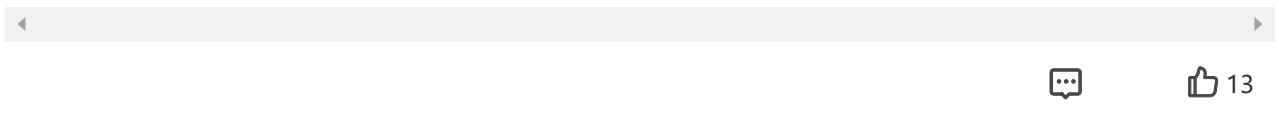


樱花落花

2020-08-03

Redis的List底层使用压缩列表本质上是将所有元素紧挨着存储，所以分配的是一块连续的内存空间，虽然数据结构本身没有时间复杂度的优势，但是这样节省空间而且也能避免一些内存碎片；

作者回复: 赞！所以不同的设计选择都是有背后的考虑。



刘忽悠

2020-08-03

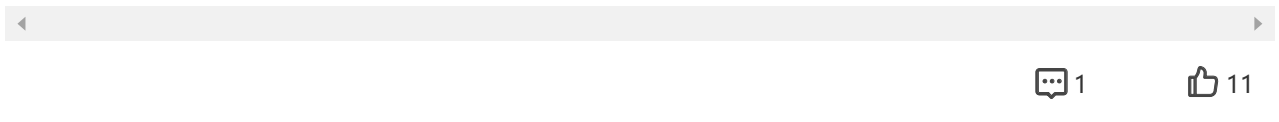
redis底层的数据压缩搞的很细致，像sds，根据字节长度划分的很细致，另外采用的c99特性的动态数组，对短字符串进行一次性的内存分配；跳表设计的也很秀，简单明了，进行范围查询很方便；dict的扩容没细看，但是看了一下数据结构，应该是为了避免发生扩容的时候出现整体copy；

个人觉得老师应该把sds，dict等具体数据结构的源码也贴上

展开 ∨

作者回复: dict的渐进式rehash是为了避免扩容时的整体拷贝，这会给内存带来较大压力。

SDS我们后面还会再聊：)



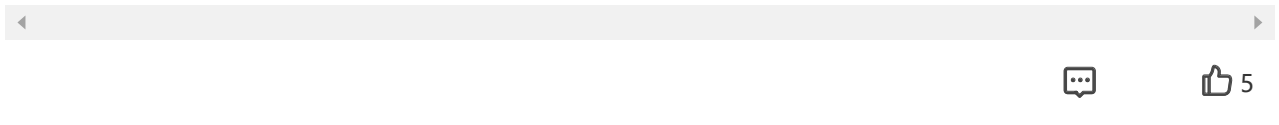
刘忽悠

2020-08-03

至于问题答案，采用压缩列表或者是整数集合，都是数据量比较小的情况，所以一次能够分配到足够大的内存，而压缩列表和整数集合本身的数据结构也是线性的，对cpu的缓存更友好一些，所以真正的执行的时间因为高速缓存的关系，速度更快

展开 ∨

作者回复: 提个小问题哈，如果压缩列表也是随机跳着访问，对CPU缓存还友好不？



**dingjiayi**

2020-08-05

老师好，请问 rehash 期间，新的 操作请求（增删改查）到达，redis 是如何处理的？



3



4

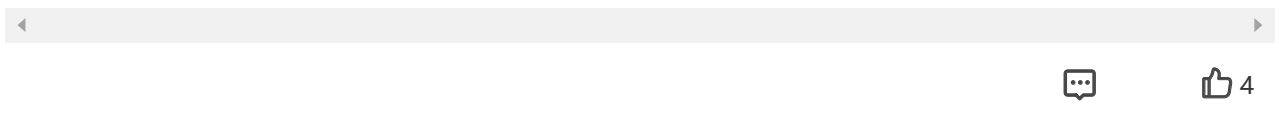


**米虫**

2020-08-04

要是加餐中来一偏，一个redis指令的执行过程，那大局观就更深刻了。

作者回复: 我们后面的课程会涉及这个过程，再耐心等等哈



**陈靖**

2020-08-10

老师的项目，是要把这些数据结构全部撸一遍吗？

展开 ∨



2



**那么、神秘美丽**

2020-08-05

引用---简单来说就是在第二步拷贝数据时，Redis 仍然正常处理客户端请求，每处理一个请求时，从哈希表 1 中的第一个索引位置开始，顺带着将这个索引位置上的所有 entries 拷贝到哈希表 2 中；等处理下一个请求时，再顺带拷贝哈希表 1 中的下一个索引位置的 entries。

提问---渐进式 rehash 的时候，数据是同时存在于哈希表1和哈希表2吗？这时候，我要...

展开 ∨



2



**yyl**

2020-08-04

问题：渐进式rehash依赖新的请求，如果接下来没有接收到新请求，rehash就不进行了吧？



4



2



**曾弼麟**

2020-08-03

intset 和 ziplist 如果直接使用确实是时间复杂度上不是很高效，但是结合Redis的使用场景，大部分Redis的数据都是零散碎片化的，通过这两种数据结构可以提高内存利用率，但是为了防止过度使用这两种数据结构Redis其实都设有阈值或者搭配使用的，例如：ziplist是和quicklist一起使用的，在quicklist中ziplist是一个个节点，而且quicklist为每个节点的大小都设置有阈值避免单个节点过大，从而导致性能下降

展开 ∨



2

**Geek\_c37e49**

2020-08-14

我是小白一个

读写一篇有两个疑问

- 1、以前了解数据结构的时候学习的都是数组的访问时间复杂度应该是 $O(1)$ ，不知道这篇文章里面的 $O(n)$ 是咋来的？
- 2、每次学习哈希表扩容的时候，大家都是讲的首先2倍的扩容。请问这个是经验值还是...

展开 ∨



1

**夜空中最亮的星 (华仔...)**

2020-08-13

想提几个问题：

- 1 redis 在接受了一个键值数据是怎么一步步存储的？
- 2 是把键和值 做哈希运算 得到哈希值 然后按一定规律 放入哈希桶？
- 3 一个哈希桶 可以存储多个键值对，那么这多对数据又是如何 存储和查找的？

展开 ∨



1

**牛牛**

2020-08-13

老师、早安,

今天的问题 @Kaito 同学回答的很好、尝试回答下: 数组上随机访问是否对cpu缓存友好？

数组对cpu缓存友好的原因是: cpu预读取一个cache line大小的数据, 数组数据排列紧凑、相同大小空间保存的元素更多, 访问下一个元素时、恰好已经在cpu缓存了. 如果是随机...

展开 ∨



1



yeek  
2020-08-11

渐进式 rehash操作，从 ht0 到 ht1 ，当ht1中的hash冲突也变大了，此时需要rehash，但是ht0 中还有数据没有迁移，此时redis如何操作？

展开 ∨



1



游弋云端

2020-08-11

Redis 采用了渐进式 rehash，那么什么时候进行新的全局Hash表的切换呢？当旧的Hash表格的负载因子达到一定值的时候？

展开 ∨



1



dingjiayi

2020-08-10

我觉得 rehash 部分的 “拷贝” 一次，改为 “移动” 更好



1



2020-08-06

读完这篇文章，才发现用了这么久的Redis，从来都没有真正去关注和学习它的底层东西，醍醐灌顶～

专栏后面还会讲到 bitmap、hyperloglog、streams 这三个值类型的使用吗？不知道实际使用场景是怎么样

展开 ∨

作者回复: 后面课程会讲的，一起学习交流：)



1



范闲

2020-08-17

在小数据量的情况下，提高内存使用率。  
同时数组和压缩列表的访问效率在小数据的时候并不低。

展开 ∨



1



Lywane

2020-08-17

rehash不仅全局有，单独的值为HASH类型的数据也会有吧



**x**

2020-08-16

redis的hash扩容跟golang中的map其实扩容机制竟然都是"渐进式扩容"！

作者回复: 我记得Golang里面的map扩容，好像是每次迁2个buckets，也是有key进行增、删或改时进行迁移。

