Redis数据结构

简单动态字符串

Simple Dynamic String 简称SDS

组成

- buf：字节数组，保存实际数据。为了表示字节数组的结束，Redis 会自动在数组最后加一个"\0"，这就会额外占用 1 个字节的开销。
- len：占4个字节，表示buf的已用长度，不包括最后一个\0
- alloc：占4个字节，表示buf实际分配的长度，一般大于len

优点

1. O(1) 时间内获取字符串的长度
2. 可以有效的防止缓冲区溢出
3. 减少修改字符串时带来的内存重新分配次数

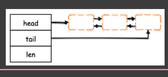
缺点

- 额外多了8字节的空间

链表

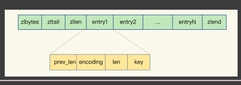
Redis 和 Linux 内核一样，也是采用双向链表实现，并没有使用单链表来节省内存

list 这一结构体主要包含了 3 个成员: 头指针、尾指针，以及双向链表所包含的节点数量



链表结构的应用非常广泛，包括数据量较小时的列表、发布订阅、慢查询和监视器等

压缩列表



组成

- zlbytes：列表长
- zltail：列表尾偏移量
- zllen：entry个数
- zlend：列表结束
- entry：列表元素，连续空间
  - prev_len，表示前一个 entry 的长度。prev_len 有两种取值情况：1 字节或 5 字节。取值 1 字节时，表示上一个 entry 的长度小于 254 字节。虽然 1 字节的值能表示的数值范围是 0 到 255，但是压缩列表中 zlend 的取值默认是 255，因此，就默认用 255 表示整个压缩列表的结束，其他表示长度的地方就不能再用 255 这个值了。所以，当上一个 entry 长度小于 254 字节时，prev_len 取值为 1 字节，否则，就取值为 5 字节。
  - len：表示自身长度，4 字节；
  - encoding：表示编码方式，1 字节；
  - content：保存实际数据。

压缩列表，ziplist 由上图所示组件构成，其目的就是为了节省内存。当一个 LIST 中只包含少量的元素，并且每一个元素要么是小正数，要么是长度很短的字符串，Redis 就会采用压缩列表来实现

理解压缩列表的一个关键点就在于，entry 中保存的数据大小是不固定的，也就是说，对于一个 ziplist 而言，我们既可以往里面扔 int，也可以往里面儿扔 string

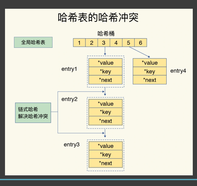
RPUSH ziplist 1 2 3 "Hello" "World"

entry 中保存的是数据，而非指针

因此，如果我们向 ziplist 中添加、删除、查询等操作时，平均时间复杂度均为 O(N)，并且添加和删除这两个操作的最坏时间复杂度为 O(N^2)，不过没关系，反正 ziplist 只会在小数据量时使用，即使是 O(N^2) 也不会有太大的性能问题

哈希表

采用哈希表实现，并且处理哈希冲突的方式为拉链法

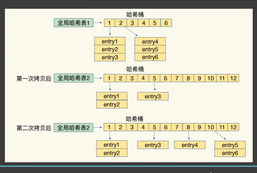


渐进式 rehash

为什么使用渐进式 rehash?

Redis 底层网络模型为单线程+epoll，也就是说，发送至 Redis 的命令是串行执行的

若采用直接扩容的方式，那么当遇到巨大的哈希表时，数据的复制将花费很长时间，势必会阻塞其它命令的执行，从而导致整体性能下降

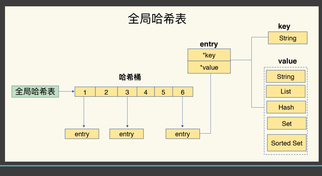


在渐进式 rehash 过程中，会同时存在 ht[1]、ht[2] 两张哈希表，并且同时对外提供服务

过程

1. 当我们新增一个 key 时，新的 key-value 将会保存在 ht[2] 中，也就是那个容量更大的哈希表中
2. 当我们查找、更新以及删除某一个 key 时，会在两张哈希表上进行，而不仅仅只是对 ht[1] 或者是 ht[2] 中的某一个哈希表进行操作

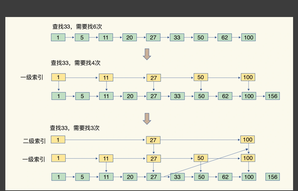
将扩容过程均摊，避免了耗时操作，保证了数据的快速访问。



Redis 本身的数据库也是使用哈希表实现的，所以才能在 O(1) 时间内查找到某一个 key

跳表

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



跳跃表是实现有序集合底层数据结构之一，当有序集合中的元素较多时，Redis 就会选择跳跃表来实现该集合

整数数组

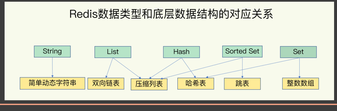
整数集合的底层实现为数组，并且元素在该数组中以有序、无重复的方式排列，通过名字我们就知道它能够实现集合，并且是整数集合

整数集合中我们可以保存 3 种类型的整数，分别为 16 位、32 位和 64 位的整数

这么实现的原因还是那个初衷, 节省内存。因为在一般情况下，程序不知道用户会保存多大的整数，所以为了避免整数溢出，可能会直接使用 64 位来保存。如果用户保存的都是小整数的话，那么就会有很多的内存浪费

Redis 出于这个考虑对 intset 进行动态编码，如果当前集合能够用 int16_t 来保存，那就用 int16_t。如果用户后续新增了一个 32 位整数，那么 Redis 会对 intset 进行升级，改用 int32_t 来保存

Redis数据类型



对应关系

数据类型	聚合统计	排序统计	二值状态统计	基数统计
Set	支持差集、交集、并集计算	不支持	不支持	精确统计，大数据量时，效率低，内存开销大
Sorted Set	支持交集、并集计算	支持		
Hash	不支持	不支持		
List	不支持	支持		不支持，元素没有去重
Bitmap	与、或、异或计算	不支持	支持，大数据量时，效率高，省内存	精确统计，大数据量时，内存开销大于HyperLogLog
HyperLogLog	不支持	不支持	不支持	概率统计，大数据量时，非常节省内存

使用场景

集合类型

特殊类型

- 记录经纬度信息使用GEO
- 时序数据使用 RedisTimeSeries