

Redis HyperLogLog

Redis 在 2.8.9 版本添加了 HyperLogLog 结构。

Redis HyperLogLog 是用来做基数统计的算法，HyperLogLog 的优点是，在输入元素的数量或者体积非常非常大时，计算基数所需的空间总是固定 的、并且是很小的。

在 Redis 里面，每个 HyperLogLog 键只需要花费 12 KB 内存，就可以计算接近 2^64 个不同元素的基 数。这和计算基数时，元素越多耗费内存就越多的集合形成鲜明对比。

但是，因为 HyperLogLog 只会根据输入元素来计算基数，而不会储存输入元素本身，所以 HyperLogLog 不能像集合那样，返回输入的各个元素。

什么是基数？

比如数据集 {1, 3, 5, 7, 5, 7, 8}，那么这个数据集的基数集为 {1, 3, 5 ,7, 8}, 基数(不重复元素)为5。 基数估计就是在误差可接受的范围内，快速计算基数。

实例

以下实例演示了 HyperLogLog 的工作过程：

```
redis 127.0.0.1:6379> PFADD runoobkey "redis"

1) (integer) 1

redis 127.0.0.1:6379> PFADD runoobkey "mongodb"

1) (integer) 1

redis 127.0.0.1:6379> PFADD runoobkey "mysql"

1) (integer) 1

redis 127.0.0.1:6379> PFCOUNT runoobkey

(integer) 3
```

Redis HyperLogLog 命令

下表列出了 redis HyperLogLog 的基本命令：

序号 命令及描述	
1	PFADD key element [element ...] 添加指定元素到 HyperLogLog 中。

2	PFCOUNT key_[key...] 返回给定 HyperLogLog 的基数估算值。
3	PFMERGE destkey sourcekey_[sourcekey...] 将多个 HyperLogLog 合并为一个 HyperLogLog

← Redis Zscan 命令

Redis Pfadd 命令 →

✎ 点我分享笔记