https://mp.weixin.qq.com/s/5eedsvW-z8k8YzdTK5SvmA

# Redis 如何从海量数据中查询出某一个 Key？

## 1 考察知识点

本题考察的知识点有以下几个：

1、Keys 和 Scan 的区别

2、Keys 查询的缺点

3、Scan 如何使用？

4、Scan 查询的特点

## 2 解答思路

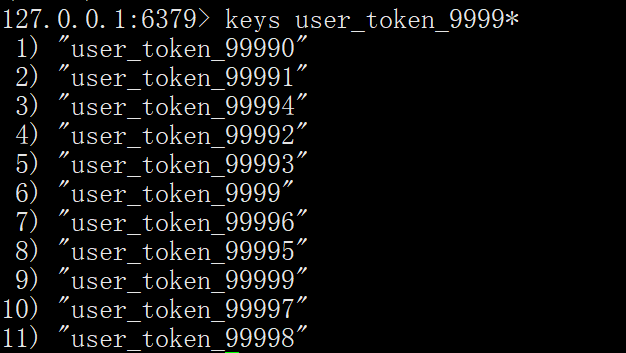
1、Keys 查询存在的问题

2、Scan 的使用

3、Scan 的特点

## 3 Keys 使用相关

### Keys 用法如下



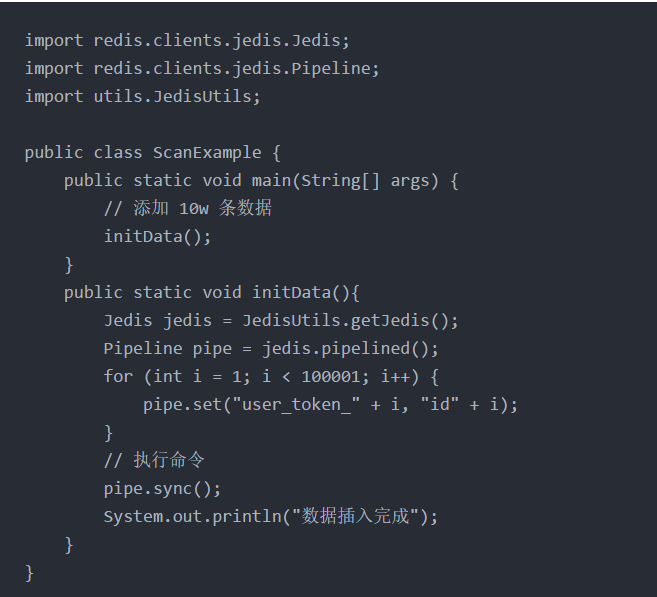
### 2）Keys 存在的问题

1、此命令没有分页功能，我们只能一次性查询出所有符合条件的 key 值，如果查询结果非常巨大，那么得到的输出信息也会非常多；

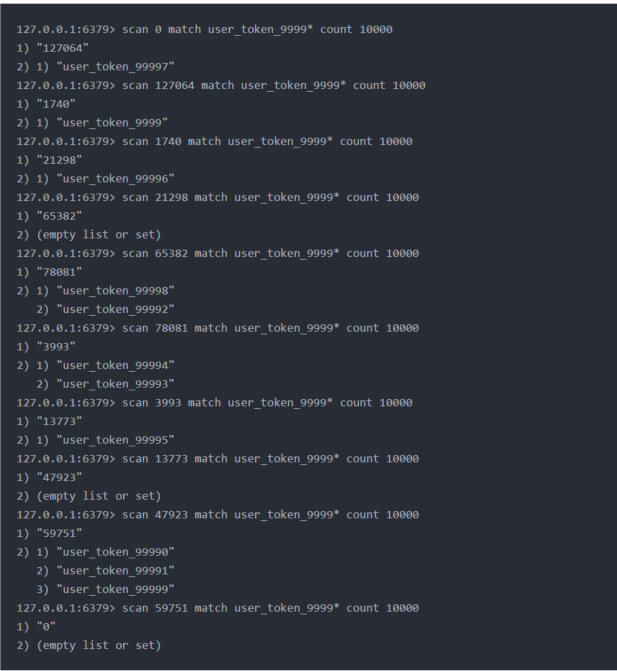
2、keys 命令是遍历查询，因此它的查询时间复杂度是 o(n)，所以数据量越大查询时间就越长。

## 4 Scan 使用相关

我们先来模拟海量数据，使用 Pipeline 添加 10w 条数据，Java 代码实现如下：



我们来查询用户 id 为 9999\* 的数据，Scan 命令使用如下：



scan命令介绍:[http://doc.redisfans.com/key/scan.html#match](http://doc.redisfans.com/key/scan.html" \l "match)

从以上的执行结果，我们看出两个问题：

1、查询的结果为空，但游标值不为 0，表示遍历还没结束；

2、设置的是 count 10000，但每次返回的数量都不是 10000，且不固定，这是因为 count 只是限定服务器单次遍历的字典槽位数量 (约等于)，而不是规定返回结果的 count 值。

**相关语法：scan cursor [MATCH pattern] [COUNT count]**

其中：

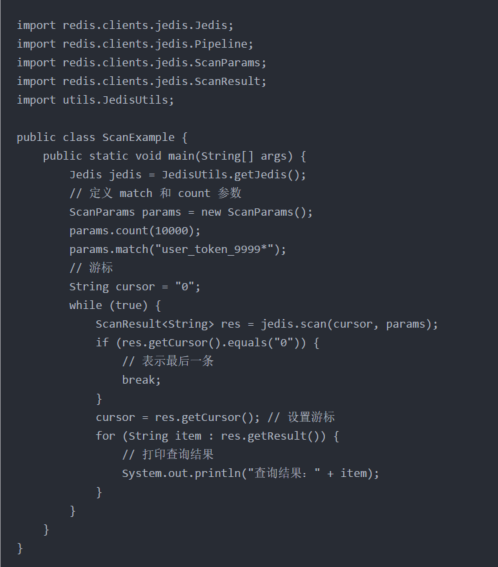
cursor：光标位置，整数值，从 0 开始，到 0 结束，查询结果是空，但游标值不为 0，表示遍历还没结束；

match pattern：正则匹配字段；

count：限定服务器单次遍历的字典槽位数量 (约等于)，只是对增量式迭代命令的一种提示 (hint)，并不是查询结果返回的最大数量，它的默认值是 10。

## 5 Scan 代码实战

本文我们使用 Java 代码来实现 Scan 的查询功能，代码如下：



以上程序执行结果如下：



## 6 总结

通过本文我们了解到，Redis 中如果要在海量的数据数据中，查询某个数据应该使用 Scan，Scan 具有以下特征：

1、Scan 可以实现 keys 的匹配功能；

2、Scan 是通过游标进行查询的不会导致 Redis 假死；

3、Scan 提供了 count 参数，可以规定遍历的数量；

4、Scan 会把游标返回给客户端，用户客户端继续遍历查询；

5、Scan 返回的结果可能会有重复数据，需要客户端去重；

6、单次返回空值且游标不为 0，说明遍历还没结束；

7、Scan 可以保证在开始检索之前，被删除的元素一定不会被查询出来；

8、在迭代过程中如果有元素被修改， Scan 不保证能查询出相关的元素。