

Hbase 学习（十） HBase Snapshots

HBase Snapshots允许你对一个表进行快照（即可用副本），它不会对Region Servers产生很大的影响，它进行复制和 恢复操作的时候不包括数据拷贝。导出快照到另外的集群也不会对Region Servers产生影响。 下面告诉你如何使用Snapshots功能

1.开启快照支持功能，在0.95+之后的版本都是默认开启的，在0.94.6+是默认关闭

```
<property>
  <name>hbase.snapshot.enabled</name>
  <value>true</value>
</property>
```

2.给表建立快照，不管表是启用或者禁用状态，这个操作不会进行数据拷贝

```
$ ./bin/hbase shell
hbase> snapshot 'myTable', 'myTableSnapshot-122112'
```

3.列出已经存在的快照

```
$ ./bin/hbase shell
hbase> list_snapshots
```

4.删除快照

```
$ ./bin/hbase shell
hbase> delete__snapshot 'myTableSnapshot-122112'
```

5.从快照复制生成一个新表

```
$ ./bin/hbase shell
hbase> clone_snapshot 'myTableSnapshot-122112', 'myNewTestTable'
```

6.用快照恢复数据，它需要先禁用表，再进行恢复

```
$ ./bin/hbase shell
hbase> disable 'myTable'
hbase> restore_snapshot 'myTableSnapshot-122112'
```

提示：因为备份（replication）是系统日志级别的，而快照是文件系统级别的，当使用快照恢复之后，副本会和master出于不同的状态，如果你需要使用恢复的话，你要停止备份，并且重置bootstrap。

如果是因为不正确的客户端行为导致数据丢失，全表恢复又需要表被禁用，可以采用快照生成一个新表，然后从新表中把需要的数据用map-reduce拷贝到主表当中。

7.复制到别的集群当中

该操作要用hbase的账户执行，并且在hdfs当中要有hbase的账户建立的临时目录（hbase.tmp.dir参数控制）

采用16个mappers来把一个名为MySnapshot的快照复制到一个名为srv2的集群当中

```
$ bin/hbase class org.apache.hadoop.hbase.snapshot.tool.ExportSnapshot -snapshot MySnapshot -copy-to hdfs://srv2:8020/hbase -mappers 16
```