

# HBase HBCK 简介

---

作者: 康凯森

日期: 2016-04-03

分类: HBase ( ../tag/HBase.html)

---

## hbck是做什么的

---

hbck 是一个检查和修复表，region一致性和完整性的工具

## 什么是region的一致性

---

- HDFS的region目录下的.regioninfo 文件 有效
- meta表行的数据和.regioninfo 文件的数据一致
- 每个region部署在唯一的regionserver,并被分配给状态合适的master

## hbck如何检查region的一致性

---

region一致性是通过检查hbase:meta表，region server上部署的 region，HDFS上的.regioninfo 文件三者是否全部一致来判断的。

## hcbk如何修复region的一致性

---

( region的一致性修复要求HBase在线以便hbck可以和master和regionserver沟通。 )

通过hbase:meta表和分配region来修复region的一致性

( 大多数region的一致性信息是短暂的 )

## 什么是表的一致性

---

扫描HDFS中表的每个region目录中的 .regioninfo 文件，每个表的一致性可以凭此来验证。

## hbck如何检查表的一致性

---

表的一致性是通过表的每一个region的所有可能rowkey来判断的。具体查看是否有空的，向后（不一致，过时）的region，region之间是否有空洞，region之间是否有重叠。

## hcbk如何修复表的一致性

---

对于表的一致性修复：会扫描HDFS中表的每个region目录中的 .regioninfo 文件。

如果有任何孤立的region（没有.regioninfo 文件）或者空洞，新的region就会被建立。

向后的region和空的退化的region（endkey==startkey）会被清除。

如果有任何重叠的region，一个新的region会被建立，并且所有数据都会合并到新的region。

## hcbk如何修复表的完整性

---

( 表的完整性修复仅仅需要通过HDFS离线完成, HBase的master和regionserver可以不必运行。 ) 具体是通过改造hbase:meta表, 合并和新建region来完成的

---

## 赏

康凯森 (<https://www.bcmeng.com/>)