

Cassandra nodetool常用命令介绍

简介

nodetool是cassandra自带的外围工具，通过JMX可以动态修改当前进程内存数据，注意cassandra是无主对等架构，默认的命令是操作本机当前进程，例如repair，如果需要做全集群修复，需要在每台机器上执行对应的nodetool命令。



如果想及时了解Spark、Hadoop或者Hbase相关的文章，欢迎关注微信公众号：iteblog_hadoop

执行nodetool help命令可获得详细帮助信息，如下：

```
[root@Cassandra8c32GTest005 cassandra]# bin/nodetool help
usage: nodetool [(-pwf <passwordFilePath> | --password-file <passwordFilePath>)]
      [(-h <host> | --host <host>)] [(-u <username> | --username <username>)]
      [(-pw <password> | --password <password>)] [(-p <port> | --port <port>)]
      <command> [<args>]
```

The most commonly used nodetool commands are:

assassinate	Forcefully remove a dead node without re-replicating any data. Use as a last resort if you cannot remove node
bootstrap	Monitor/manage node's bootstrap process
cleanup	Triggers the immediate cleanup of keys no longer belonging to a node. By default, clean all keyspaces
clearsnapshot	Remove the snapshot with the given name from the given keyspaces. If no snapshotName is specified we will remove all snapshots
compact	Force a (major) compaction on one or more tables or user-defined compaction on given SSTables
compactionhistory	Print history of compaction
compactionstats	Print statistics on compactions
decommission	Decommission the *node I am connecting to*
describecluster	Print the name, snitch, partitioner and schema version of a cluster
describering	Shows the token ranges info of a given keyspace
disableautocompaction	Disable autocompaction for the given keyspace and table
disablebackup	Disable incremental backup
disablebinary	Disable native transport (binary protocol)
disablegossip	Disable gossip (effectively marking the node down)
.....	

命令非常多，本文中，我们也不会做简单英文翻译，重点说一些运维中常用到的命令工具，未涉及到的命令可以在日常工作中通过使用，会慢慢加深了解。

常用命令介绍

nodetool status：集群所有node 心跳状态

```
[root@Cassandra8c32GTest005 cassandra]# bin/nodetool status
Datacenter: datacenter1
=====
Status=Up/Down
|/ State=Normal/Leaving/Joining/Moving
-- Address      Load      Tokens   Owns (effective)  Host ID                               Rack
UN 192.168.1.105 1.63 MiB  256      100.0%           0892685d-c79f-4e5b-b9ac-80ae1fb2a5f6 rack1
UN 192.168.1.106 1.65 MiB  256      100.0%           6520eeee-5802-48b2-8ff2-3f89d7ebd5e3 rack1
```

nodetool info: 当前节点进程全局信息

```
[root@Cassandra8c32GTest005 cassandra]# bin/nodetool info
ID : 0892685d-c79f-4e5b-b9ac-80ae1fb2a5f6
Gossip active : true
Thrift active : true
Native Transport active: true
Load : 1.63 MiB
Generation No : 1565767843
Uptime (seconds) : 159165
Heap Memory (MB) : 630.35 / 7922.00
Off Heap Memory (MB) : 0.01
Data Center : datacenter1
Rack : rack1
Exceptions : 0
Key Cache : entries 172, size 19.91 KiB, capacity 100 MiB, 254 hits, 472 requests, 0.538 recent hit rate, 14400 save period in seconds
Row Cache : entries 0, size 0 bytes, capacity 0 bytes, 0 hits, 0 requests, NaN recent hit rate, 0 save period in seconds
Counter Cache : entries 0, size 0 bytes, capacity 50 MiB, 0 hits, 0 requests, NaN recent hit rate, 7200 save period in seconds
Chunk Cache : entries 18, size 1.12 MiB, capacity 480 MiB, 1000 misses, 2706 requests, 0.630 recent hit rate, NaN microseconds miss latency
Percent Repaired : 100.0%
```

Token : (invoke with -T/--tokens to see all 256 tokens)

加减机器，扩容，缩容

集群的扩容是非常常用的功能，加减机器了。先看看加机器,如果配置正确，节点正常启动后，会自动加入集群

参考如下文章，加节点 https://docs.datastax.com/en/dse/6.7/dse-admin/datastax_enterprise/operations/opsAddNodeToCluster.html

nodetool cleanup: 在老节点上，分别执行，删除非自己管理的数据

机器过保下线：

nodetool decommission:当前节点下线，会将当前节点数据拷贝到其他副本

nodetool removenode: 从其他副本节点拷贝数据到数据重分布后的目标节点，有数据不一致风险，用于当前节点不能重新拉起，提供数据读取服务。

nodetool assassinate:强制移除节点，无任何数据拷贝

加减盘后，让sst重分布，rebalance磁盘

relocatesstables：磁盘加减盘后，relocatesstables可让sst重新分布

nodetool netstats 监控数据迁移进度

数据安全,备份

nodetool enablebackup: 开启增量备份，data目录增加一个sstable文件的话，在backup目录会创建一个硬链接，可以异步备份到s3或者oss上

nodetool garbagecollect：主动做回收，用于磁盘空间告警场景。

nodetool snapshot：对全部的keyspace都做snapshot，按当前时间截取快照名，在table的data Dir/snapshots/\${ts}/存放数据文件

refresh：

上述备份后的sst如何restore呢，手动导入sstable到正确位置后，执行refresh，让进程可见

数据正确性及其他运维命令

repair：修数据用，让多副本数据一致

verify：数据巡检，检查sst内数据和crc是不是一致。

resetlocalschema：从集群其他节点同步schema，各个节点上表schema数据必须是一致的。

nodetool flush: memtable刷至本地磁盘，生成sstable

nodetool compact:对指定的sstables做一次major compaction

sstabledump工具往往需要上面工具生成的sst作为输入

rebuild_index：如果二级索引与base表不一致，可重建

排查问题相关

getendpoints:

如果集群比较大，不可能扫描所有日志，可通过partitionKey找到后端的node，方便定位问题。

```
[root@Cassandra8c32GTest004 cassandra]# bin/nodetool getendpoints ycsb usertable user7562145027629296129
```

192.168.1.105

192.168.1.106

tablehistograms: 查看表partition数据统计，可检查出某partition是否过大

```
[root@Cassandra8c32GTest004 cassandra]# bin/nodetool tablehistograms ycsb usertable ycsb/usertable histograms
```

Percentile	SSTables	Write Latency (micros)	Read Latency (micros)	Partition Size (bytes)	Cell Count
50%	0.00	0.00	0.00	1109	10
75%	0.00	0.00	0.00	1109	10
95%	0.00	0.00	0.00	1109	10
98%	0.00	0.00	0.00	1109	10
99%	0.00	0.00	0.00	1109	10
Min	0.00	0.00	0.00	125	0
Max	0.00	0.00	0.00	1109	10

toppartitions: 当节点负载比较高，输出一个top

partitions可以拿到当前访问比较热的partition，前端可以降低压力。

结语

其他命令多为调整运行参数，不是那么常用，需要的时候，通过帮助手册，搜索关键字就能找到对应的命令，不再赘述了。

钉钉群交流

为了营造一个开放的 Cassandra 技术交流，我们建立了微信群和钉钉群，为广大用户提供专业的技术分享及问答，定期在国内开展线下技术沙龙，专家技术直播，欢迎大家加入。



如果想及时了解Spark、Hadoop或者HBase相关的文章，欢迎关注微信公众号：iteblog_hadoop

钉钉群入群链接：<https://c.tb.cn/F3.ZRTY0o>

本博客文章除特别声明，全部都是原创！
原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。
本文链接: **【】**（ ）