

若泽数据@HBase生产调优参数

Author: 若泽数据@J哥 生产HBase生产经验总结, 如有疑问及时与本人交流

1. Linux参数

1.1 进程 文件数

```
echo "* soft nofile 960000" >> /etc/security/limits.conf
echo "* hard nofile 960000" >> /etc/security/limits.conf
echo "* soft nproc 960000" >> /etc/security/limits.conf
echo "* hard nproc 960000" >> /etc/security/limits.conf
```

重新登录, 检查是否生效

```
[root@ruozedata001 ~]# ulimit -a
open files                      (-n) 960000
max user processes              (-u) 960000
```

1.2 网络、内核、进程能拥有的最多内存区域

```
echo "net.core.somaxconn=32768" >> /etc/sysctl.conf
echo "kernel.threads-max=196605" >> /etc/sysctl.conf
echo "kernel.pid_max=196605" >> /etc/sysctl.conf
echo "vm.max_map_count=393210" >> /etc/sysctl.conf
```

#生效

```
sysctl -p
```

1.3 swap

```
more /etc/sysctl.conf | vm.swappiness
echo vm.swappiness = 0 >> /etc/sysctl.conf
```

#生效

```
sysctl -p
```

1.4 关闭大页面

```
echo never >> /sys/kernel/mm/transparent_hugepage/enabled
echo never >> /sys/kernel/mm/transparent_hugepage/defrag
```

1.5 vm.min_free_kbytes

```
vm.min_free_kbytes = 1048576 (1G, 8G或者更多, 取决于机器内存)
```

1.6 NUMA

```
vm.zone_reclaim_mode = 0
```

2. HDFS参数

```
dfs.namenode.handler.count : 256  
dfs.datanode.handler.count : 128  
dfs.datanode.max.transfer.threads: 12288  
dfs.socket.timeout 1800000ms  
dfs.client.socket-timeout 1800000ms
```

3. Zookeeper参数

略

4. HBase参数

regionserver gc参数:

```
-XX:+UseG1GC
-XX:G1NewSizePercent=30
-XX:G1MaxNewSizePercent=60
-XX:+ParallelRefProcEnabled
-XX:InitiatingHeapOccupancyPercent=65
-XX:-ResizePLAB
-XX:MaxGCPauseMillis=500
-XX:+UnlockDiagnosticVMOptions
-XX:+G1SummarizeConcMark
-XX:G1HeapRegionSize=32m
-XX:G1HeapWastePercent=20
-XX:ConcGCThreads=16
-XX:ParallelGCThreads=38
-XX:MaxTenuringThreshold=2
-XX:G1MixedGCCountTarget=64
-XX:G1OldCSetRegionThresholdPercent=5
-XX:MetaspaceSize=300M
-Dcom.sun.management.jmxremote.port=39999
-Dcom.sun.management.jmxremote.ssl=false
-Dcom.sun.management.jmxremote.authenticate=false
```

dfs.socket.timeout 1800000ms (有意思的参数, 务必要加, 否则扛不住写)

dfs.client.socket-timeout 1800000ms

zookeeper.session.timeout : 600000(10m, 默认3m)

hbase.hregion.max.filesize : 10G

hbase.hregion.memstore.flush.size : 256m

hbase.hregion.memstore.block.multiplier : 6

hbase.regionserver.global.memstore.upperLimit(hbase.regionserver.global.memstore.size): 0.45

hbase.regionserver.global.memstore.lowerLimit(hbase.regionserver.global.memstore.size.lower.limit): 0.91

hfile.block.cache.size : 0.35

hbase.regionserver.global.memstore.upperLimit+hfile.block.cache.size 不能超过0.8, 有oom风险, 约定俗成的。

hbase.hstore.compactionThreshold : 6 (默认3)

hbase.hstore.blockingStoreFiles : 21 (默认7)

hbase.hregion.majorcompaction : 0

hbase.regionserver.handler.count : 128

hbase.regionserver.metahandler.count : 64

regionserver heap memory: 48G

计算集群region数量的公式：

```
((RS Xmx) * hbase.regionserver.global.memstore.size) / (hbase.hregion.memstore.flush.size * (# column families))  
==>48G*0.45/256=86个
```

5. 官网

<http://hbase.apache.org/book.html#performance>