# hadoop集群优化

### linux参数

1. 文件句柄数限制，用户进程数限制

查看当前配置：ulimit -a

修改参数：/etc/security/limits.d/90-nproc.conf

\* soft nofile 32768

\* hard nofile 65536

\* soft noproc 10240

\* hard noproc 10240

2、关闭swap分区，降低使用swap的可能性

查看当前配置：cat /proc/sys/vm/swappiness

修改参数：/etc/sysctl.conf

vm.swappiness = 0

立即生效：sysctl –p

1. socket连接backlog限制

查看当前配置：cat /proc/sys/net/core/somaxconn

修改参数：/etc/sysctl.conf

net.core.somaxconn = 10240

1. 禁用noatime和nodiratime属性，可以禁止记录最近一次访问时间戳

文件挂载时设置这两个属性可以明显提高性能

1. 预读缓冲区，默认为256，可以提高顺序读文件的性能

blockdev --setra 1024 /dev/sda

1. 磁盘RAID和LVM，避免在DataNode的机器上执行RAID和LVM操作

### JVM参数

NameNode内存计算公式

节点数 \* 单个节点磁盘容量（MB）/ 块大小（MB） \* 副本数

例如：3个节点，每个节点2TB，块大小128MB，3个副本

3 \* 2 \* 1024 \* 1024 / 128 \* 3 = 16GB

### hadoop参数

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 文件 | 参数名 | 参数值 | 默认值 | 说明 |
| core-site.xml | ipc.server.listen.queue.size | 10240 | 128 | backlog大小 |
| io.compression.codecs | org.apache.hadoop.io.compress.GzipCodec,  org.apache.hadoop.io.compress.DefaultCodec,  org.apache.hadoop.io.compress.BZip2Codec,  org.apache.hadoop.io.compress.SnappyCodec |  |  |
| fs.trash.interval | 10080 | 0 | 回收站文件保留时长（分钟） |
| fs.trash.checkpoint.interval | 60 | 0 |  |
| hdfs-site.xml | dfs.blocksize | 128MB | 128MB |  |
| dfs.namenode.handler.count | 30 | 10 | log(n) \* 20 |
| dfs.datanode.handler.count | 20 | 10 |  |
| dfs.datanode.du.reserved | 10737418240 | 0 | 50G |
| dfs.datanode.max.transfer.threads | 8192 | 4096 |  |
| mapred-site.xml | mapreduce.task.io.sort.mb | 512 | 100 |  |
| mapreduce.task.io.sort.factor | 100 | 10 |  |
| mapreduce.map.sort.spill.percent |  |  |  |
| mapreduce.task.timeout |  |  |  |
| mapred.child.java.opts | -Xmx1024m |  |  |
| mapreduce.reduce.shuffle.parallelcopies | 50 | 5 |  |