# 角色分配

|  |  |
| --- | --- |
| 主机 | 服务进程 |
| master01 | ResourceManager、NameNode、HMaster、hive、DFSZKFailoverController、kylin |
| master02 | AzkabanSingleServer、NameNode、JobHistoryServer、HMaster、DFSZKFailoverController、ResourceManager |
| node01 | NodeManager、HRegionServer、JournalNode、QuorumPeerMain、DataNode |
| node02 | NodeManager、HRegionServer、JournalNode、QuorumPeerMain、DataNode |
| node03 | NodeManager、HRegionServer、JournalNode、QuorumPeerMain、DataNode |

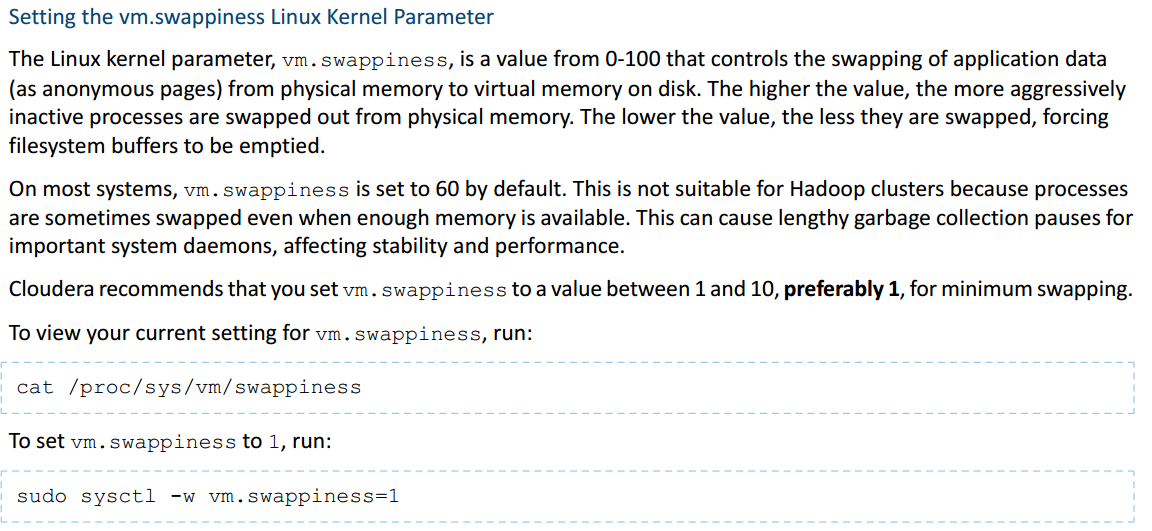
# 操作系统

## 系统安装

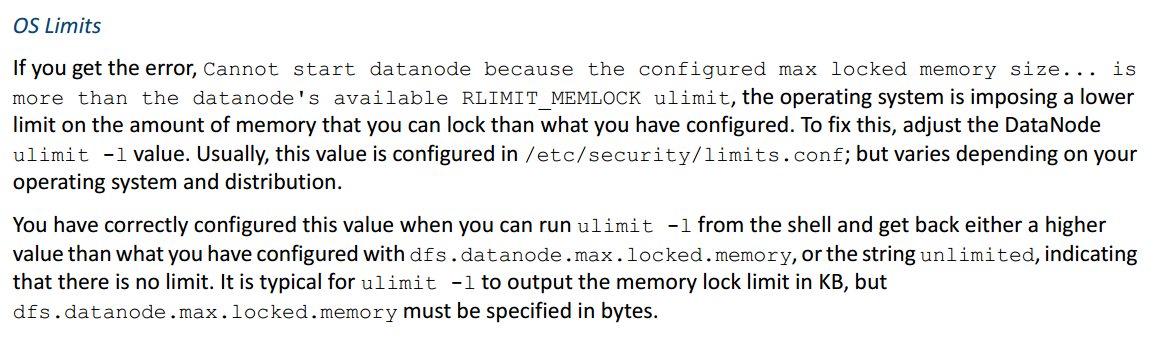
cpu、内存、文件系统

## 内核参数

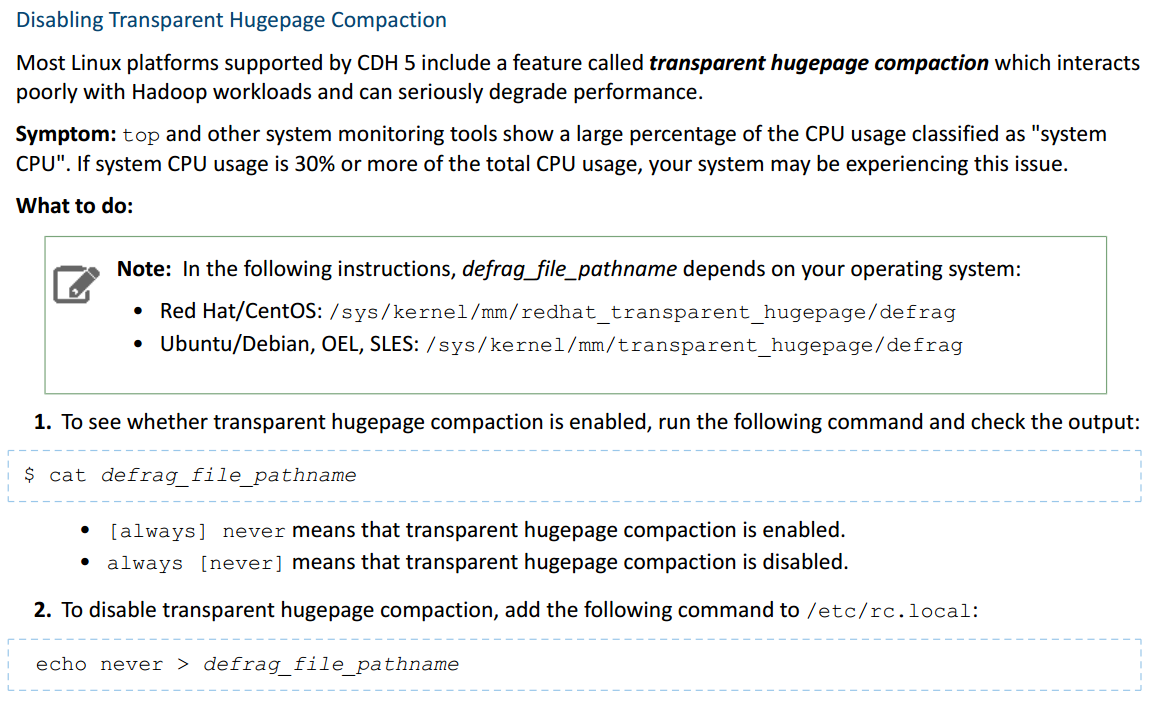
### 禁用swap



### 修改ulimt



### 关闭 THP



## 网络配置

Hadoop、Hive、HBase、zookeeper、Kylin、mysql、azkaban所需要端口

防火墙规则

<http://www.zhixing123.cn/ubuntu/40649.html>

<http://blog.csdn.net/wulantian/article/details/46341043>

# 集群

## 安装Ansible

**在master节点上安装**

yum install -y epel-release.noarch

yum install -y ansible

如出现一下错误信息

**"msg": "Aborting, target uses selinux but python bindings (libselinux-python) aren't installed!"**

**则在各节点上安装**

yum install -y libselinux-python

## 创建hadoop用户并配置免密码

### 创建用户

在master01上执行

#echo csdn.net | openssl passwd -1 -stdin

$1$Ljkoi/W3$EcrRquLD6w/vMefdZEg8..

#ansible -i all\_hosts all -m user -a 'name=hadoop password="$1$Ljkoi/W3$EcrRquLD6w/vMefdZEg8.."'

### 免密码

Master01:

ssh-keygen

ssh-copy-id master01

ansible-playbook -i all\_hosts playbooks/authorize\_keys.yml -e "master\_host=master01"

Master02

ssh-keygen

ssh-copy-id master02

ansible-playbook -i all\_hosts playbooks/authorize\_keys.yml -e "master\_host=master02"

## 安装zookeeper

ansible-playbook -i production playbooks/zookeeper.yml --become-user=root --become-method=su -b --ask-su-pass

## 安装hadoop

ansible-playbook -i production playbooks/hadoop.yml --become-user=root --become-method=sudo -b --ask-su-pass

**华为提供设备是16 Core 32G**

**CPU： 给 OS和服务进程预留 4 Core ,给YARN提供12Core**

**内存：Hbase 12G 、yarn 10G、服务进程 6G、OS 4G**

注意配置 YARN的CPU和内存

在production文件中的

yarn\_memory\_mb=8192 (华为部署配置到10G)

yarn\_cpu\_vcores=4 (华为部署配置到12 Core)

## 安装HBase

ansible-playbook -i production playbooks/hbase.yml --become-user=root --become-method=sudo -b --ask-su-pass

注意配置hbase heapsize

在production文件中的

hbase\_heapsize=4096 (华为部署配置到12G)

## 安装Hive

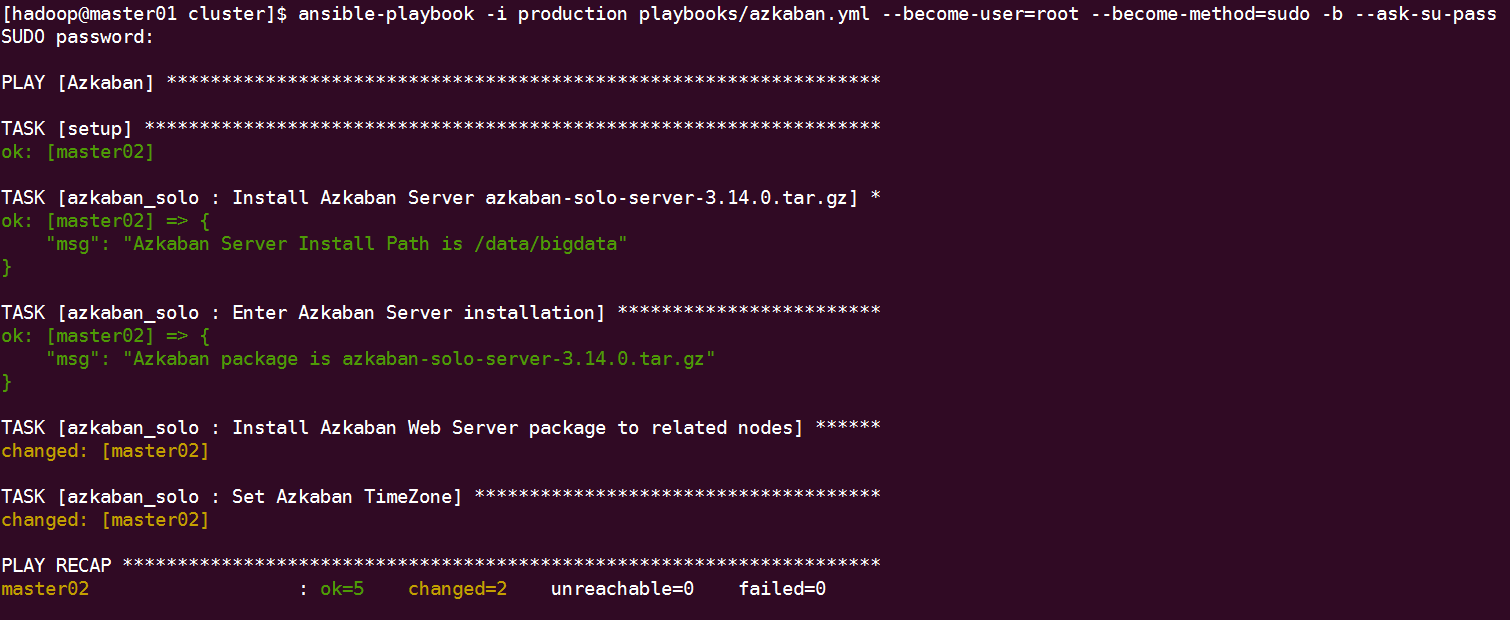
ansible-playbook -i production playbooks/hive2.yml --become-user=root --become-method=sudo -b --ask-su-pass

## 安装Kylin

ansible-playbook -i production playbooks/kylin.yml --become-user=root --become-method=sudo -b --ask-su-pass

## 安装Azkaban

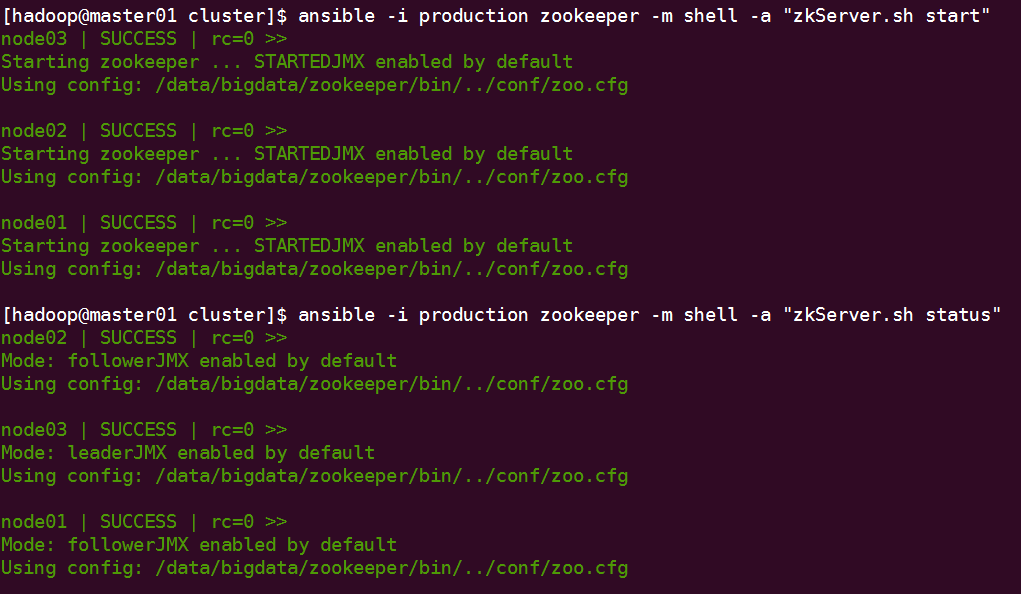
ansible-playbook -i production playbooks/azkaban.yml --become-user=root --become-method=sudo -b --ask-su-pass



## 启动zookeeper

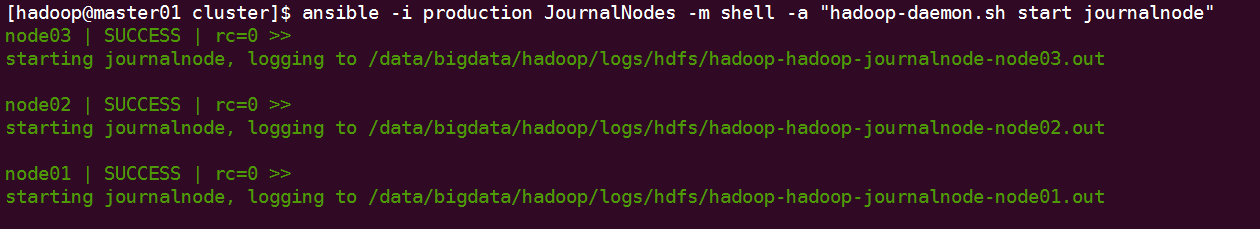
ansible -i production zookeeper -m shell -a "zkServer.sh start"

ansible -i production zookeeper -m shell -a "zkServer.sh status"



## 启动journalnode

ansible -i production JournalNodes -m shell -a "hadoop-daemon.sh start journalnode"



## 格式化HDFS

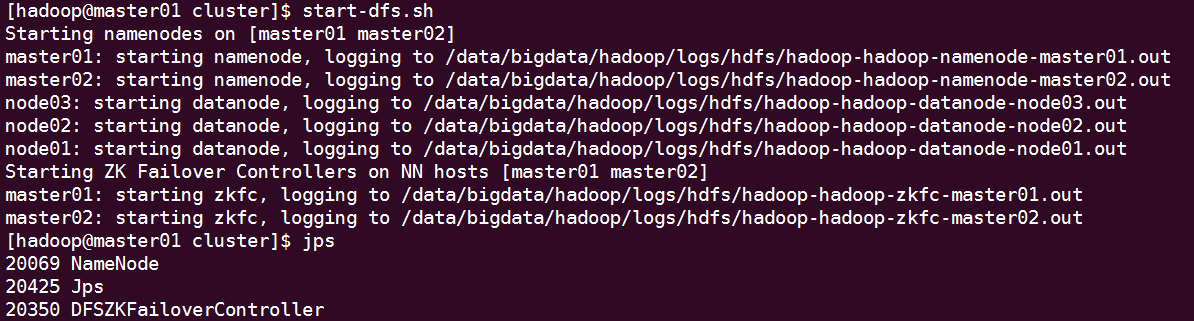
hdfs namenode –format

## 格式化ZK

hdfs zkfc –formatZK

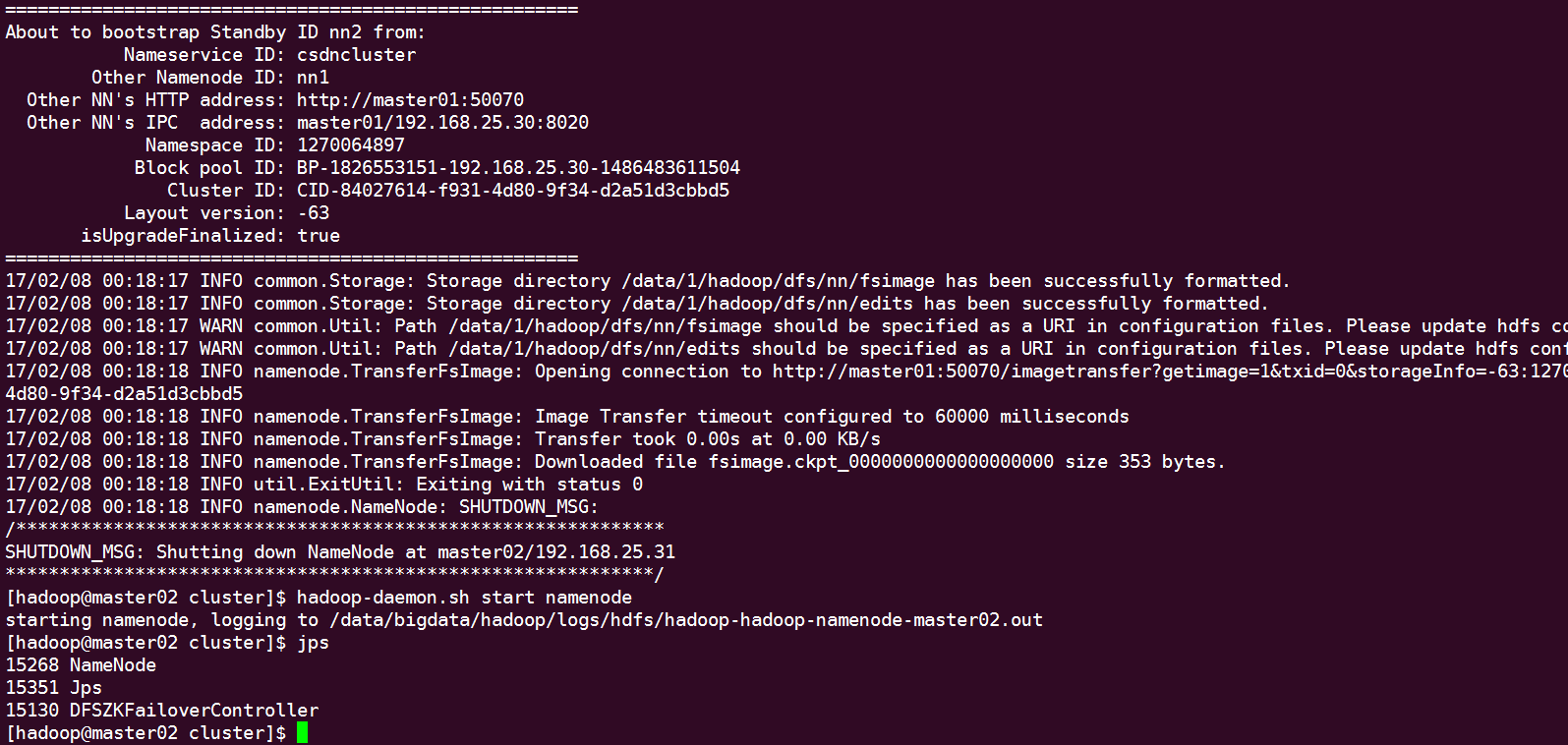
## 启动HDFS

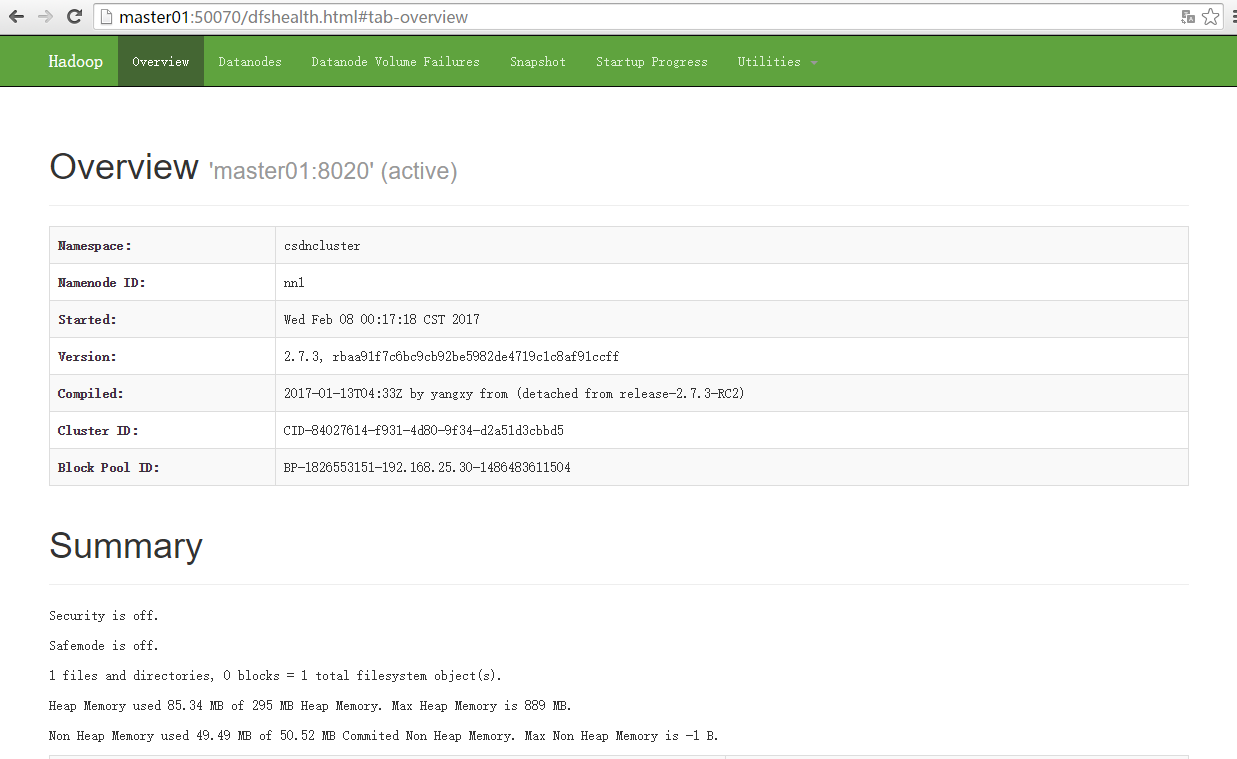
start-dfs.sh



hdfs namenode -bootstrapStandby

hadoop-daemon.sh start namenode

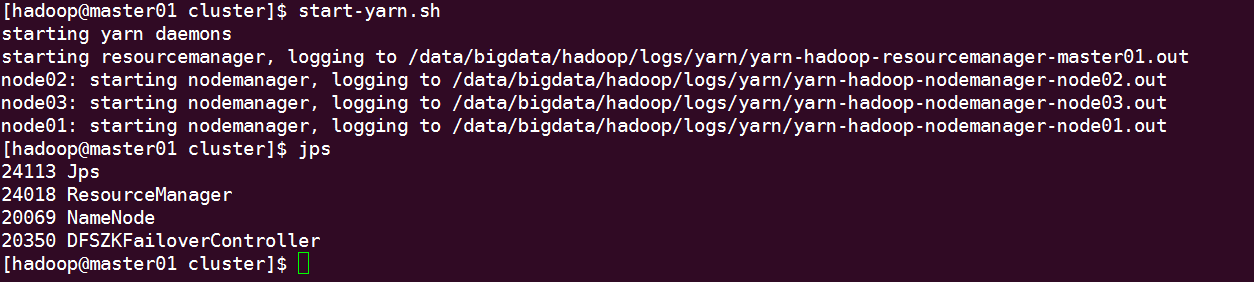




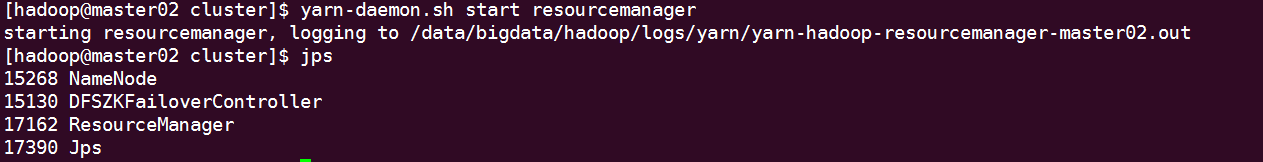


## 启动YARN

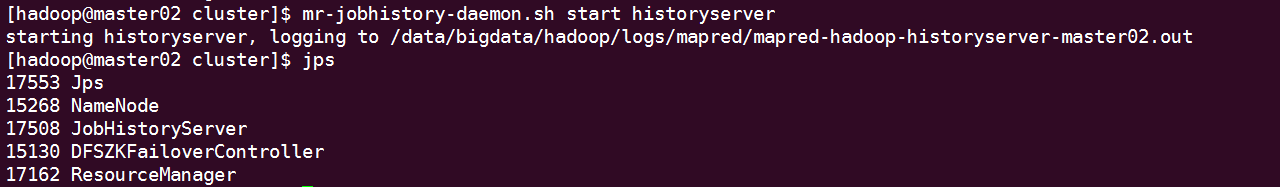
start-yarn.sh

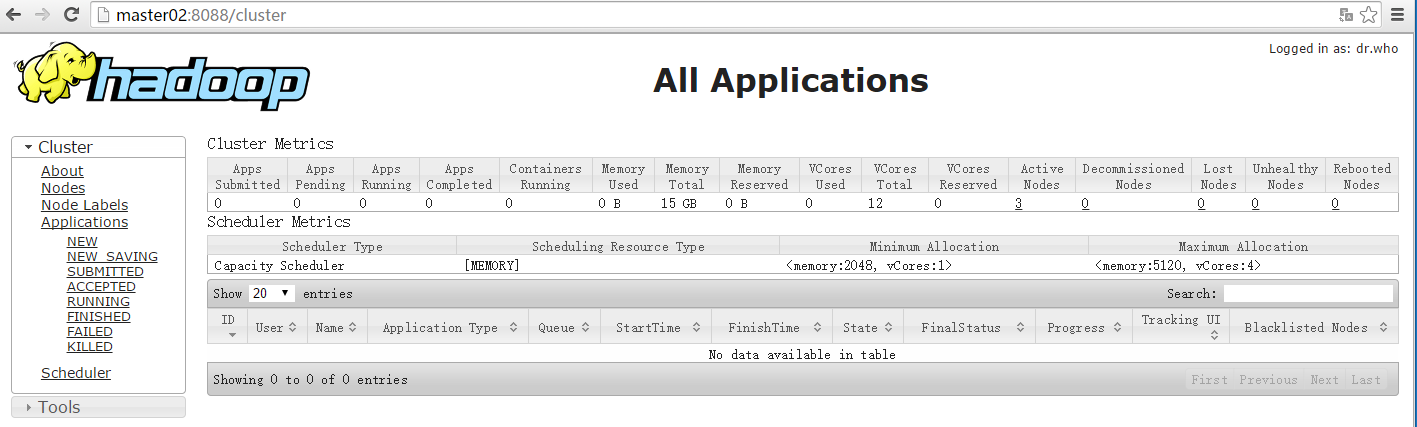


yarn-daemon.sh start resourcemanager



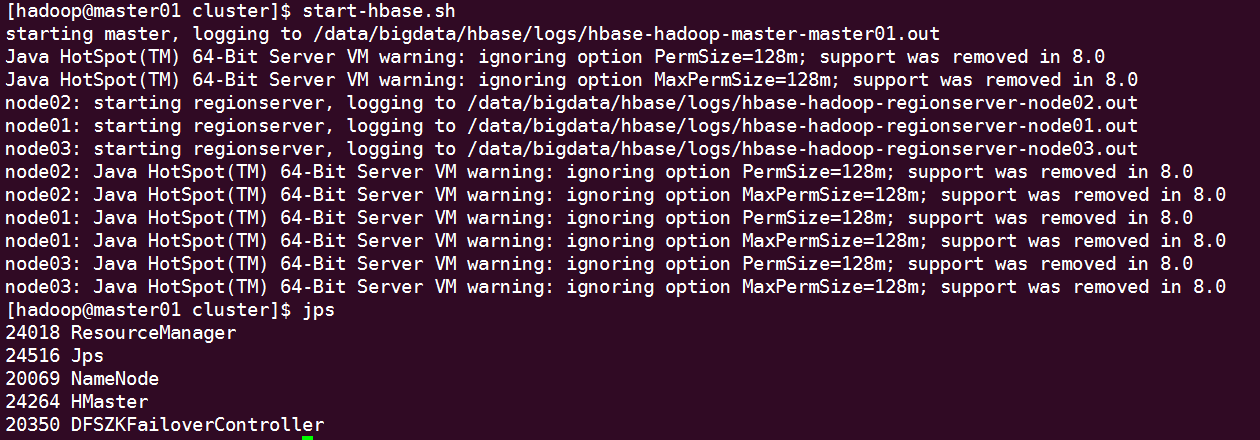
mr-jobhistory-daemon.sh start historyserver



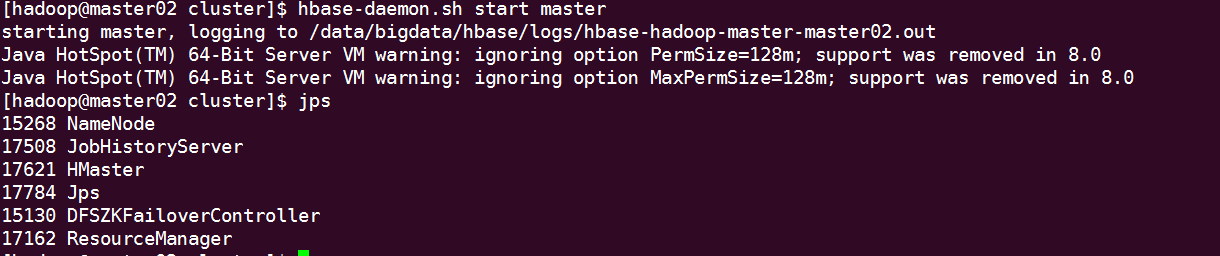


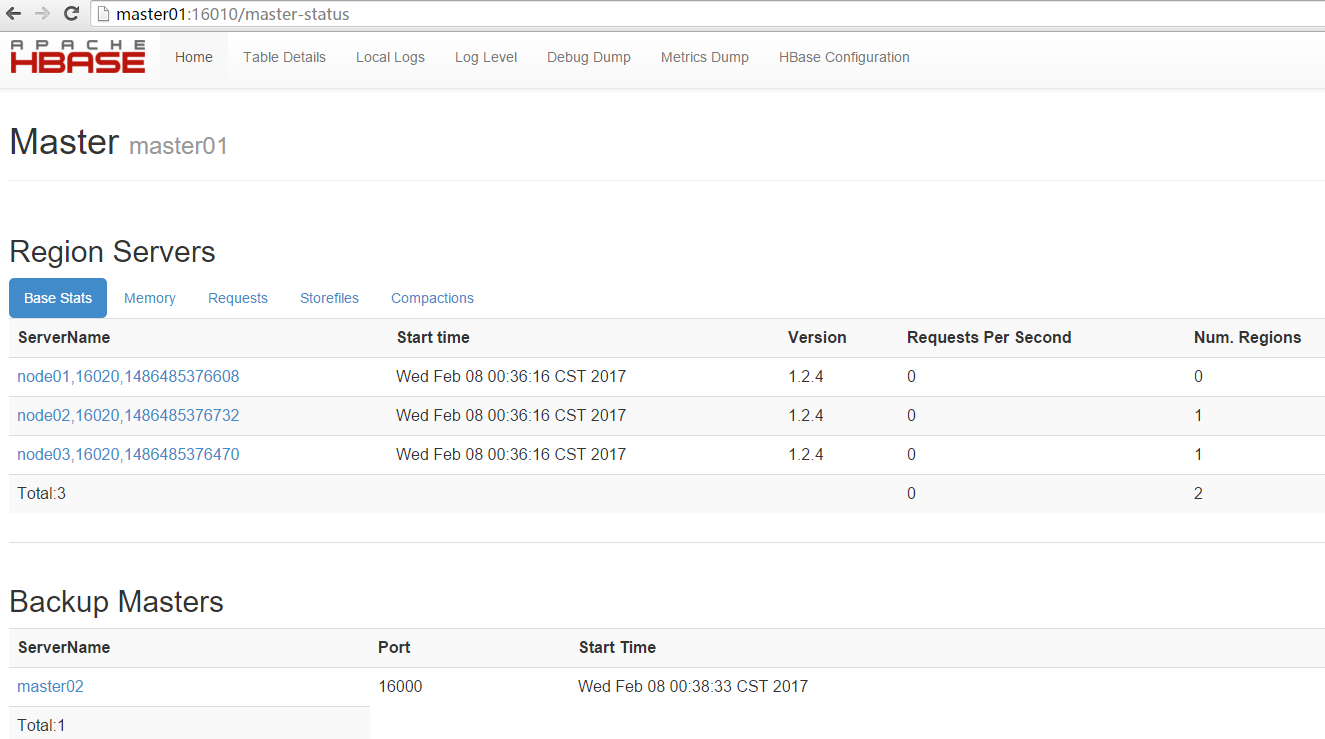
## 启动HBase

start-hbase.sh



hbase-daemon.sh start master

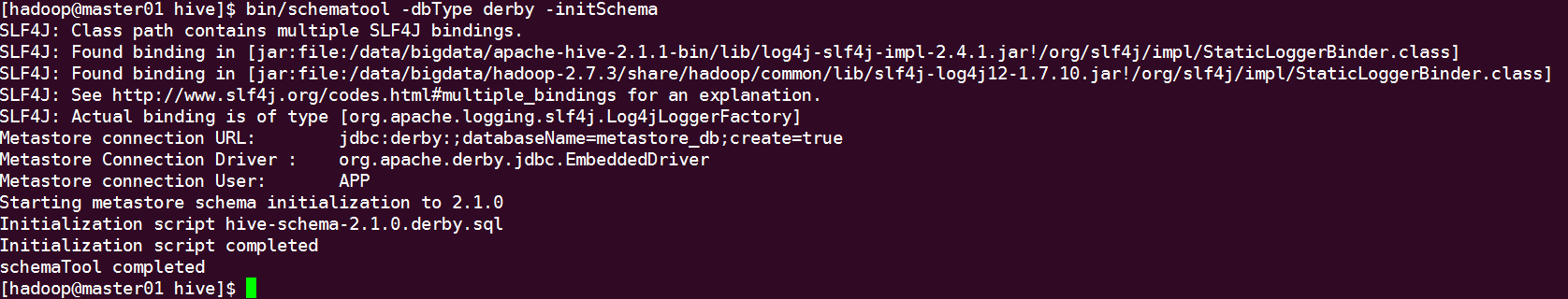




## 启动Hive

初始化hive schema

bin/schematool -dbType derby -initSchema

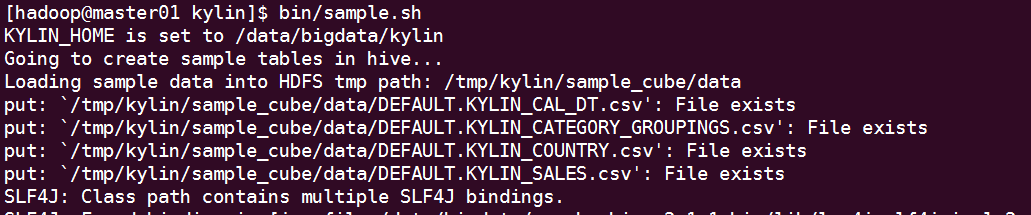


启动 metastore服务

nohup bin/hive --service metastore >> metastore.out 2>&1 &

## 启动kylin

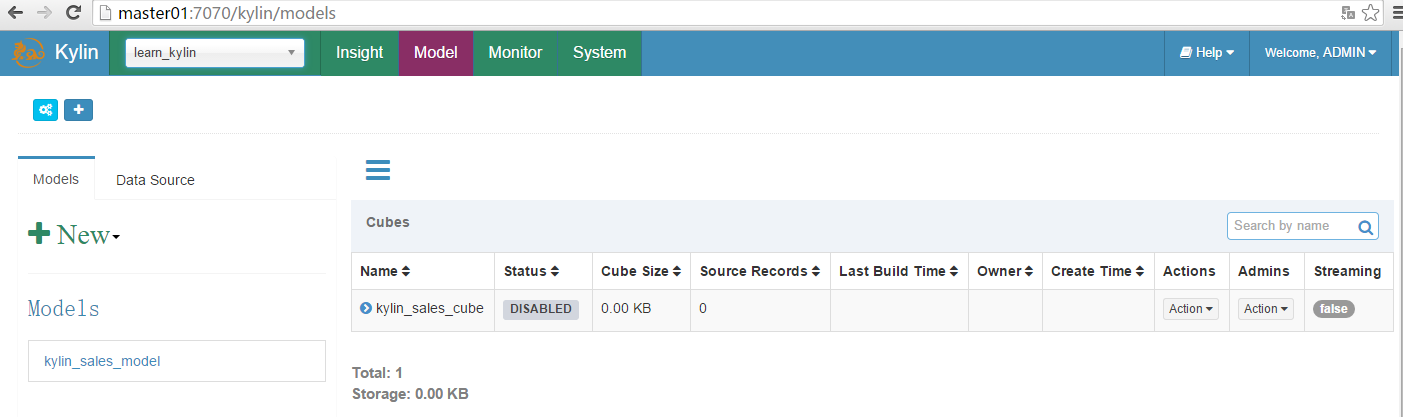
执行kylin的样例



启动kylin

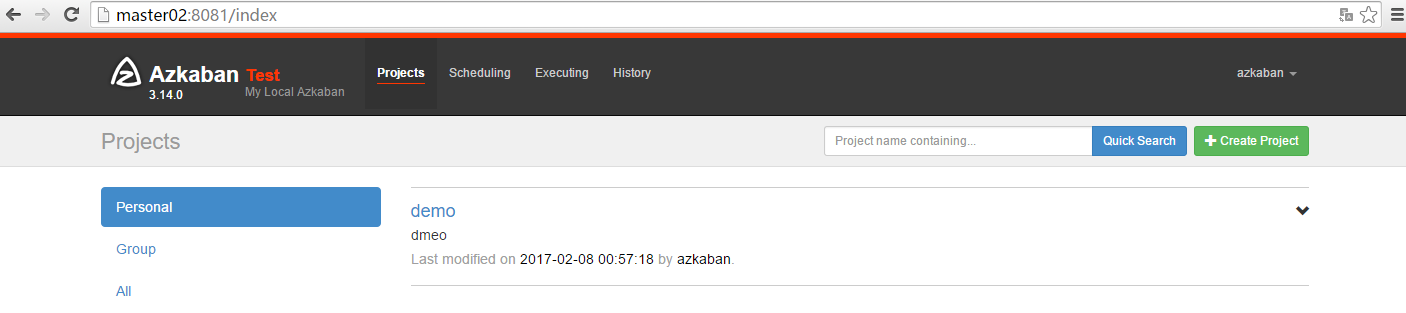
bin/kylin.sh start

用户名：ADMIN 密码：KYLIN



## 启动azkaban

bin/azkaban-solo-start.sh



# DataSee

# 附录

## DNS设置

#Vi /etc/resolv.conf

nameserver 114.114.114.114

nameserver 8.8.8.8

## Centos 7.2 设置“vm.swappiness=0” 无法开机自动生效的问题

<http://www.bubuko.com/infodetail-1933665.html3>.

find /usr/lib/tuned -name '\*.conf' -type f -exec grep "vm.swappiness" {} \+

