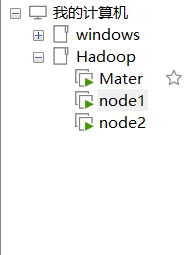
创建虚拟机，master，node1，node2



**对master，node1，node2配置相应的静态ip和免密登录，ip和hostname映射关系**

vi /etc/sysconfig/network-scripts/ifcfg-ens33 配置静态ip

基于虚拟网络编辑器的子网ip设置如下：

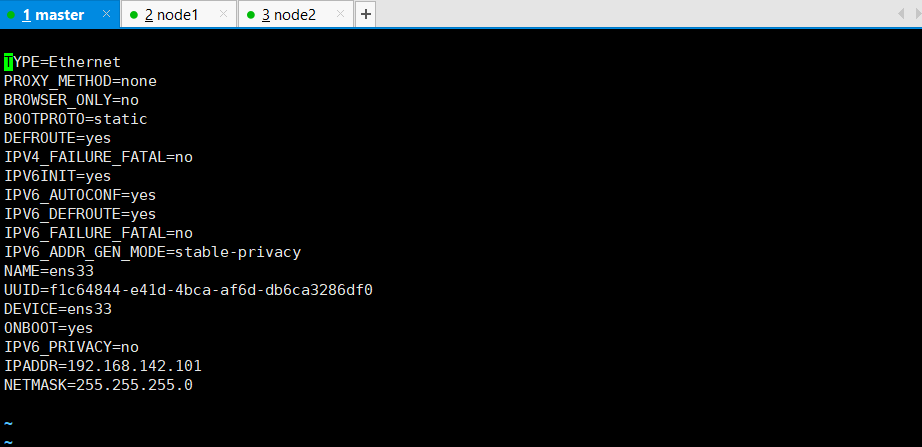
192.168.142.101 master

192.168.142.102 node1

192.168.142.103 node2

其他：更改BOOTPROTO=static

更改ONBOOT=yes



配完后重启网络服务：service network restart

映射（配置到每一个主机）：

vi /etc/hosts 加入

192.168.142.101 master

192.168.142.102 node1

192.168.142.103 node2

使用ssh-keygen -t rsa 然后一直回车

cd /.ssh/

ssh-copy-id -i id\_rsa.pub -p22 root@node1

ssh-copy-id -i id\_rsa.pub -p22 root@node2

在node1和node2中输入

cat ~/.ssh/id\_rsa.pub >> ~/.ssh/authorized\_keys 将公钥保存到本地

然后可以使用Xshell进行操作，用Xftp传输文件

**配置java和hadoop环境**

将hadoop和jdk上传到master中的/opt/soft/

解压hadoop和jdk

tar -xzvf hadoop-2.10.1.tar.gz

tar -xzvf jdk-8u211-linux-x64.tar.gz

然后可以用rm rf 文件名\*\*\* 删除压缩包

解压完后，将java和hadoop路径写进环境变量 vi /etc/profile

添加：export JAVA\_HOME=/opt/soft/jdk1.8.0\_211

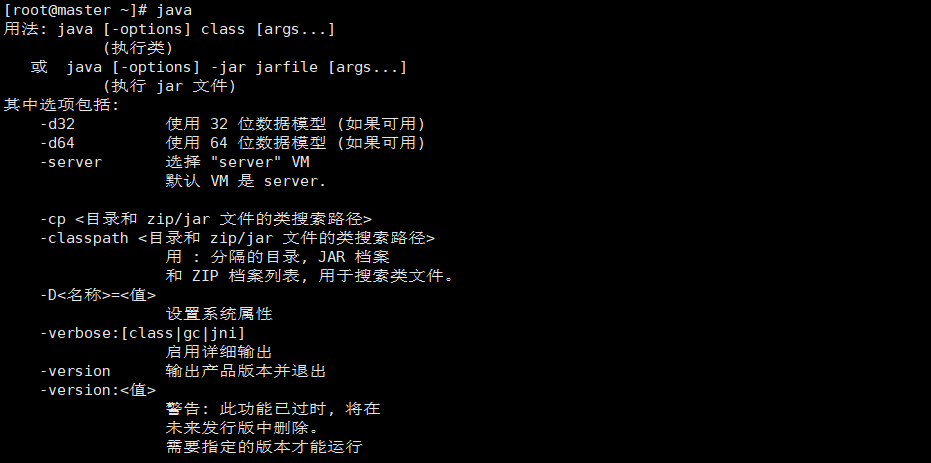
export PATH=$JAVA\_HOME/bin:$PATH

export HADOOP\_HOME=/opt/soft/hadoop-2.10.1

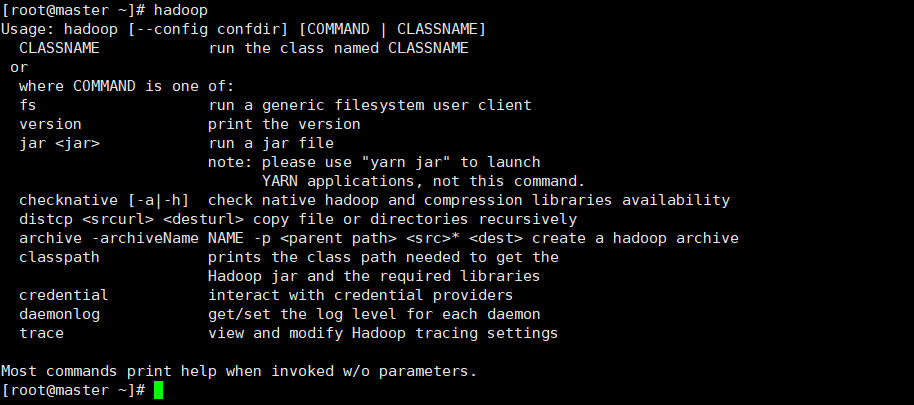
export PATH=$HADOOP\_HOME/bin:$HADOOP\_HOME/sbin:$PATH

退出source /etc/profile 刷新配置文件

使用java命令进行测试：



使用hadoop：



配置完成。

**写hadoop配置文件**

1. **core-site.xml**

<configuration>

<property>

<name>fs.defaultFS</name>

<value>hdfs://master:9000</value>

</property>

<property>

<name>hadoop.tmp.dir</name>

<value>/opt/soft/hadoop\_tmp</value>

</property>

</configuration>

1. **hdfs-site.xml**

<configuration>

<property>

<name>dfs.replication</name>

<value>2</value>

</property>

<property>

<name>dfs.namenode.name.dir</name>

<value>/opt/soft/hadoop-2.10.1/data/namenode</value>

</property>

<property>

<name>dfs.datanode.data.dir</name>

<value>/opt/soft/hadoop-2.10.1/data/datanode</value>

</property>

</configuration>

1. **yarn-site.xml**

<configuration>

<!-- Site specific YARN configuration properties -->

<!-- NodeManager上运行的附属服务。需配置成mapreduce\_shuffle，才可运行MapReduce程序 -->

<property>

<name>yarn.nodemanager.aux-services</name>

<value>mapreduce\_shuffle</value>

</property>

<property>

<description>The hostname of the ResourceManager.</description>

<name>yarn.resourcemanager.hostname</name>

<value>master</value>

</property>

<property>

<name>yarn.resourcemanager.address</name>

<value>${yarn.resourcemanager.hostname}:8032</value>

</property>

<property>

<name>yarn.resourcemanager.scheduler.address</name>

<value>${yarn.resourcemanager.hostname}:8030</value>

</property>

<property>

<name>yarn.resourcemanager.webapp.address</name>

<value>${yarn.resourcemanager.hostname}:8088</value>

</property>

<property>

<name>yarn.resourcemanager.webapp.https.address</name>

<value>${yarn.resourcemanager.hostname}:8090</value>

</property>

<property>

<name>yarn.resourcemanager.resource-tracker.address</name>

<value>${yarn.resourcemanager.hostname}:8031</value>

</property>

<property>

<name>yarn.resourcemanager.admin.address</name>

<value>${yarn.resourcemanager.hostname}:8033</value>

</property>

</configuration>

1. **mapreduce-site.xml**

<configuration>

<!-- 通知框架MR使用YARN -->

<property>

<name>mapreduce.framework.name</name>

<value>yarn</value>

</property>

<property>

<name>mapreduce.jobhistory.address</name>

<value>master:10020</value>

</property>

<property>

<name>mapreduce.jobhistory.webapp.address</name>

<value>master:19888</value>

</property>

</configuration>

将配置文件一并发给node1和node2

scp -r 目录 root@node1 目录 ----传文件夹用-r 传文件不用-r

要将hadoop jdk 和环境变量发送，环境变量也可以重新配

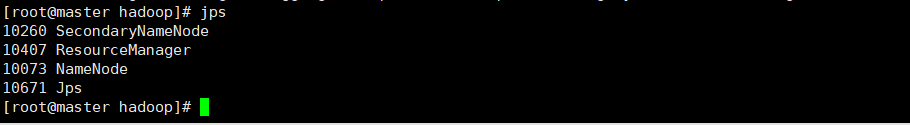
**格式化namenode**

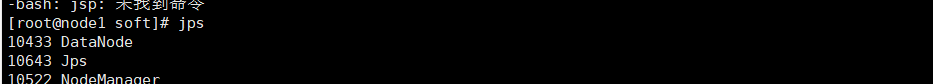
hadoop namenode -format

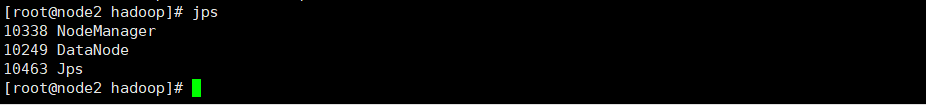
**启动所有hadoop服务**

start-all.sh

用jps命令看是否成功启动



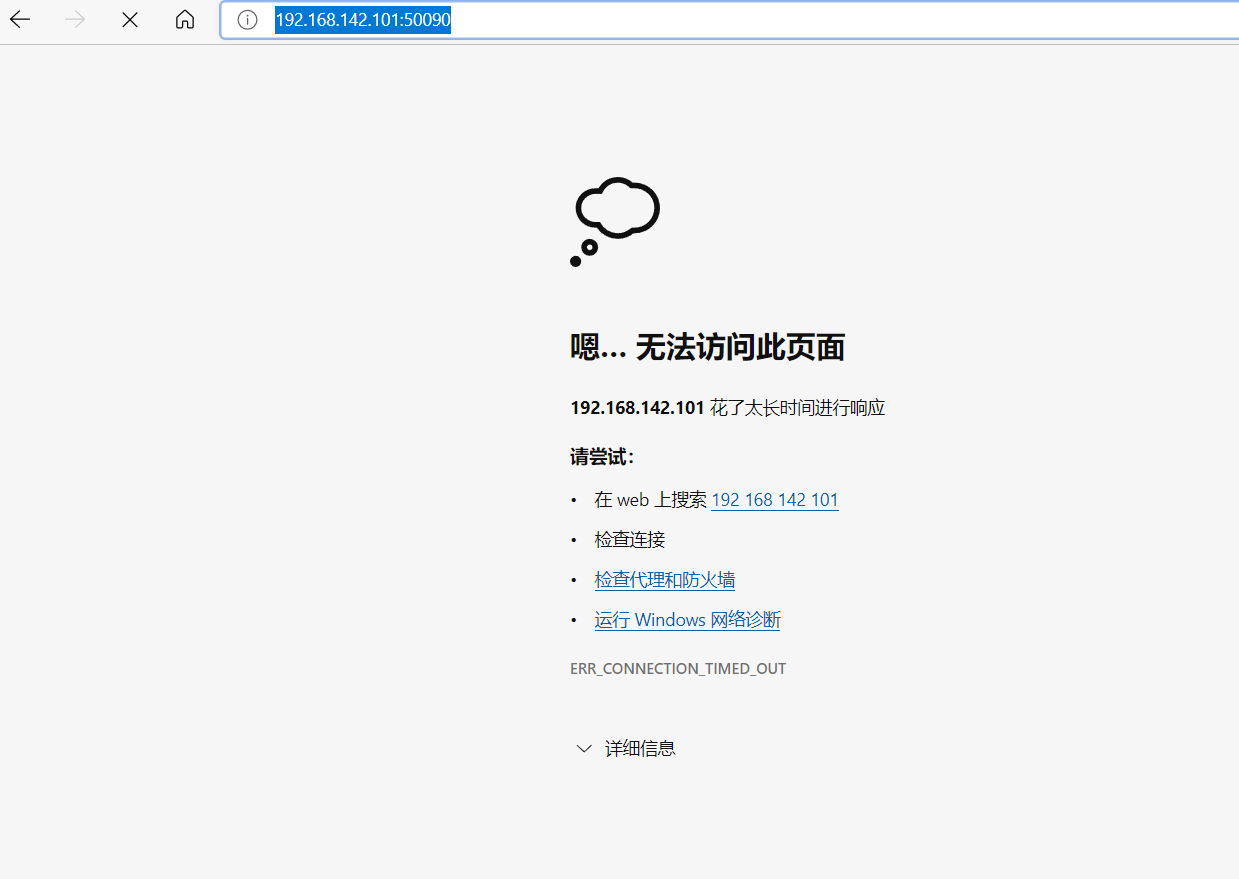




看样子已经成功启动。

关闭防火墙进行web查看：

发现访问不了



在hdfs-site.xml中加入：

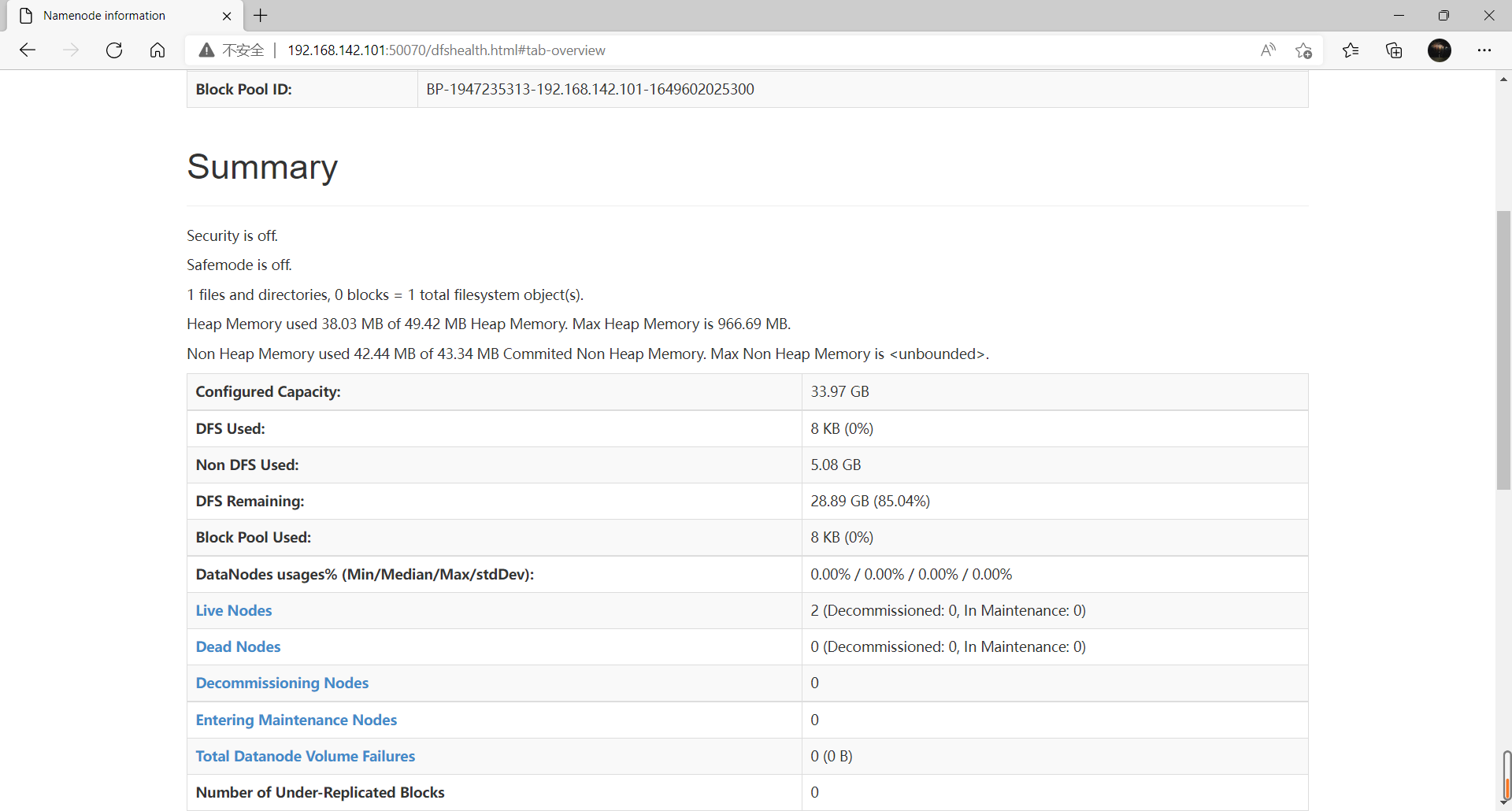
<property>

<name>dfs.namenode.http-address</name>

<value>master:50070</value>

</property>

重启hdfs



已经上去了，节点：

