Hadoop 学习之集群安装(上)

将配置的伪分布模式的虚拟机克隆两份并更名为centos\_node2和centos\_note3

分别打开三个虚拟机，第一个基本不需要配置了，现在主要配置centos\_note2和centos\_note3的配置。

1. 首先修改IP地址和主机名

因为centos是hadoop1和192.168.100.11，在这里设置centos\_note2和centos\_note3的主机名和IP地址分别为：hadoop2/192.168.100.12和hadoop3/192.168.100.13。并把主机名和ip地址绑定

Note2的修改主机名：#vim /etc/sysconfig/network 下的HOSTNAME=hadoop2，note3的相应的改为hadoop3

再通过#hostname hadoop2。设置临时的主机名那么机器就不用重新启动了。

Note2的修改ip地址：在右上角点编辑链接，然后编辑将ipv4改为手动，地址192.168.100.12 掩码：255.255.255.0 网关：192.168.100.1。

Note3的改为相应的ip

绑定：#vim /etc/hosts

Ip hostname

保存并重启网络服务：

#service network restart

在宿主机上ping以下看是否三台机器都能ping通。一般情况都能通的。只要和宿主机上的VMnet1网关在同一网段一般没问题。

通过pietty远程连接三台虚拟机，进行下一步的操作

1. 将note2和note3中配置数据重置
   1. 删掉远程连接ssh免登陆：

#rm –rf /root/.ssh/\*

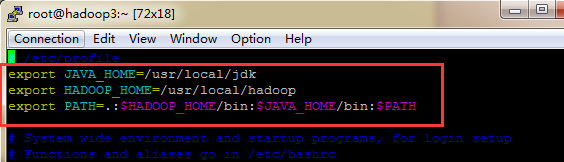
* 1. 删掉原来的hadoop安装程序

#rm –rf /usr/local/\*(如果该目录下还有其他的有用数据，则只删除hadoop的即可)

* 1. 删掉环境变量中的hadoop和jdk的配置

#vim /etc/profile

将JAVA\_HOME和HADOOP\_HOM的那三行都删掉



#source /etc/profile 使得环境变量配置文件重新生效

1. 重新配置三台虚拟机，搭建集群环境前的系统环境
   1. 配置每台机器上的ssh免密码登录(以主机名为hadoop1的centos为例)

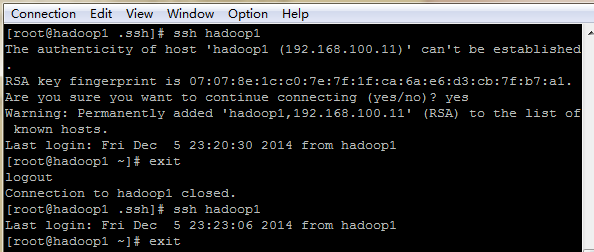
#ssh-keygen –t rsa 一路回车，在/root/.ssh/下生成id\_rsa.pub

生成公钥信息

#cp id\_rsa.pub authorized\_keys (不能写错否则就不能免密码登录了)

#登录本机，免密码登录

#ssh hadoop1 就能免密码登录了



其他两台也是同样的操作

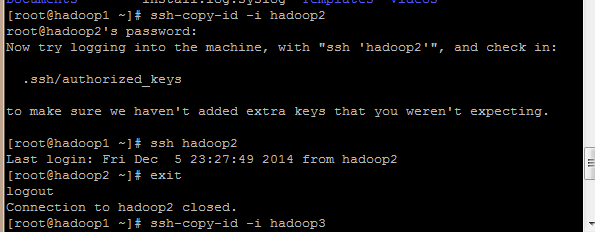
* 1. 三台机器间的免密码登录

将自己的公钥复制到允许需要免密码登录的其他主机上

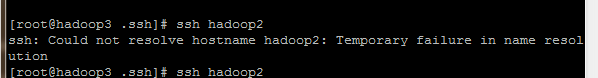
#ssh-copy-id -i hadoop2

输入hadoop2的密码

下次就可以免密码登录了。



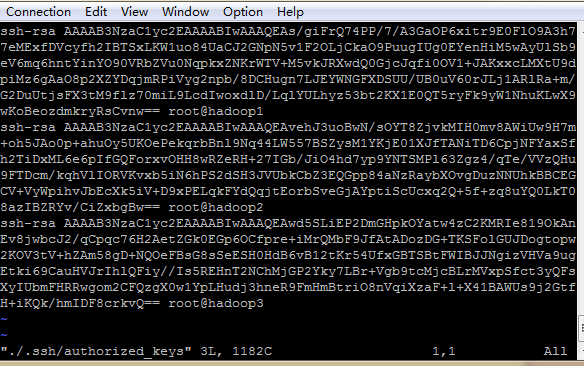
如果主机名和ip没有绑定会出现这样的错误



在hadoop2和hadoop3上分别执行

##ssh-copy-id -i ip(hadoop1的ip)

完事后hadoop2和hadoop3就可以免密码登录hadoop1了。

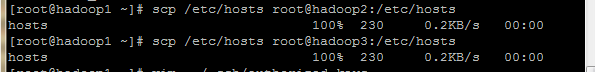


将hadoop1下的ssh免密码许可文件复制到hadoop2和hadoop3上

#scp /root/.ssh/authorized\_keys [root@hadoop2:/root/.ssh/](mailto:root@hadoop2:/root/.ssh/)

如果之前做过hadoop1到相应主机的免密码登录就不需要密码，没做过就需要输入密码。分别复制完成后，现在三台虚拟机之间就可以互相免密码登录了，但还不能使用主机名登录，现在讲ip和主机名绑定的hosts文件也在三台虚拟机之间同步

#scp /etc/hosts root@hadoop2:/etc/hosts



这样三台虚拟机之间就可以通过主机名相互免密码登录了。

* 1. 为hadoop2和hadoop3复制安装文件和配置文件

前提:为了保持干净和便于查找错误，将hadoop下的logs和tmp目录删除

#cd /usr/local/hadoop/

#rm –rf ./{logs/,tmp/}

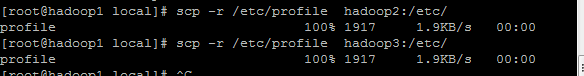
将jdk和hadoop安装文件复制都hadoop2和hadoop3

# scp -r /usr/local/{jdk,hadoop} hadoop2:/usr/local/

# scp -r /usr/local/{jdk,hadoop} hadoop3:/usr/local/

将profile配置文件复制到hadoop2和hadoop3

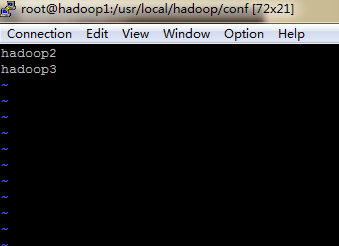
# scp -r /etc/profile hadoop2:/etc/



在每个节点使profile生效:

#source /etc/profile

1. 搭建集群环境
   1. 修改hadoop1下的hadoop的配置文件slaves，将另两台主机名加进去作为从节点,hadoop1作为主节点。每个主机名占一行

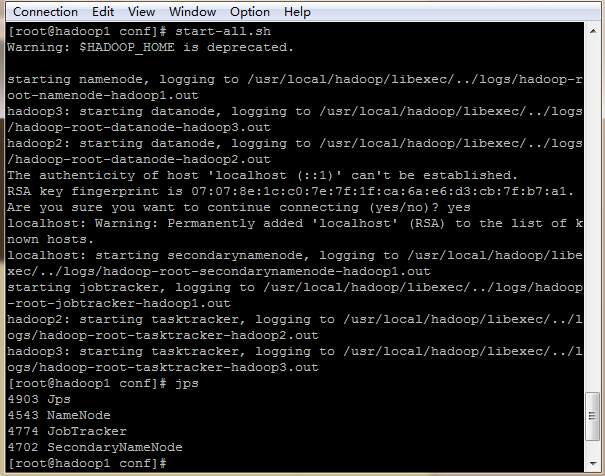


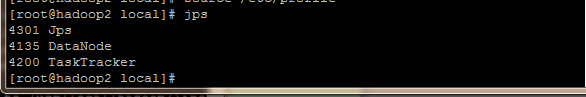
* 1. 启动hadoop集群环境
     1. 先格式化namenode

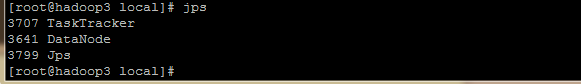
#hadoop namenote –format

* + 1. 启动hadoop集群

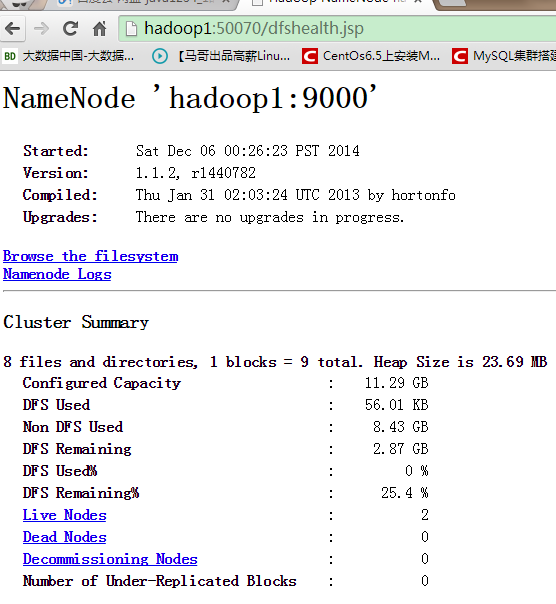
#start-all start







这样在主节点上就启动了是哪个进程NameNode、JobTracker、SecondarynameNode。在从节点启动了DataNode和TaskTracker



点击Live Bodes

