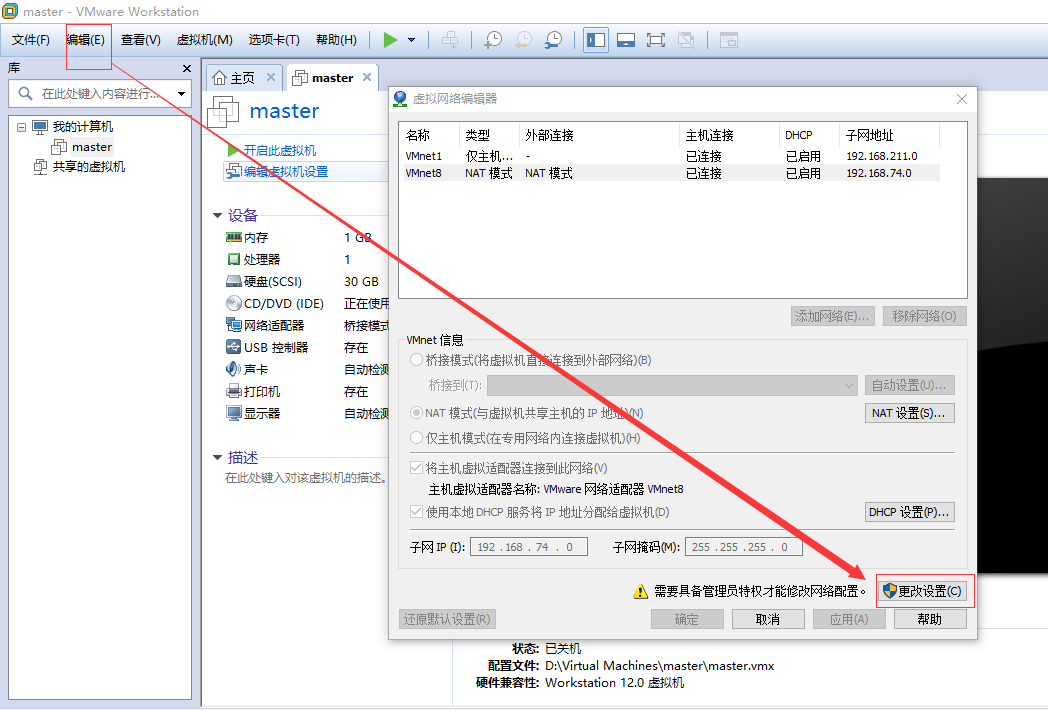
Root 密码：123456

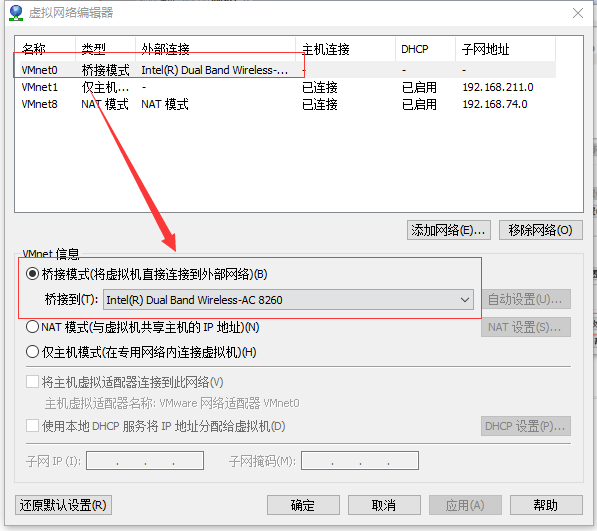
Geming 密码: 123456

# 虚拟机安装配置

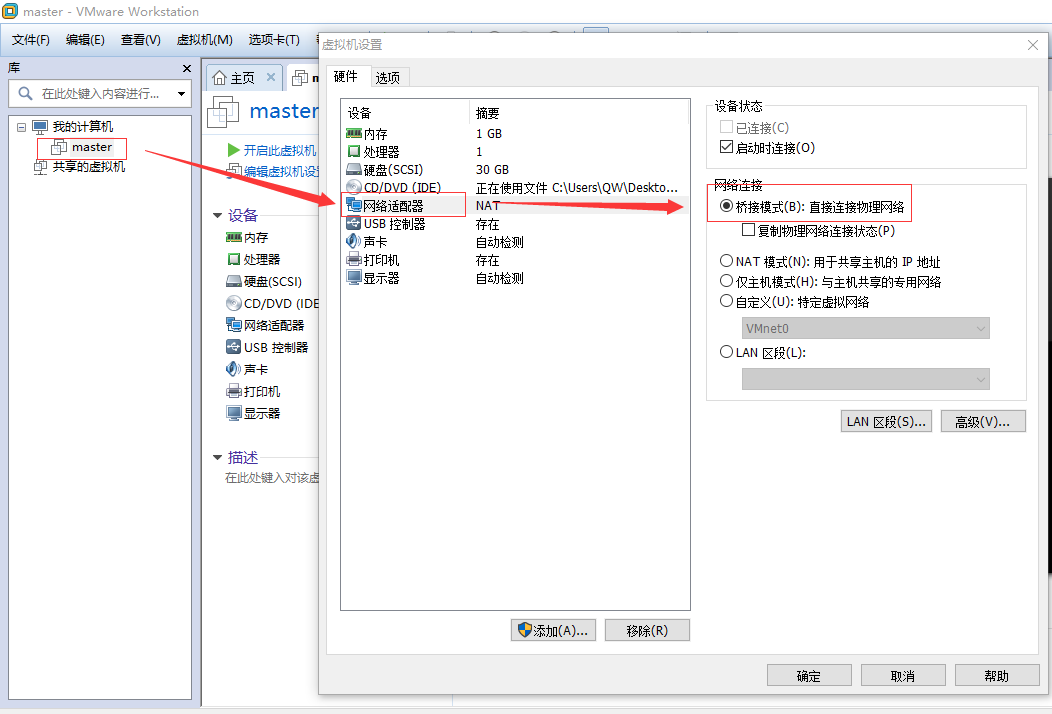
首先VMware->编辑->虚拟网络编辑器 打开网络编辑器。



单击更改设置，打开wmnet0 桥接网络选项：

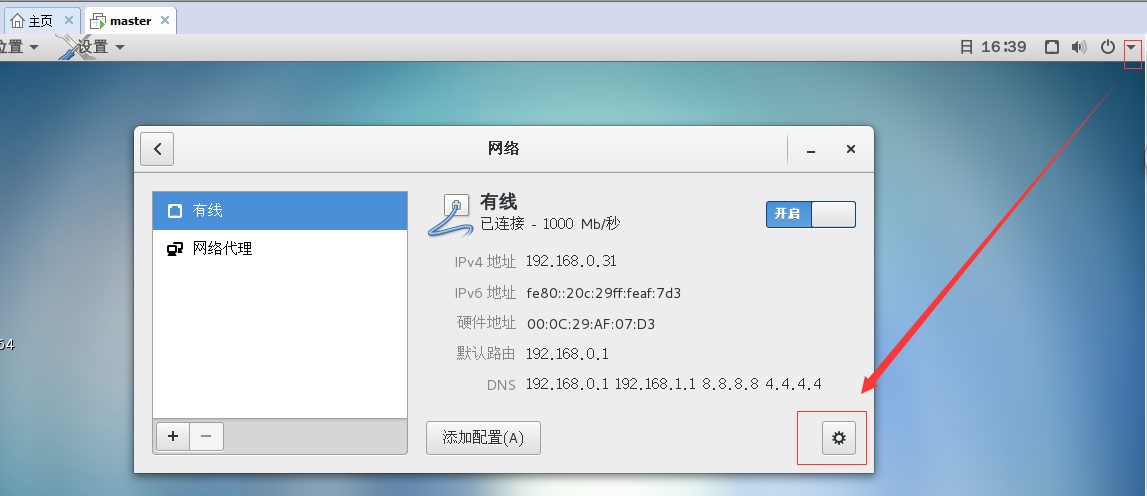


把桥接到的网卡设置成正确的网卡。（根据实际连接的网卡，可能是无线的，也可能是有线的）。然后设置虚拟机的网络适配器，如下：



在虚拟机的设置中，将网络模式改成桥接模式。

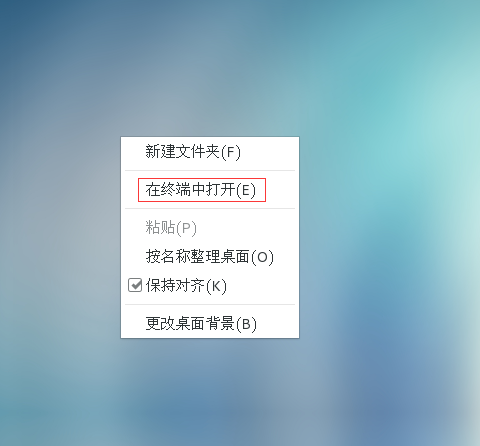
启动虚拟机，以root用户登录CentOS-7



打开网络配置，如下图：设置ip地址、dns。根据实际情况设置，可先ipconfig 查看一下物理电脑的ip地址，再设置。

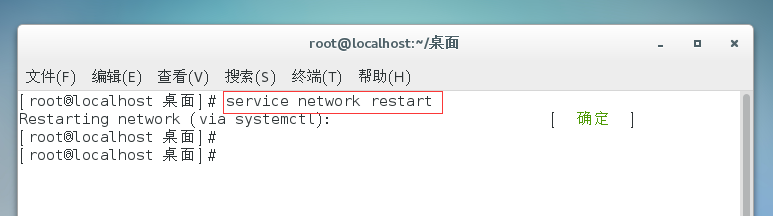


在ipv4 下面配置完ip地址 和 dns 。“应用” 退出。

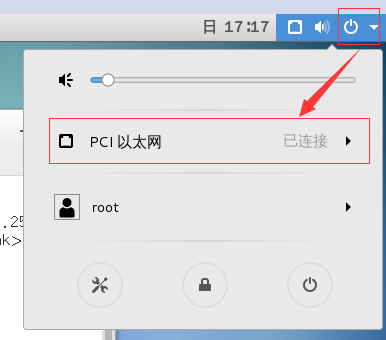


回到桌面，鼠标右键打开上面的弹出菜单，选择“在终端中打开”

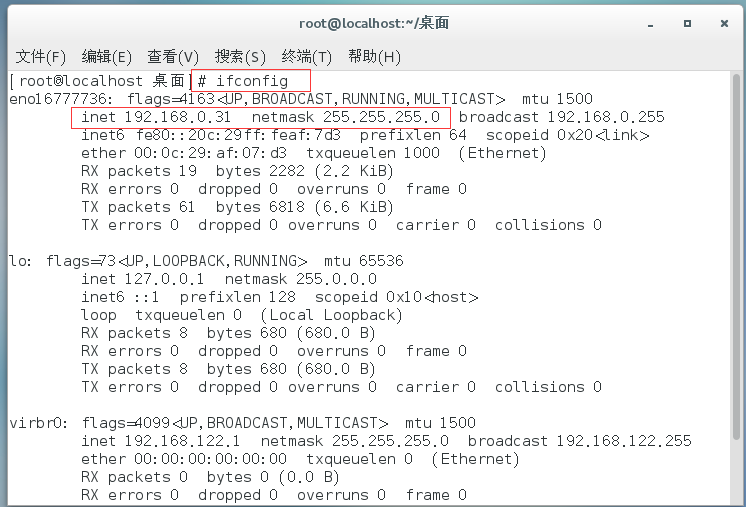
输入下面的命令，重启一下网卡（**可不进行该步骤**）：



也可以通过下图来启动和查看网络状态



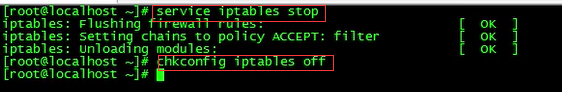
或者在终端中输入ifconfig查看网络状态



然后就是物理机和虚拟机互相ping一下，看看网络是否畅通。

# CentOS-7安装配置

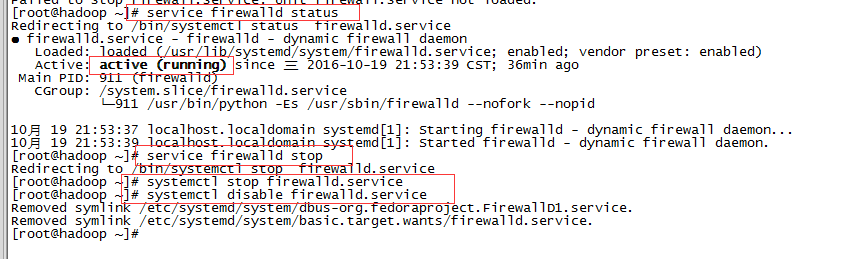
~~关闭防火墙~~

~~~~

~~Service iptables stop 关闭防火墙~~

~~Chkconfig iptables off 永久关闭防火墙~~

上述关闭防火墙的方式只适应centos6，对centos7来说得用下面的方法



查看防火墙状态

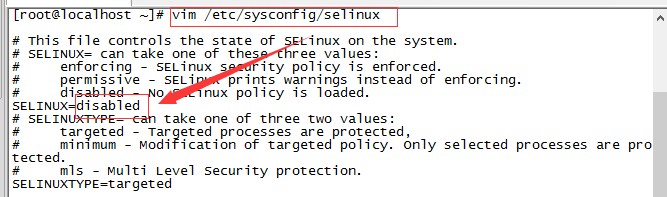
得知防火墙启动了

关闭防火墙

禁用防火墙

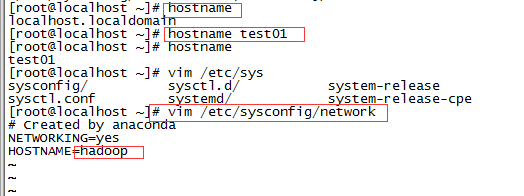
具体可参考技术积累/Hadoop/在windows下用浏览器无法访问虚拟机中Linux的hadoop管理页面.pdf (由于防火墙导致不能访问的。)

关闭selinux （ 是[美国国家安全局](http://baike.baidu.com/view/14954.htm)（NSA）对于[强制访问控制](http://baike.baidu.com/view/4420215.htm)的实现，是 Linux历史上最杰出的新安全子系统。NSA是在Linux社区的帮助下开发了一种访问控制体系，在这种访问控制体系的限制下，进程只能访问那些在他的任务中所需要文件。）如下图：



按下i进入插入模式，改成disabled，按下esc进入正常模式，输入“:wq”进入命令模式，并保存退出

改变主机名称



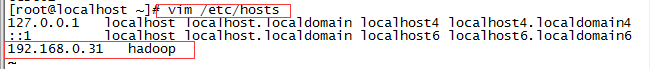
Hostname：查看主机名

Hostname test01 : 临时改变主机名为 test01

Vim /etc/sysconfig/network

把hostname 改成hadoop ，永久改变主机名为hadoop

添加ip与主机名的映射

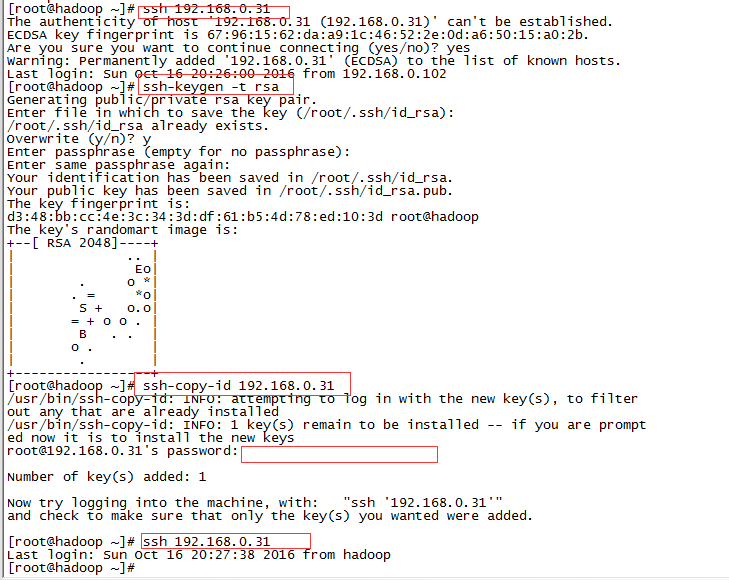


Reboot 重启centos

C:\Users\QW\AppData\Roaming\Tencent\Users\747679105\QQ\WinTemp\RichOle\64ODFR7)~_[@6HTY`L}~5NR.png

再输入hostname试试。

**设置ssh无密码连接。**



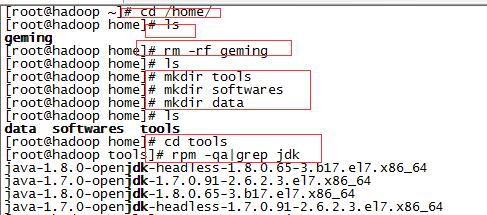
Ssh 192.168.0.31 进行ssh连接，一般情况下需要输入密码。

Ssh-keygen –t rsa 生成密码公钥，输入几个回车即可

Ssh-copy-id 192.168.0.31 复制密码公钥到192.168.0.31 主机上

Ssh 192.168.0.31 再进行ssh连接就不需要密码了，直接可以连接上了。

# JDK安装配置



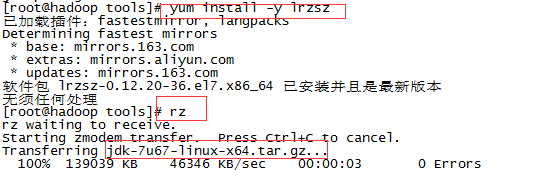
在home文件夹下 创建文件夹tools、softwares、data。

检查是否安装过jdk。

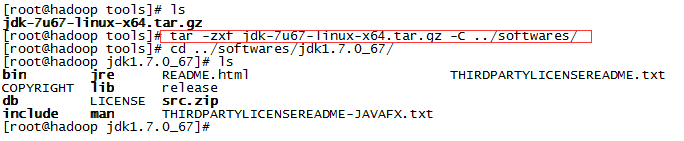
将已经安装的jdk一个一个全部卸载掉。上面有4个。



安装lrzsz ,用于上传文件。Rz 上传jdk文件到cent-os中。



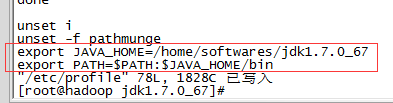
解压到softwares下



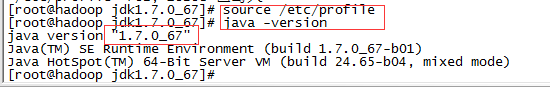
配置java环境变量

C:\Users\QW\AppData\Roaming\Tencent\Users\747679105\QQ\WinTemp\RichOle\CKRLPV`2P~COXRSDG9~T05R.png

Pwd 获得java路径，然后copy路径。 编辑profile文件



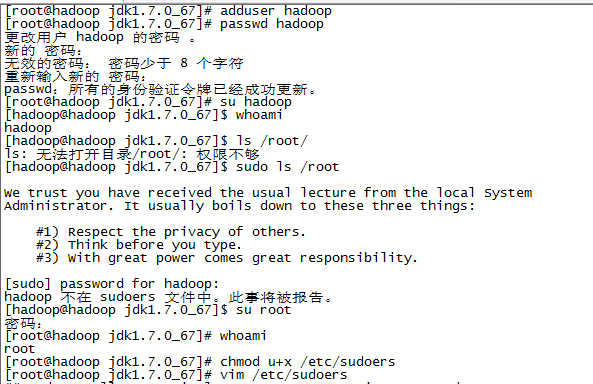
添加这么两行



刷新配置文件，查看java版本。

安装成功！

# 新用户安装配置



添加新用户：hadoop

设置它的密码：123456

切换用户到hadoop

查看当前用户

访问root 文件夹（权限不够）

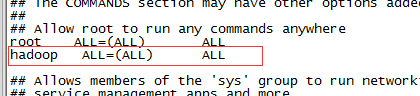
用sudo模式访问root文件夹（还是被拒绝）

切换用户回root

查看当前用户

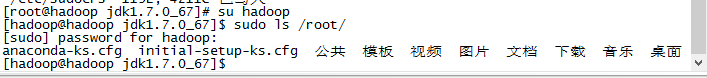
设置etc/sudoers 为可写

编辑它



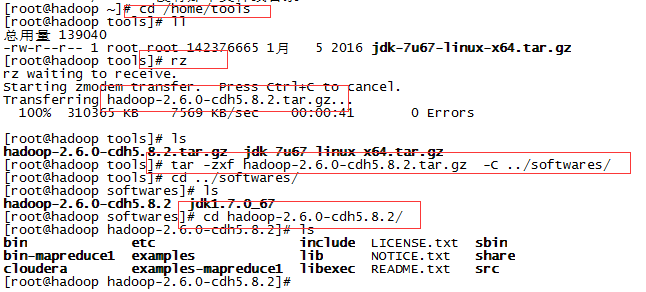
添加这么一行，将用户hadoop加入sudo 命令可执行用户列表中。

在这里需要输入:wq! 强制保存退出才行，因为这个文件本身是只读的。



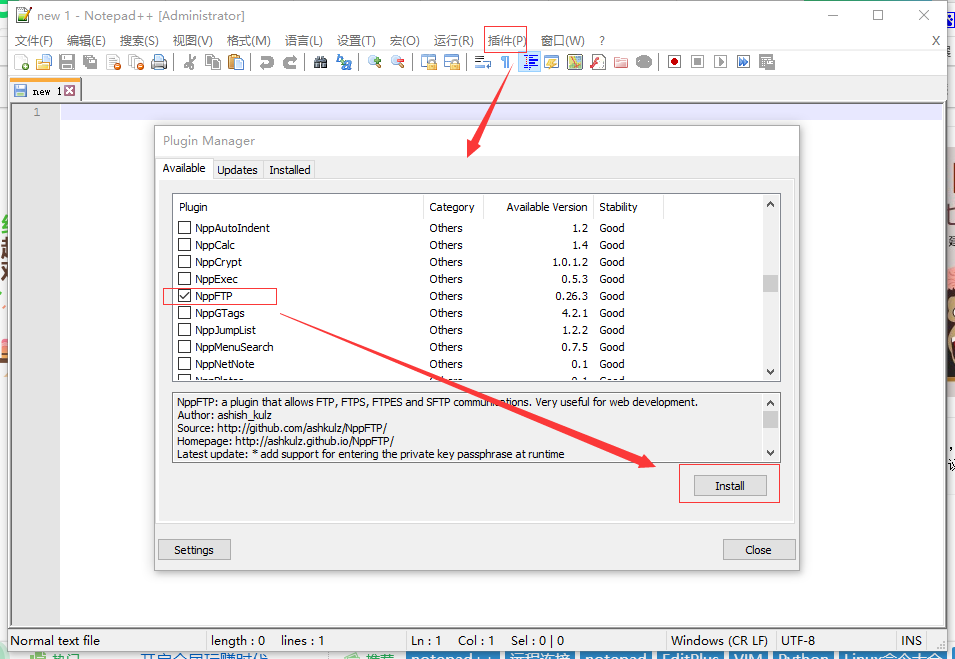
成功执行了sudo命令，可以访问root文件夹了。

# 安装配置hadoop



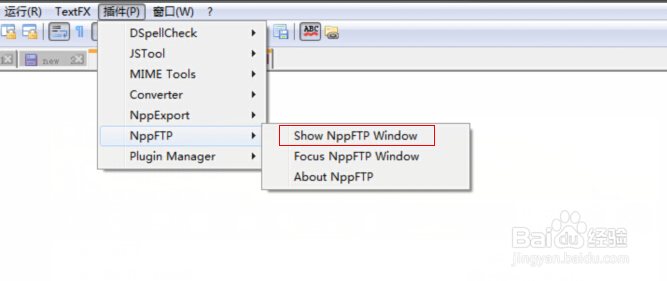
上传hadoop文件 ，并解压。

**通过notpad++ 远程连接linux ，并更改配置文件**

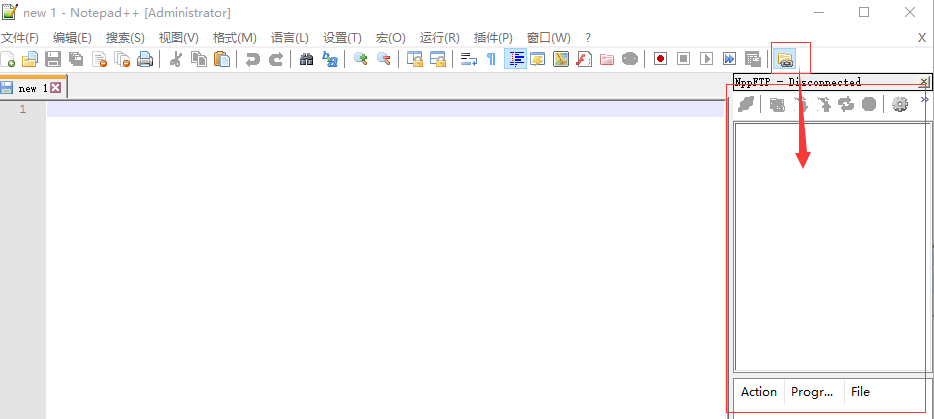


插件 -> 插件管理 ，安装插件 NppFTP。

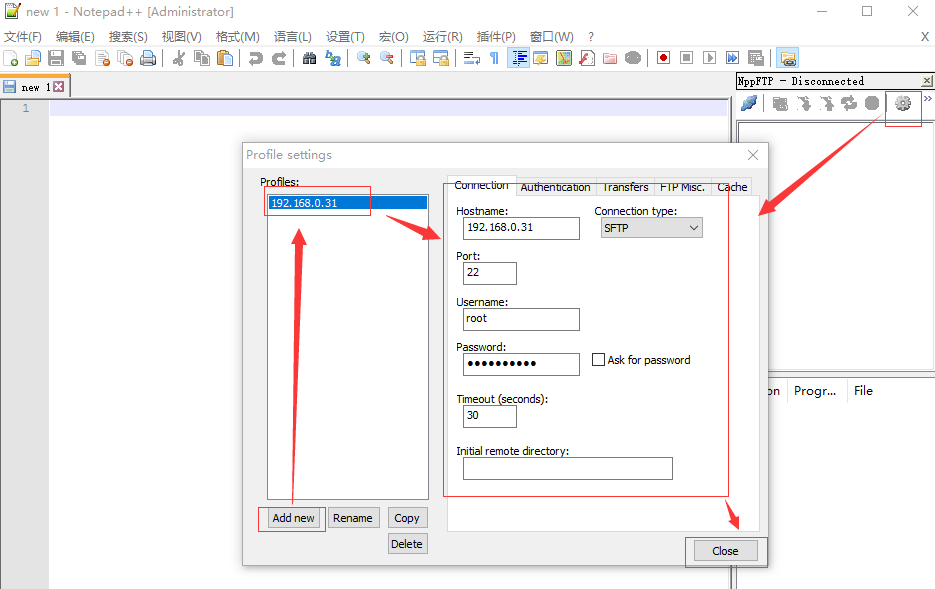
通过下面两种方式打开连接窗口：



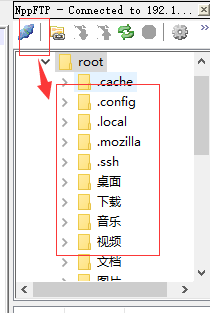
或者



然后单击设置 -> profile setting 打开下图



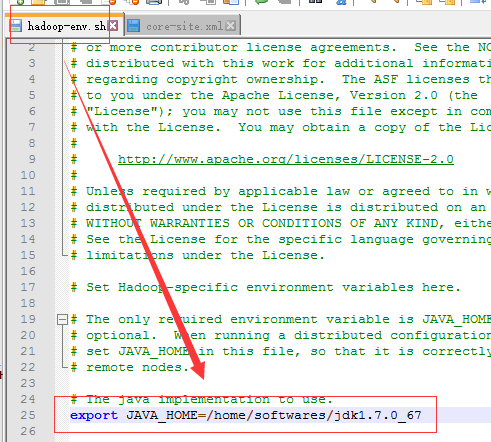
然后按下图连接linux



Echo 显示环境变量内容

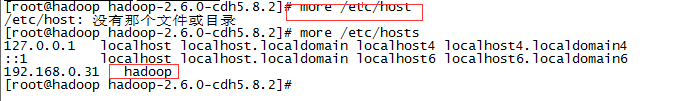
C:\Users\QW\AppData\Roaming\Tencent\Users\747679105\QQ\WinTemp\RichOle\84TF8K)){OUYM`)%8B@FPO1.png

拷贝 java\_home 环境变量内容



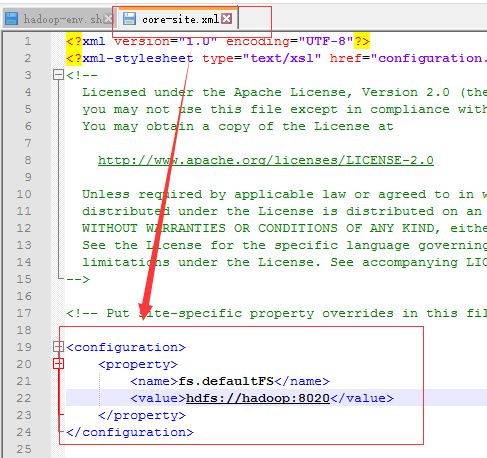
粘贴到图上所示处。上面的文件所在地址为：

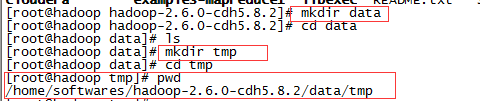
C:\Users\QW\AppData\Roaming\Tencent\Users\747679105\QQ\WinTemp\RichOle\S8U7}89I$FEEBO2W8L_]%BR.png



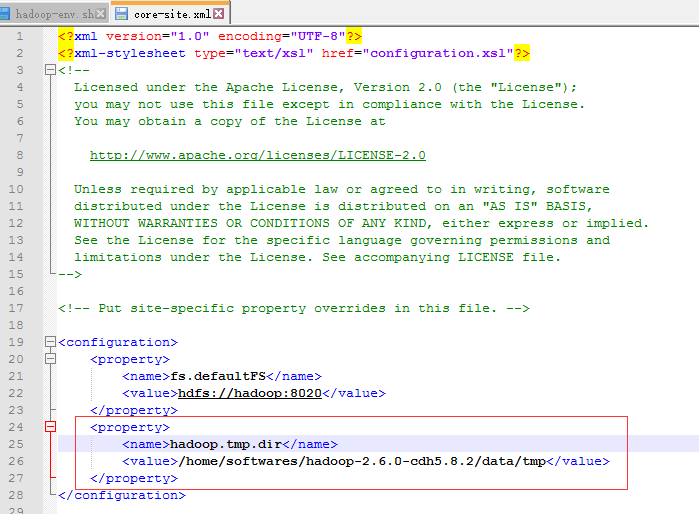
查看host文件，知道 主机名是hadoop。

按下图修改指定文件。



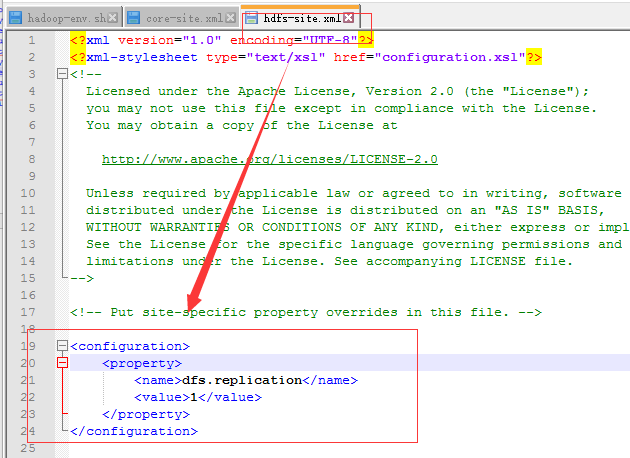


创建data/tmp 文件夹，并取路径。



按上图添加配置。

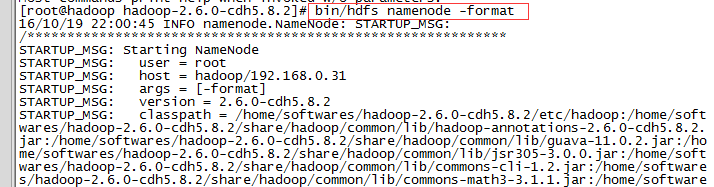
然后继续按下图配置



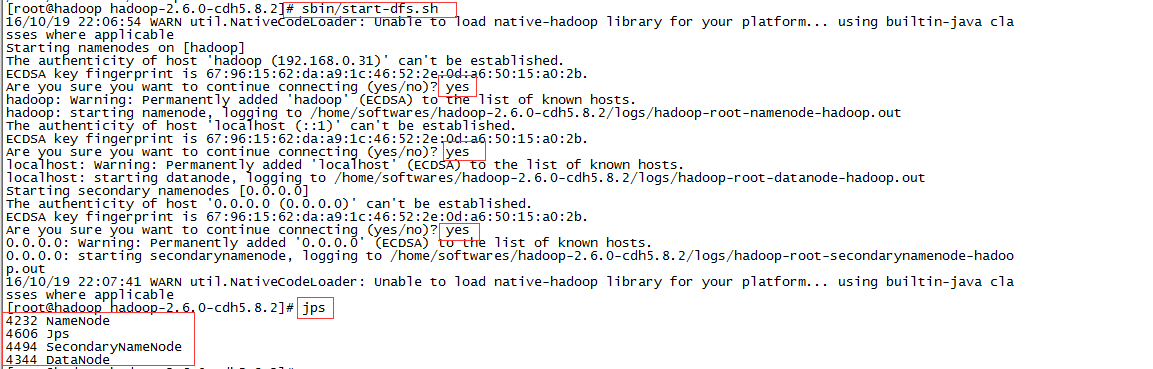
然后到/home/softwares/hadoop-2.6.0-cdh5.8.2 下执行格式化HDFS 命令：

bin/hdfs namenode -format

如下图：

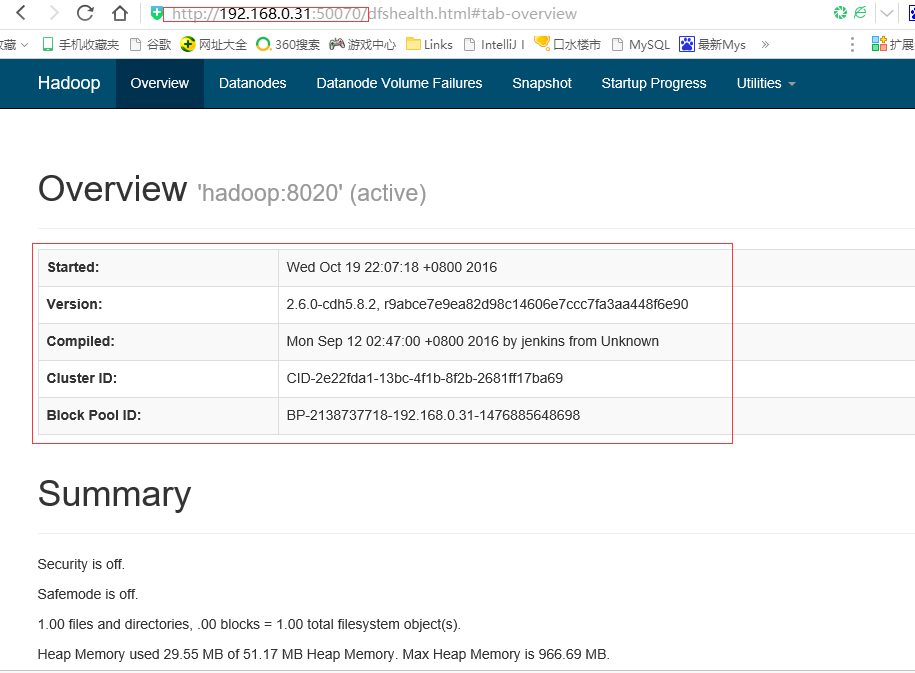


然后执行命令：sbin/start-dfs.sh 启动namenode 和 datanode 守护进程



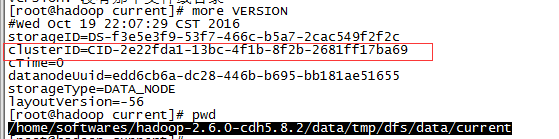
输入3个yes之后启动成功。可输入jps命令查看服务有没有起来。如果没有起来可以去logs文件夹下查看错误日志。

在浏览器中输入网址：<http://192.168.0.31:50070/> 查看NameNode 管理页面 （如果防火墙没开的话）



上面的cluster id （集群id） 跟 存放在

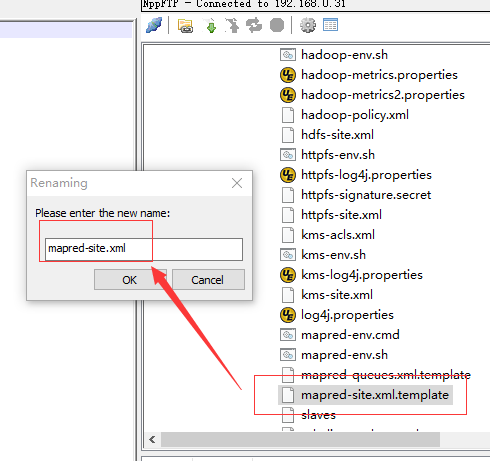
/home/softwares/hadoop-2.6.0-cdh5.8.2/data/tmp/dfs/data/current 下的VERSION 配置文件中的集群id 一样。如下图



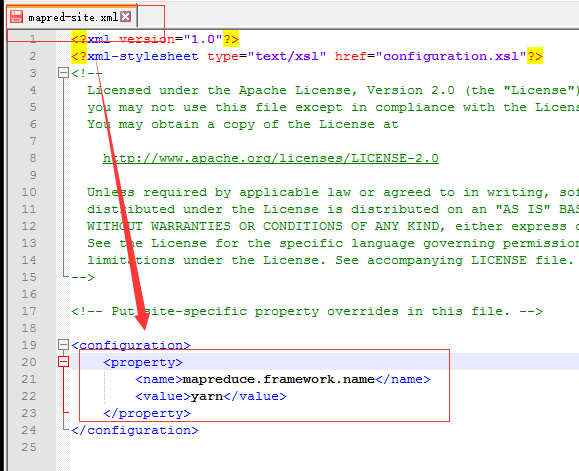
Block pool id （块id）跟存在该文件夹下的块文件名一样，如下图：

C:\Users\QW\AppData\Roaming\Tencent\Users\747679105\QQ\WinTemp\RichOle\0E6WZR0{18``ZKA(~Z{2RCA.png

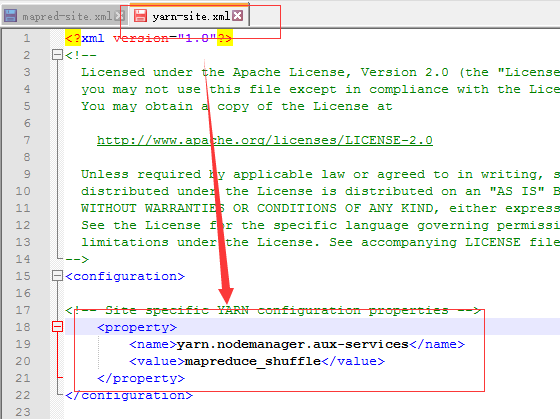
# 配置YARN：



按上图更改文件名

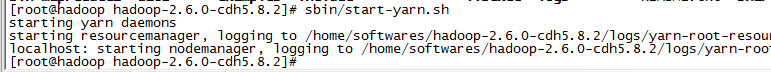


添加配置项

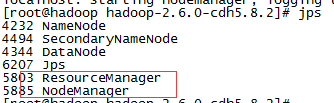


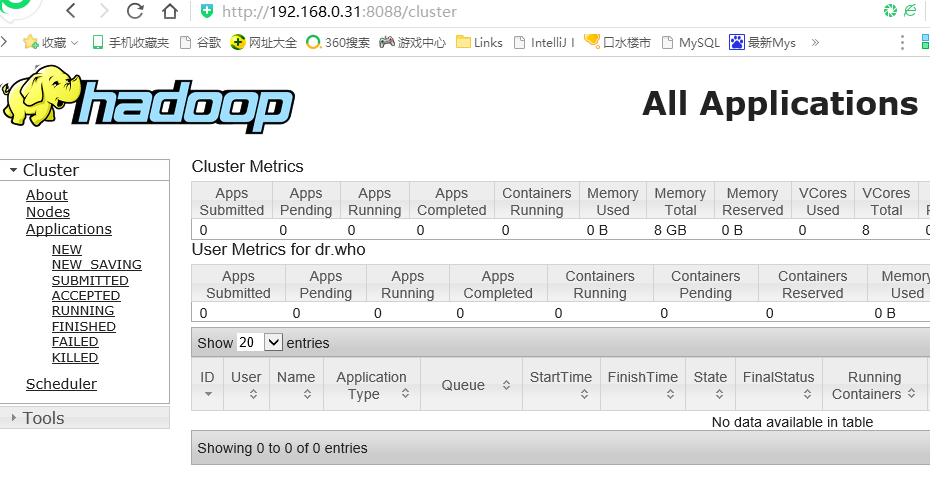
继续添加配置项

sbin/start-yarn.sh 命令启动yarn 相关的ResourceManager 守护进程和NodeManager 守护进程



Jps 查看一下是否启动成功。





输入网址查看yarn配置界面

安装配置可参考网址

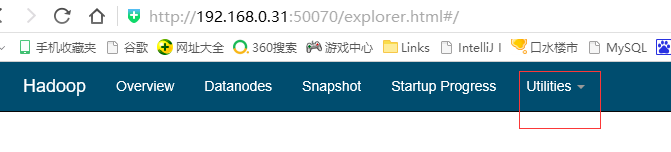
<http://hadoop.apache.org/docs/r2.6.5/hadoop-project-dist/hadoop-common/SingleCluster.html>

# 使用hadoop

上传数据文件到hadoop中的hdfs中：

C:\Users\QW\AppData\Roaming\Tencent\Users\747679105\QQ\WinTemp\RichOle\VN68(CX@[MXR8M{B1~(`VU2.png

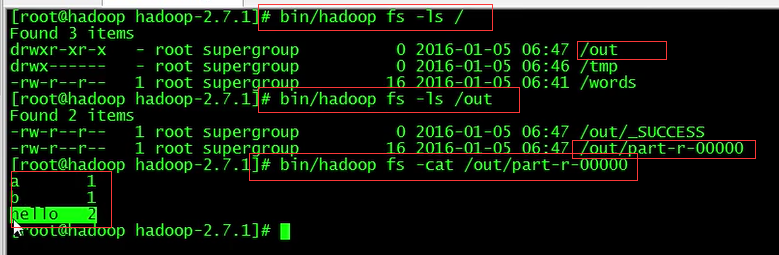
通过执行hadoop fs –put 命令，把/home/data/下的words 这个文件上传到 hdfs中，也叫words。



上传成功的文件可在utilties中的browse the file system 中看到并下载。

C:\Users\QW\AppData\Roaming\Tencent\Users\747679105\QQ\WinTemp\RichOle\69L69@UD[I4Q9FUN037AVUJ.png

对words文件进行词频数统计，结果输出到out文件中



查看out的统计结果