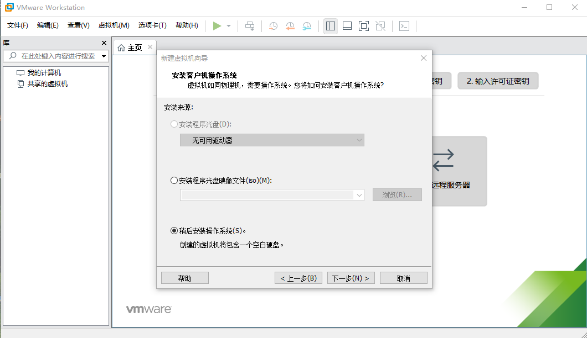
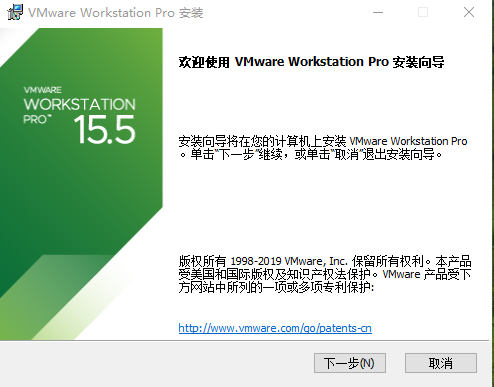
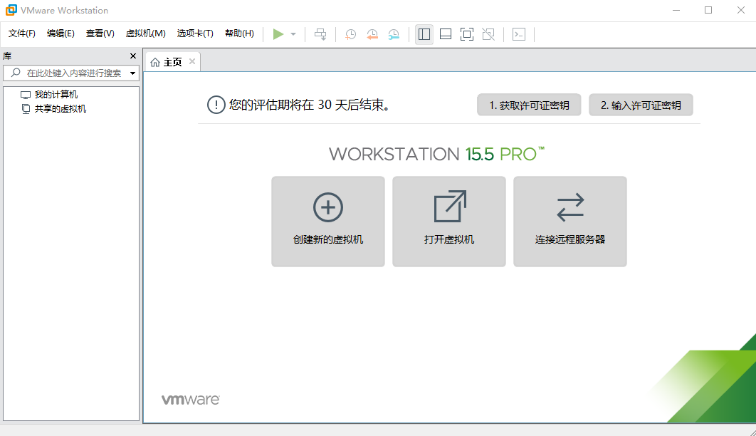
# 搭建Hadoop伪分布式大数据环境

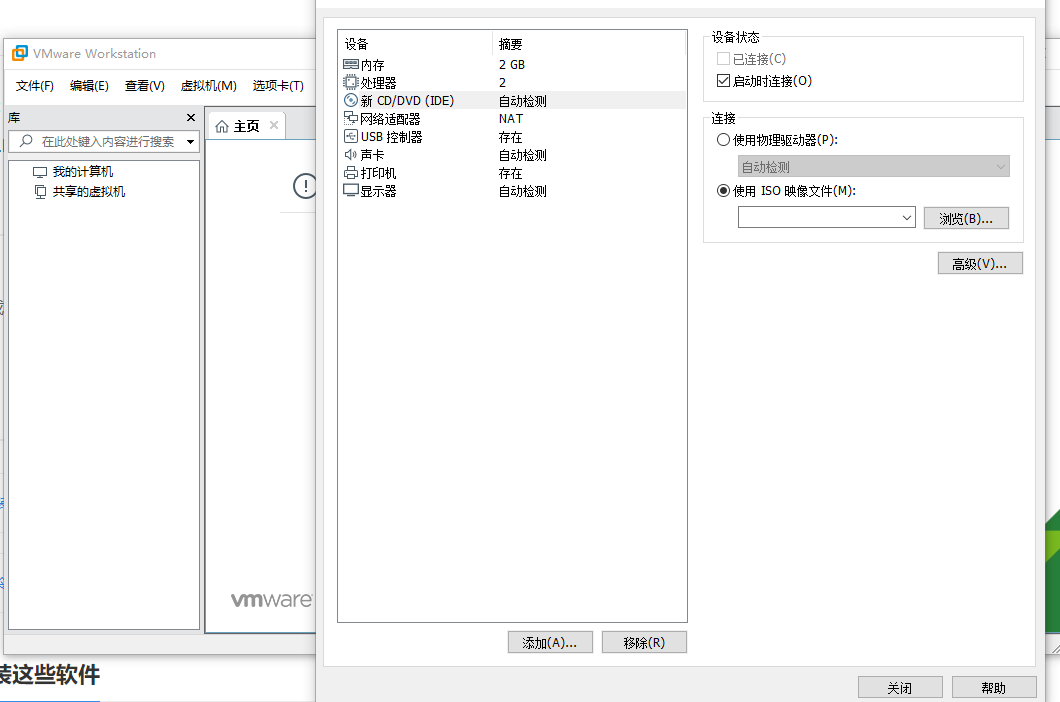
### 安装VMWare或VirtualBox虚拟机

在vm官网下载





1. 在虚拟机环境中安装Linux操作系统



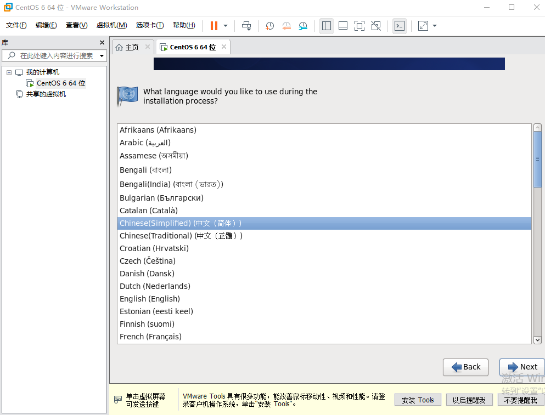
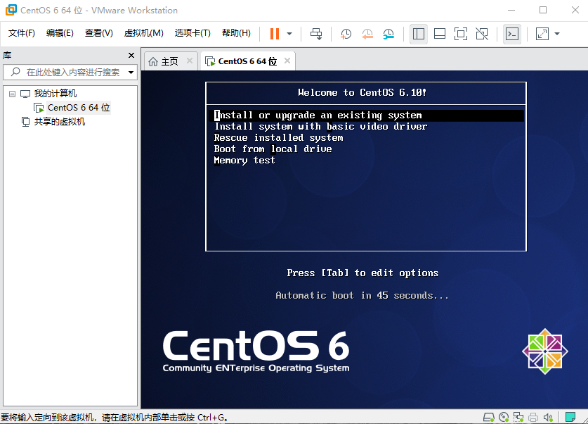
在虚拟机的dvd中放入linux系统文件

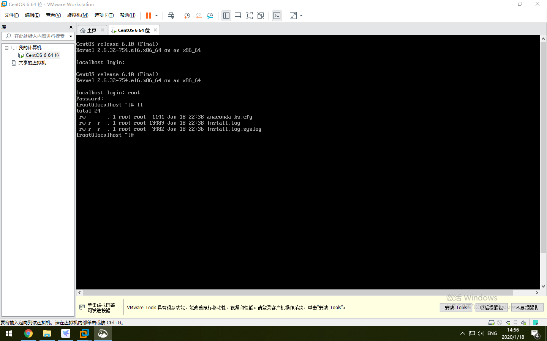
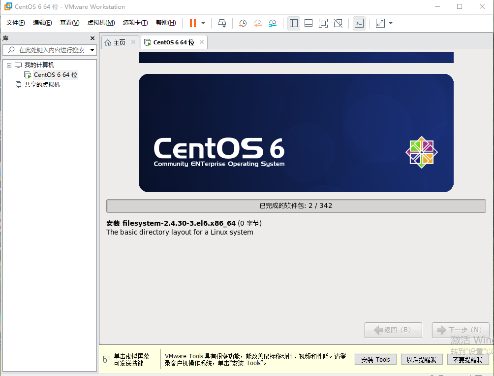
安装linux系统

下载centos6

<http://mirrors.cn99.com/centos/6.10/isos/x86_64/CentOS-6.10-x86_64-bin-DVD1.iso>

进入安装界面，

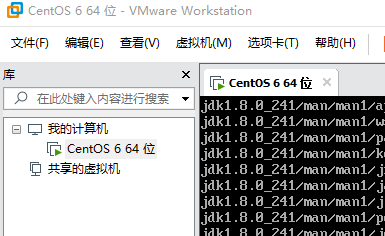




唯一需要注意的图形化界面不需要，选择mini版本安装即可

### 虚拟机中linux的网络问题

1.设置nat 在编辑中 点击 虚拟网络编辑器 得到如下图





2.进入虚拟机的虚拟网络编辑器

先点更改设置，

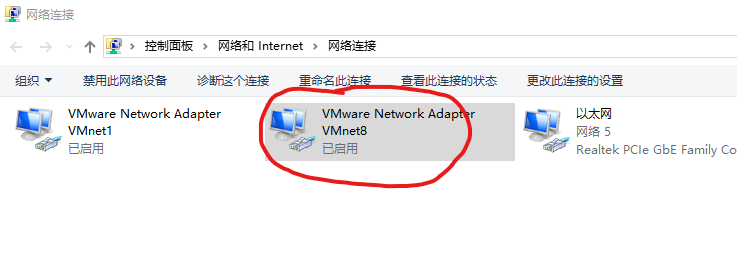
再分别进行 nat设置和dhcp设置



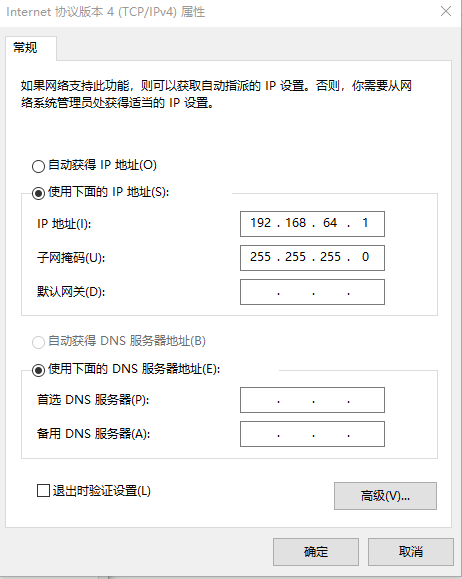
3.打开windows的控制面板 进入网络和共享中心



点击更改适配器设置，打开vm 网卡的属性

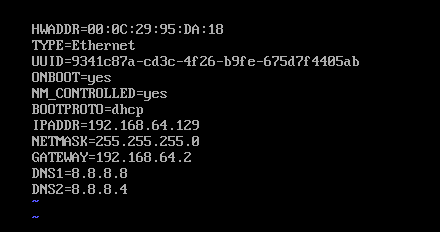


设置如下



4.进入linux

vi /etc/sysconfig/network-scripts/ifcfg-eth0 设置下列参数



保存之后执行命令 service network restart 重启网络连接

5.关闭linux防火墙

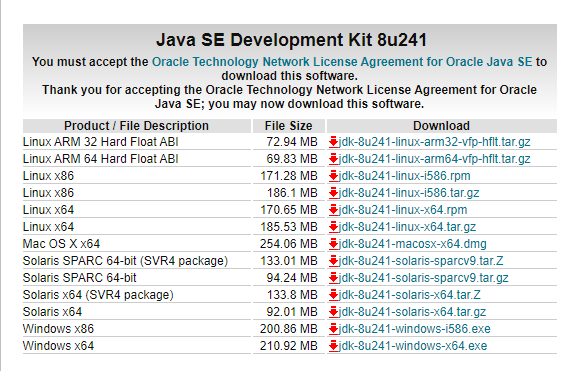
命令：chkconfig --level 2345 iptables off  
或者 chkconfig iptables off

6.下载xftp，xshell 对虚拟机远程操作

此时主机ip地址通过ifconfig命令可知。

### 在Linux系统中安装JDK并完成环境变量配置

1.下载jdk linux 64位的

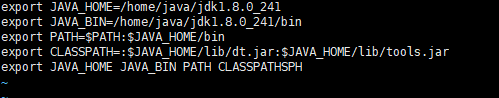


把下好的jdk通过xftp传入linux中

解压

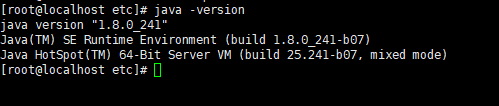
批注 2020-01-18 170846

修改 /etc/profile 增加以下内容



刷新

批注 2020-01-18 171937



安装成功

### 安装Hadoop，实现伪分布式配置

安装hadoop

在官网下载

<https://mirrors.tuna.tsinghua.edu.cn/apache/hadoop/common/hadoop-2.7.7/hadoop-2.7.7.tar.gz>

通过xftp传入linux

解压

批注 2020-01-18 172116

**配置hadoop-env.sh （在 Hadoop安装目录下的 etc/Hadoop 中）**

# The java implementation to use.

export JAVA\_HOME=/home/java/jdk1.8.0\_241

**配置core-site.xml（在 Hadoop安装目录下的 etc/Hadoop 中）**

<configuration>

<property>

        <name>fs.defaultFS</name>

        <value>hdfs://192.168.64.128:9000</value>

    </property>

    <property>

        <name>hadoop.tmp.dir</name>

        <value>/home/hadoop/hadoop-2.7.7/tmp</value>

    </property>

</configuration>

格式化 namenode**（在 Hadoop安装目录下）**

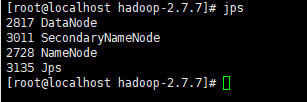
bin/hdfs namenode –format

批注 2020-01-18 173556

启动

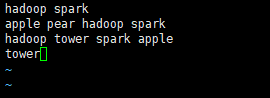
批注 2020-01-18 173635

验证 输入jps 查看进程



### 执行wordcount程序，统计指定目录下的词频

创建一个文件如下

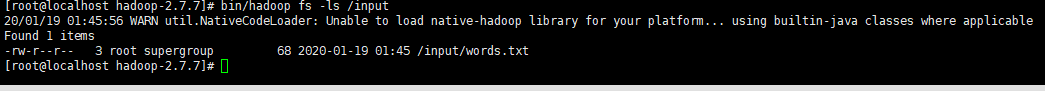


将此文件传到hdfs上

（提前建好目录 输入命令 bin/hdfs dfs –mkdir /input）

批注 2020-01-18 174538

查看



执行wordcount程序

批注 2020-01-18 175746

查看结果

