

大数据分析之环境部署

- 一、课程引入
- 二、课堂目标
- 三、需要准备好哪些软件?
- 四、安装vmware并配置网络
 - 4.1 安装vmware
 - 4.2 配置vmware网络步骤
- 五、安装CentOS7.5操作系统
 - 5.1 新建3台虚拟机
 - 5.2 编辑虚拟机
 - 5.3 安装CentOS7操作系统
 - 5.4 使用xshell连接虚拟机
- 六、CentOS7系统的基本常识
- 七、常用的Linux命令
- 八、虚拟机基本配置
- 九、环境测试
- 十、总结与作业

大数据分析之环境部署

一、课程引入

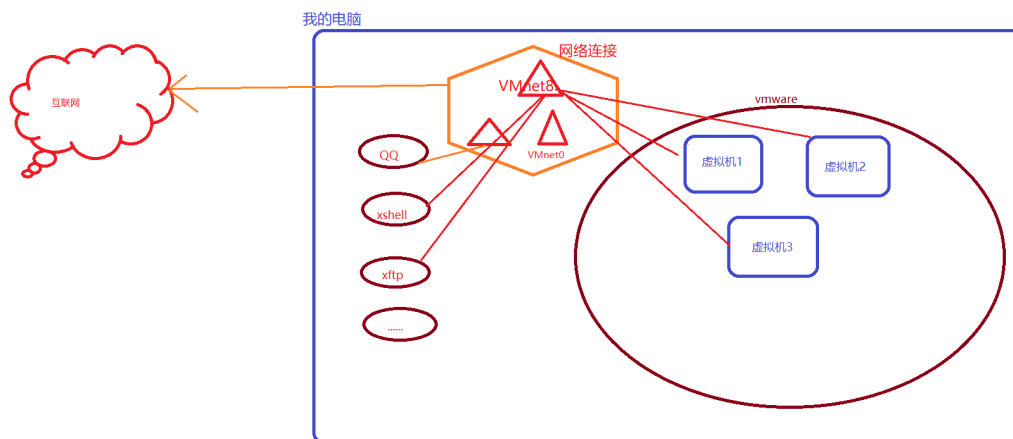
- 回顾已经学过的技能:Excel/MySQL/Tableau/Power BI/Python
- 大数据分析和它们有什么不同?
- 为什么需要在自己电脑上部署环境?
- 课程介绍①大数据分析之环境部署②大数据分析之Hadoop学习③大数据分析之Hive学习
- 行业案例: 企业级HiveSQL的应用

二、课堂目标

1. 理解大数据分析的开发环境
2. 尝试在自己的电脑上搭建好开发环境
3. 掌握CentOS7系统常见的操作

三、需要准备好哪些软件?

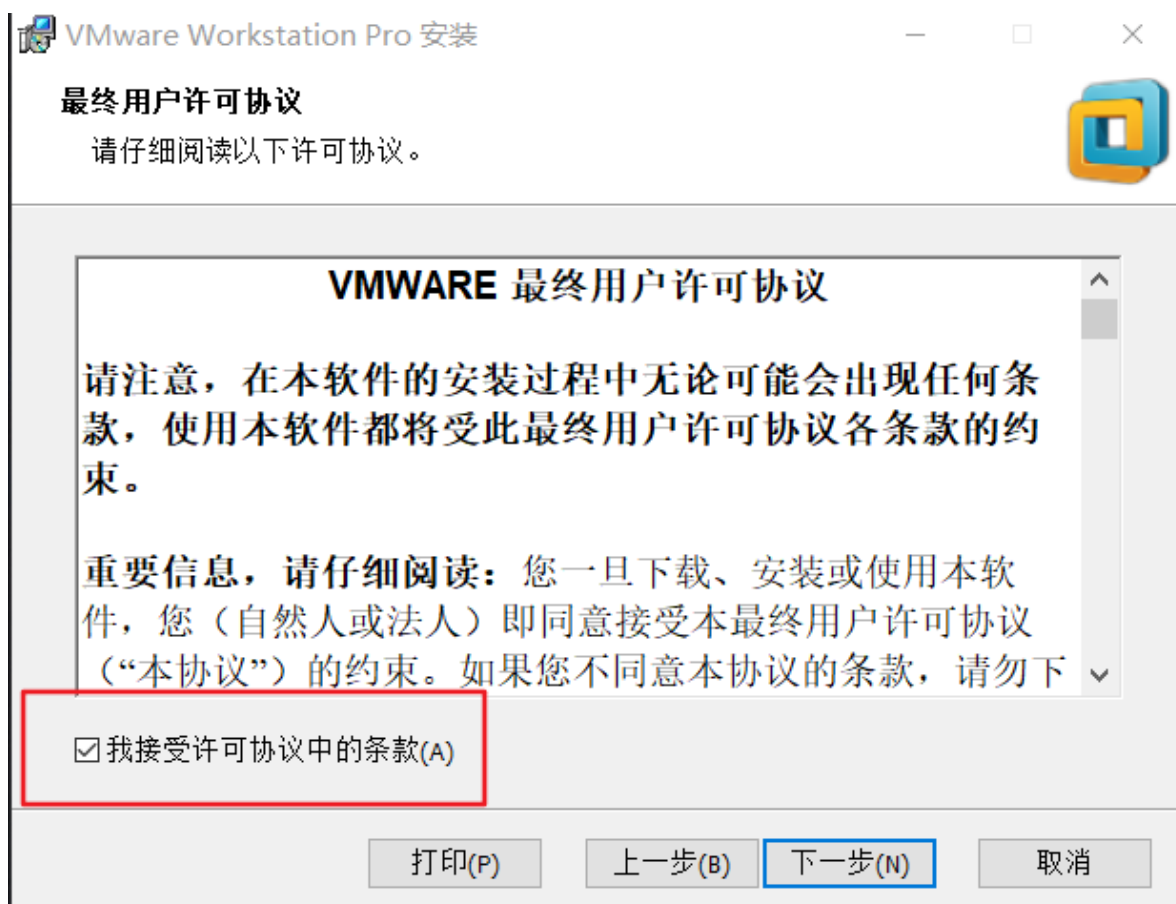
- VMware
- Xshell6
- Xftp6
- CentOS7系统
- hadoop(服务器上)
- jdk(服务器上)
- hive(服务器上)



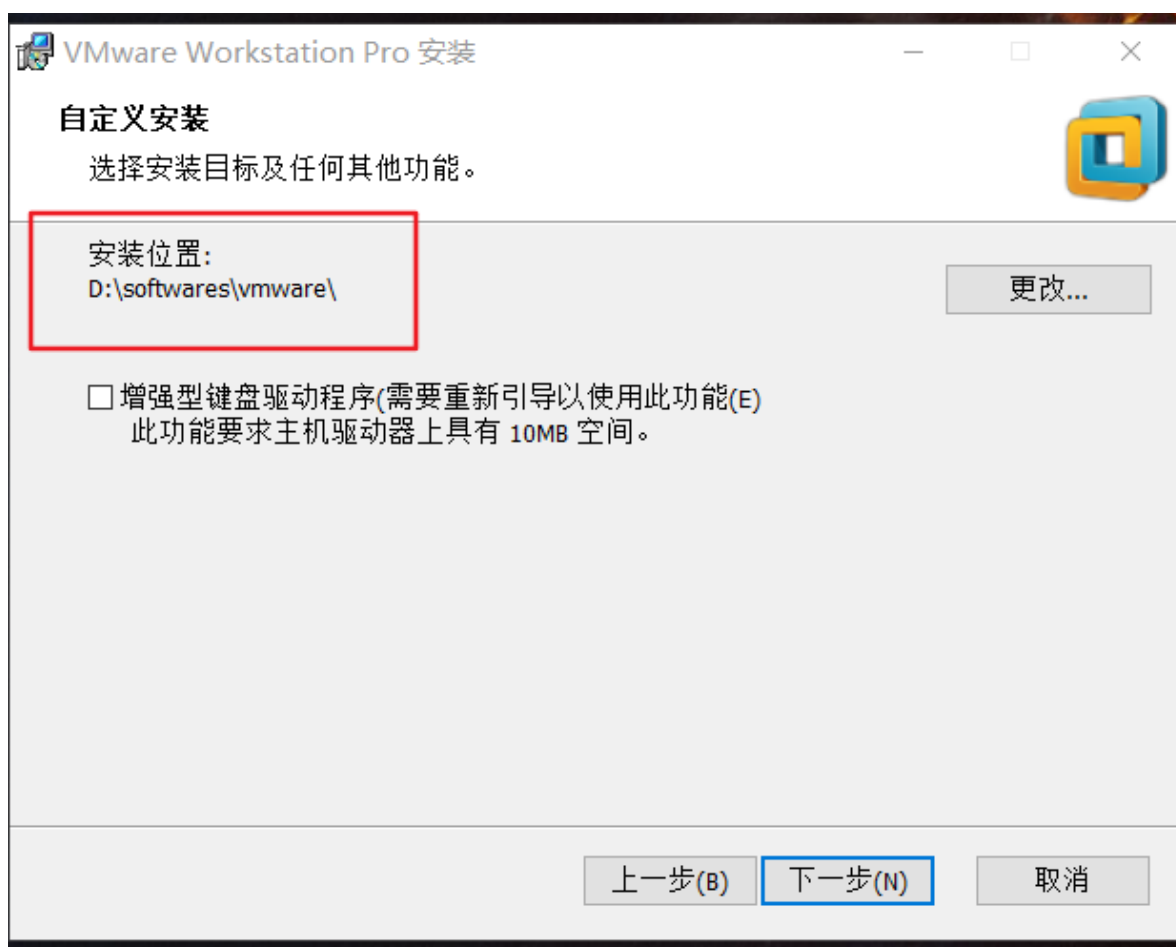
四、安装vmware并配置网络

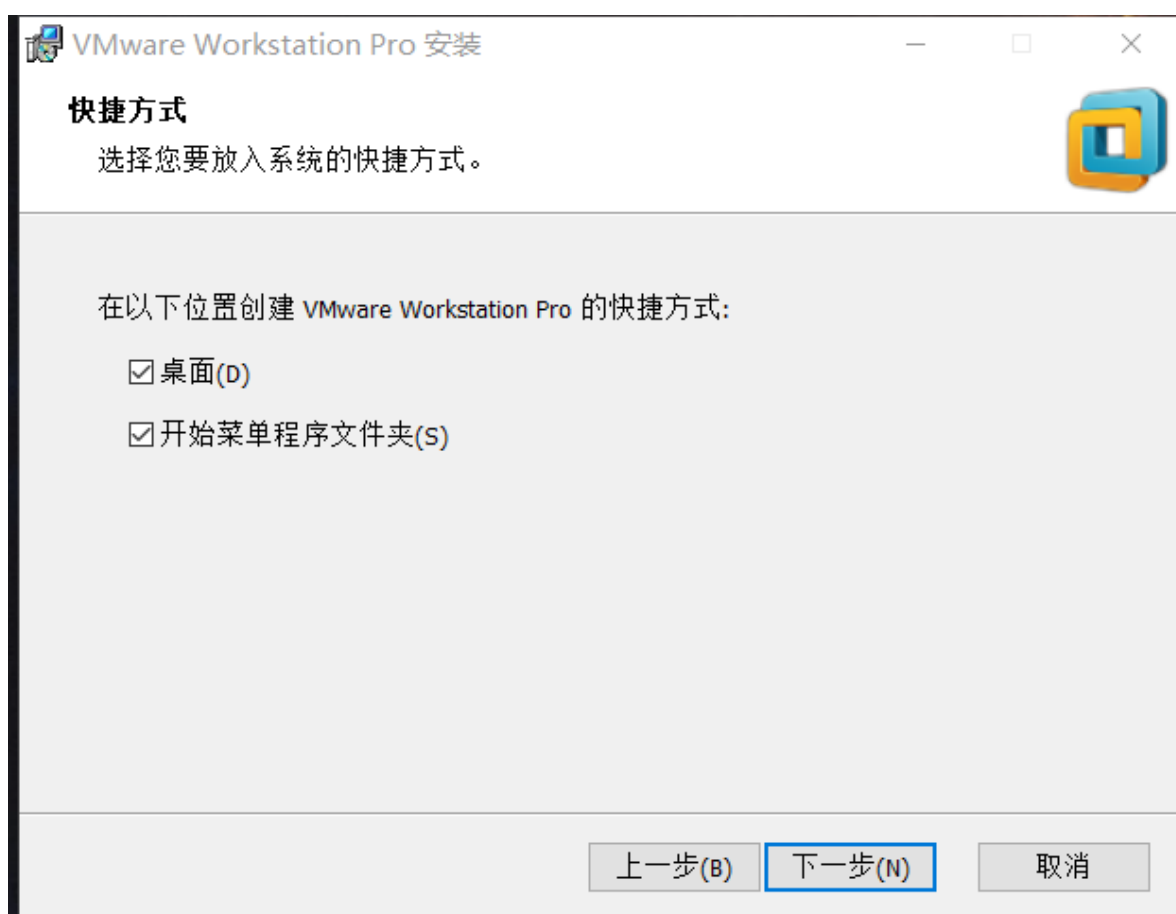
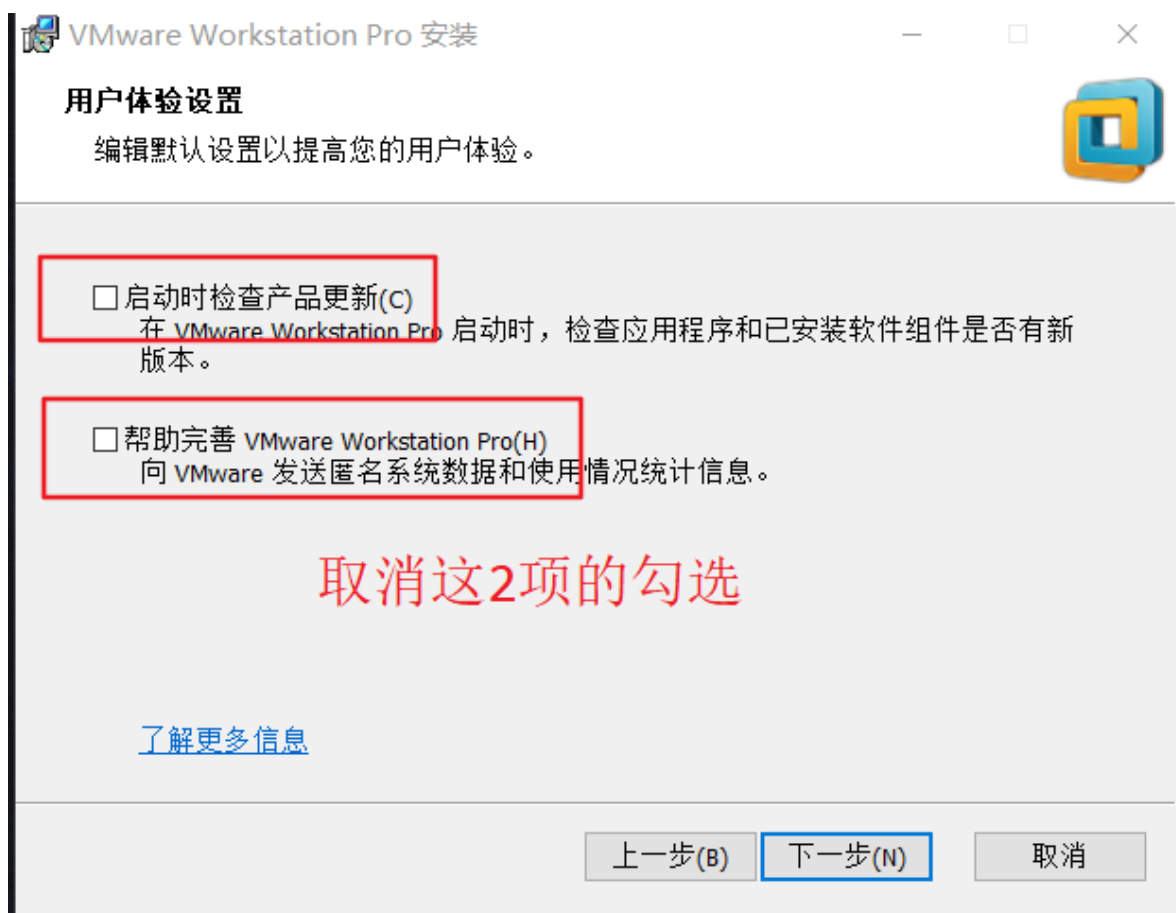
4.1 安装vmware





这里自己选择一个安装位置即可







4.2 配置vmware网络步骤



网络 共享

连接时使用:



VMware Virtual Ethernet Adapter for VMnet8

配置(C)...

此连接使用下列项目(O):

- ☒ Microsoft 网络客户端
- ☐ VMware Bridge Protocol
- ☒ Microsoft 网络的文件和打印机共享
- ☒ Internet 协议版本 4 (TCP/IPv4)
- ☐ Microsoft 网络适配器多路传送器协议
- ☒ Microsoft LLDP 协议驱动程序
- ☒ Internet 协议版本 6 (TCP/IPv6)
- ☒ 链路层拓扑发现响应程序

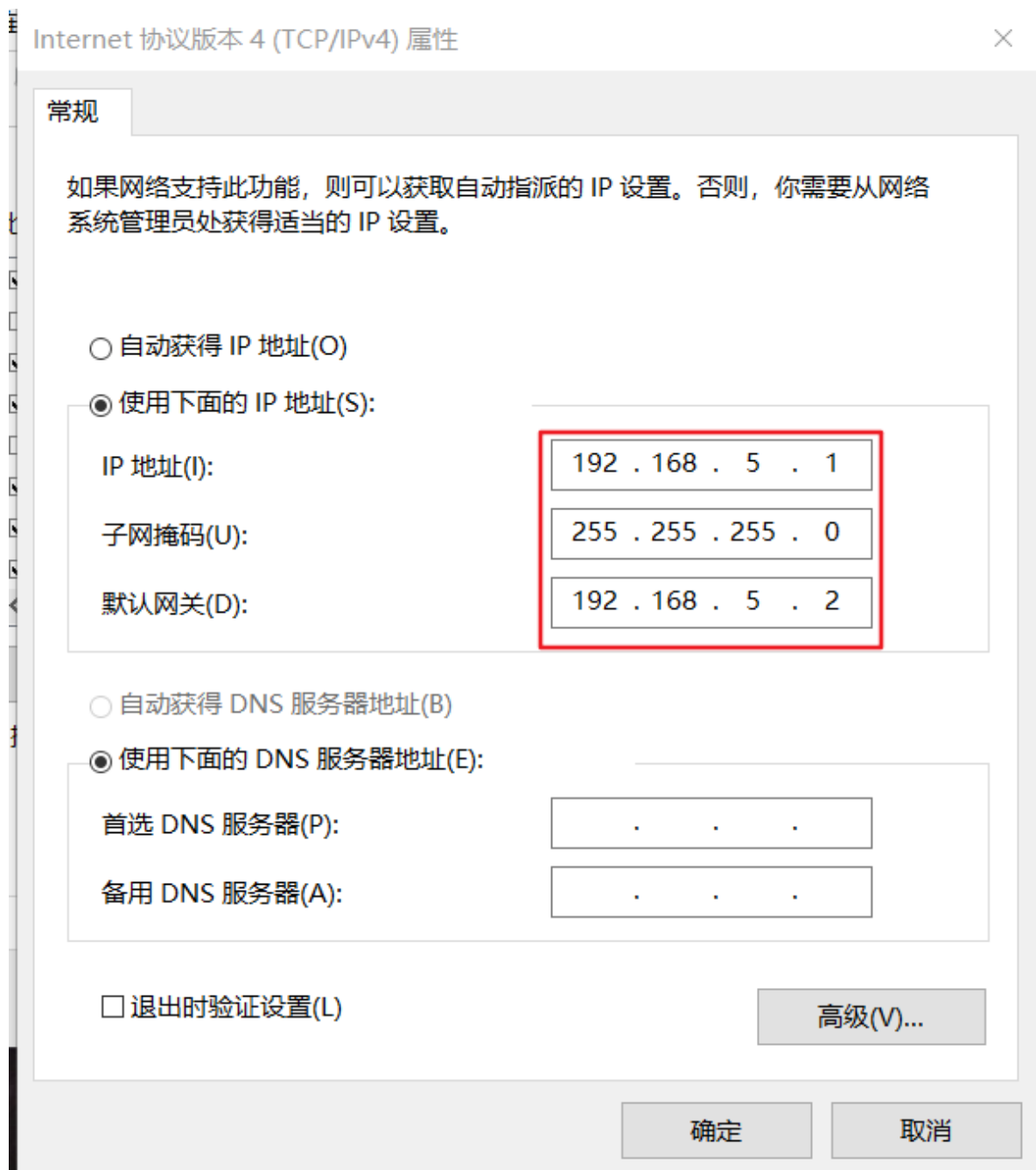
安装(N)...

卸载(U)

属性(R)

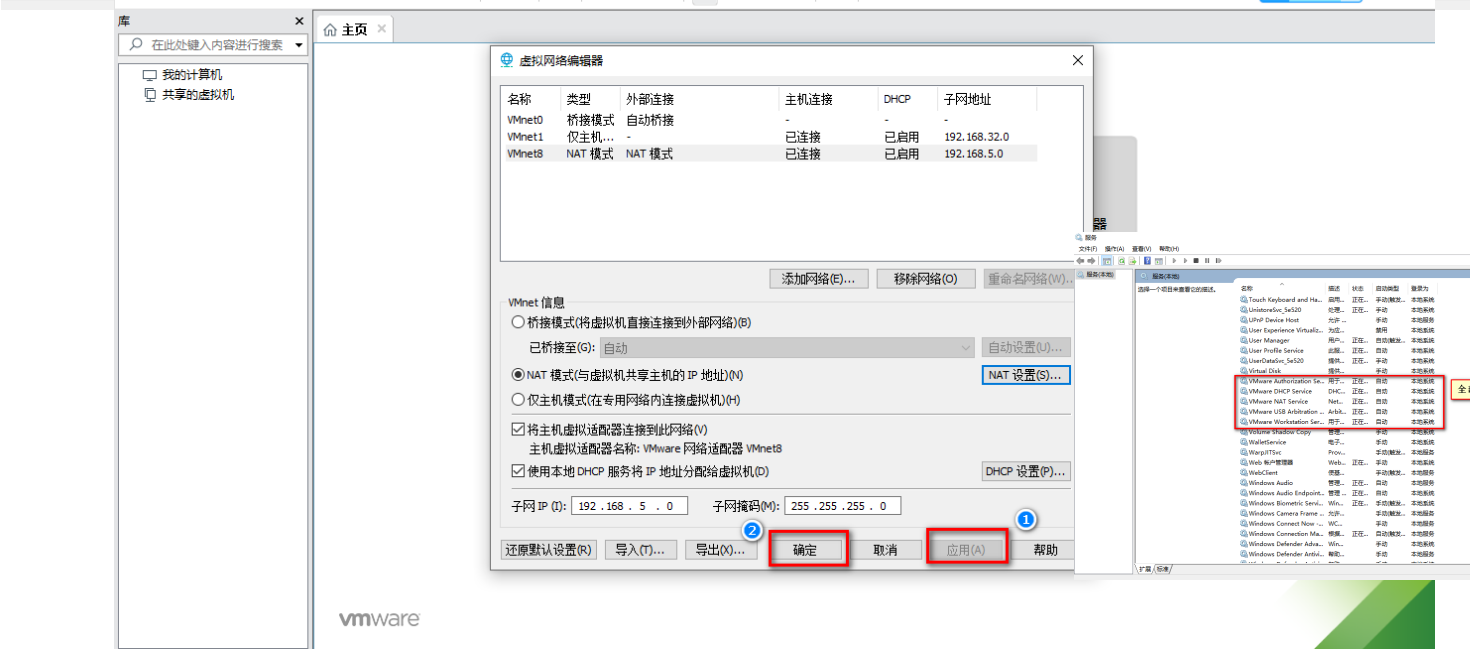
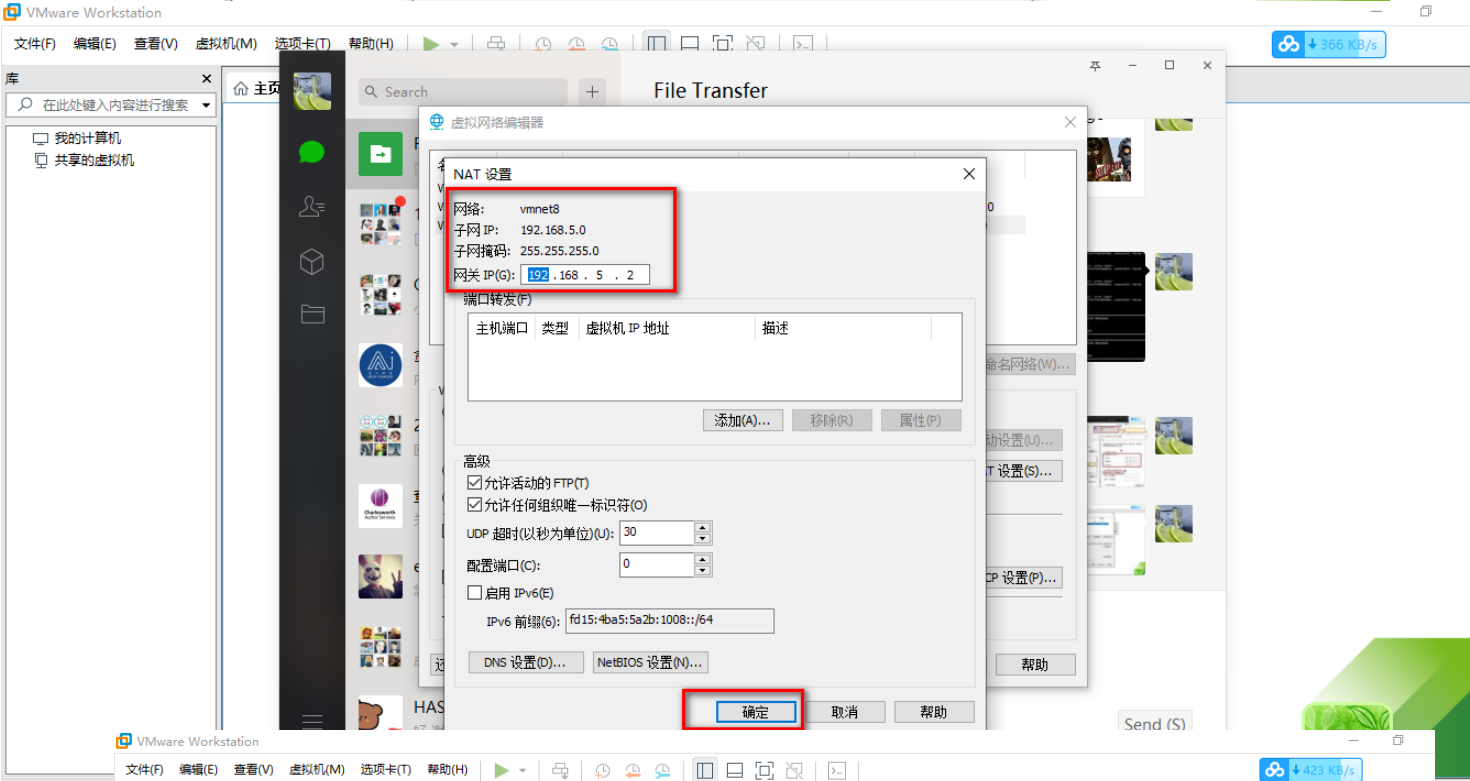
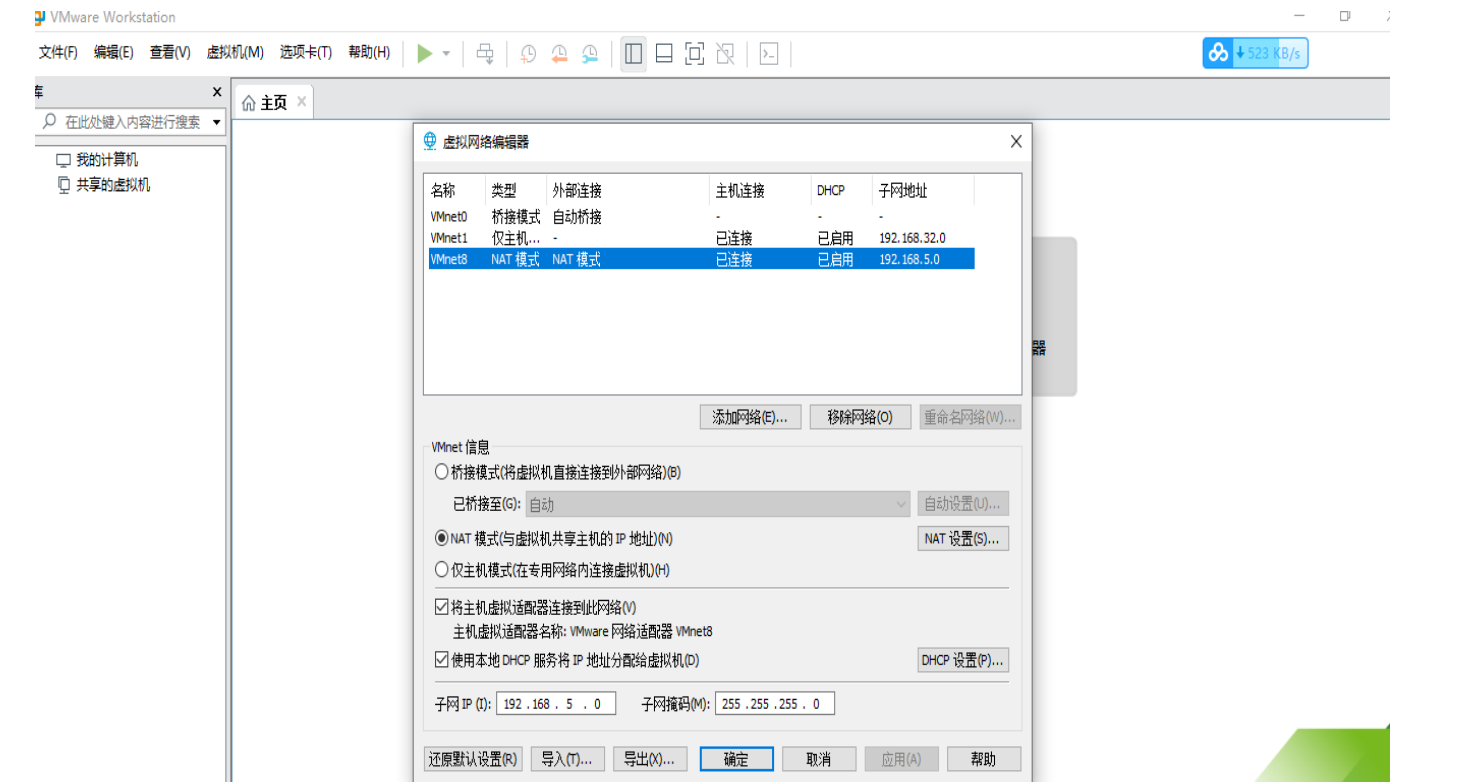
描述

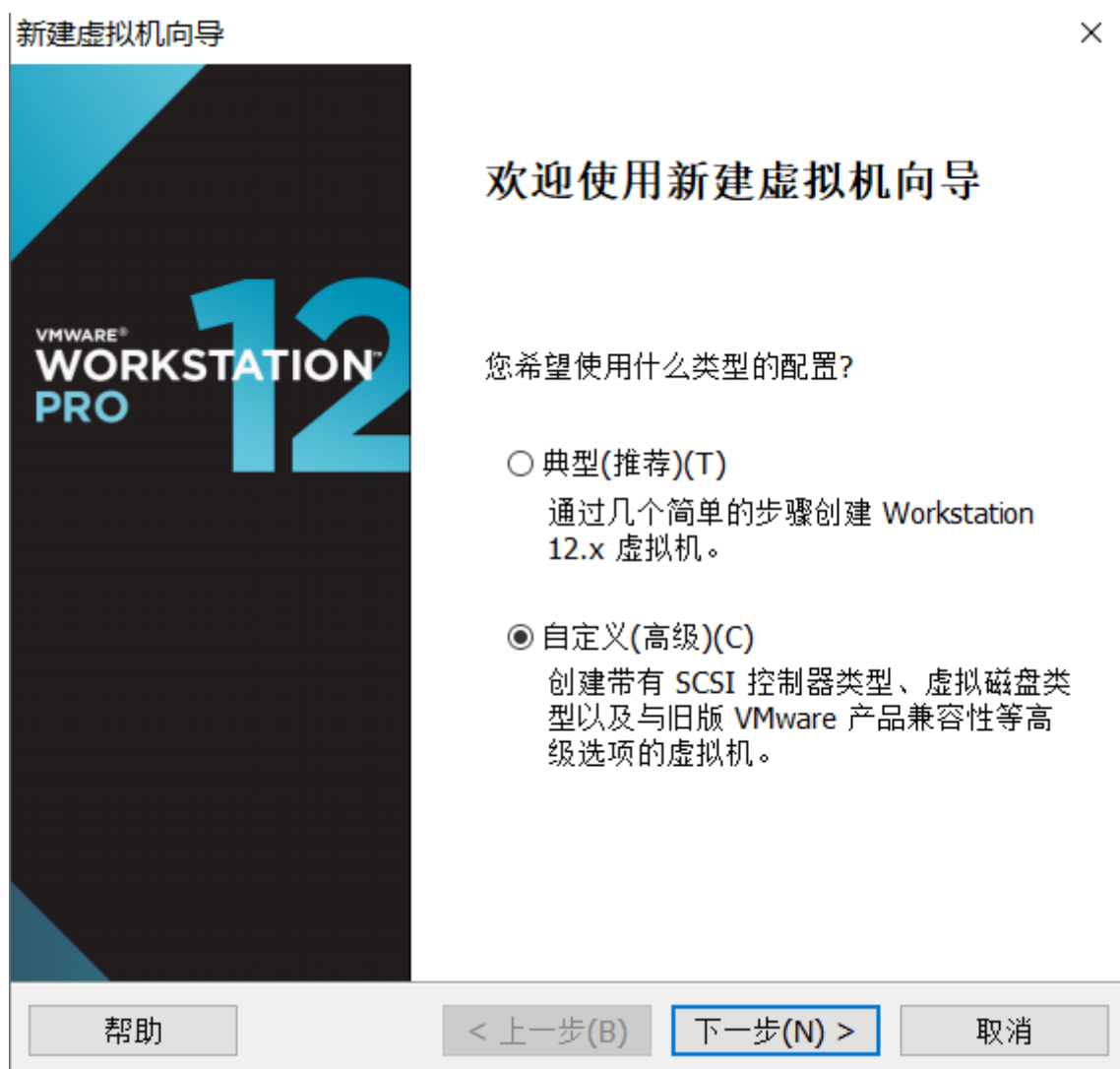
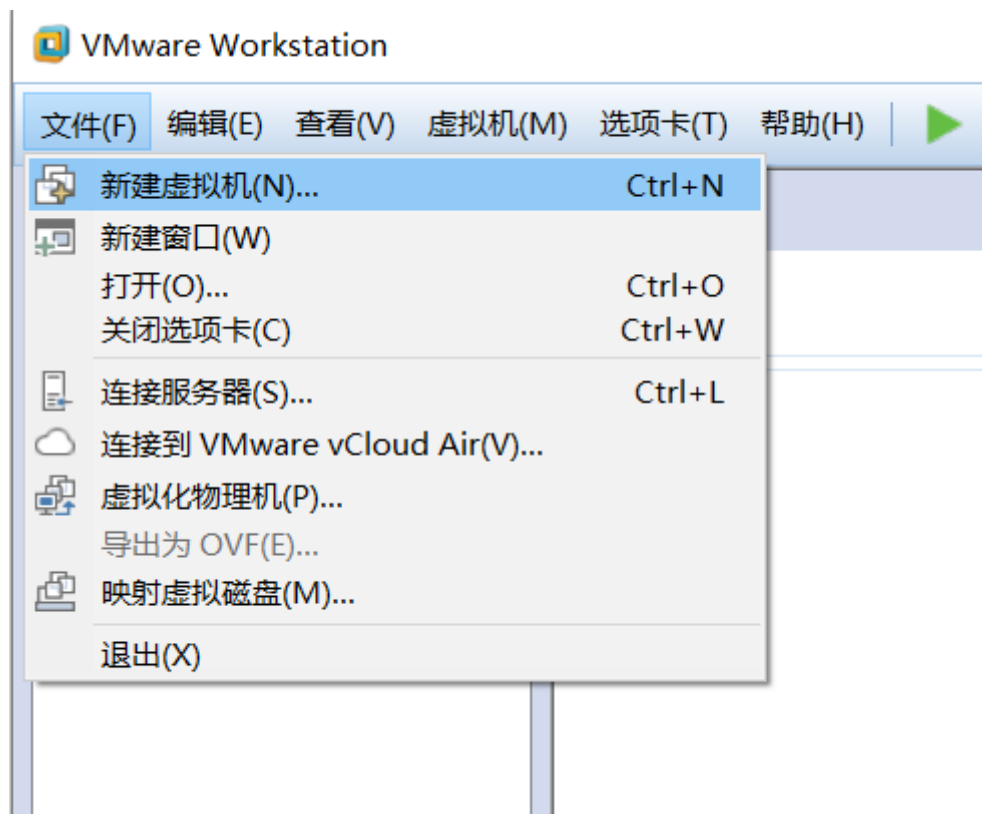
传输控制协议/Internet 协议。该协议是默认的广域网络协议，用于在不同的相互连接的网络上通信



五、安装CentOS7.5操作系统

5.1 新建3台虚拟机





选择虚拟机硬件兼容性

该虚拟机需要何种硬件功能？

虚拟机硬件兼容性

硬件兼容性(H): Workstation 12.x ▾

兼容: ☒ ESX Server(S)

兼容产品:

Fusion 8.x
Workstation 12.x

^
v

限制:

64 GB 内存
16 个处理器
10 个网络适配器
8 TB 磁盘大小

^
v

帮助 < 上一步(B) 下一步(N) > 取消

安装客户机操作系统

虚拟机如同物理机，需要操作系统。您将如何安装客户机操作系统？

安装来源：

☐ 安装程序光盘(D):

无可用驱动器

☐ 安装程序光盘映像文件(iso)(M):

D:\softpackages\CentOS-7-x86_64-DVD-1810.iso

浏览(R)...

☒ 稍后安装操作系统(S)。

创建的虚拟机将包含一个空白硬盘。

帮助

< 上一步(B)

下一步(N) >

取消

选择客户机操作系统

此虚拟机中将安装哪种操作系统？

客户机操作系统

- ☐ Microsoft Windows(W)
- ☒ Linux(L)
- ☐ Novell NetWare(E)
- ☐ Solaris(S)
- ☐ VMware ESX(X)
- ☐ 其他(O)

版本(V)

CentOS 64 位

帮助

< 上一步(B)

下一步(N) >

取消

命名虚拟机

您要为此虚拟机使用什么名称？

虚拟机名称(V):

node201

位置(L):

D:\data\VMachines\node201

浏览(R)...

在“编辑”>“首选项”中可更改默认位置。

第一台虚拟机取名为node201，
第二台和第三台为node202， node203

< 上一步(B)

下一步(N) >

取消

新建虚拟机向导

✕

处理器配置

为此虚拟机指定处理器数量。

处理器

处理器数量(P):

2

▼

每个处理器的核心数量(C):

2

▼

总处理器核心数量:

4

帮助

< 上一步(B)

下一步(N) >

取消

此虚拟机的内存

您要为此虚拟机使用多少内存？

指定分配给此虚拟机的内存量。内存大小必须为 **4 MB** 的倍数。

64 GB -
32 GB -
16 GB -
8 GB -
4 GB -
2 GB -
1 GB -
512 MB -
256 MB -
128 MB -
64 MB -
32 MB -
16 MB -
8 MB -
4 MB -

此虚拟机的内存(M):

1024 MB

最大推荐内存:
13716 MB

推荐内存:
1024 MB

客户机操作系统最低推荐内存:
512 MB

帮助

< 上一步(B)

下一步(N) >

取消

网络类型

要添加哪类网络？

网络连接

☐ 使用桥接网络(R)

为客户机操作系统提供直接访问外部以太网网络的权限。客户机在外部网络上必须有自己的 IP 地址。

☒ 使用网络地址转换(NAT)(E)

为客户机操作系统提供使用主机 IP 地址访问主机拨号连接或外部以太网网络连接的权限。

☐ 使用仅主机模式网络(H)

将客户机操作系统连接到主机上的专用虚拟网络。

☐ 不使用网络连接(T)

帮助

< 上一步(B)

下一步(N) >

取消

选择 I/O 控制器类型

您要使用何种类型的 SCSI 控制器？

I/O 控制器类型

SCSI 控制器：

- ☐ BusLogic(U) (不适用于 64 位客户机)
- ☒ LSI Logic(L) (推荐)
- ☐ LSI Logic SAS(S)

帮助

< 上一步(B)

下一步(N) >

取消

新建虚拟机向导

×

选择磁盘类型

您要创建何种磁盘?

虚拟磁盘类型

☐ IDE(I)

☒ SCSI(S) (推荐)

☐ SATA(A)

帮助

< 上一步(B)

下一步(N) >

取消

选择磁盘

您要使用哪个磁盘？

磁盘☒ **创建新虚拟磁盘(V)**

虚拟磁盘由主机文件系统上的一个或多个文件组成，客户机操作系统会将其视为单个硬盘。虚拟磁盘可在一台主机上或多台主机之间轻松复制或移动。

☐ **使用现有虚拟磁盘(E)**

选择此选项将重新使用之前配置的磁盘。

☐ **使用物理磁盘(适用于高级用户)(P)**

选择此选项将为虚拟机提供直接访问本地硬盘的权限。

帮助

< 上一步(B)

下一步(N) >

取消

指定磁盘容量

磁盘大小为多少？

最大磁盘大小(GB)(S):

50

针对 CentOS 64 位 的建议大小: 20 GB

☐ 立即分配所有磁盘空间(A)。

分配所有容量可以提高性能，但要求所有物理磁盘空间立即可用。如果不立即分配所有空间，虚拟磁盘的空间最初很小，会随着您向其中添加数据而不断变大。

☐ 将虚拟磁盘存储为单个文件(O)☒ 将虚拟磁盘拆分成多个文件(M)

拆分磁盘后，可以更轻松地在计算机之间移动虚拟机，但可能会降低大容量磁盘的性能。

帮助

< 上一步(B)

下一步(N) >

取消

指定磁盘文件

您要在何处存储磁盘文件？

磁盘文件

将为每 **2 GB** 容量的虚拟磁盘创建一个磁盘文件。除第一个文件之外，每个文件的文件名称将根据此处所提供的文件名称自动生成。

已准备好创建虚拟机

单击“完成”创建虚拟机。然后可以安装 CentOS 64 位。

将使用下列设置创建虚拟机：

名称：	node201
位置：	D:\data\VMachines\node201
版本：	Workstation 12.x
操作系统：	CentOS 64 位
硬盘：	50 GB, 拆分
内存：	1024 MB
网络适配器：	NAT
其他设备：	4 个 CPU 核心, CD/DVD, USB 控制器, 打印机, 声卡

自定义硬件(C)...

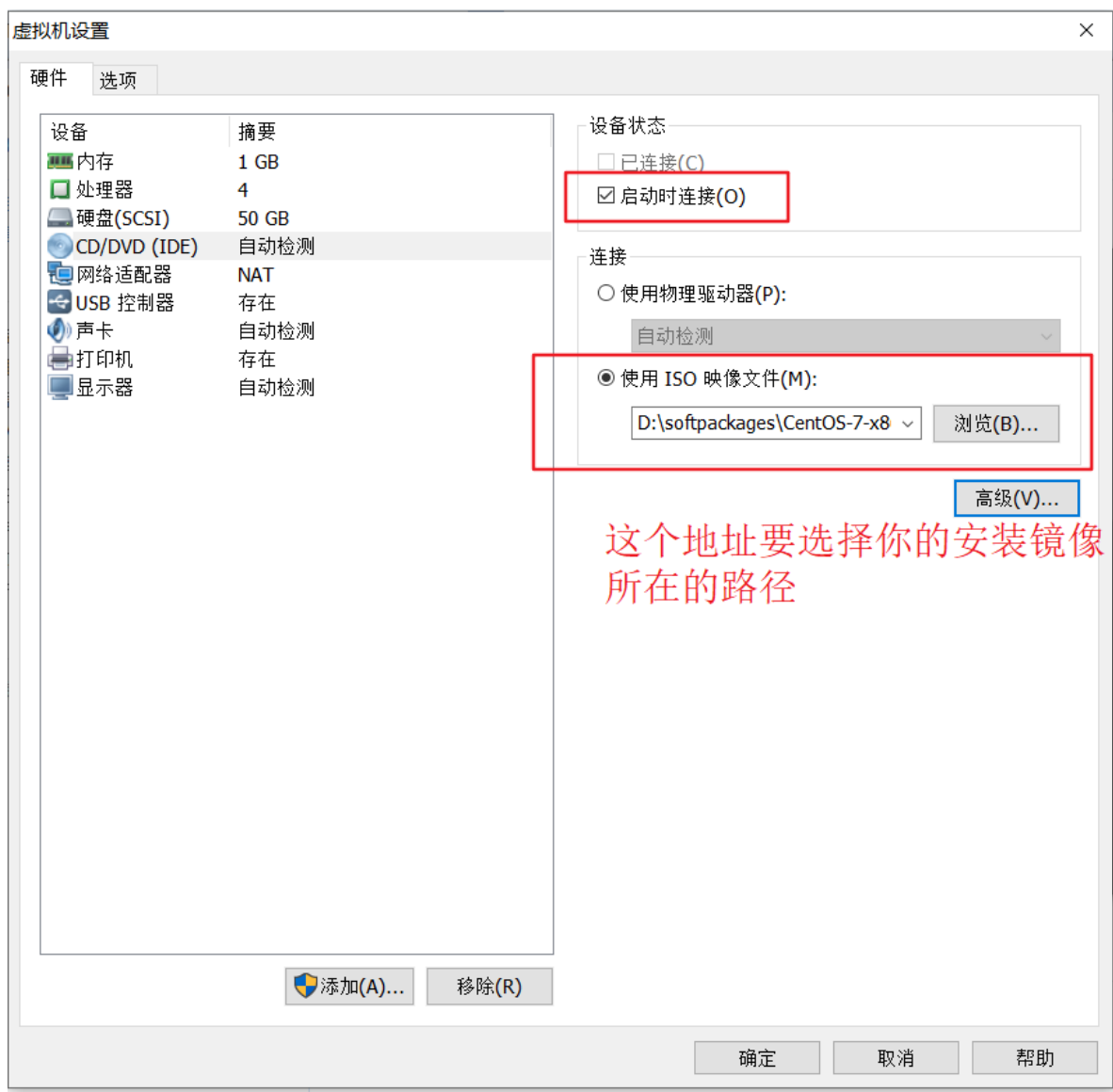
< 上一步(B)

完成

取消

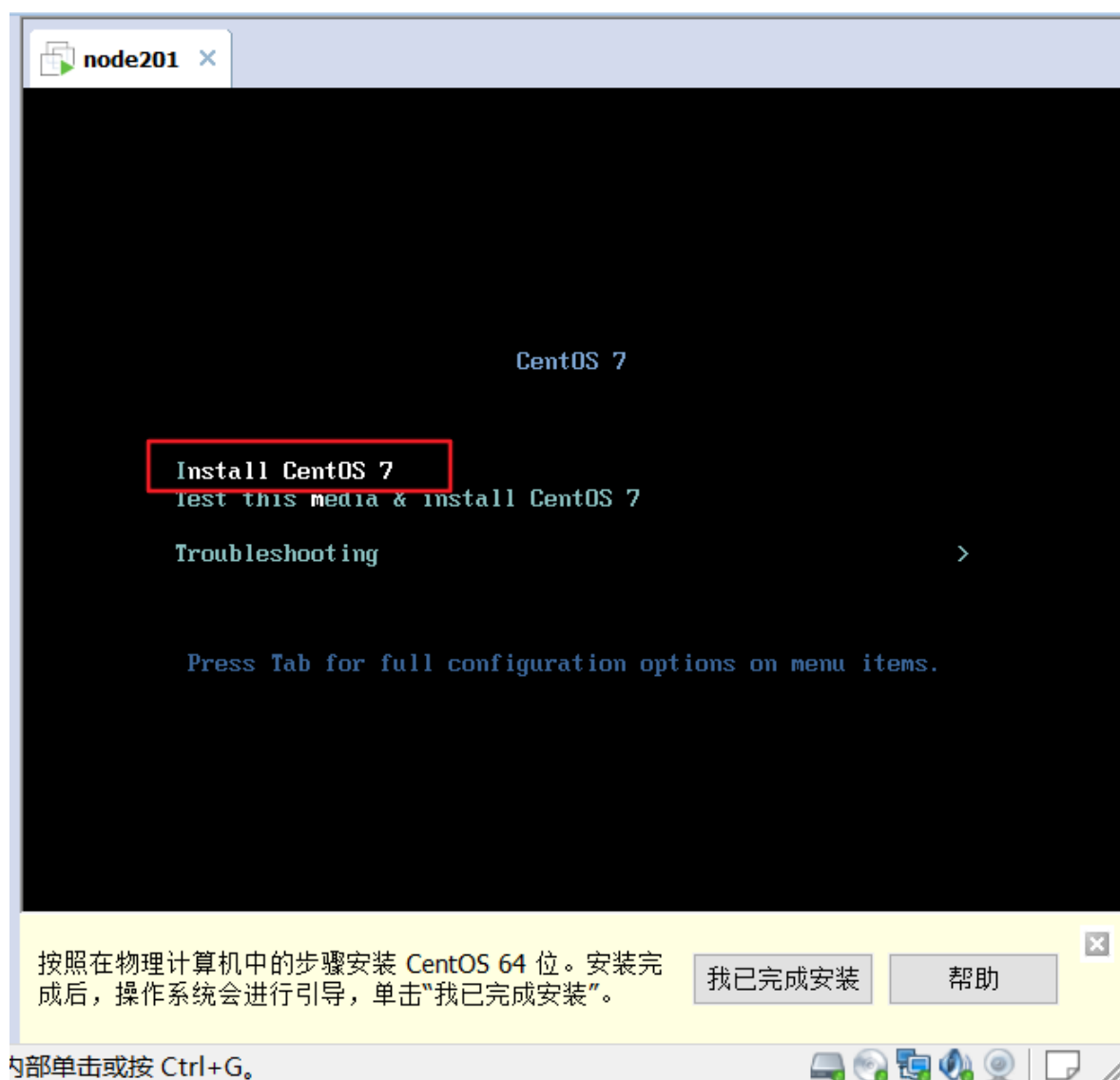
5.2 编辑虚拟机

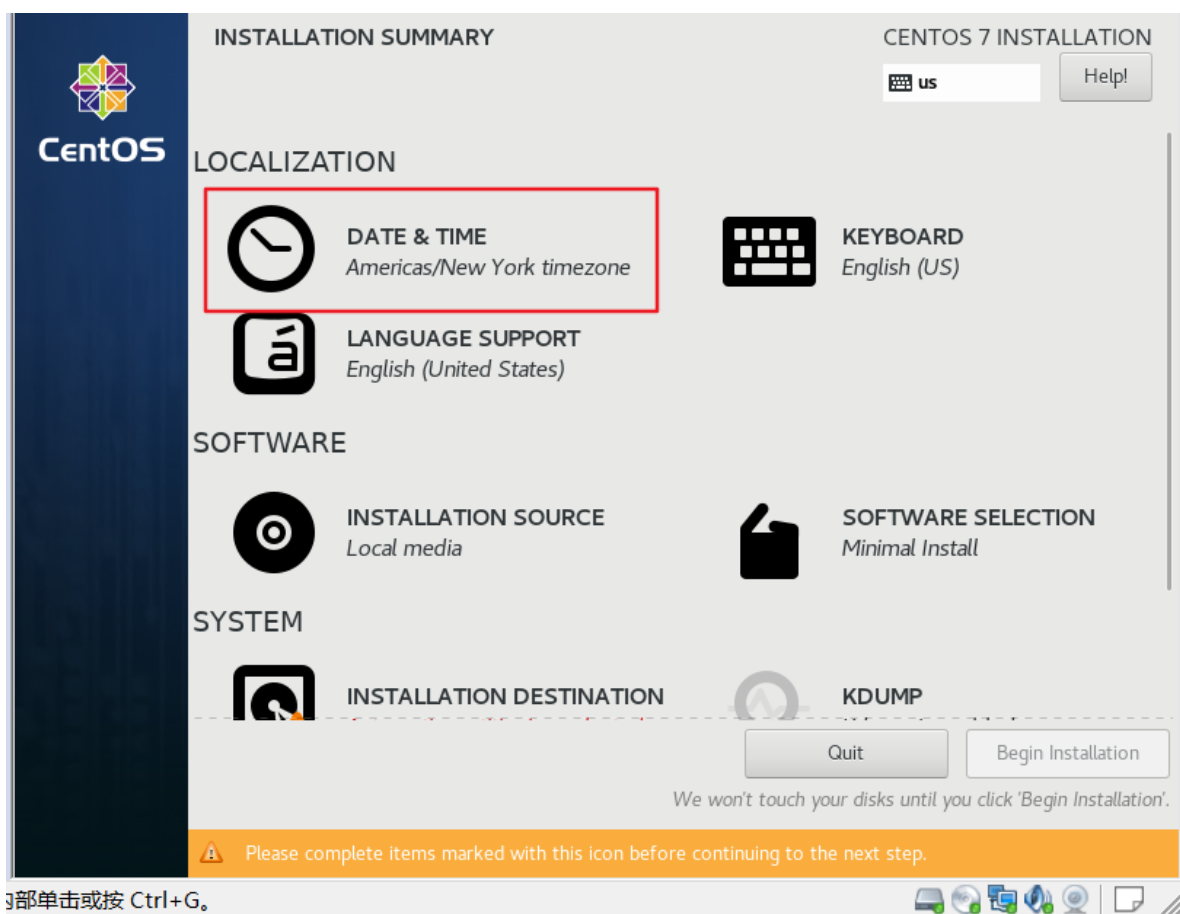
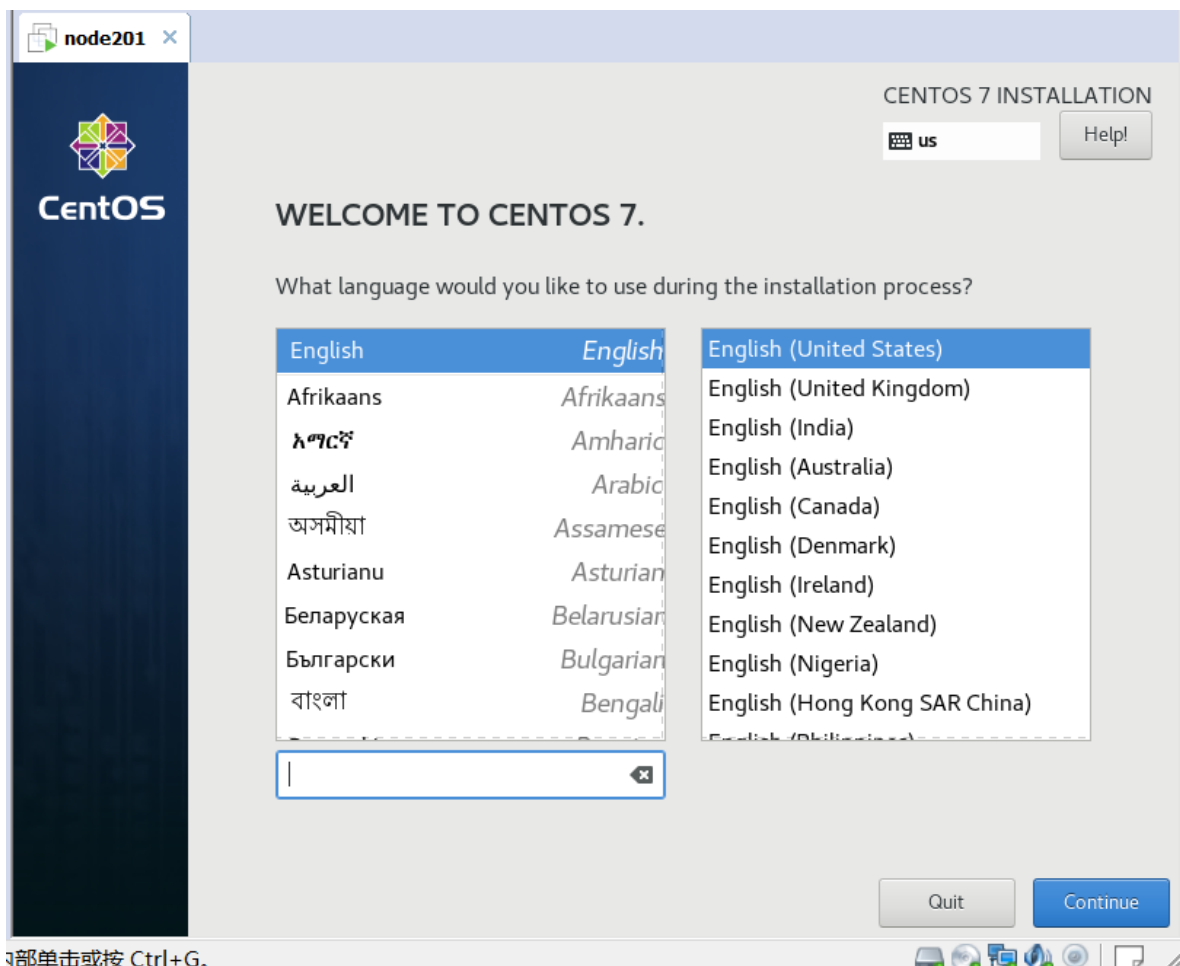


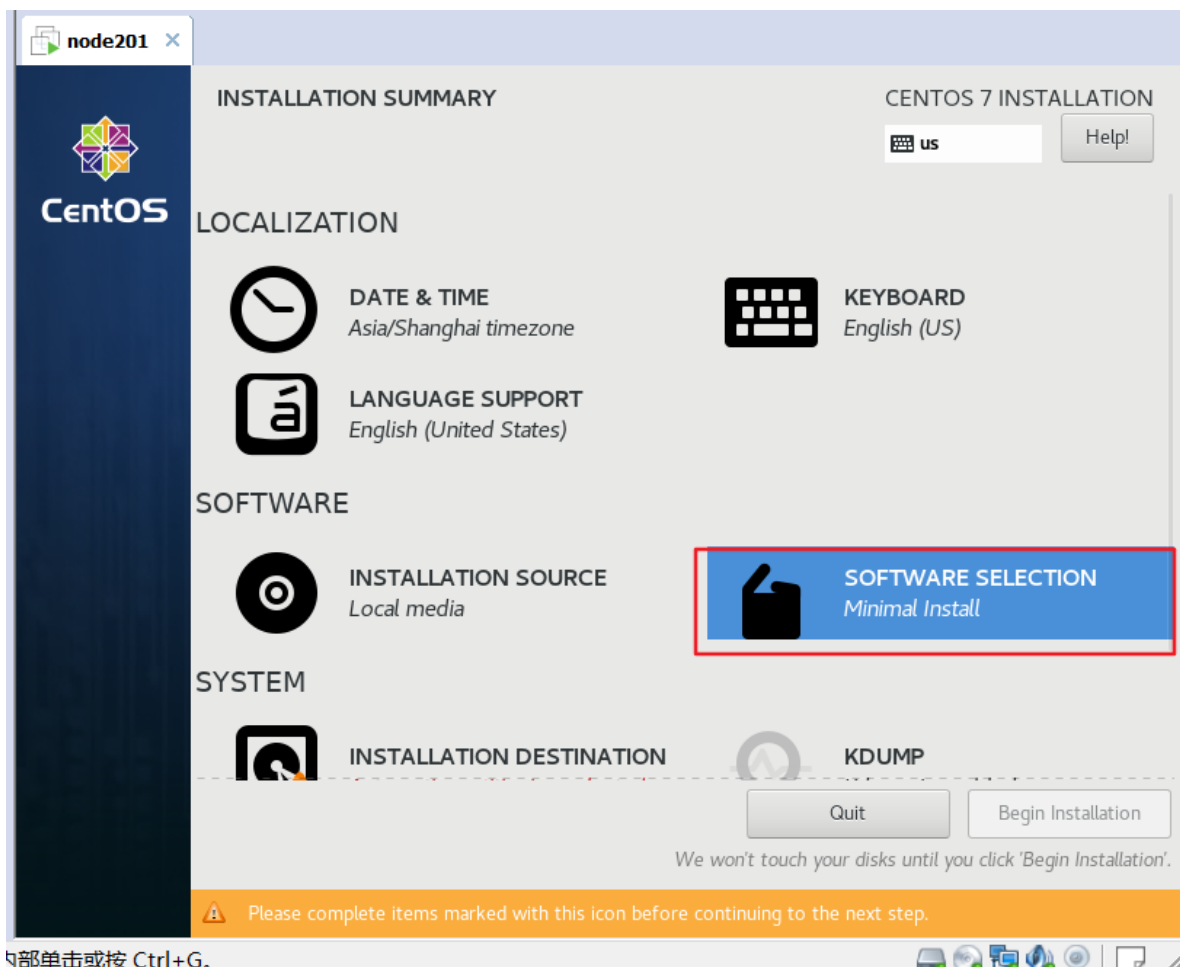
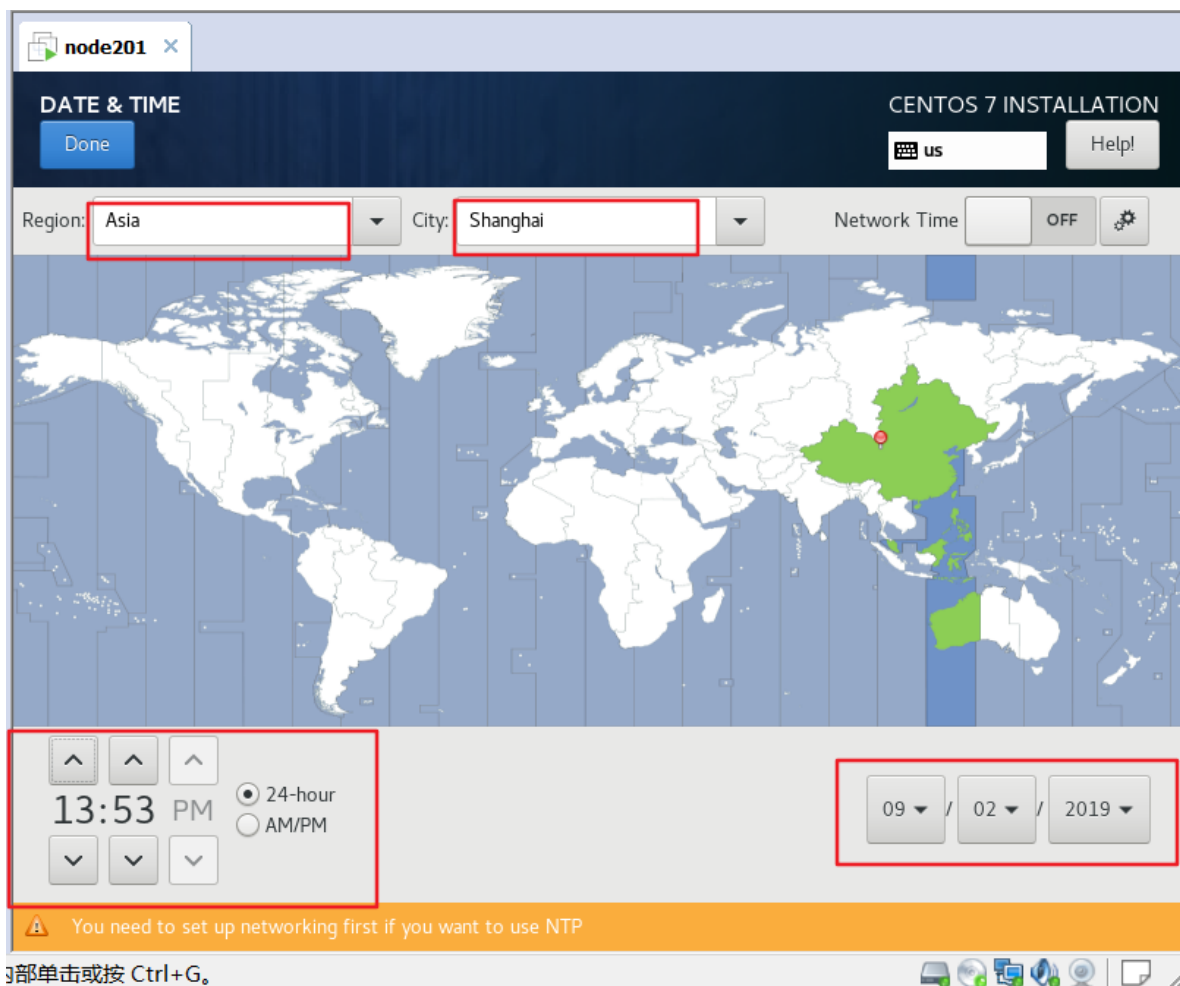


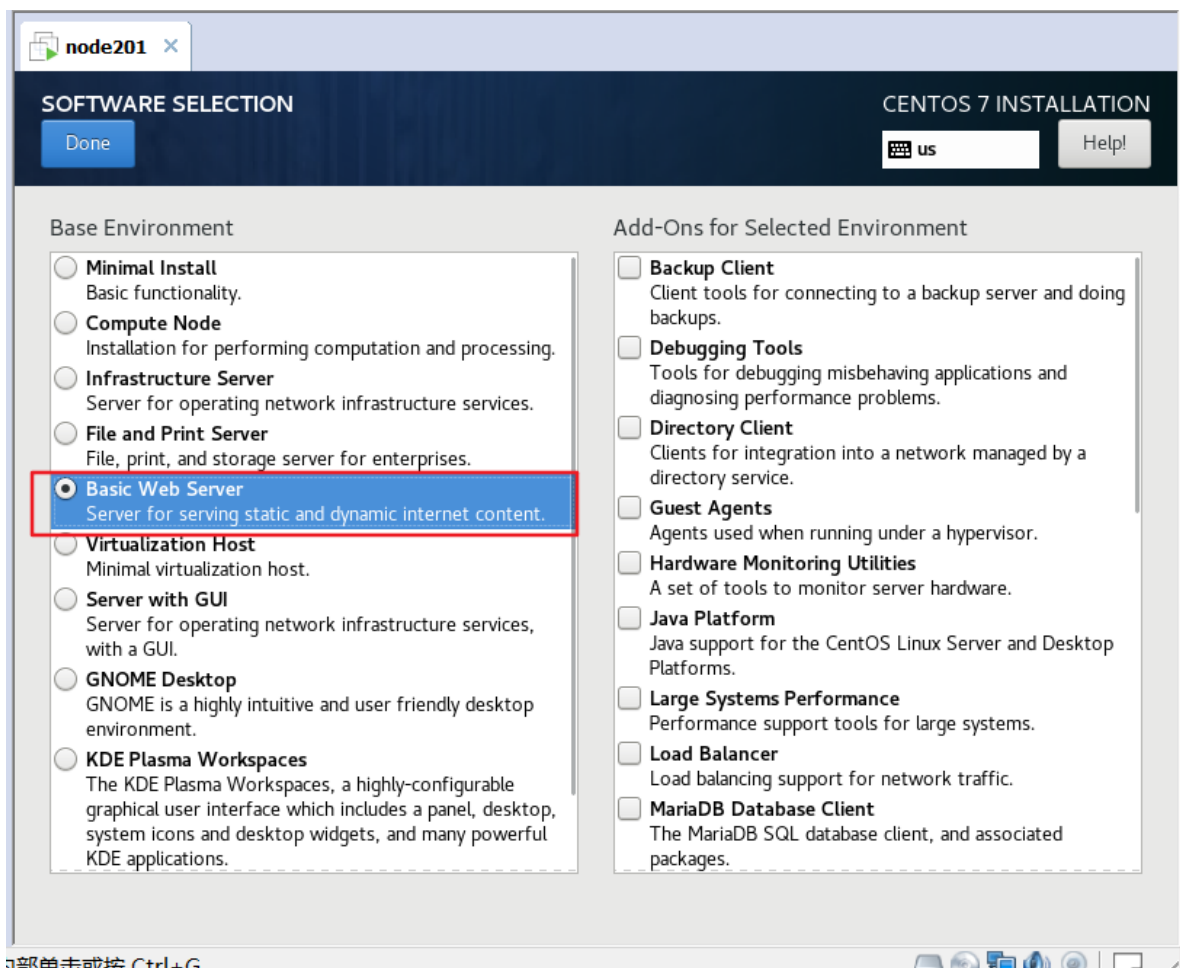
5.3 安装CentOS7操作系统



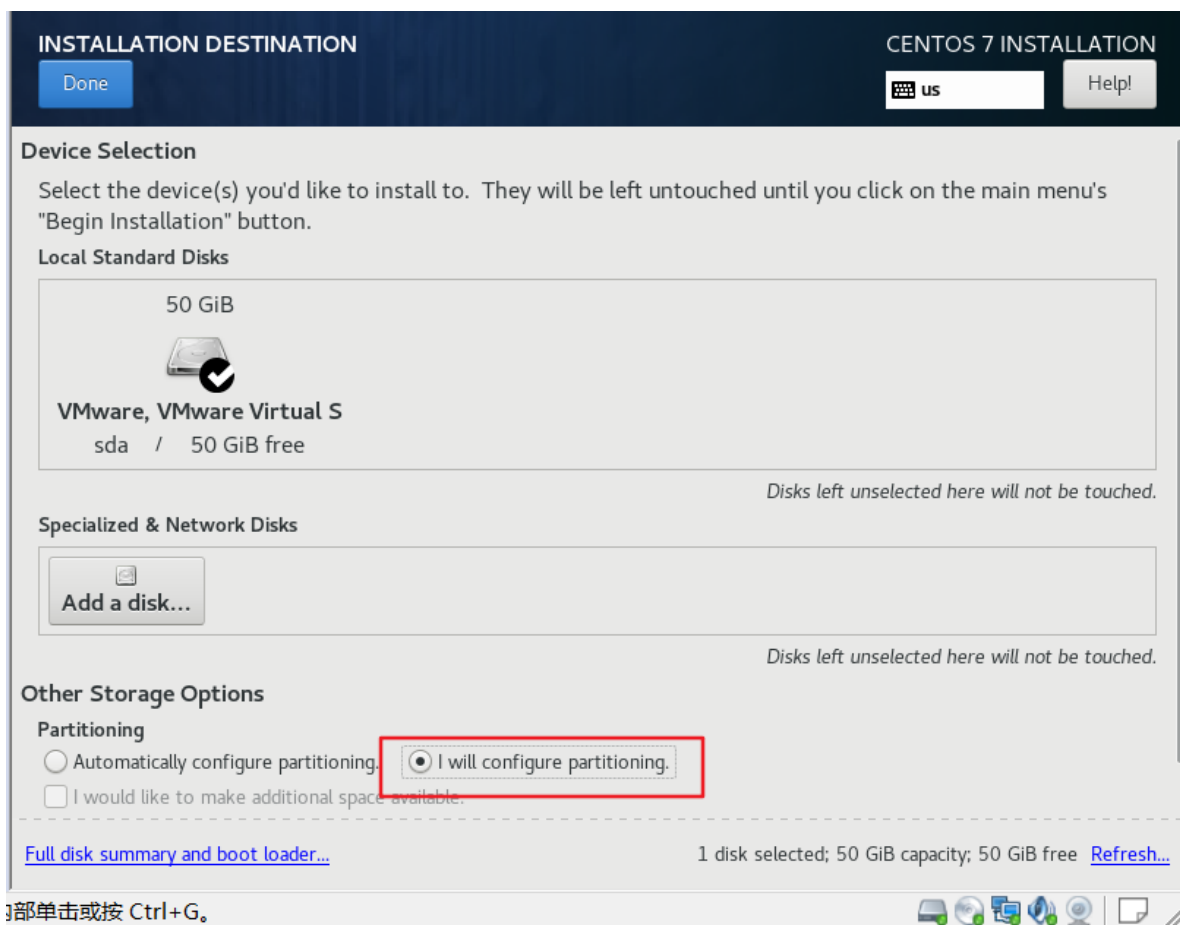




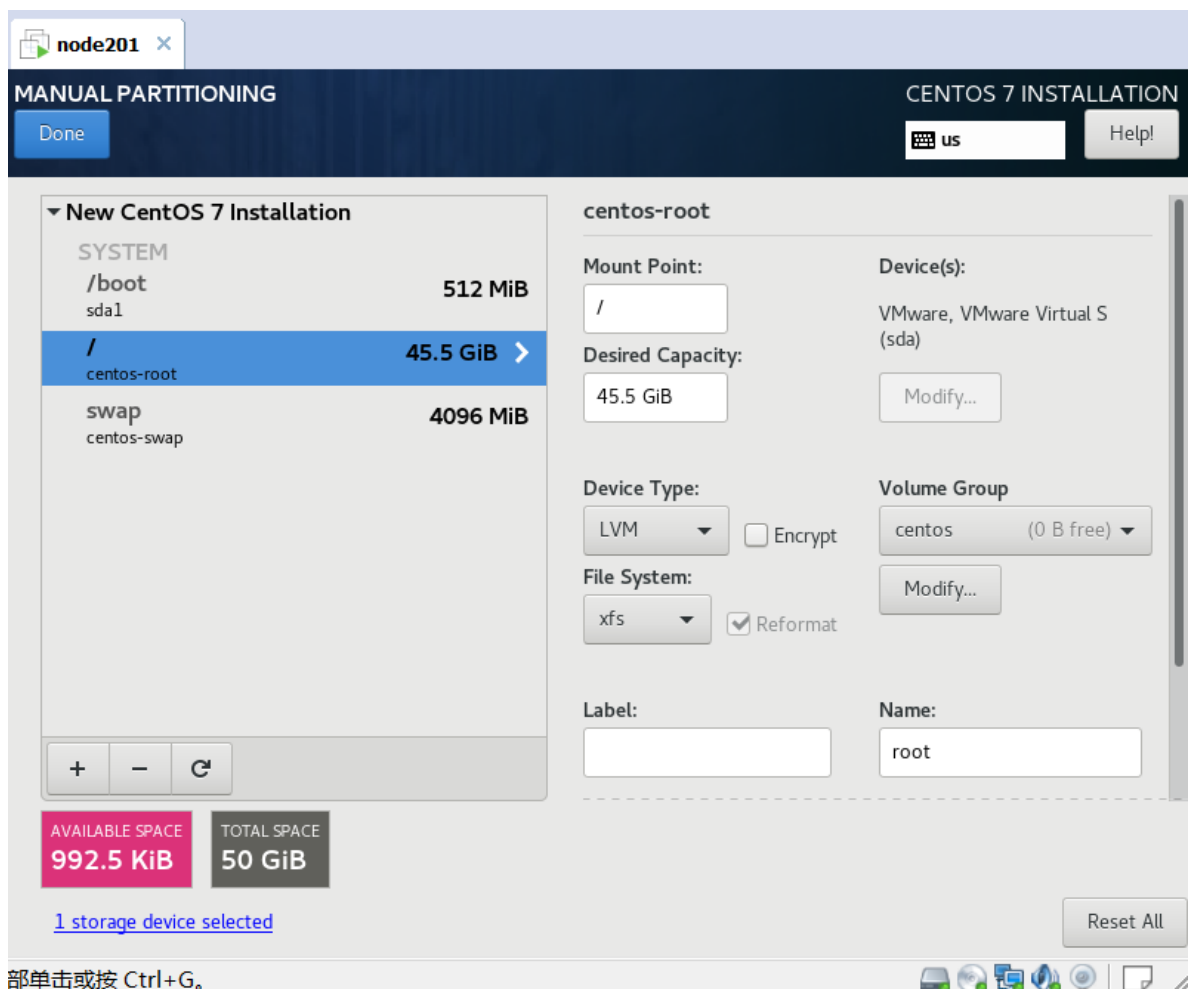
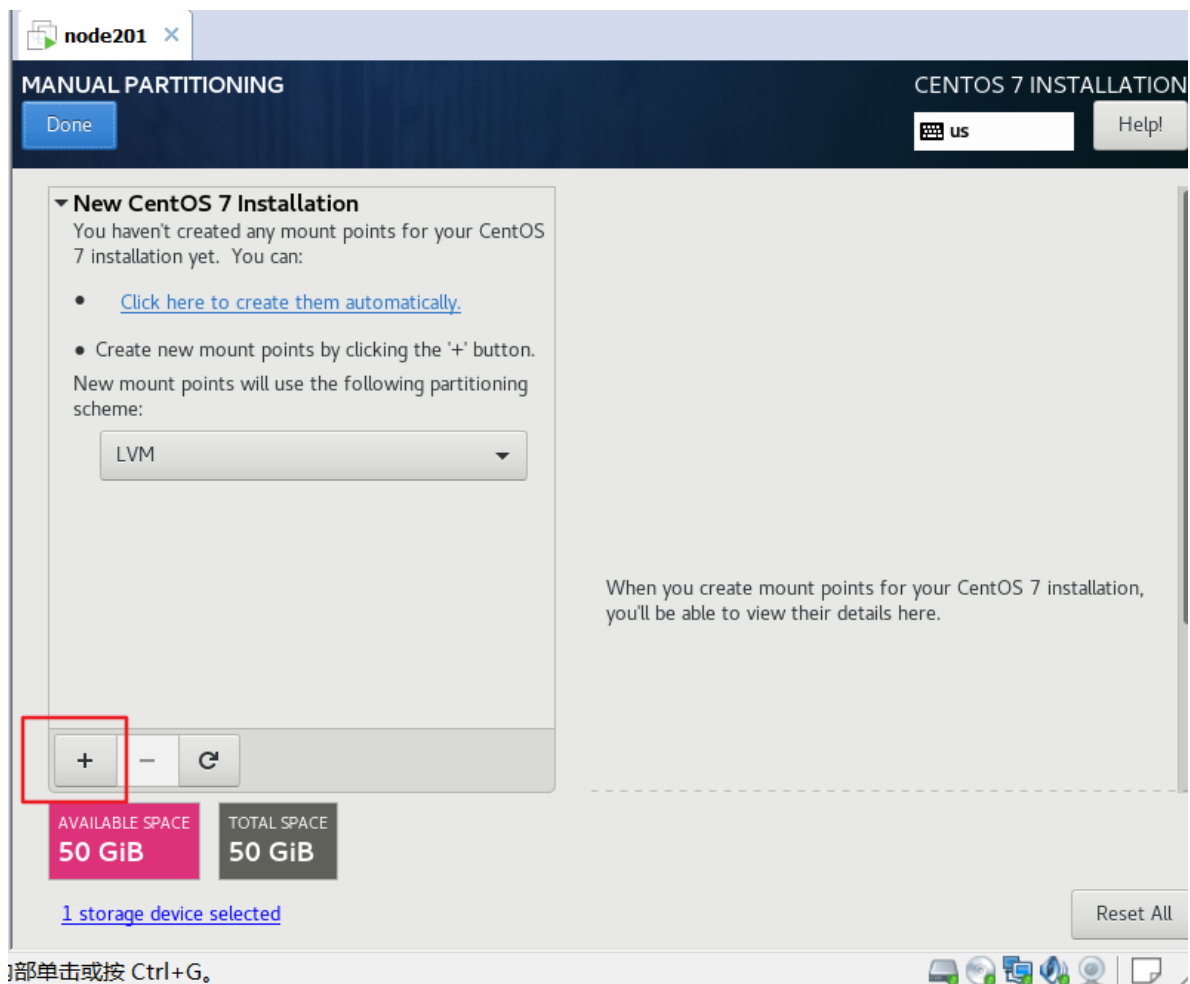


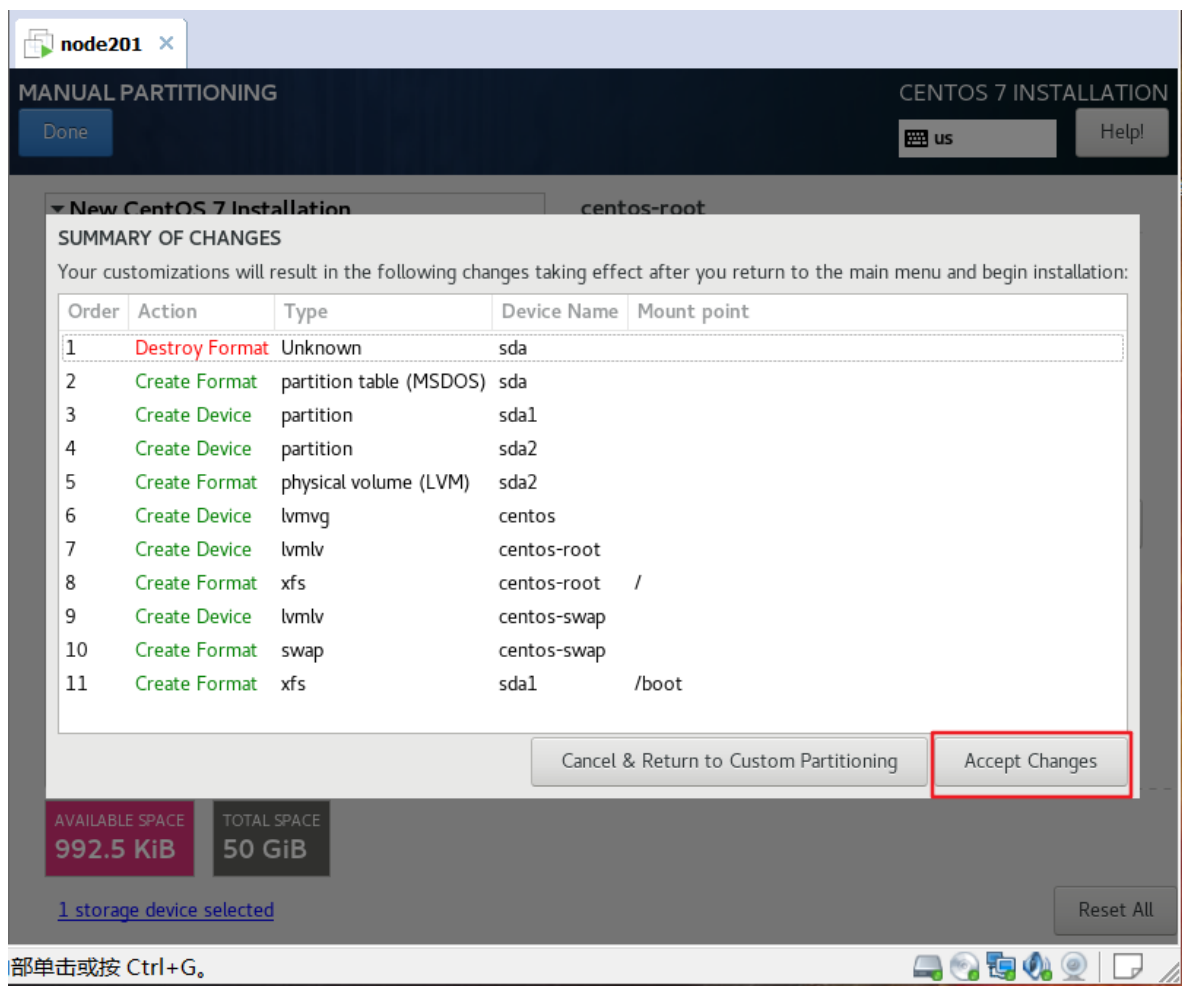


全部单击或按 Ctrl+G

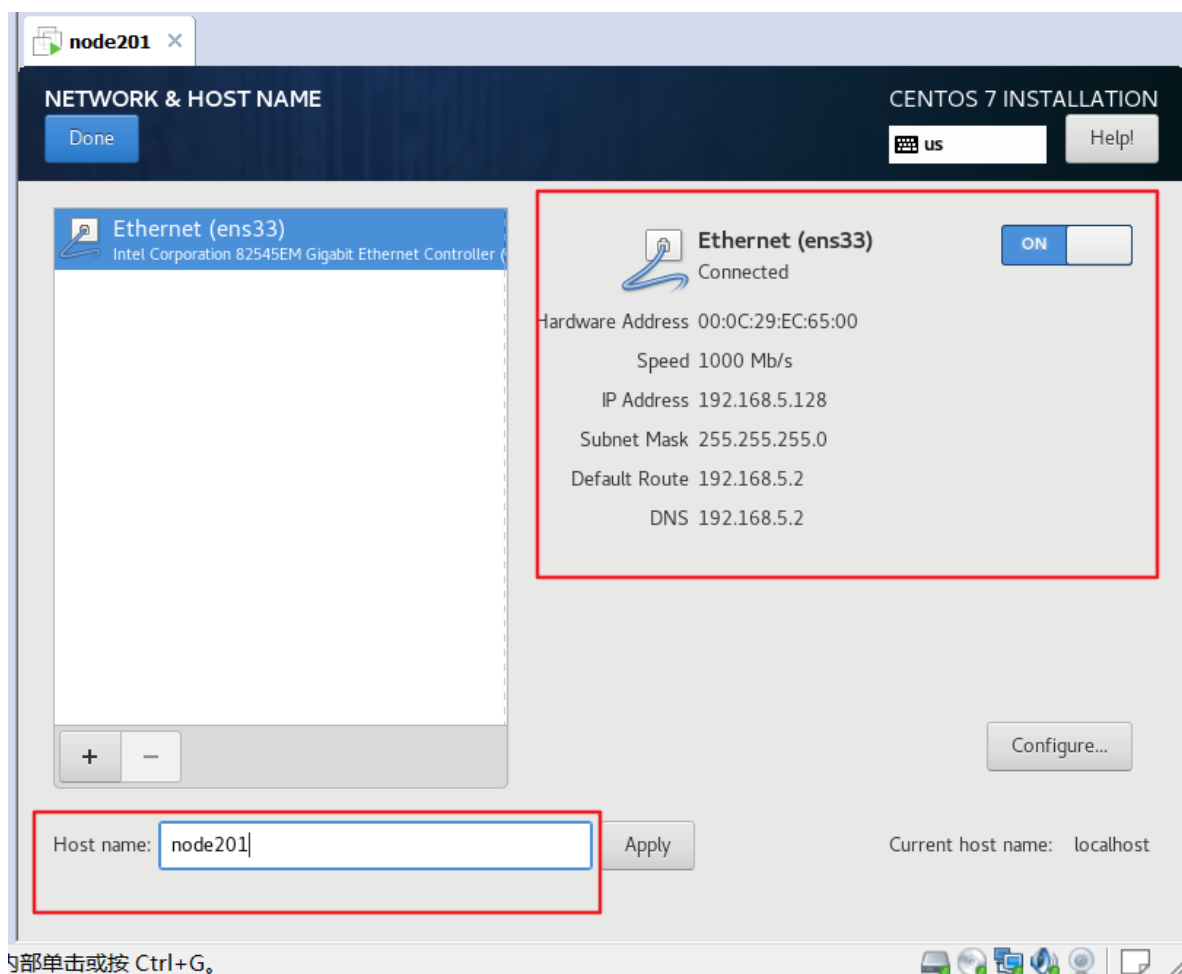


全部单击或按 Ctrl+G。

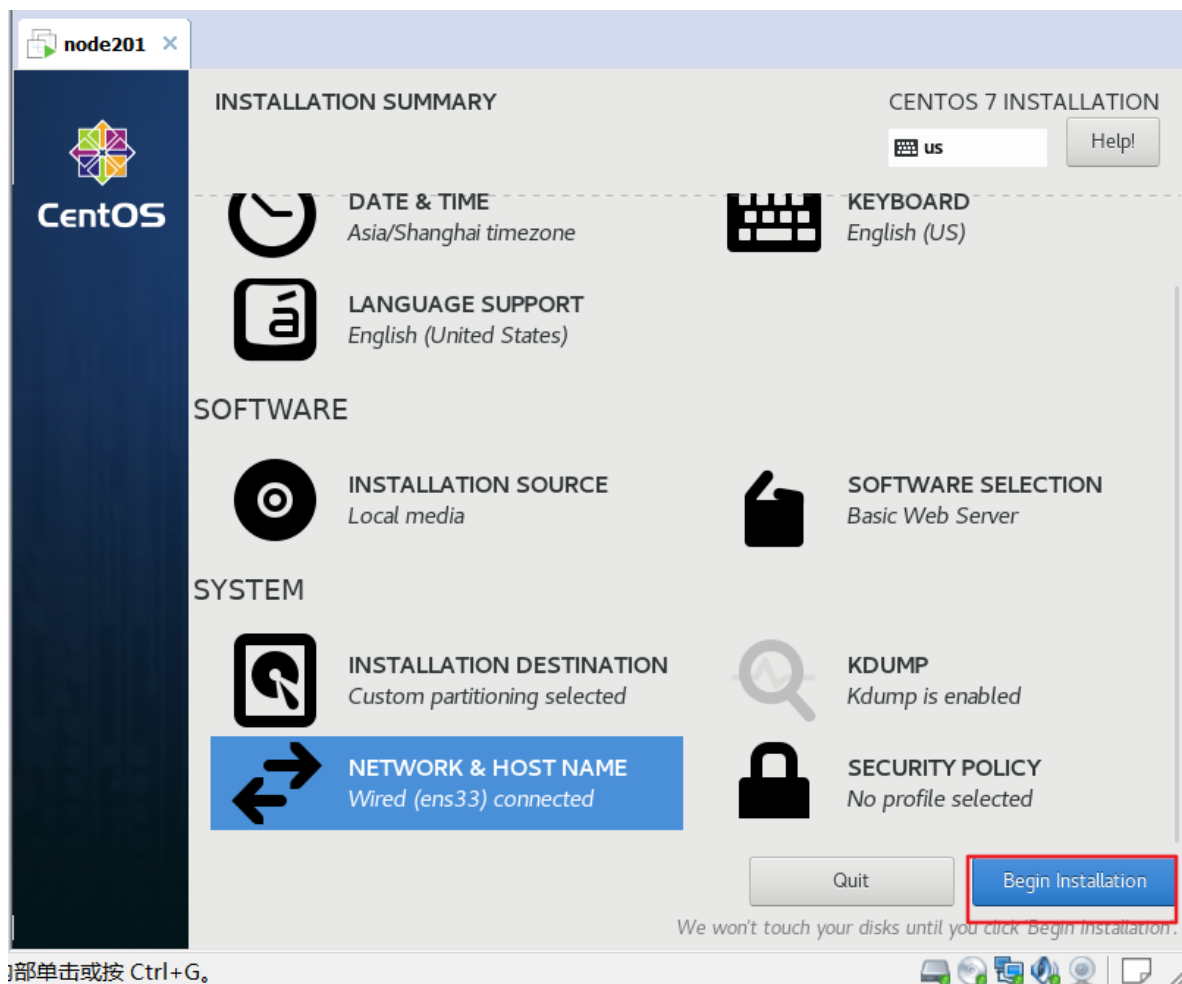




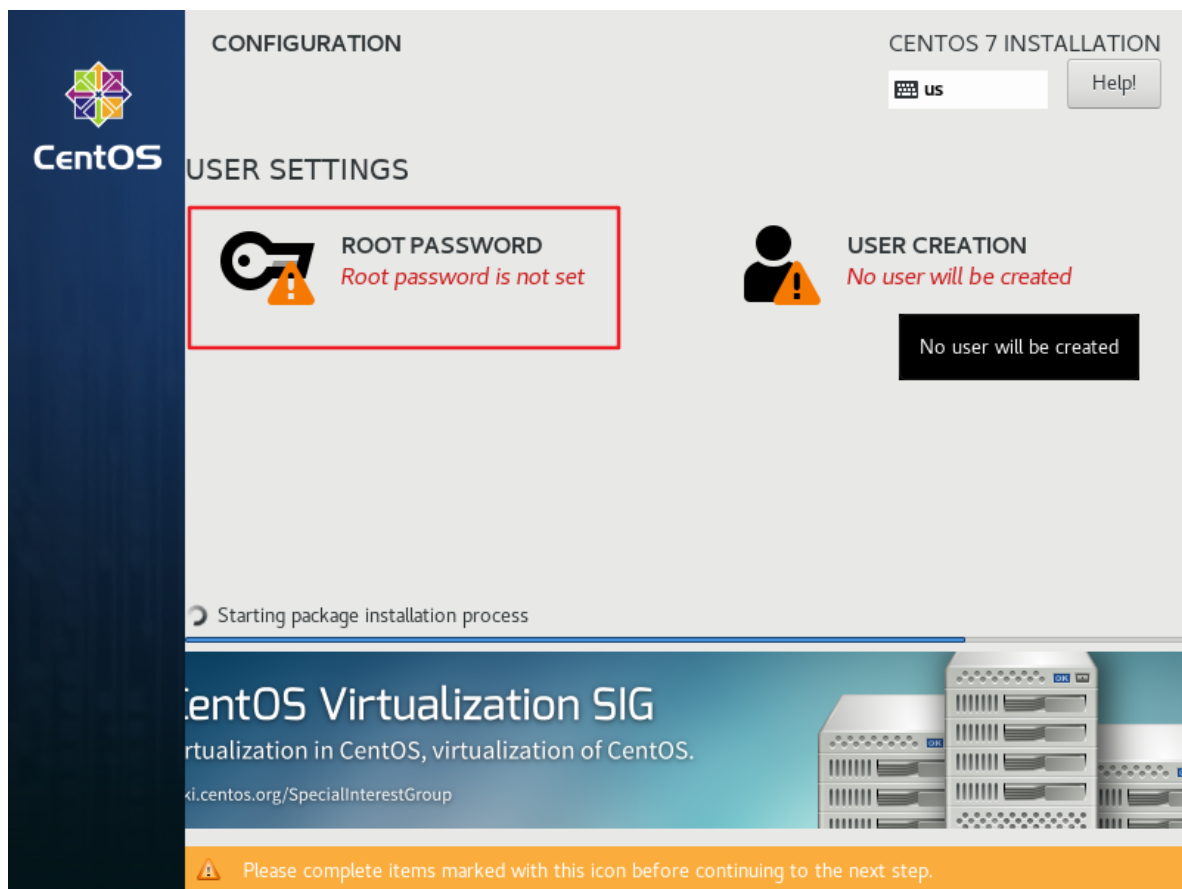
部单击或按 Ctrl+G。

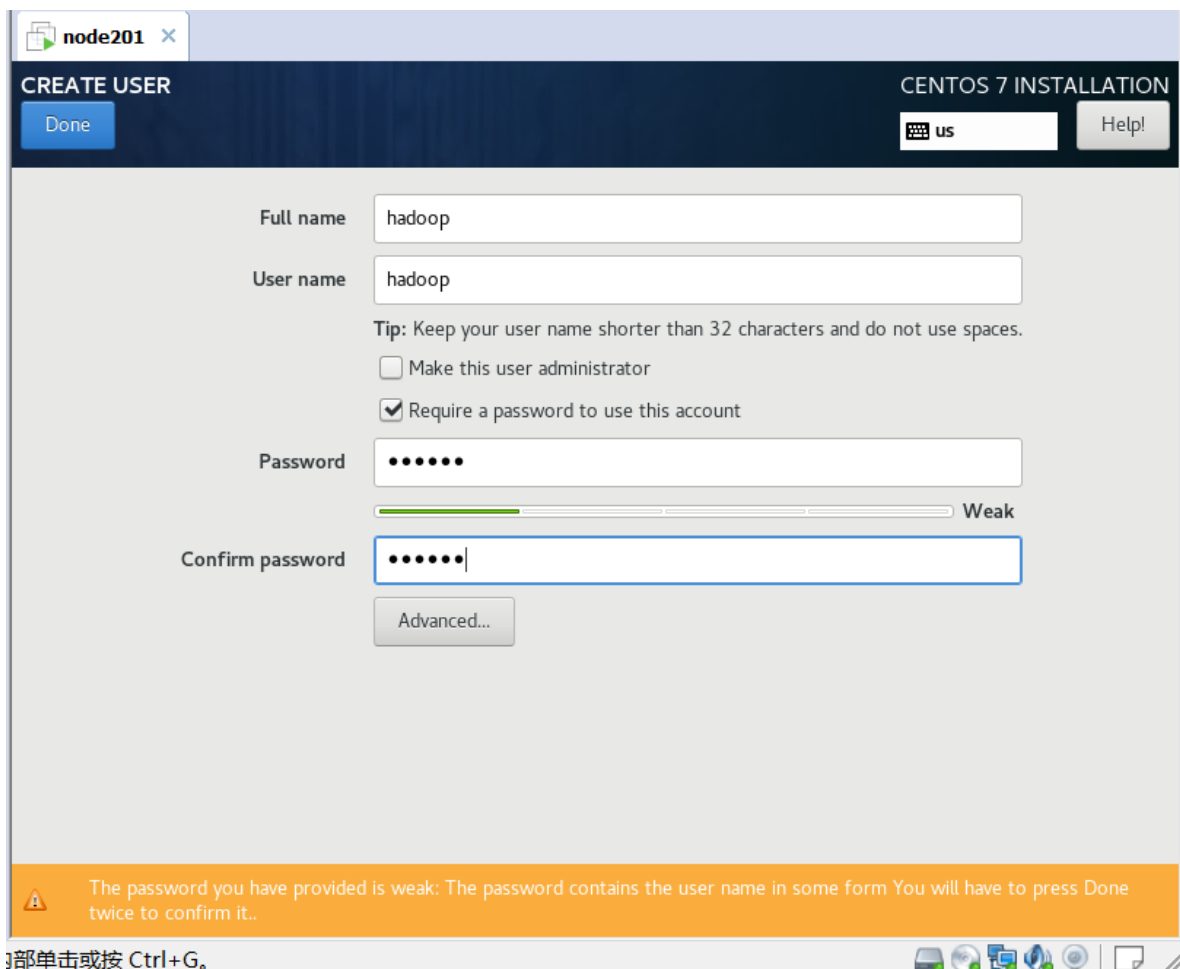
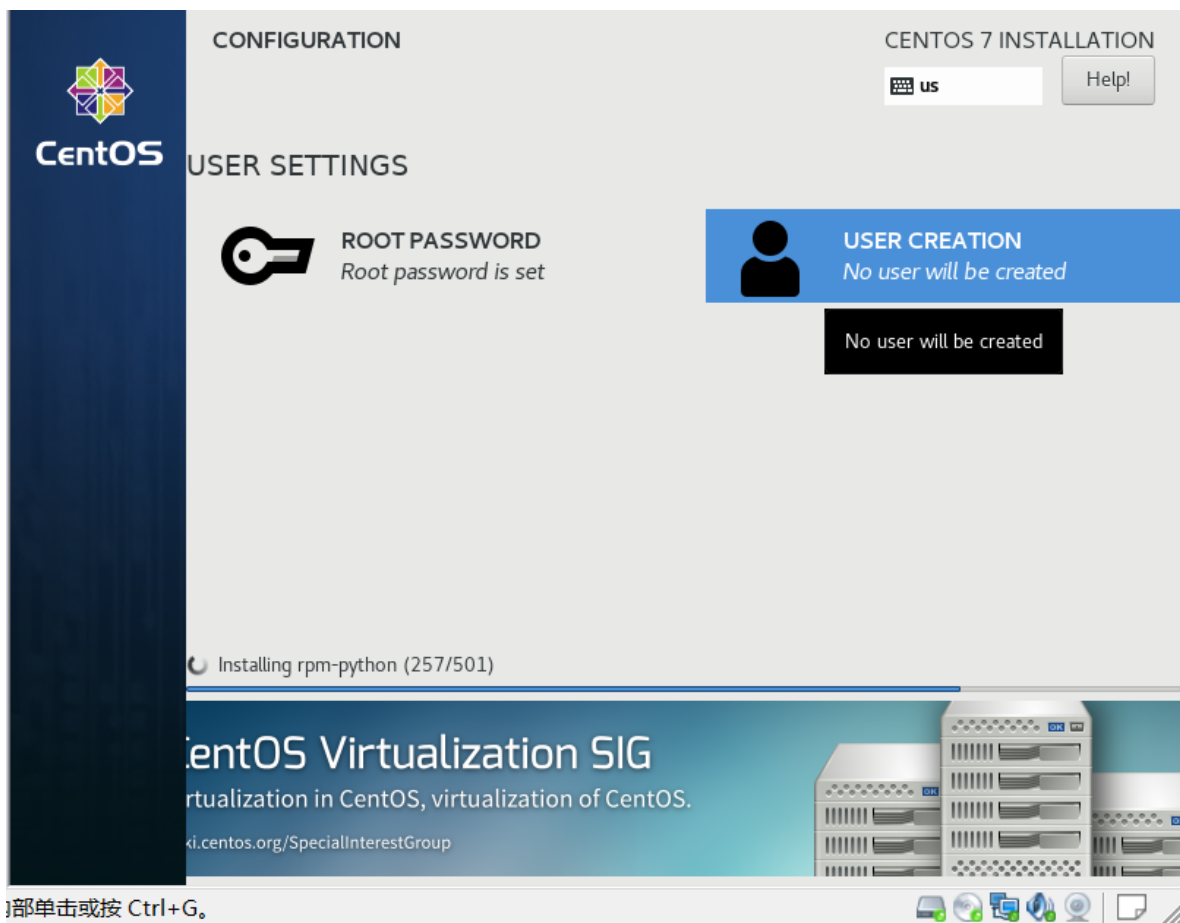


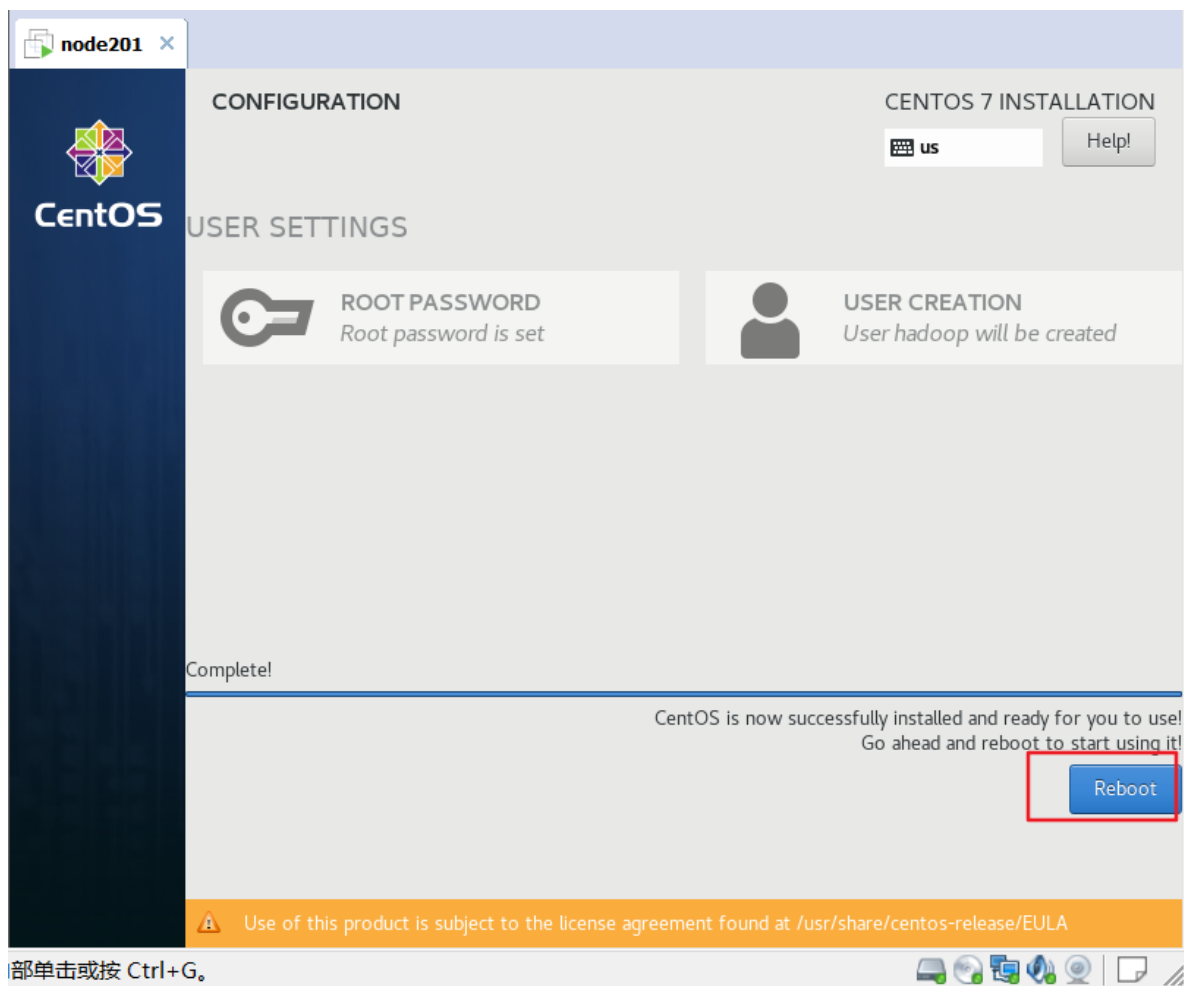
部单击或按 Ctrl+G。



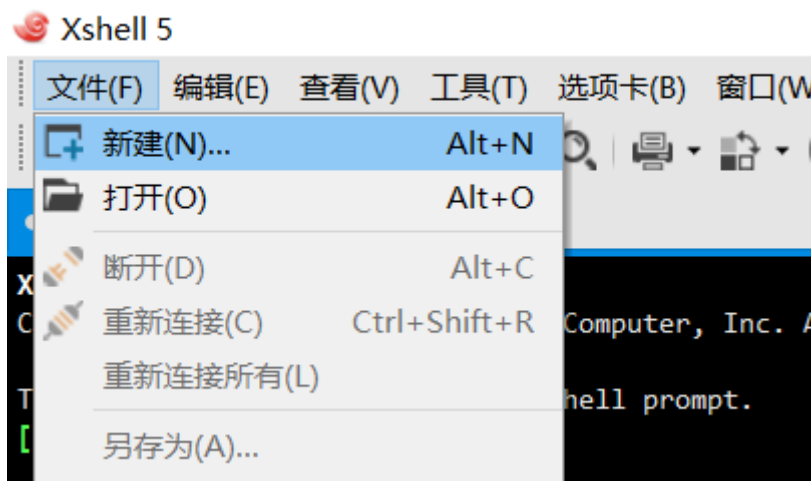
部单击或按 Ctrl+G。

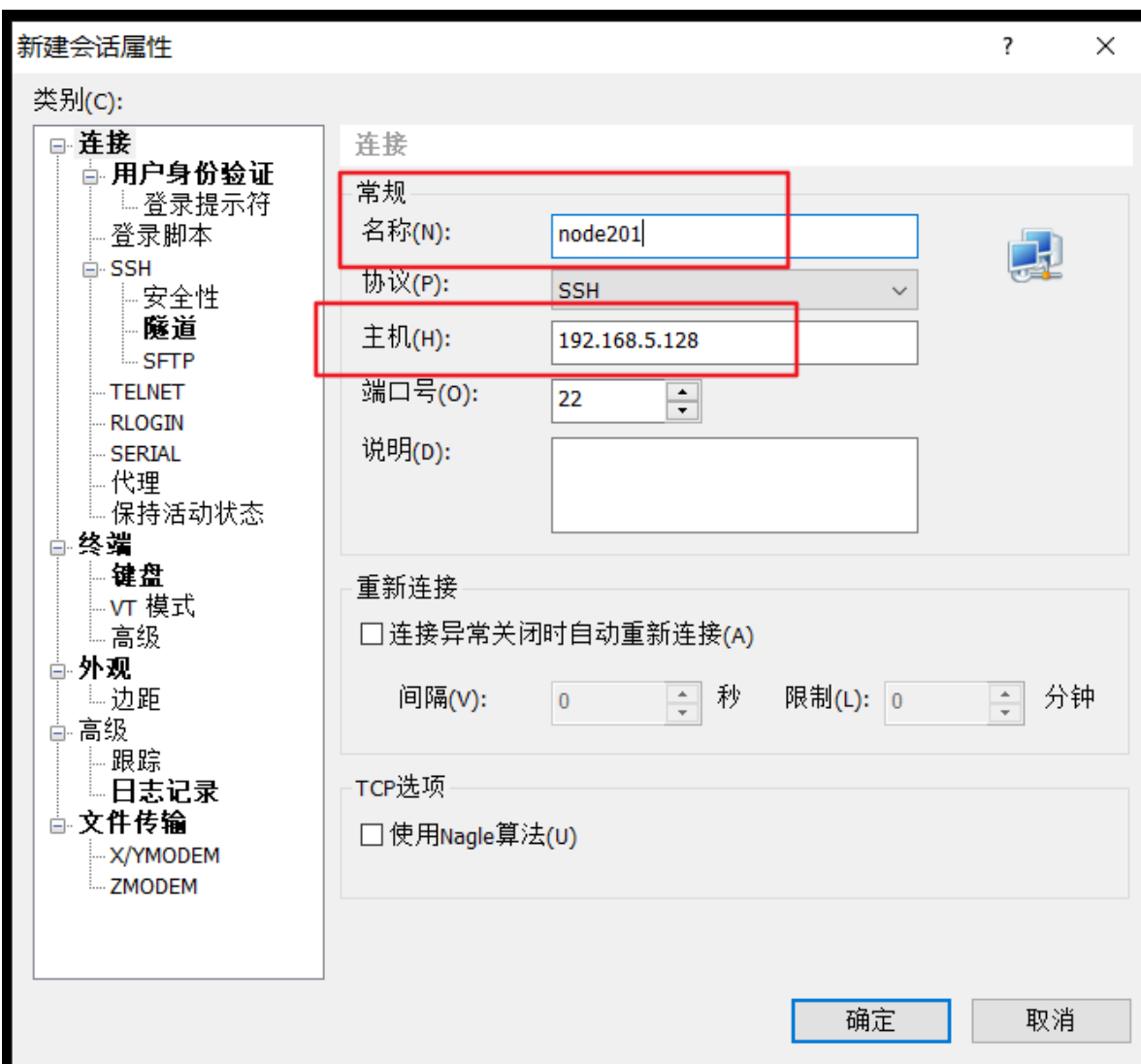






5.4 使用xshell连接虚拟机





类别(C):

- [-] 连接
 - [-] 用户身份验证
 - 登录提示符
 - 登录脚本
 - [-] SSH
 - 安全性
 - 隧道
 - SFTP
 - TELNET
 - RLOGIN
 - SERIAL
 - 代理
 - 保持活动状态
- [-] 终端
 - 键盘
 - VT 模式
 - 高级
- [-] 外观
 - 边距
- [-] 高级
 - 跟踪
 - 日志记录
- [-] 文件传输
 - X/YMODEM
 - ZMODEM

连接 > 用户身份验证

请选择身份验证方法和其它参数。

会话属性中此部分是为了登录过程更便捷而提供的。如果需要安全性很高的状态的话建议您空出此字段。

方法(M):

Password

设置(S)...

用户名(U):

root

密码(P):

••••

用户密钥(K):

<无>

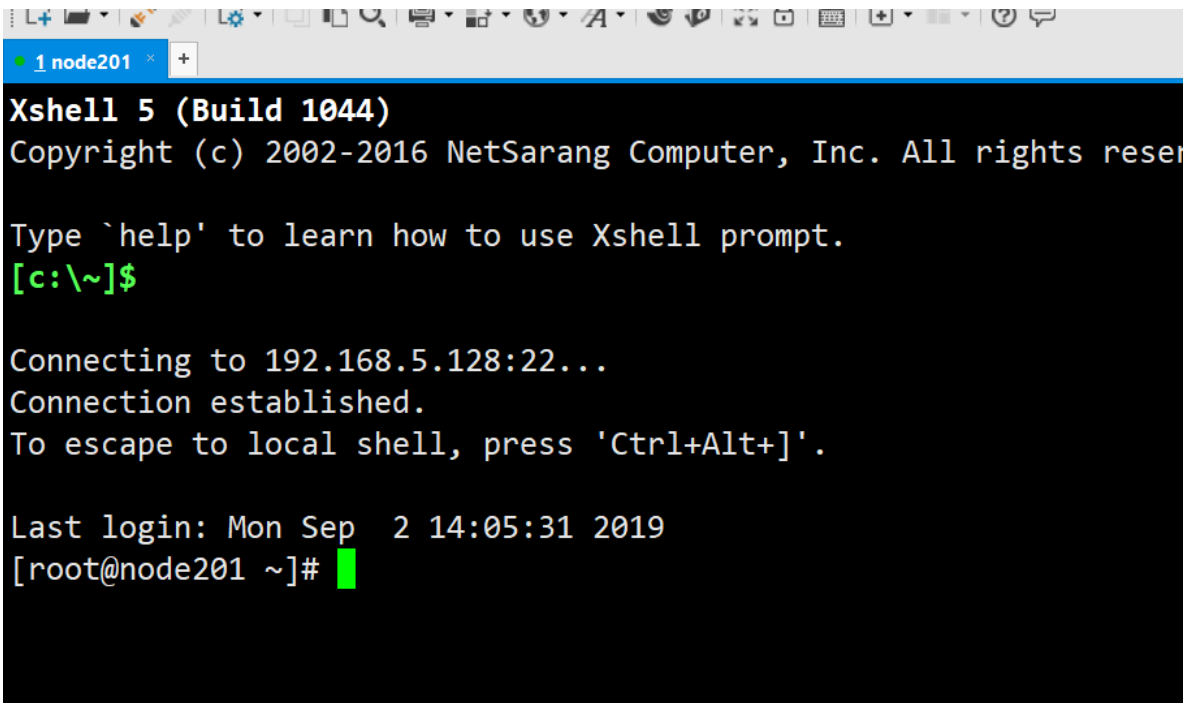
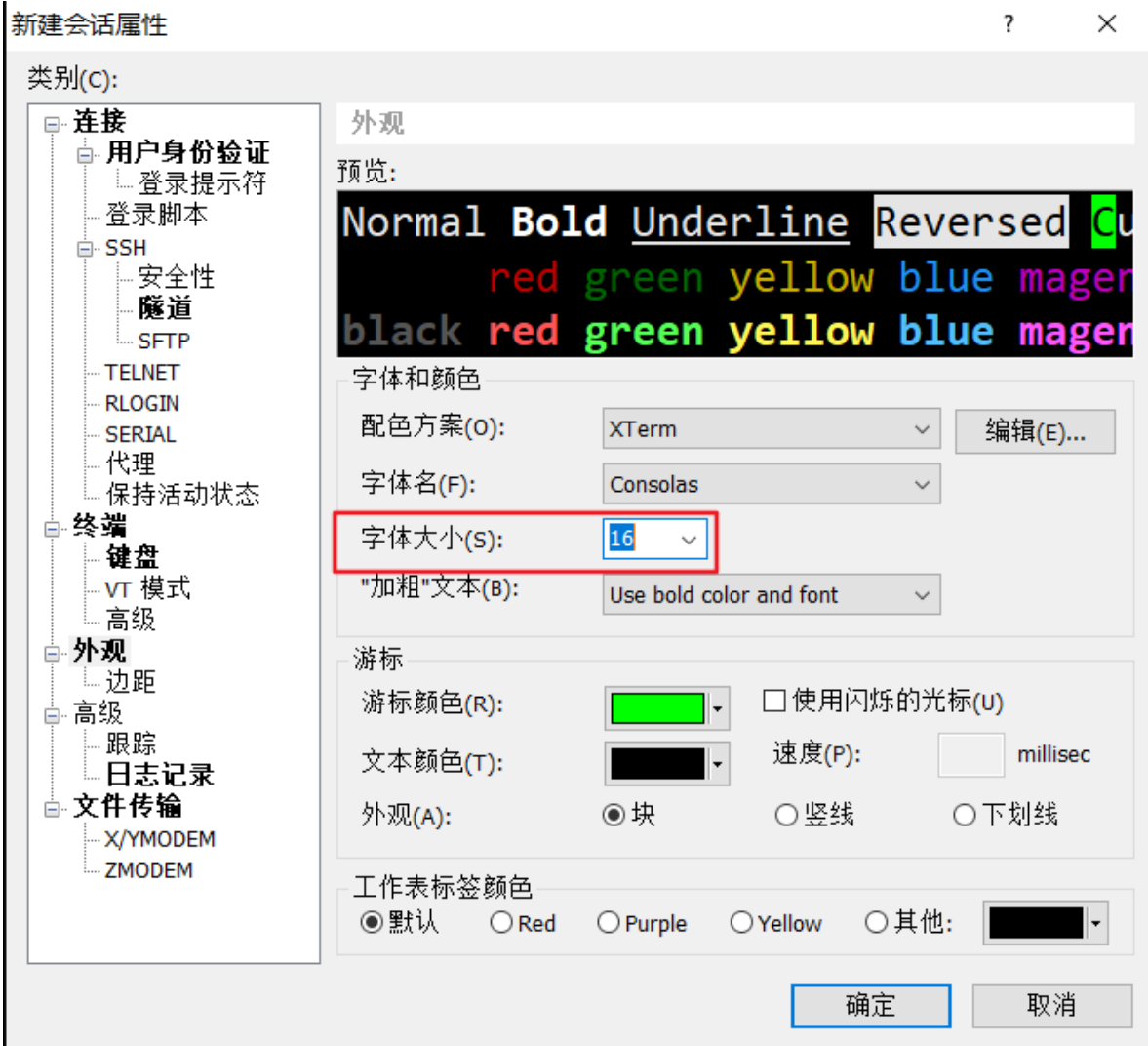
浏览(B)...

密码(A):

注释: 公钥和Keyboard Interactive仅在SSH/SFTP协议中可用。

确定

取消



六、CentOS7系统的基本常识

1. CentOS7是一个Linux操作系统，能够同时登陆多个用户。这些用户的文件互不相通。
2. Linux系统中，我们通过命令来和主机进行交互。
3. root用户是Linux系统中的超级管理员。我们一般使用普通的用户进行操作。
4. 普通用户的操作受到很多限制，各个用户之间有一层权限保证操作的安全性。

5. 学习好linux系统，关键是掌握命令的使用。命令的基本格式：命令 -选项 参数

七、常用的Linux命令

cd：切换目录

1. cd /etc/ 表示进入/etc这个目录中
2. cd ../ 表示返回上一级目录
3. cd 进入当前用户的家目录

su - 用户名：切换用户

ls：列出目录下的文件

ll：列出目录下的文件

ifconfig：查看网卡的配置信息

mkdir：新建一个目录/文件夹

touch：新建一个文件

vim：文件编辑器

1. 一般模式
2. 插入模式
3. 命令模式

rm -rf 文件：删除文件

八、虚拟机基本配置

九、环境测试

十、总结与作业

1. 根据老师的视频来部署环境
2. 熟悉常见的linux命令
3. 记录自己在安装过程中的问题