

# 大数据处理应用实践教材

(内部使用)

小象学院

All rights reserved



## 目录

第一	-章 虚拟机和系统安装	9
S	等一节 虚拟机 VMware 安装	9
1.	下载 VMWARE	9
2.	打开 VMWARE 安装文件	9
3.	在启动的安装界面选择下一步	9
4.	一直选择下一步,直到提示安装完成	9
5.	安装完成在桌面双击 VMWARE 快捷方式图标显示激活页面	.10
S	5二节 新建虚拟机	.10
1.	新建虚拟机	.10
2.	选择自定义安装,点击下一步	.10
3.	虚拟机硬件兼容性选择窗口,不需要操作,继续点击下一步	, 11
4.	给虚拟机安装操作系统	. 11
5.	选择稍后手动安装的操作系统类型和版本	.12
6.	设置虚拟机名称和虚拟机文件存储位置	.12
7.	处理器配置	.13
8.	虚拟机内存上设置	.13
9.	设置虚拟机网络类型	.14
10.	剩下的配置按照推荐的配置一直点下一步完成	.14
11.	虚拟机可使用的最大磁盘容量	.16
12.	设置虚拟机磁盘文件路径和名称	.17
13.	创建完成	.17
14.	创建虚拟机完成会显示如下界面	.18



	第三节 在虚拟机中安装 centos 操作系统	20
1.	. 下载 ISO 系统映像文件	20
2.	. 添加系统镜像文件路径	20
3.	. 安装系统镜像	21
4.	. 选择安装方式	21
5.	. 安装完成界面,点击 NEXT 按钮	22
6.	. 系统配置	23
	6.1 选择系统语言,简体中文	23
	6.2 选择键盘布局,选择"美式键盘"	23
	6.3 选择存储设备,"基本存储设备"	24
	6.4 是否清空新建的虚拟磁盘中的数据,选择 Y"是,忽略所有数据"	24
	6.5 设置主机名	24
	6.6 设置时区,"亚洲/上海"	25
	6.7 设置 root 用户密码,root123	25
	6.8 磁盘分区划分	25
	6.9 引导程序配置,默认即可,点击下一步	27
	6.10 选择桌面系统类型	27
	6.11 经过一段时间等待,安装成功,重启	28
7.	. 重启之后,用户信息配置	28
	7.1 欢迎页面右下角"前进"按钮	28
	7.2 创建普通用户 Hadoop	29
	7.3 设置日期和时间	30
	7.4 在机器崩溃的时候是否把内存中的数据 dump 下来,选择不启用。点击右下角完成按	安钮,
	重启	30
	7.5 重启完成,到此虚拟机和操作系统安装完成	31
鲁	ち一音 CFNTOS 网络设置	31



	第一节 网络设置	31
1.	查看虚拟网关 IP 地址	31
2.	设置虚拟机网络适配器	33
3.	设置虚拟机 IP、子网掩码、网关、DNS	33
4.	设置虚拟机 IP 地址与主机名 HOSTNAME 映射	34
5.	设置安装虚拟机电脑的虚拟网卡 VMNET8 的 IP 地址、网关、子网掩码	35
	第二节 使用 CRT 远程连接虚拟机	37
1.	下载 CRT 软件	37
2.	配置 CRT 远程设置	37
3.	保存密码	39
4.	输入登录的用户名	39
	第三节 虚拟机 iptables 防火墙设置	40
1.	查看 IPTABLES 状态	40
2.	开启/关闭 IPTABLES	40
3.	查看 IPTABLES 自动启动状态	40
4.	设置 IPTABLES 开机启动/不启动	40
	第四节 虚拟机上外网注意事项	41
1.	安装虚拟机的电脑服务中所有关于 VMWARE 的服务开启	41
2.	虚拟机的 IP、网关、DNS、子网掩码配置正确	42
3.	选中一台虚拟机->鼠标右键设置->"网络适配器"要选择自定义"VMNET8(NA	AT 模式)"42
第	5三章 克隆虚拟机	42
1.	NODE 节点鼠标右键管理->克隆	42
2.	从当前状态开始克隆,会将当前源虑拟机中的已安装和配置的环境克隆下来	43



3.	选择"创建完整克隆"44
4.	创建新的虚拟机名称,克隆虚拟机文件存储位置44
5.	点击"完成"之后,等待进度条完成44
6.	创建完成45
7.	启动新克隆的虚拟机 NODE245
8.	克隆 NODE03
第四	]章  SSH 免密码登录46
1.	使用 ROOT 用户登录到 NODE01 机器,首先进入 ROOT 用户的家目录下46
2.	进入到.SSH 目录下(注意.SSH 是隐藏文件)47
3.	在.SSH 目录下执行如下命令生成私钥(ID_RSA)和公钥文件(ID_RSA.PUB)47
4.	执行完成会生成私钥(ID_RSA)和公钥文件(ID_RSA.PUB)47
5.	通过 SSH-COPY-ID 命令将 NODE01 这台机器 ROOT 用户的公钥文件(ID_RSA.PUB)
文件	- 内容拷贝到 NODE02 和 NODE03 两台机器47
6.	在 NODE02 和 NODE03 的 ROOT 用户的家目录下,进入到.SSH 目录48
7.	在 NODE01 通过 SSH 登录到 NODE02 和 NODE03, 此时不在需要输入密码48
第五	C章 常用命令和系统命令48
穿	5一节 常用命令和快捷键48
1.	停止进程
2.	清屏48
3.	搜索历史命令48
4.	文件名自动补全48
5.	进入目录49
6.	进入用户家目录49



7.	进入最近上一次访问的目录	49
8.	返回上一级目录	49
9.	查看目录文件列表	49
10.	查看文件全部内容	49
11.	查看文件末尾 NUM 行	49
12.	查看文件头部 NUM 行	49
13.	查看文件行数	50
14.	查看文件中字符个数	50
15.	复制文件	50
16.	重命名文件	50
17.	查看历史命令	50
18.	软链接	50
19.	查找文件	50
匀	<b>第二节 系统管理命令</b>	52
1.	查看本地时间	52
2.	联网更新系统时间	52
3.	查看主机名	52
4.	修改主机名(重启后永久生效)	52
5.	修改 IP(重启后永久生效)	52
6.	修改/设置 IP 和主机名映射	53
7.	统计文件或文件夹的大小	53
8.	查看磁盘的空间	53
9.	关机	53



10.	重启	53
11.	查看命令帮助	53
12.	ECHO 使用	53
13.	SERVICE 后台服务管理	53
14.	查看自动启动服务	54
第六	六章 用户管理	55
1.	添加新组	55
2.	添加新用户	55
3.	设置用户密码	55
4.	给用户添加组	55
5.	添加新用户同时添加组	55
6.	查看用户组	56
7.	删除用户	56
8.	删除组	56
9.	将用户从组中删除	56
10.	切换用户	56
11.	给普通用户赋予 ROOT 系统权限	56
第七	七章  文件管理	58
1.	创建文件	58
2.	删除文件	58
3.	创建文件夹	58
4.	删除文件夹及所有子文件夹	58
5.	文件或文件夹权限	58



6.	修改文件权限5	9
7.	查看文件内容6	0
8.	VIM 文本编辑器	0
第八	章 压缩解压命令6	2
1.	ZIP 压缩/解压缩,压缩文件后缀名.ZIP6	2
2.	GZIP 压缩/解压缩,压缩文件后缀名.GZ6	2
3.	BZIP2 压缩/解压缩,压缩文件后缀名.BZ2	2
4.	打包/解包,打包文件后缀名.TAR6	2
5.	一次性完成打包和 GZIP 压缩的操作,文件后缀名.TAR.GZ6	2
6.	一次性完成打包和 BZIP2 压缩的操作,文件后缀名.TAR.BZ2	3
第九	.章  LINUX 软件安装6	4
1.	使用 YUM 安装文件传输工具 LRZSZ6	4
2.	手动安装 JDK6	4
3.	TELNET 安装6	5
第十	-章 CENTOS 安装 MYSQL6	6
1.	安裝前检查6	6
2.	开始安装6	7
3.	启动 MYSQL 服务6	7
4.	设置开机启动6	7
5.	查看开机启动设置是否成功6	7
6.	登录 MYSQL6	7
7.	建立远程 ROOT 用户6	8
8.	删除 USER 表中的干扰项	8





# 第一章 虚拟机和系统安装

## 第一节 虚拟机VMware安装

#### 1. 下载 VMware

链接: https://pan.baidu.com/s/1Or5iaasRyGAkw6R3iIgKWQ 提取码: rh2u

## 2. 打开 VMware 安装文件

#### 3. 在启动的安装界面选择下一步

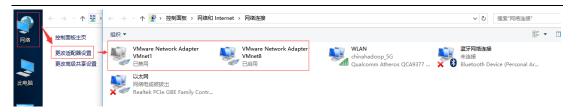
如下图所示:



#### 4. 一直选择下一步,直到提示安装完成

安装完成在桌面选中网络图标,点击鼠标右键->属性->找到更改适配器设置选项点击->查看在网络连接列表中包含 VMware Network Adapter VMnet1 和 VMware Network Adapter VMnet8 则表示 VMware 安装成功。





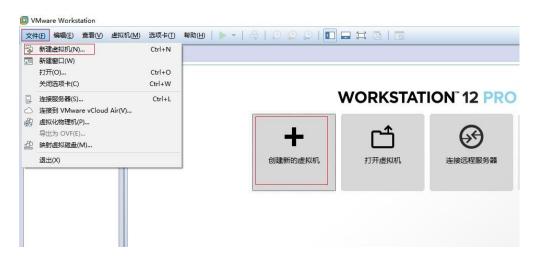
# 5. 安装完成在桌面双击 VMware 快捷方式图标显示激活 页面

#### 注册码/key: 5A02H-AU243-TZJ49-GTC7K-3C61N

如果注册码无效,可以在百度搜索关键词"vmware workstation 12 密钥"将会找到答案。

## 第二节 新建虚拟机

#### 1. 新建虚拟机

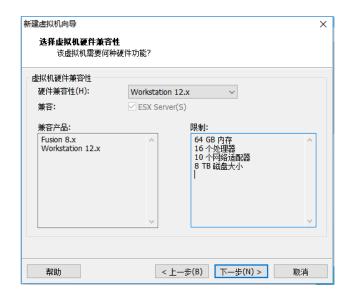


#### 2. 选择自定义安装,点击下一步





# 3. 虚拟机硬件兼容性选择窗口,不需要操作,继续点击下一步



## 4. 给虚拟机安装操作系统

有三个选项,分别是 1)使用光驱从系统光盘安装; 2)从系统光盘映像文件自动安装(iso文件),缺点是无法进行个性化设置; 3)先创建虚拟机,跳过系统



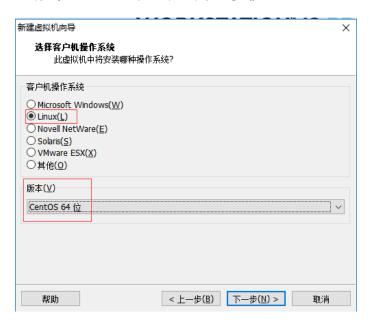
安装,稍后手动安装。

在这里选择第三个选项, 先不安装操作系统, 稍后手动安装, 点击下一步按钮。



#### 5. 选择稍后手动安装的操作系统类型和版本

这里选择 Linux, 版本 Centos64 位,点击下一步按钮。

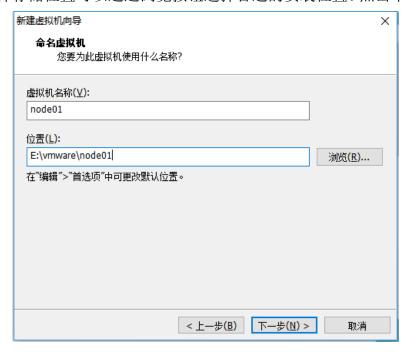


#### 6. 设置虚拟机名称和虚拟机文件存储位置

虚拟机名称可以自定义,建议最好是自己比较简单的名字,容易记忆。

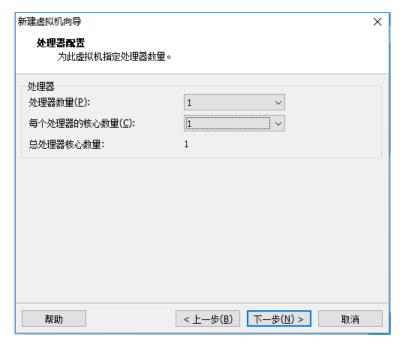


虚拟机文件存储位置可以通过浏览按钮选择合适的安装位置。点击下一步按钮。



## 7. 处理器配置

处理器配置都是虚拟出来的,按照默认配置就可以。如果机器性能强也可以根据需求自定义选择。



## 8. 虚拟机内存上设置



滑动下图中左侧的滑块和右侧的上下调节按键都可以设置内存大小。



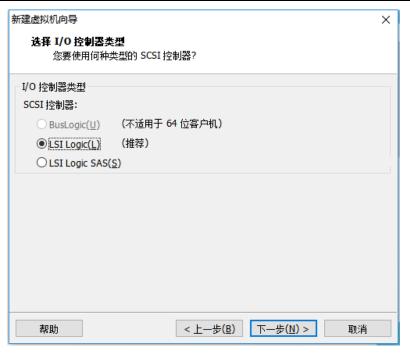
#### 9. 设置虚拟机网络类型

选择 NAT 类型,虚拟机可以连接到外网,桥接比较复杂而且具有局限性,所以选择 NAT。



## 10.剩下的配置按照推荐的配置一直点下一步完成











## 11.虚拟机可使用的最大磁盘容量

注意: 千万不要将"立即分配所有磁盘空间"勾选上,否则会在硬盘上立即划分出一个固定的空间给虚拟机使用,如果不勾选会根据需要自动调节,最大上限是设置的值。





## 12.设置虚拟机磁盘文件路径和名称

使用推荐默认选项即可,继续点下一步。



#### 13.创建完成

详细列出虚拟机的配置信息,点击完成。





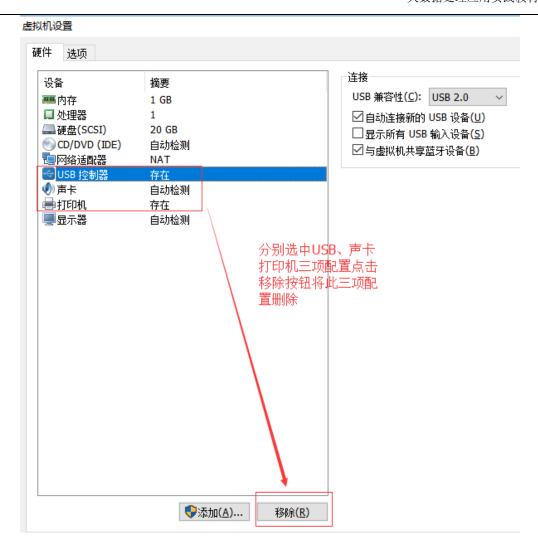
## 14.创建虚拟机完成会显示如下界面

接下来删除一些我们不需要的配置,点击编辑虚拟机设置



在虚拟机设备详情界面中选择 USB 控制器、声卡、打印机分别点击移除按钮将这三项没用的设备删除。







## 第三节 在虚拟机中安装centos操作系统

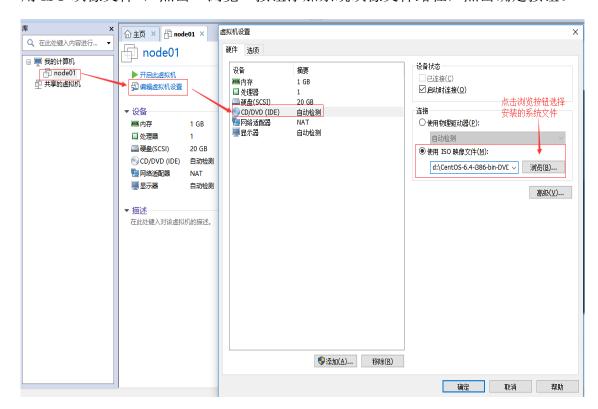
## 1. 下载 ISO 系统映像文件

链接: https://pan.baidu.com/s/1g\_VWyWUIqoYvmZuN9Ztd\_Q

提取码: o6ia

#### 2. 添加系统镜像文件路径

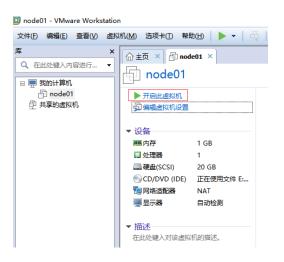
在 vmware 最左侧虚拟机机器列表中,选择需要安装系统的虚拟机,点击"编辑虚拟机设置"按钮,在弹出的虚拟机设备列表中选中 CD/DVD(IDE),选中"使用 ISO 映像文件",点击"浏览"按钮添加系统映像文件路径,点击确定按钮。





#### 3. 安装系统镜像

点击"开启此虚拟机"按钮



注意:此处可能出现的问题

使用 vmware 虚拟机安装操作系统出现错误提示"二进制转换与此平台上的长模式不兼容"错误提示的解决办法

重启电脑,在开机黑屏的时候持续按 F1 或 F2 或 Delete 进入 BOIS,(不同电脑品牌不同主板使用的进入 BOIS 的命令不一样,如果这几个键不能进入 BOIS 请根据个人电脑情况查询如何进入 BOIS);进入 BOIS -> Security -> Virtualization -> Intel(R) Virtualization Technology(有的电脑此项设置可能有所不同,请百度查找一下) 使用键盘的上下键和回车键将 Disabled 修改成 Enabled,按 F10 保存退出。

#### 4. 选择安装方式

选择第一个"Install or upgrade an existing system",按回车键选择。

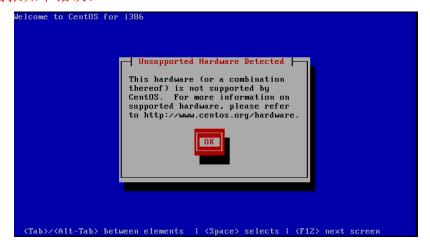




经过一段时间之后会弹出如下提示窗口,是否检测安装盘是否有问题,点 Skip 按钮跳过。

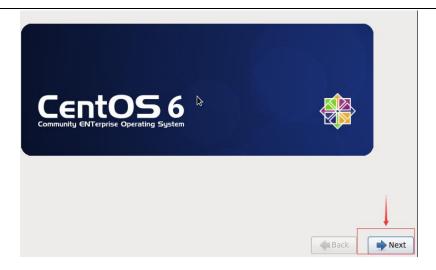


注意:安装的操作系统如果与之前设置的版本不同(64 位系统安装了 32 位的系统)会报如下错误。



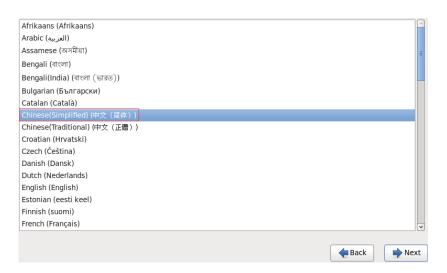
## 5. 安装完成界面,点击 next 按钮





## 6. 系统配置

## 6.1 选择系统语言,简体中文



6.2 选择键盘布局,选择"美式键盘"





#### 6.3 选择存储设备,"基本存储设备"



# 6.4 是否清空新建的虚拟磁盘中的数据,选择 Y "是,忽略所有数据"



## 6.5 设置主机名

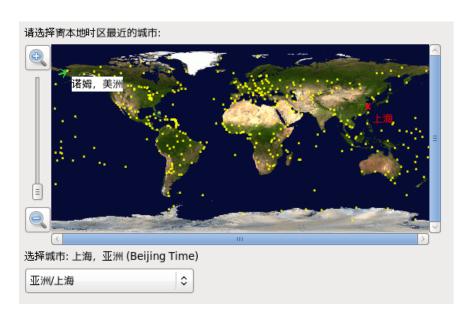


	100		_	
	Ш			
	L	3		
90	(	30		
1111		-	-	_

请为这台计算机命名。该主机名会在网络中定义这台计算机。

主机名: node01

## 6.6 设置时区,"亚洲/上海"



#### 6.7 设置 root 用户密码, root123

根帐号被用来管理系统。请为根用户输入一个密码。	
根密码 (P) : •••••	
确认 (C) : ••••••	

A	<b>危霧密码</b>
您的密码不够多	安全:它基于字典单词

## 6.8 磁盘分区划分

25



选择使用所有空间,并将左下角的"查看并修改分区布局"选中,如果系统默认给我们分的我们自己觉得不好,可以自己修改。



#### 磁盘划分情况:

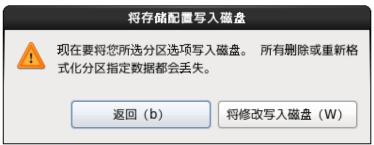


硬盘驱动器下的 sda 表示只有一块盘,如果有多块盘会出现 sdb、sdc等, sda1 是启动分区,sda2 是存储分区。sda2 又分成了两个逻辑分区 lv\_root 和 lv\_swap,这两个分区可以调整大小,并且通过外挂盘可以增加逻辑分区容



量。点击下一步会出现格式化的警告,点击格式化,再点击"将修改写入磁盘"。





#### 6.9 引导程序配置,默认即可,点击下一步



#### 6.10 选择桌面系统类型

分别有如下系统类型可以选择:桌面版、精简桌面版、最小化安装、基本服务器、数据库服务器、web 服务器、虚拟主机、软件开发工作站。



CentOS 默认安装是最小安装。您现在可以选择一些另外的软件。
● Desktop
Minimal Desktop
○ Minimal
O Basic Server
O Database Server
○ Web Server
O Virtual Host
Software Development Workstation
请选择您的软件安装所需要的存储库。
✓ CentOS
→ (A)添加额外的存储库 📝 修改库 (M)
或者.
● 以后自定义 (I) ○ 现在自定义 (C)

#### 6.11 经过一段时间等待,安装成功,重启

祝贺您,您的 CentOS 安装已经完成。

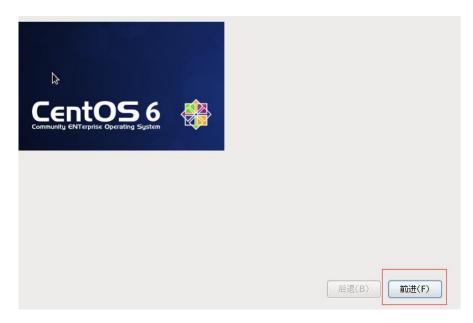
请重启以便使用安装的系统。请注意:可使用更新以确定您的系统正常工作,且建议在重 后后安装这些更新。

## 7. 重启之后,用户信息配置

#### 7.1 欢迎页面右下角"前进"按钮







## 7.2 创建普通用户 Hadoop



创建用户
您必须为您的系统创建一个常规使用的(非管理)'用户名'。要创建系统'用户名',请提供以下所需信息。
用户名(U): hadoop
全名(e): hadoop
密码 (P) : ·····
确认密码(m): ・・・・・
如果您需要使用网络验证,比如 Kerberos 或者 NIS ,请点击"使用网络登录"按钮。
使用网络登录(L)
如果您需要在创建该用户时有更多控制[指定主目录和(/或者)UID],请点击高级按钮。
高級(A)

由于密码简单会出现提示,点击"是"



#### 7.3 设置日期和时间



7.4 在机器崩溃的时候是否把内存中的数据 dump 下来,选择



#### 不启用。点击右下角完成按钮, 重启



#### 7.5 重启完成, 到此虚拟机和操作系统安装完成



## 第二章 Centos 网络设置

## 第一节 网络设置

## 1. 查看虚拟网关 IP 地址

VMware 的工具栏"编辑"->"虚拟网络编辑器"



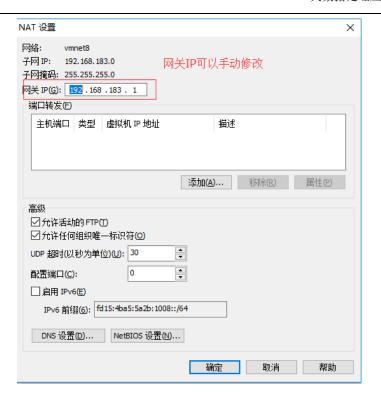


虚拟网络编辑器选择 VMnet8 显示类型为 NAT 模式,点击"NAT 设置"选项查看和修改网关 IP,如果不是管理员用户权限将无法修改,可以在虚拟网络编辑器的右下角点击更改设置,具有修改权限。



NAT 设置界面中的网关 IP 可以手动修改,将所有需要连通的虚拟机的网关 IP 设置为该网关 IP,





#### 2. 设置虚拟机网络适配器

选中网络适配器选项->在对话框的右侧选择"自定义:特定虚拟网络"值为"VMnet8(NAT 模式)"。



## 3. 设置虚拟机 IP、子网掩码、网关、DNS



首先在虚拟机桌面右键打开终端,首先输入 su root 命令切换到 root 用户。 然后在终端输入命令: vim /etc/sysconfig/network-scripts/ifcfg-eth0 输入 i 进入编辑模式,然后对文件内容进行修改。

修改内容如下:

DEVICE=eth0

TYPE=Ethernet

ONBOOT=yes #是否开机启用

BOOTPROTO=static #IP 地址设置为静态

IPADDR=192.168.183.100 #IP 地址根据虚拟网关的子网 IP 网段自定义

NETMASK=255.255.255.0 #与虚拟网关的子网掩码相同

GATEWAY=192.168.183.1 #设置网关

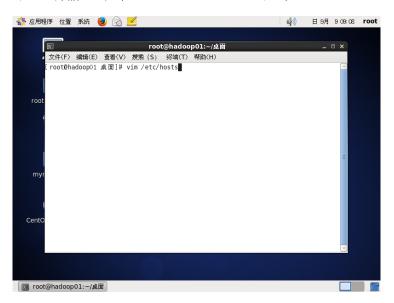
DNS1=192.168.183.1 #设置 DNS

#### 注意: 网关和 DNS 要设置, 否则无法上外网

修改完成后,键盘 Esc 键退出编辑模式,然后输入":wq"保存退出。 然后在终端输入命令重启网络: service network restart

# 4. 设置虚拟机 IP 地址与主机名 hostname 映射

在终端输入命令: vim /etc/hosts, 回车



输入 i 进入编辑模式, 将光标移动到文件末尾(进入编辑模式后, 终端最



下面会显示"--插入--")

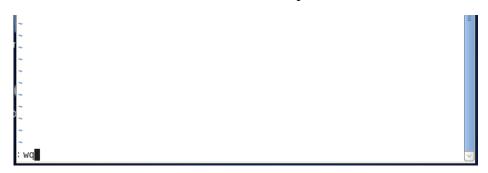
在文件最后一行添加如下 IP 与主机名的映射

192.168.183.100(虚拟机的 ip) node1(虚拟机的主机名)

- \*获取虚拟机主机名命令: hostname
- \*获取虚拟机 IP 命令: ifconfig



按"Esc"键退出编辑模式,输入": wq",然后回车保存并退出。

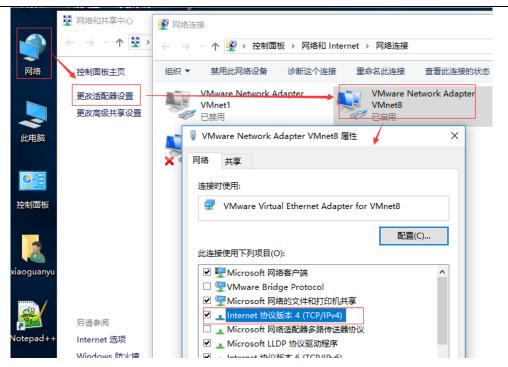


## 5. 设置安装虚拟机电脑的虚拟网卡 VMnet8 的 IP 地址、

#### 网关、子网掩码

控制面板->网络和 Internet->更改适配器设置->选择 VMnet8 右键属性->鼠标双击 IPv4 选项修改 IP、网关、子网掩码。



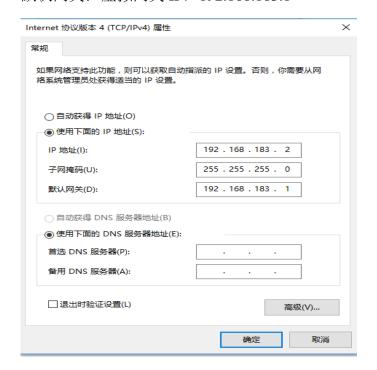


IP 地址:根据 vmware 虚拟网关 IP 网段设置与网关 IP 不同的地址,如

192.168.183.2

子网掩码:与虚拟网关的子网掩码相同 255.255.255.0

默认网关: 虚拟网关 IP, 192.168.183.1





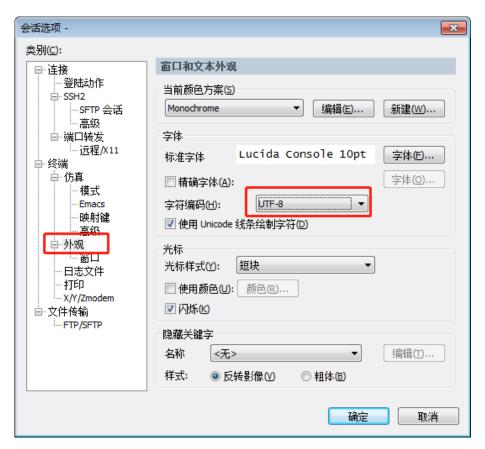
# 第二节 使用CRT远程连接虚拟机

### 1. 下载 CRT 软件

链接: https://pan.baidu.com/s/1pq1NTVMDMOn9JHFd2vNISA 提取码: hiuw



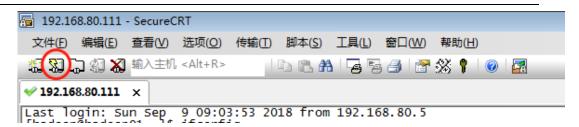
打开软件后,先修改一下软件字符集:选项-->会话选项-->外观-->字符编码-->UTF-8



### 2. 配置 CRT 远程设置

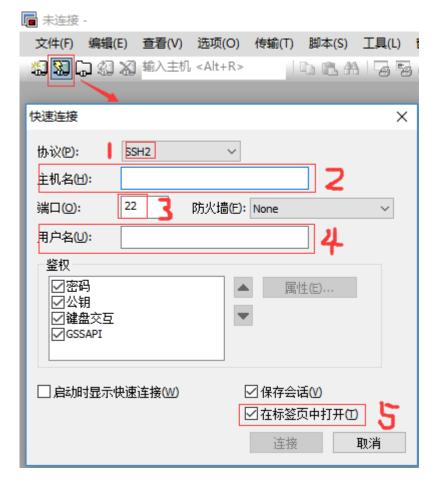
新建一个远程连接(第二排第二个选项新建快速连接)





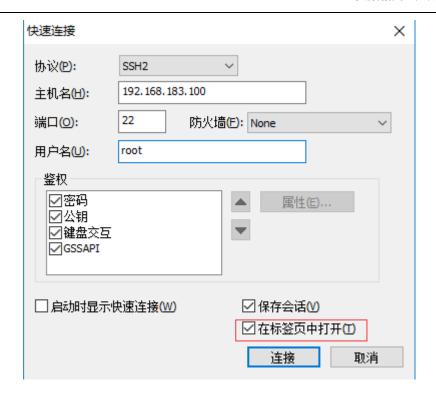
需要注意的5个地方:

- (1) 协议选择默认的 SSH2;
- (2) 主机名填写需要远程连接的服务器 IP 或者主机名,如果填写主机名要 在本机的 hosts 中配置 IP 与主机名的映射;
- (3) 端口默认 22:
- (4) 填写远程访问的用户名
- (5) 勾选在标签页中打开,这样在打开多个会话时会显示在同一个界面的 不同标签页,方便操作,如果不希望这样的显示方式可以不勾选。



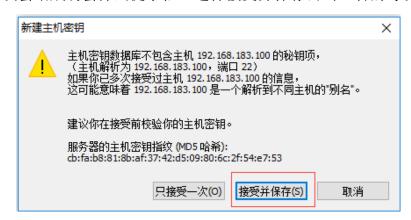
样例配置如下:



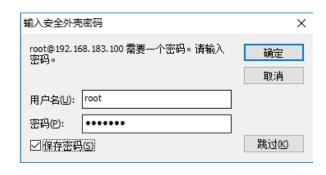


### 3. 保存密码

第一次会话成功会弹出提示框,选择接受并保存密码,省的每次都提示。



### 4. 输入登录的用户名





# 第三节 虚拟机iptables防火墙设置

开启防火墙会阻止某些外部的访问,比如访问 tomcat 的 8080 端口,因此 我们需要关闭防火墙,并设置防火墙开机不启动。

# 1. 查看 iptables 状态

service iptables status

# 2. 开启/关闭 iptables

service iptables start

service iptables stop

# 3. 查看 iptables 自动启动状态

chkconfig iptables --list

# 4. 设置 iptables 开机启动/不启动

chkconfig iptables on

chkconfig iptables off

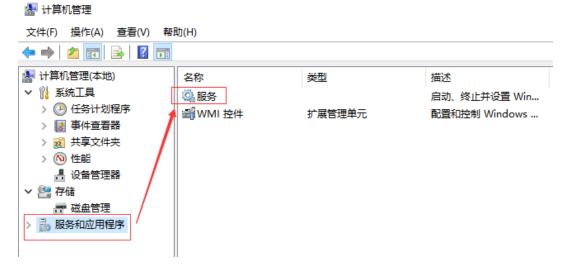


# 第四节 虚拟机上外网注意事项

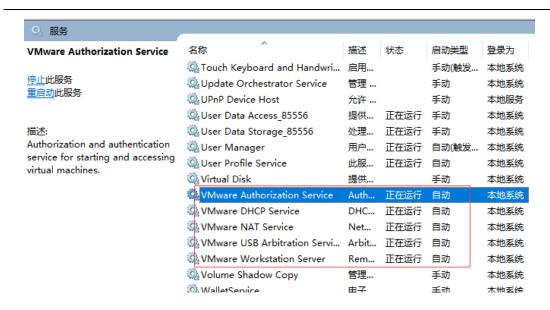
# 1. 安装虚拟机的电脑服务中所有关于 VMware 的服务开启

开启方法: 我的电脑右键管理->服务和应用程序->服务 找到 VM 开头的相关服务项全部开启。









# 2. 虚拟机的 IP、网关、DNS、子网掩码配置正确

上述配置都配置完成发现无法 ping 外网

需要设置网关

route add default gw 192.168.183.1

添加 nameserver

vim /etc/resolv.conf

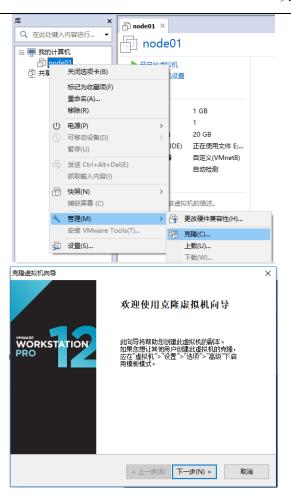
nameserver 192.168.183.1

3. 选中一台虚拟机->鼠标右键设置->"网络适配器"要选择自定义"VMnet8(NAT 模式)"

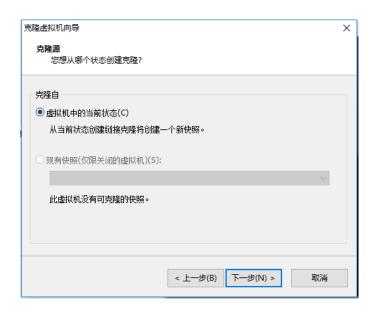
# 第三章 克隆虚拟机

1. node 节点鼠标右键管理->克隆





# 2. 从当前状态开始克隆,会将当前源虚拟机中的已安装和配置的环境克隆下来

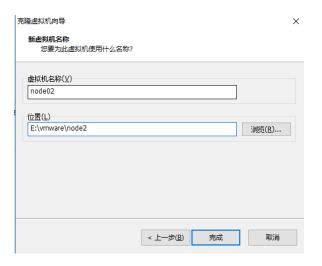




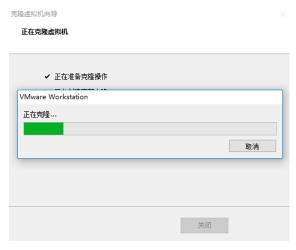
# 3. 选择"创建完整克隆"



# 4. 创建新的虚拟机名称,克隆虚拟机文件存储位置

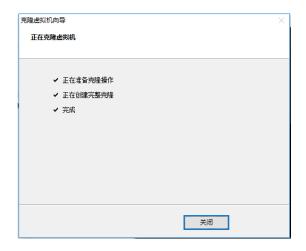


# 5. 点击"完成"之后,等待进度条完成





### 6. 创建完成



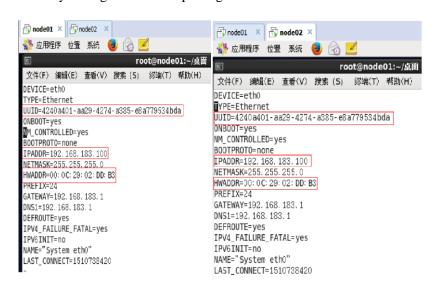
# 7. 启动新克隆的虚拟机 node2

问题: 克隆之后物理地址就会重复,会生成一个新的网卡 eth1,出现网络问题,需要切换到 eth0 网络

解决办法:

删除 ifcfg-eth0 文件中的 UUID 和 HWADDR,修改 IP 地址

vim /etc/sysconfig/network-script/ifcfg-eth0

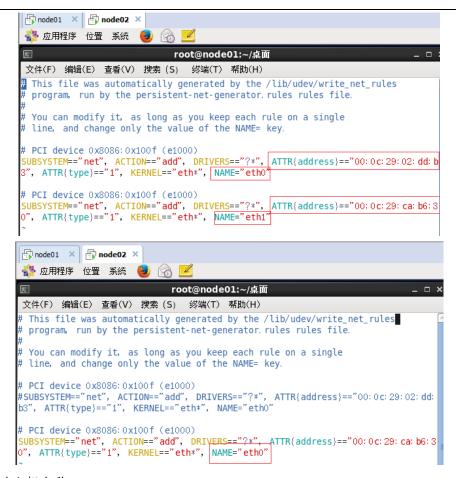


将网络从 eth1 修改成 eth0

vim /etc/udev/rules.d/70-persistent-net.rules

直接注释掉 eth0,将 eth1 的名字改为 eth0





修改主机名称

vim /etc/sysconfig/network

HOSTNAME=node02

重启完成

# 8. 克隆 Node03

操作同 node02 克隆方式相同,只是 ip 和主机名不同而已!

# 第四章 SSH 免密码登录

1. 使用 root 用户登录到 node01 机器,首先进入 root 用户的家目录下



cd /root

2. 进入到.ssh 目录下(注意.ssh 是隐藏文件)

cd .ssh

3. 在.ssh 目录下执行如下命令生成私钥(id\_rsa)和公钥文件(id\_rsa.pub)

执行命令 ssh-keygen -t rsa

```
[root@node01 .ssh]# ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/root/.ssh/id_rsa):
```

遇到需要输入的提示一直按回车

4. 执行完成会生成私钥(id\_rsa)和公钥文件(id\_rsa.pub)

```
[root@node01 .ssh]# ]]
总用量 16
-rw------ 1 root root 1572 12月 6 07:13 authorized_keys
-rw----- 1 root root 1675 12月 5 22:42 id_rsa
-rw-r--r-- 1 root root 393 12月 5 22:42 id_rsa.pub
-rw-r--r-- 1 root root 1980 12月 8 09:57 known_hosts
```

公钥内容如下: (注意这是我自己 node01 机器的 root 用户公钥,每个人机器上的都不一样,不需要跟我的一致)

[root@node01 .ssh]# cat id\_rsa.pub
ssh-rsa AAAAB3NzaC1yc2EAAAABIwAAAQEAwzLi38qousV0wFZVQL3Q5UJfyDwvq7JBChqiMJGmM41RUPUkpksfgFWJHK7GZwHJ9W
6v2QqPxspnomVHj/frxod8g561/yDZ/BXUTd0BKcsxRh4wRxqLEIluiZ8+4B5GsMJzO+oHrKpmzSHt0ugm2SJccWmBZKwTFIdf+jUV
gImk9uGx///MF1h+W1QRTynNbCwoUypPFDC0g6cCLdfU80fyAfzkXTg24oGf0BvrWsUbZtwXBwK9j4cc6JYwYhcZBTM1pRjzi6IxpS
34/rJyaxsF7foogvRAH8XvBJy/gwzoMX+uy8nPVPXD8292FmOvwZ5J0AOGapKTnhOxQATX410== root@node01

5. 通过 ssh-copy-id 命令将 node01 这台机器 root 用户的公钥文件(id\_rsa.pub)文件内容拷贝到 node02 和 node03 两台机器

ssh-copy-id -i node02 ssh-copy-id -i node03



# 6. 在 node02 和 node03 的 root 用户的家目录下,进入到.ssh 目录

cd /root/.ssh

在 authorized\_keys 文件中已经添加了 node01 的公钥文件内容,可以和第 4 步 node01 机器生成的 root 用户公钥文件内容对比下,应该是一样的

[root@node02 .ssh]# cat authorized\_keys
ssh-rsa AAAAB3NzaC1yc2EAAAAB1wAAAQEAwzLi38qousv0wFZVQL3Q5UJfyDwvq7JBChqiMJGmM41RUPUkpkSfgFWJHK7GZwHJ9w
6v2QqPxspnomVHj/frxod8g561/yDZ/BXUTd0BkcsxRh4wRXqLEIIuiZ8+4B5GsMJzO+oHrKpmzSHt0ugm2SJccWmBZKwTFIdf+jUv
gmlk9uGx///MF1h+W1QRTynNbCwoUypFDC0g6cCLdfU80fyAfzkXTg24oGf0BvrWsUbZtwXBwK9j4cc6JYwYhcZBTM1pRjzi6IxpS
34/rJYaxsF7foogvRAH8XvBJv/qwzoMX+uY8nPVPXD829zFmOvwZ5J0AQapKTnhQxqATX41Q== root@node01

7. 在 node01 通过 ssh 登录到 node02 和 node03, 此时不在需要输入密码

ssh node02 这样就可以直接登录到 node02 机器

# 第五章 常用命令和系统命令

# 第一节 常用命令和快捷键

1. 停止进程

ctrl + c

2. 清屏

ctr1 + 1

3. 搜索历史命令

ctrl + r

4. 文件名自动补全



Tab 键

### 5. 进入目录

cd path

### 6. 进入用户家目录

cd∼

# 7. 进入最近上一次访问的目录

cd -

### 8. 返回上一级目录

cd ..

### 9. 查看目录文件列表

1s: 查看目录下文件列表 (不显示隐藏文件)

1s -a: 查看目录下所有文件列表(显示隐藏文件)

1s -1: 查看目录下所有文件的详细信息列表,该命令等价于 11

1s -la: 查看目录下所有文件详细信息列表(显示隐藏文件)

### 10. 查看文件全部内容

cat filename

### 11. 查看文件末尾 num 行

tail -num filename

# 12. 查看文件头部 num 行

head -num filename



# 13.查看文件行数

wc -1 filename

### 14.查看文件中字符个数

wc -w filename

### 15.复制文件

- cp filename ./newfilename
- cp filename /otherdir
- 使用通配符\*复制多个文件
- cp hello\*. txt /otherdir 复制 hello 开头的 txt 文件到其他指定路径

### 16.重命名文件

mv filename newfilename

### 17.查看历史命令

history

结合管道查找指定命令

history | grep rm 在历史命令中查找执行 rm 的操作

### 18. 软链接

#### 创建软链接

linux 下的软链接类似于 windows 下的快捷方式

ln -s targetfile linkfile 为 targetfile 文件创建软链接 linkfile 删除软链接

rm -rf linkfile

### 19.查找文件



1) 按文件名称查找,可以使用统配符\*

在/datadir 路径下查找 hello 开头的文件,并列出详细信息 find /datadir -name "hello\*" -ls

2) 按照文件所属用户查找

在/home/hadoop 路径下查找属于用户 hadoop 的文件

find /home/hadoop -user hadoop -ls



# 第二节 系统管理命令

### 1. 查看本地时间

date

### 2. 联网更新系统时间

 → 设置时区为上海,禁止使用 utc vi /etc/sysconfig/clock
 ZONE="Asia/Shanghai"
 UTC=false #设置为 false

➤ linux 的时区设置为上海
ln -sf /usr/share/zoneinfo/Asia/Shanghai /etc/localtime

- ➤ 对准时间,需要先安装 ntp 服务器 yum install ntp ntpdate -u ntp.api.bz
- ➤ 设置硬件时间和系统时间一致并校准 /sbin/hwclock --systohc
- ➤ 重启 reboot

### 3. 查看主机名

hostname

### 4. 修改主机名(重启后永久生效)

vim /ect/sysconfig/network

### 5. 修改 IP(重启后永久生效)

vim /etc/sysconfig/network-scripts/ifcfg-eth0



# 6. 修改/设置 IP 和主机名映射

vim /etc/hosts

# 7. 统计文件或文件夹的大小

du -sh /bigdata/data

### 8. 查看磁盘的空间

df -h

### 9. 关机

halt

### 10.重启

reboot

# 11.查看命令帮助

command - help

# 12.echo 使用

- ➤ 在显示器上显示文字 echo helloworld
- ➤ 将显示内容输出到文件中(覆盖) echo helloworld > testfile
- ➤ 将显示内容追加输出到文件的末尾 echo hellohaha >>testfile

### 13.service 后台服务管理

> service --status-all 查看系统中所有后台服务



- > service network status 查看指定服务的状态
- > service network stop 停止指定服务
- > service network start 启动指定服务
- > service netework restart 重启指定服务

### 14.查看自动启动服务

chkconfig

```
[root@node01 ~] # chkconfig
NetworkManager 0:关闭 1:关闭 2:启用 3:启用 4:启用 5:启用 6:关闭
abrt-ccpp 0:关闭 1:关闭 2:关闭 3:启用 4:关闭 5:启用 6:关闭
abrtd 0:关闭 1:关闭 2:关闭 3:启用 4:关闭 5:启用 6:关闭
acrid 0:关闭 1:关闭 2:关闭 3:启用 4:关闭 5:启用 6:关闭
acrid 0:关闭 1:关闭 2:点用 3:启用 4:启用 5:启用 6:关闭
acrid 0:关闭 1:关闭 2:点用 3:启用 4:启用 5:启用 6:关闭
autof 0:关闭 1:关闭 2:点用 3:启用 4:启用 5:启用 6:关闭
autof 0:关闭 1:关闭 2:点用 3:启用 4:启用 5:启用 6:关闭
autof 0:关闭 1:关闭 2:点用 3:启用 4:启用 5:启用 6:关闭
blk-availability 0:关闭 1:关闭 2:关闭 3:启用 4:启用 5:启用 6:关闭
blk-availability 0:关闭 1:关闭 2:关闭 3:启用 4:启用 5:启用 6:关闭
certmonger 0:关闭 1:关闭 2:关闭 3:启用 4:启用 5:启用 6:关闭
crond 0:关闭 1:关闭 2:关闭 3:启用 4:启用 5:启用 6:关闭
crond 0:关闭 1:关闭 2:点用 3:启用 4:启用 5:启用 6:关闭
crond 0:关闭 1:关闭 2:点用 3:启用 4:启用 5:启用 6:关闭
crond 0:关闭 1:关闭 2:启用 3:启用 4:启用 5:启用 6:关闭
firstboot 0:关闭 1:关闭 2:点闭 3:点用 4:启闭 5:关闭 6:关闭
dnsmasq 0:关闭 1:关闭 2:关闭 3:点用 4:启闭 5:关闭 6:关闭
htcacheclean 0:关闭 1:关闭 2:关闭 3:关闭 4:关同 5:关闭 6:关闭
htcacheclean 0:关闭 1:关闭 2:关闭 3:关闭 4:关同 5:关闭 6:关闭
ipftables 0:关闭 1:关闭 2:关闭 3:点用 4:启用 5:启用 6:关闭
ipftables 0:关闭 1:关闭 2:关闭 3:点用 4:启用 5:点用 6:关闭
ipftables 0:关闭 1:关闭 2:关闭 3:点用 4:点用 5:点用 6:关闭
```

- ▶ chkconfig iptables --list 查看指定服务的自动启动状态
- > chkconfig iptables off 关闭指定服务的自动启动
- > chkconfig iptables on 启动指定服务的自动启动

#### 系统启动级别管理

vim /etc/inittab

```
[root@node01 ~]# vim /etc/inittab
  inittab is only used by upstart for the default runlevel.
  ADDING OTHER CONFIGURATION HERE WILL HAVE NO EFFECT ON YOUR SYSTEM.
#
  System initialization is started by /etc/init/rcs.conf
#
#
  Individual runlevels are started by /etc/init/rc.conf
  Ctrl-Alt-Delete is handled by /etc/init/control-alt-delete.conf
  Terminal gettys are handled by /\text{etc/init/tty.conf} and /\text{etc/init/serial.conf}, with configuration in /\text{etc/sysconfig/init}.
#
  For information on how to write upstart event handlers, or how
  upstart works, see init(5), init(8), and initctl(8).
#
  Default runlevel. The runlevels used are:
    0 - halt (Do NOT set initdefault to this)
1 - Single user mode
#
#
    2 - Multiuser, without NFS (The same as 3, if you do not have networking)
3 - Full multiuser mode
#
#
    4 -
         unused
#
         X11
         reboot (Do NOT set initdefault to this)
#
id:5:initdefault:
                                默认启动级别,桌面版
```



- 0: 关闭所有进程并终止系统。(不要设置)
- 1: 单用户模式,只能系统管理员进入,在该模式下处理在有登录用户时不能进行更改的文件。
- 2: 多用户的模式,但并不支持文件共享,这种模式很少应用。
- 3: 最常用的运行模式,主要用来提供真正的多用户模式,也是多数服务器的缺省模式。
- 4: 一般不被系统使用。
- 5: 桌面启动模式,如果关闭将不启动桌面程序。
- 6: 关闭所有运行的进程并重新启动系统。(不要设置)

# 第六章 用户管理

#### 1. 添加新组

groupadd groupname

### 2. 添加新用户

useradd username

### 3. 设置用户密码

passwd username

### 4. 给用户添加组

usermod -g groupname username

### 5. 添加新用户同时添加组

useradd -g groupname username



### 6. 查看用户组

groups username

### 7. 删除用户

userdel username

### 8. 删除组

groupdel groupname

### 9. 将用户从组中删除

gpasswd -d groupname username

### 10.切换用户

su username

### 11.给普通用户赋予 root 系统权限

- a) 如果普通用户想获取临时 root 用户系统权限需要 root 用户添加,为了防止误操作或者恶意操作引起的系统故障或重大事故,尽量减少使用 root 用户,因为 root 用户具有最高系统权限。普通用户想获取 root 用户的临时权限,需要 root 用户为其添加。
- b) 用 root 编辑/etc/sudoers
- c) vim /etc/sudoers
- d) 在文件为需要 sudo 权限的用户添加 一行,如给 hadoop 用户添加 sudo 权限。
- e) root ALL=(ALL) ALL
- f) hadoop ALL=(ALL) ALL
- g) hadoop 用户在需要使用 root 用户系统权限的时候在操作命令之前加上



sudo

h) sudo useradd data\_bi



# 第七章 文件管理

### 1. 创建文件

touch filename

### 2. 删除文件

rm filename (会有提示是否删除) rm -f filename (强制删除)

### 3. 创建文件夹

mkdir dirname
mkdir -p dir1/dir2/dir3 (递归创建文件夹)
mkdir dir/{dir1, dir2}在 dir 在已存在的 dir 文件夹下同时创建 dir1 和
dir2 两个文件夹

# 4. 删除文件夹及所有子文件夹

rm -r dirname (会有提示是否删除文件夹)
rm -rf dirname (强制删除文件夹,没有提示)

### 5. 文件或文件夹权限

文件夹或者文件名称前用四种字母和符号表示的一串字符串表示的文件或者文件夹类型和权限。

-rw-rw-r--

drwxrwxr-x

1rwxrwxrwx

-: 表示文件类型为文件



- d: 表示文件类型为文件夹
- 1: 小写的 L 表示符号链接
- r: 可读
- w: 可写
- x: 可执行

drwxrwxr-x 文件夹权限举例:

第一组 rwx:表示这个文件夹的拥有者对文件夹的权限:可读可写可执行

第二组 rwx:表示这个文件夹的所属组对文件夹的权限:可读可写可执行

第三组 r-x: 表示这个文件夹的其他用户对文件夹的权限: 可读,不可写,可执行

### 6. 修改文件权限

- ➤ chmod u/g/o +/- r/w/x 表示给文件的用户/组/其他添加或者取消读写 执行权限
- ▶ chmod u+x filename 表示为文件/文件夹所属用户添加可执行权限
- ▶ chmod u-x filename 表示为文件/文件夹所属用户删除可执行权限
- ▶ chmod g-rw filename 表示将文件/文件夹对所属组的 rw 权限取消
- ▶ chmod o-rw filename 表示将文件/文件夹对其他人的 rw 权限取消

文件权限的每一组可以用三位二进制数表示

rwxrwxr-x 对应的二进制 111111101 每组换算成十进制数字 775 如果想取消其他用户的执行权限 rwxrwxr--换成二进制 111111100 换算成十进制数字 774

执行赋权命令: chmod 774 finame

如果要将一个文件夹的所有内容权限统一修改,则可以使用-R参数 chmod -R 777 dirname/

注意:目录没有执行权限的时候普通用户不能进入,属主也不可以 当文件具有读写权限时,虽然没有执行权限,如果该文件父文件夹具有写



权限普通用户可以在文件夹中删除该文件,这样的删除操作属于对父级文件夹的修改。如果父文件夹没有写权限普通用户是不能再文件夹中删除文件的。

### 7. 查看文件内容

- ➤ cat filename 将文件内容全部输出到控制台
- ➤ more filename 将文件内容分页显示,按空格下翻页,按 b 字母键上翻页,按 q 字母键 退出。
- ➤ less filename 将文件内容分页显示,按空格下翻页,按 b 字母键上翻页,按上箭头(↑)上翻一行,按下箭头(↓)下翻一行。
- ➤ tail -num filename 查看文件的后 num 行
- ➤ tail -f filename
  Linux 中每一个文件有一个 inode, 文件名修改 inode 编号不变, 当文件

修改名称后 tail -f 继续跟踪原 inode 编号文件;不支持文件滚动。

- ➤ tail -F filename 按照文件名称跟踪文件,文件修改名称后,有新的文件如果使用原文件 名称,则继续跟踪该名称文件;支持文件滚动。
- ► head -num filename 查看文件头部的 num 行

### 8. Vim 文本编辑器

#### 8.1 vim test

首先会进入"一般模式",此模式只接受各种快捷键,不能编辑文件内容;



- i 从一般模式进入编辑模式,此模式下可以输入内容;
- o 从一般模式进入编辑模式并且是光标所在行的下一行开始输入内容;
- u 撤销到上一步操作,多次撤销按多次 u;

Esc 退出编辑模式,回到一般模式;

: 进入"底行命令模式",输入wq命令,回车即可保存退出;

#### 8.2 常用快捷键(在一般模式下使用)

- ▶ a 在光标后一位开始插入
- ▶ A 在该行的最后插入
- ▶ I 在该行的最前面插入
- ▶ gg 直接跳到文件的首行
- ▶ G 直接跳到文件的末行
- ▶ dd 删除行,输入 2dd,则一次性删除光标所在行一下的 2 行
- ▶ yy 复制当前行,复制多行"num+yy",如复制三行 3yy,则复制当前行和后 2 行
- ➤ p 粘贴
- ▶ v 进入字符选择模式,选择完成后,按 v 复制,按 p 粘贴
- ▶ ctrl+v 进入块选择模式,选择完成后,按 y 复制,按 p 粘贴
- ▶ shift+v 进入行选择模式,选择完成后,按y复制,按p粘贴

#### 8.3 查找并替换

在一般模式下按:冒号进入命令模式操作,输入 %s/被替换的内容/新替换的内容,如下:

%s/hello/hi 查找文件中所有 hello, 替换为 hi

#### 8.4 查找

在一般模式下输入"/查找的内容",如下:

/hi 查找文件中出现的 hi,并定位到第一个找到的地方,按 n 可以定位到下一个匹配位置(按 N 定位到上一个)



# 第八章 压缩解压命令

### 1. zip 压缩/解压缩,压缩文件后缀名.zip

```
zip tes.zip test.txt
压缩文件夹
zip -r dirtest.zip dirtest
解压缩文件/文件夹
unzip tes.zip/dirtest.zip
```

# 2. gzip 压缩/解压缩,压缩文件后缀名.gz

```
gzip test.txt
解压 gz 文件:
gzip -d test.gz
```

# 3. bzip2 压缩/解压缩,压缩文件后缀名.bz2

```
bzip2 test1
解压 bz2 文件
bzip2 -d test1.bz2
```

### 4. 打包/解包, 打包文件后缀名.tar

▶ 打包:

```
tar -cvf barball.tar ball/
可以将多个文件打在一个包里
tar -cvf files.tar file1 file2 file3
```

➤ 解包: tar -xvf barball.tar

# 5. 一次性完成打包和 gzip 压缩的操作,文件后缀名.tar.gz



#### ▶ 打压缩包:

tar -zcvf tarball.tar.gz ball/

▶ 解压缩包:

tar -zxvf tarball.tar.gz

▶ 解压缩到指定路径下

tar -zxvf tarball.tar.gz -C ./tardir

# 6. 一次性完成打包和 bzip2 压缩的操作,文件后缀名.tar.bz2

tar -jcvf bz2dir.tar.bz2 bz2dir/ 将 bz2dir.tar.bz2解压到/usr 目录下面 tar -jxvf bz2dir.tar.bz2



# 第九章 Linux软件安装

# 1. 使用 yum 安装文件传输工具 lrzsz

yum 是基于 RMP 包管理,能够从指定的服务器自动下载安装 RPM 包,并且自动处理依赖关系,方便快捷。

常用命令:

- ➤ 查看 yum 提供的 rpm 列表 yum list
- ➤ 使用管道查找 yum 提供的 java 软件包 yum list | grep java
- ➤ 使用 yum 安装 lrzsz yum install -y lrzsz.x86\_64
- ➤ 移除 jdk
  yum remove -y java-1.6.0-openjdk-1.6.0.35-1.13.7.1.el6\_6.x86\_64

\*\*\*\*\*\*\*\*\*\*

yum install -y lrzsz

rz: 从本机向服务器上传文件

sz: 从服务器下载文件到本机

### 2. 手动安装 JDK

- ➤ 下载安装包: jdk-8u151-linux-x64.tar.gz 链接: https://pan.baidu.com/s/1C9vI4vS9hWjjgJ\_rzKPv6A 提取码: 56mv
- ➤ 安装包上传到服务器,解压到安装目录: tar -zxvf jdk-8u151-linux-x64.tar.gz -C /bigdata
- ➤ 创建软链接:
  ln -s /bigdata/jdk1.8.0\_151 /usr/local/jdk



▶ 配置环境变量:

vim /etc/profile

export JAVA\_HOME=/usr/local/jdk

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

▶ 重新加载环境变量:

source /etc/profile

注意:如果系统预装了openjdk,将其卸载方法:

使用 rpm 查找安装的 jdk 版本

rpm -qa | grep jdk

移除预装的 jdk

yum -y remove java-1.6.0-openjdk-1.6.0.35-1.13.7.1.el6\_6.x86\_64

### 3. telnet 安装

➤ CentOS 下查看系统是否已安装 telnet

rpm -qa | grep telnet

- ➤ telnet 要挂在 xinetd 下,查看是否安装了 xinetd 服务 rpm -qa | grep xinetd
- ➤ 两者都没有的安装的话, yum 安装以下服务, 否则只安装一个即可 yum install xinetd telnet telnet-server -y
- ➤ 配置 telnet
- 1) 设置开机启动

chkconfig telnet on

2) 修改配置文件

vi /etc/xinetd.d/telnet

将"disable=yes"改成" disable=no"

3) 激活服务

重启 xinetd 服务

service xinetd restart

或者/etc/rc.d/init.d/xinetd restart



# 第十章 CentOS安装MySQL

### 1. 安装前检查

在安装 MySQL 之前,先检查 CentOS 系统中是否已经安装了一个 MySQL,如果已经安装先卸载,不然会导致安装新的 MySQL 失败。

查看系统之前是否已安装 MySQL:

rpm -qa | grep mysql

[user@master workspace]\$ rpm -qa | grep mysql mysql-libs-5.1.71-1.el6.x86\_64

从上图得知,CentOS6 系统自带了一个 MySQL,我们需要删除这个老版本,用 **root** 用户执行下面语句: (如果不存在老版本,则不需要执行下面删除操作)

rpm -e --nodeps mysql-libs-5.1.71-1.el6.x86\_64 \*此处的包名根据实际情况修改

```
[user@master ~]$ su root
Password:
[root@master user]# rpm -e --nodeps mysql-libs-5.1.71-1.el6.x86_64

[root@master user]# rpm -qa | grep mysql
[root@master user]#
```

上图中,我们先切换到"**root**"用户下,然后执行删除语句,删除之后,我们再次查看,发现已经成功删除了 CentOS6 自带的旧 MySQL 版本。

在删除 MySQL 的 rpm 后,还要进行一些扫尾操作,网上有两种操作。

(1) 第一种善后处理: 使用下面命令进行处理。

rm -rf /var/lib/mysql\*

rm -rf /usr/share/mysql\*

(2) 另一种善后处理: 卸载后/var/lib/mysql 中的/etc/my.cnf 会重命名为my.cnf.rpmsave, /var/log/mysqld.log 会重命名为/var/log/mysqld.log.rpmsave, 如果确定没用后就手工删除。

rm -rf /etc/my.cnf.rpmsave

rm -rf /var/log/mysqld.log.rpmsave



# 2. 开始安装

(1) 查看有没有安装过:

yum list installed | grep mysql\*

(2) 查看有没有安装包:

yum list mysql\*

(3) 安装 mysql

# wget http://repo.mysql.com/mysql-community-release-el6-5.noarch.rpm

# rpm -ivh mysql-community-release-el6-5.noarch.rpm

# yum repolist all | grep mysql

# yum install mysql-community-server -y

# 3. 启动 mysql 服务

service mysqld start 或者/etc/init.d/mysqld start

### 4. 设置开机启动

chkconfig --add mysqld chkconfig mysqld on

### 5. 查看开机启动设置是否成功

chkconfig --list | grep mysqld

mysqld θ:关闭 1:关闭 2:启用 3:启用 4:启用 5:启用 6:关闭

# 6. 登录 MySQL

设置 root 密码:

mysql\_secure\_installation

登录:

mysql -u root -p 输入密码即可。



# 7. 建立远程 root 用户

赋权:

GRANT ALL PRIVILEGES ON \*.\* TO 'root'@'%' IDENTIFIED BY '你设置的密码' WITH GRANT OPTION;

刷新:

flush privileges;

### 8. 删除 user 表中的干扰项

删除用户名为空的数据:

delete from mysql.user where user=";

刷新:

flush privileges;

# 9. 为 Hive 建立相应的 MySQL 账户,并赋予足够的权限

创建用户:

CREATE USER 'hive'@'%' IDENTIFIED BY 'hive123';

\*与hive-site.xml中的ConnectionUserName与ConnectionUserPassword相对应

授权:

GRANT ALL PRIVILEGES ON \*.\* TO 'hive'@'%' WITH GRANT OPTION;

刷新:

flush privileges;