

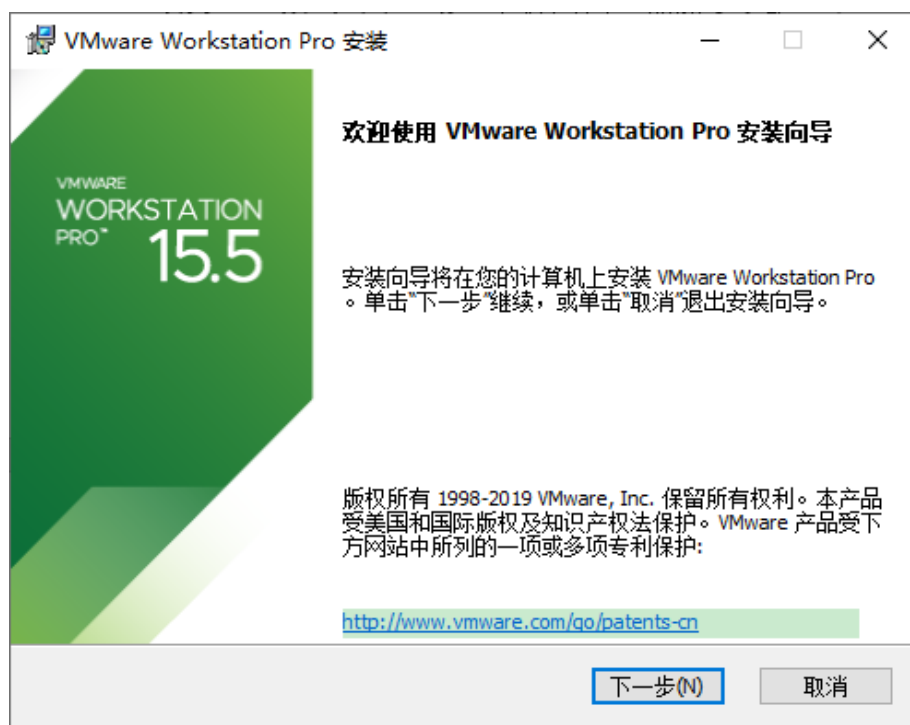


## 第1章 VM 与 Linux 的安装

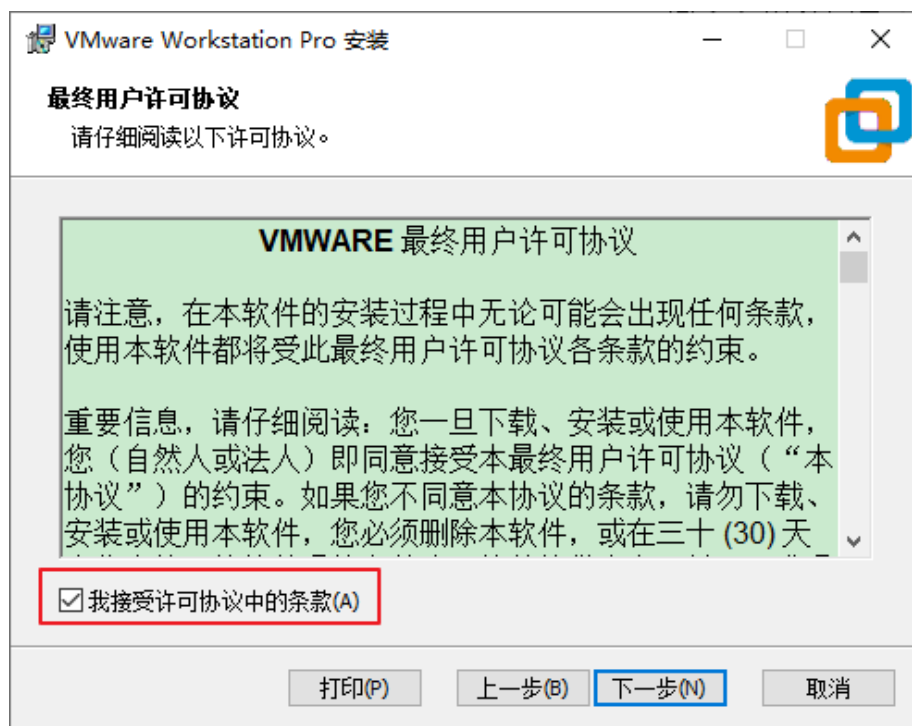
学习大数据，为了能够在本机模拟真实环境，需要在本机安装虚拟机软件(VMware)和虚拟系统（CentOS7）。相关软件在资料文件夹中，请大家下载后，进行安装。

### 1.1 VMWare 安装

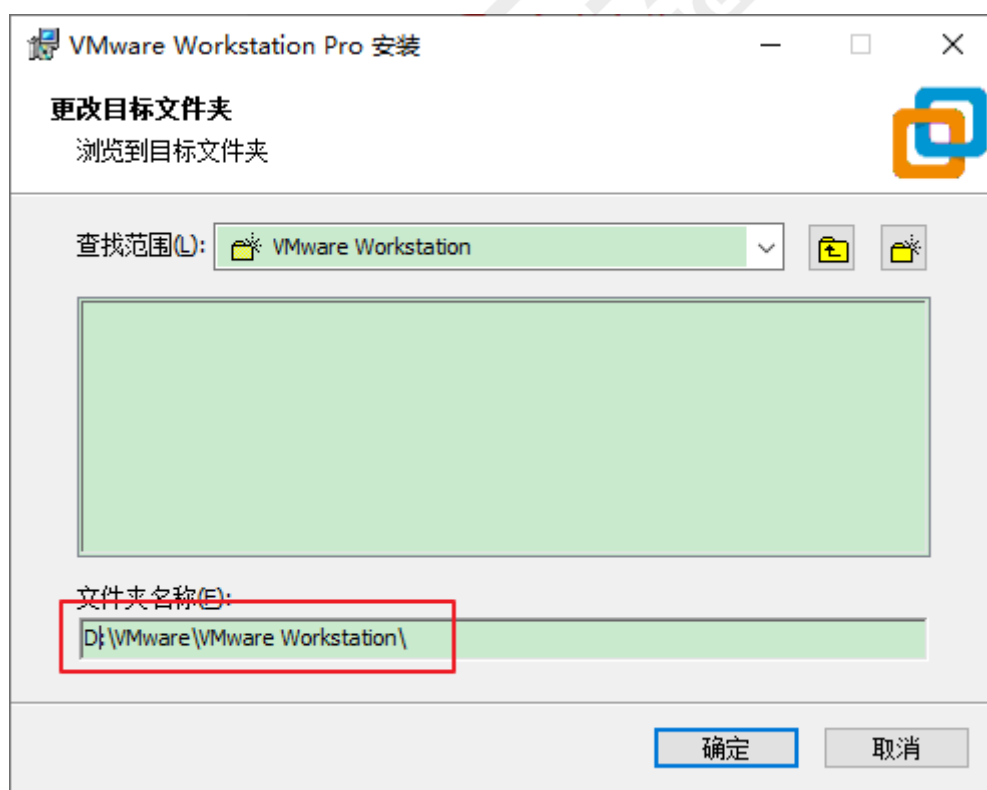
#### 1) VMware Workstation Pro 安装向导



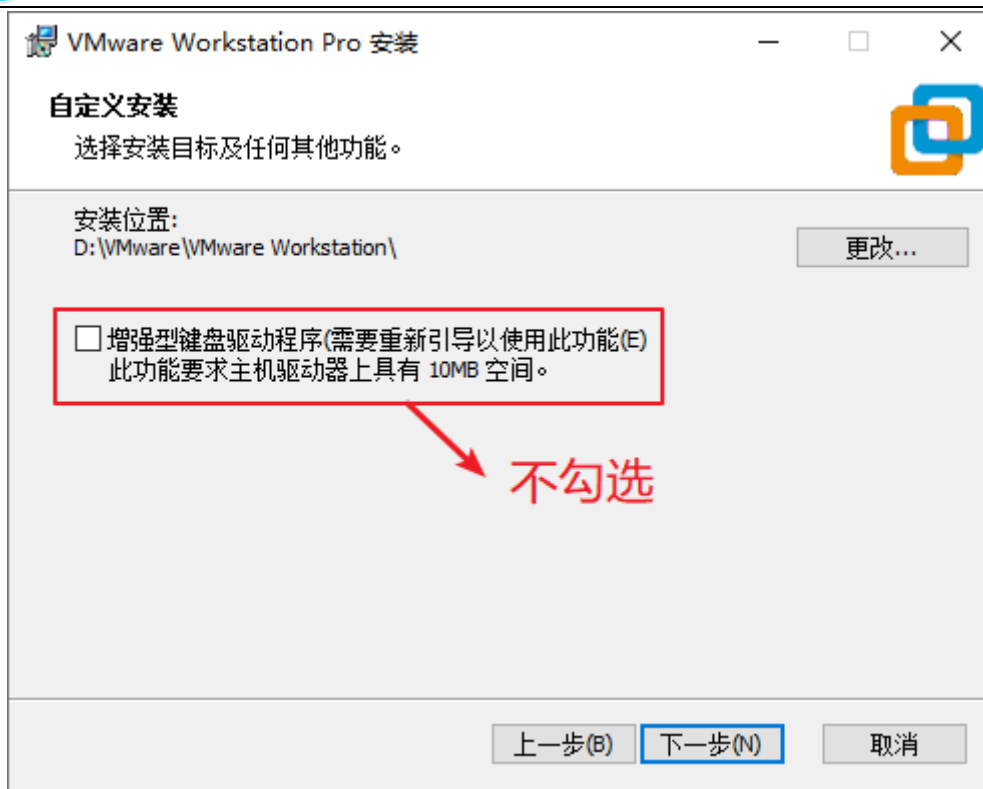
#### 2) VMware Workstation 安装的许可协议。



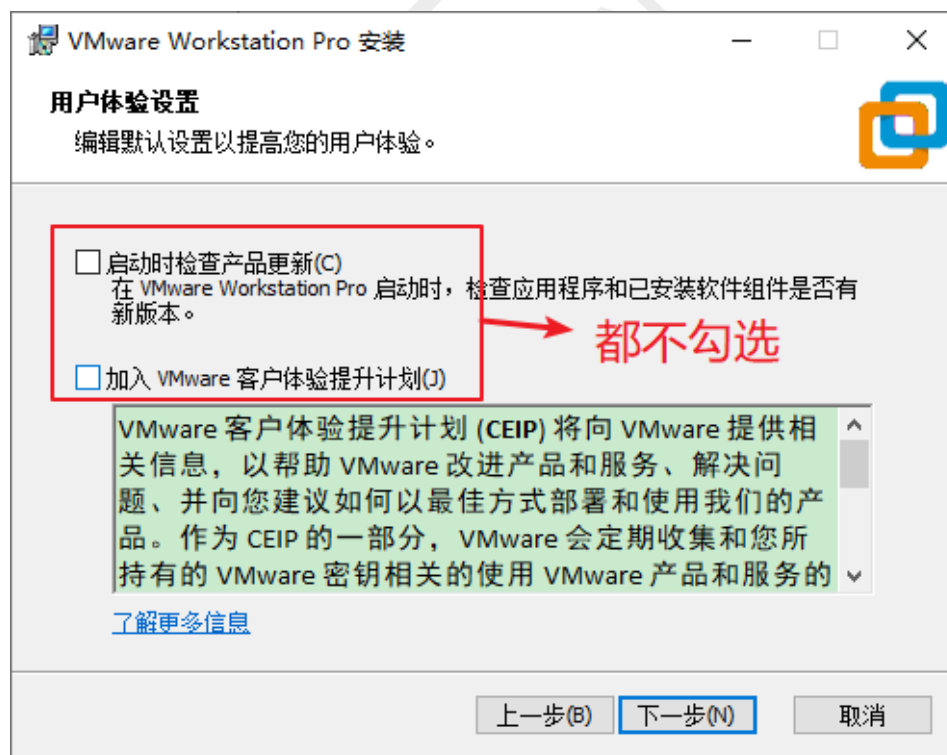
3) VMware Workstation 安装路径。



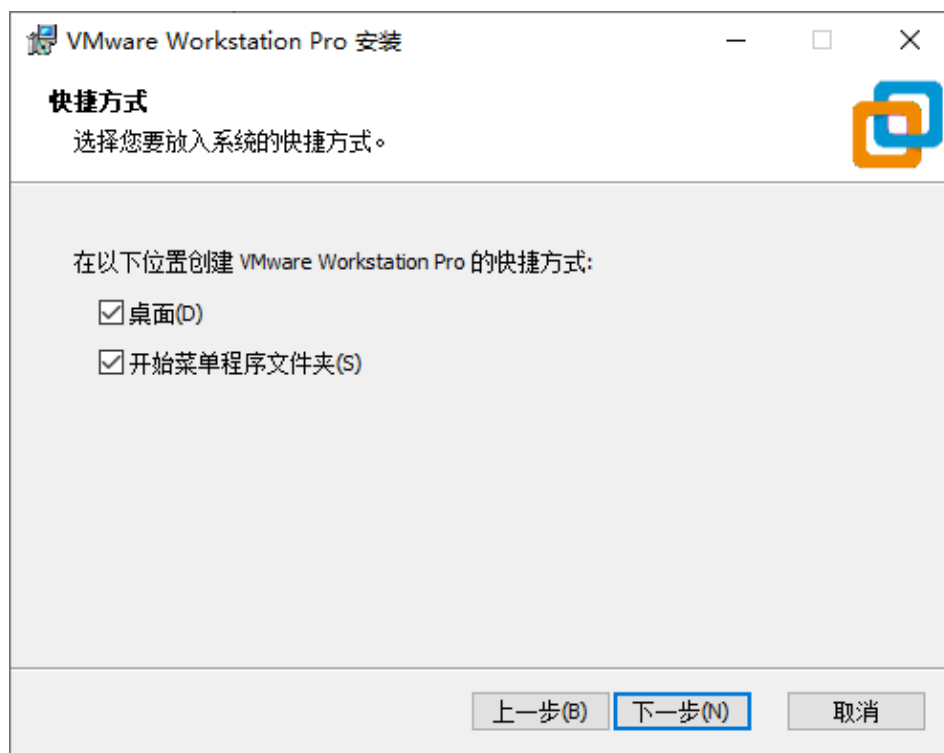
4) VMware Workstation 增强型键盘功能。



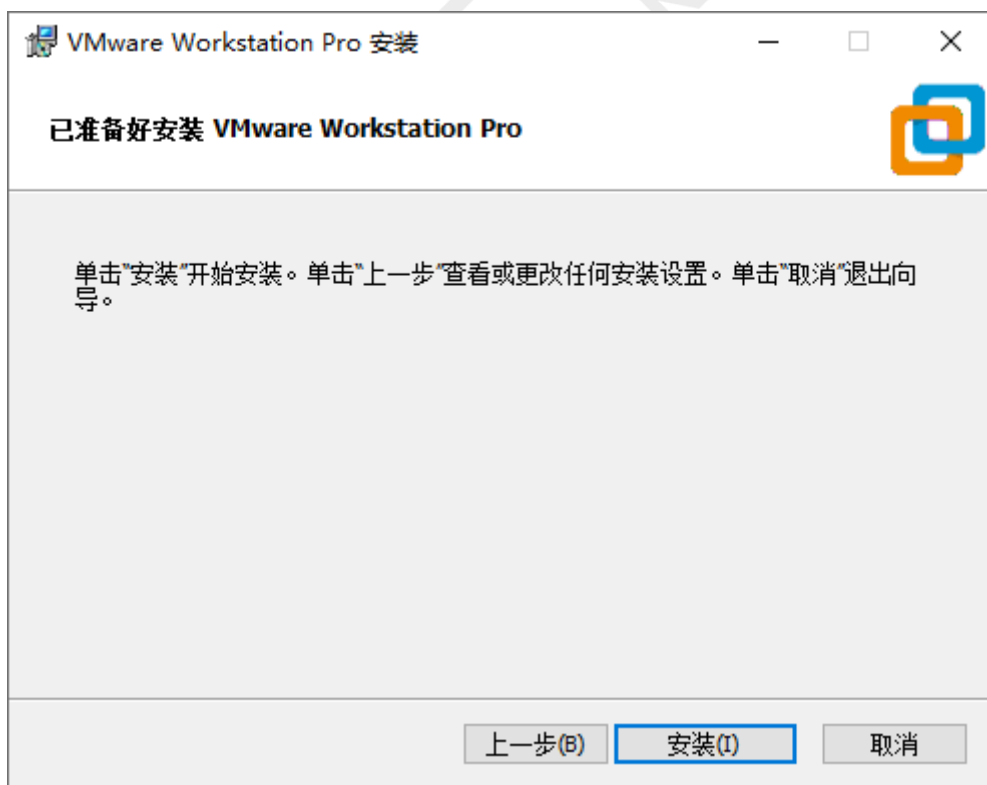
5) VMware Workstation 软件检查更新和帮助完善



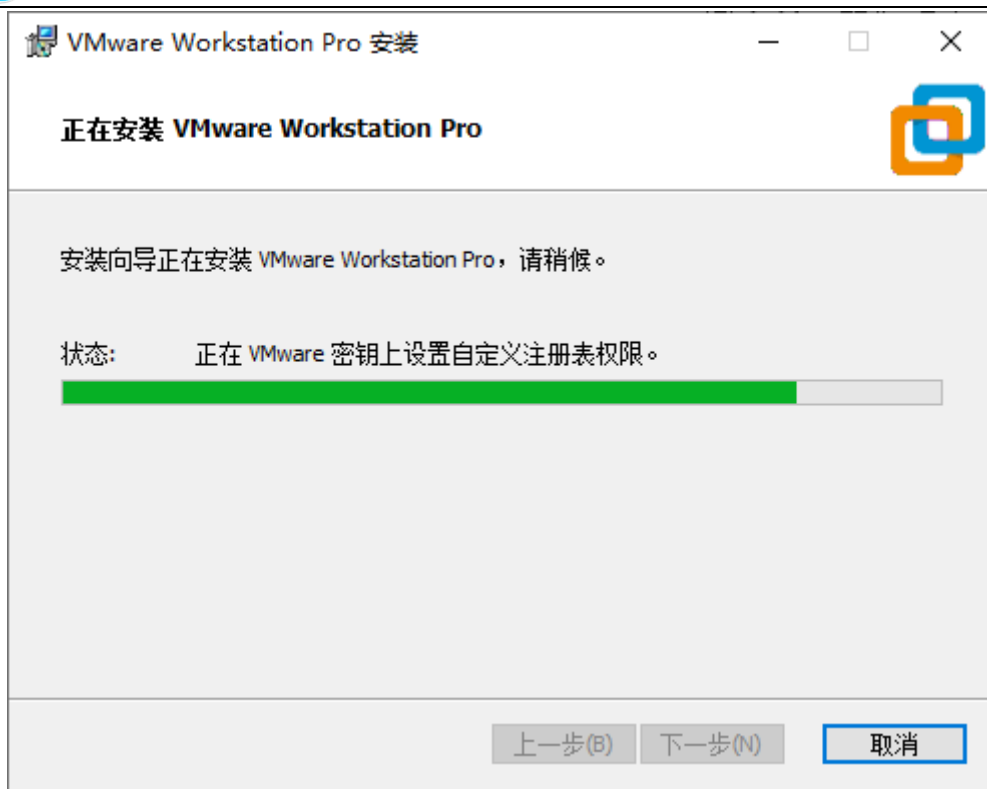
6) VMware Workstation 快捷方式。



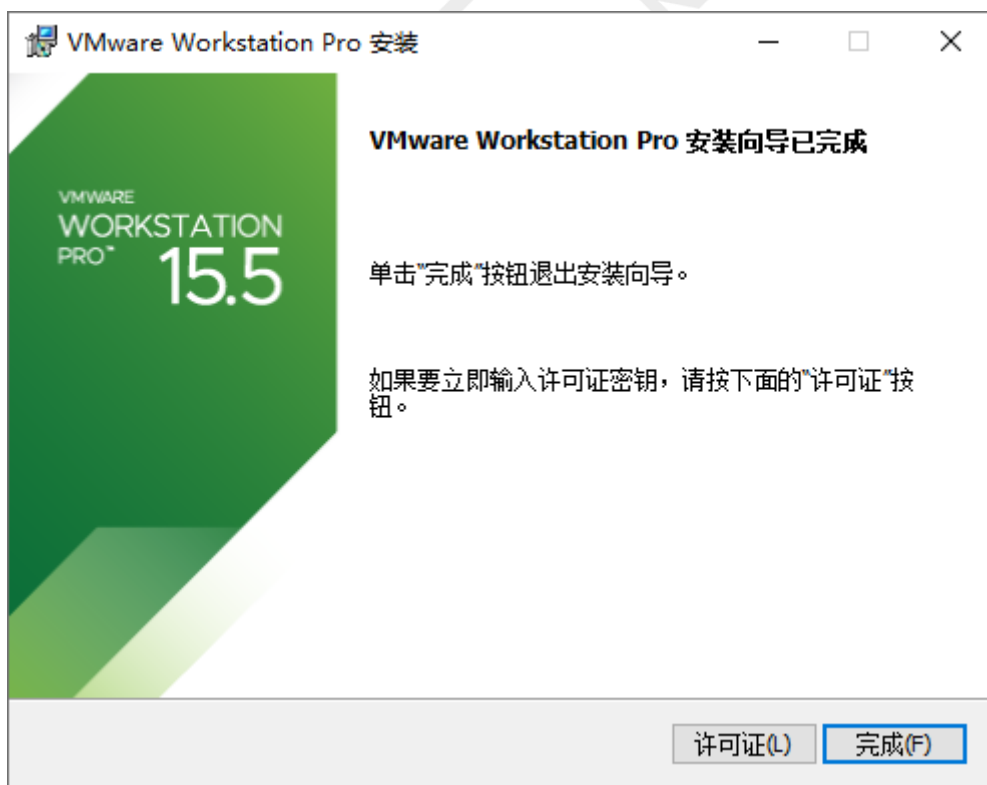
7) VMware Workstation 执行请求



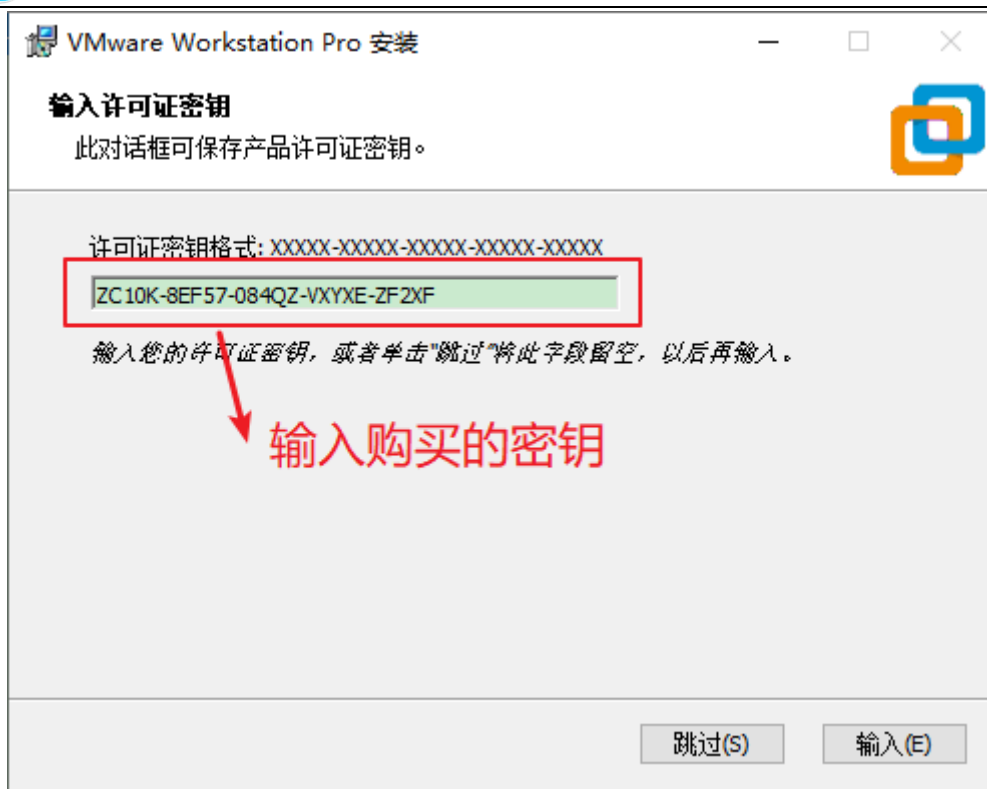
8) VMware Workstation 正在执行请求。



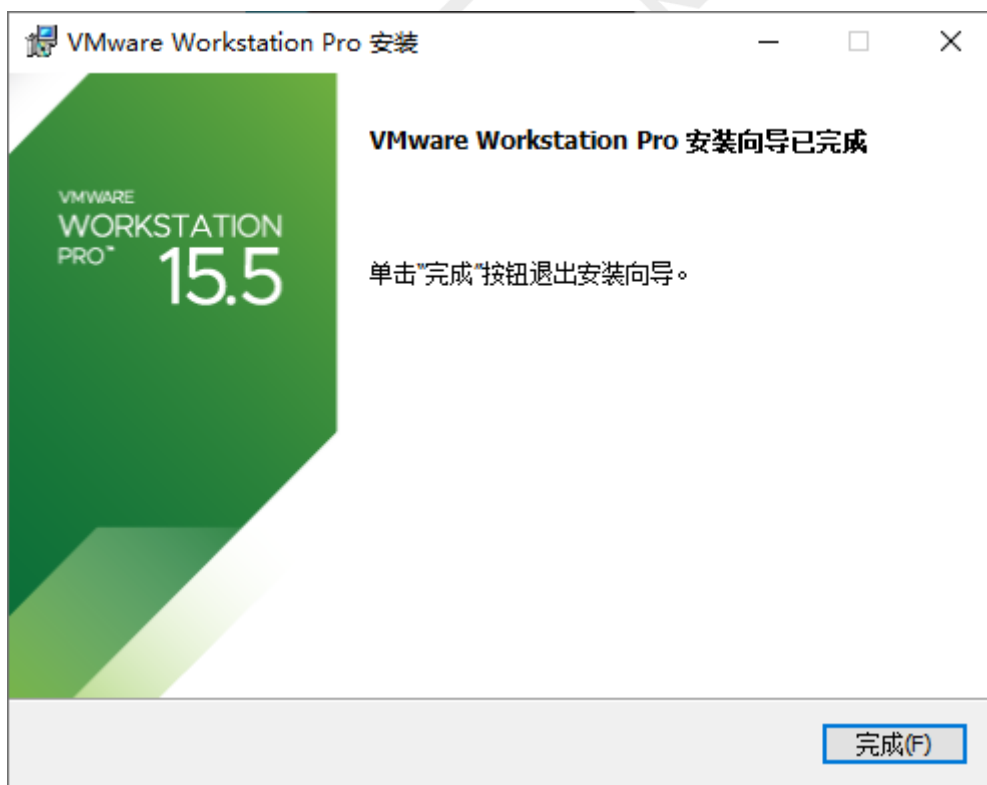
9) VMware Workstation 点击许可证。



10) VMware Workstation 输入许可证密钥。

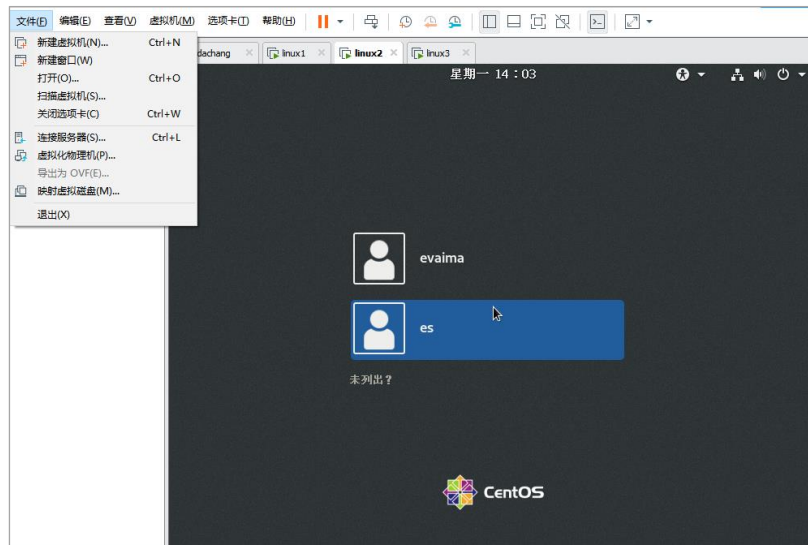


11) VMware Workstation 安装向导完成



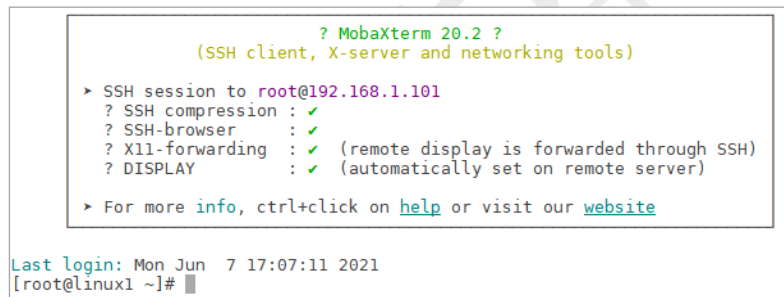
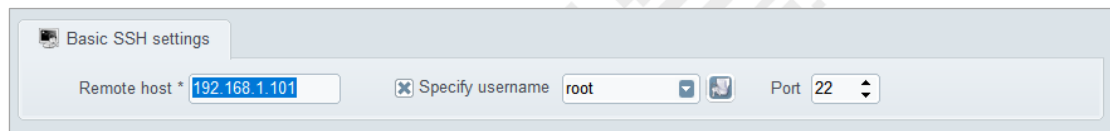
## 1.2 导入 Linux 系统

1) 点击【文件/打开】菜单，选择资料文件目录中的 3 个虚拟机，依次导入到 vmware 即可。



2) 导入成功后，启动虚拟机。使用账号 root，密码 root 登录

3) 使用任意远程工具访问虚拟机，能正常访问即可，资料文件夹中的 MobaXterm\_Installer\_v20.2 就是远程工具，也是讲师使用的工具。





## 第 2 章 基础环境测试

### 2.1 主机名

环境中的 3 台虚拟机已经配置好了对应的主机名，分别为

```
linux1
linux2
linux3
```

#### 2.2.1 查看主机名

```
[root@linux1 ~]# hostname
```

```
[root@linux1 ~]# hostname
linux1
```

#### 2.2.2 修改主机名

➤ 编辑 `/etc/hostname` 文件

```
[root@linux1 ~]# vim /etc/hostname
```

如果不想使用当前的主机名，修改为我们想要修改的主机名即可

➤ 修改 `/etc/hosts`

```
[root@linux1 ~]# vim /etc/hosts
```

添加内容

```
192.168.1.101 linux1
192.168.1.102 linux2
192.168.1.103 linux3
```

### 2.2 IP 地址

环境中的 3 台虚拟机已经配置好了对应的 IP 地址，分别为：

```
192.168.1.101 linux1
192.168.1.102 linux2
192.168.1.103 linux3
```

修改 IP 地址

```
[root@linux1 ~]# vim /etc/sysconfig/network-scripts/ifcfg-ens33
```

```
ONBOOT=yes
IPADDR=192.168.1.101
GATEWAY=192.168.1.2
DNS1=114.114.114.114
DNS2=8.8.8.8
PREFIX=24
```

### 2.3 防火墙

虚拟机中以及配置好了防火墙

# 关闭防火墙的自动启动

```
[root@linux1 ~]#systemctl disable firewalld
```

# 开启防火墙的自动启动

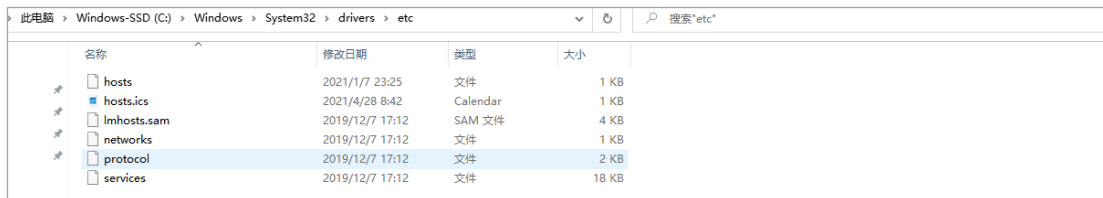
```
[root@linux1 ~]#systemctl enable firewalld
```

# 查看防火墙状态

```
[root@linux1~]#systemctl is-enabled firewalld
```

## 2.4 HOST

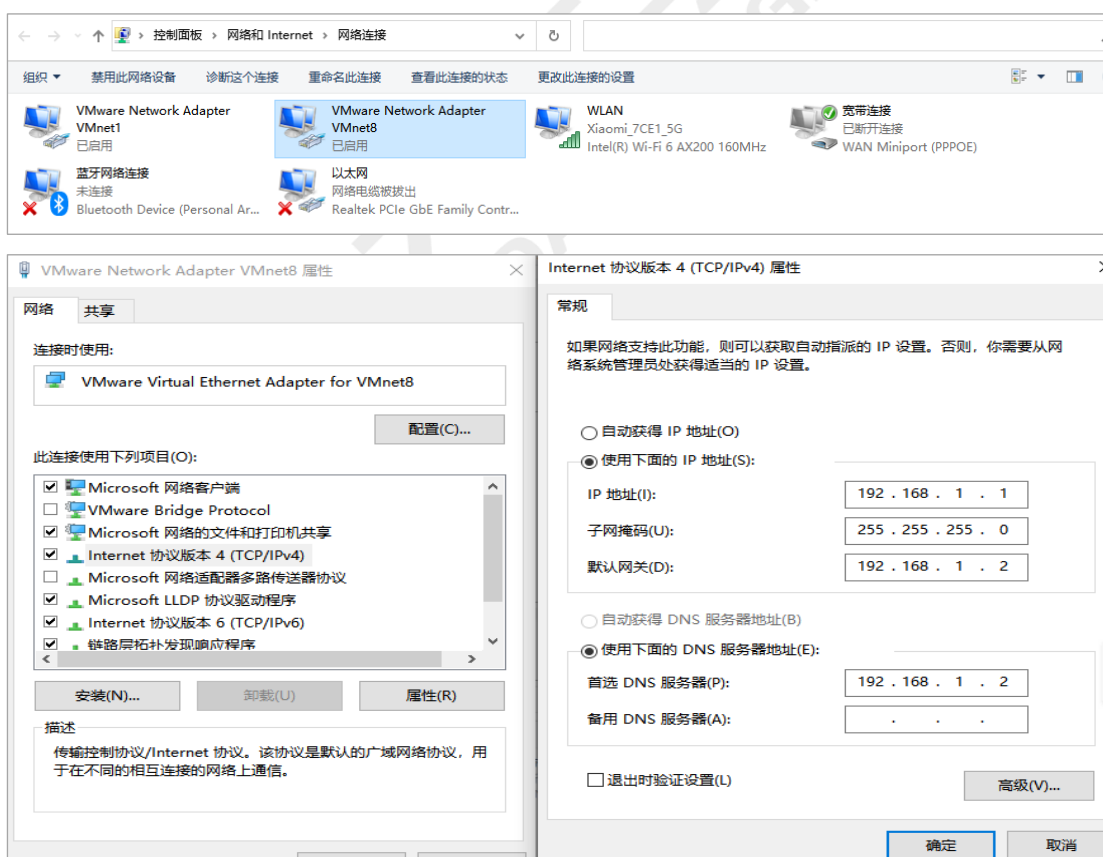
环境中的 3 台虚拟机已经配置好了对应的 IP 地址和主机名，为了能够在 Windows 环境中能正确访问，需要配置 Windows 的域名解析



修改 hosts 文件

```
# localhost name resolution is handled within DNS itself.
# 127.0.0.1 localhost
# ::1 localhost
192.168.1.100 linux0
192.168.1.101 linux1
192.168.1.102 linux2
192.168.1.103 linux3
192.168.1.104 linux4
```

修改本机 VMware 的 Ipv4 配置，配置如下图：



## 2.5 基础脚本

虚拟机中有很多软件和命令需要执行，每一次都要输入大量的指令，非常的麻烦，在掌握了基本的语法后，可以通过脚本文件直接执行相关的命令，为此，系统中封装了一些简单的脚

本文件

Name	Size (KB)	Last modified	Owner	Group	Access	Size (Bytes)
..	1	2021-02-04 18:59	root	root	-rwxrwxrwx	0
xcall	7	2021-02-04 18:59	root	root	-rwxrwxrwx	237
xinit	1	2021-02-04 18:59	root	root	-rwxrwxrwx	1814
xstart	1	2021-04-20 16:46	root	root	-rwxrwxrwx	1725
xstop	1	2021-02-04 18:59	root	root	-rwxrwxrwx	541
xsync						

**xcall** 用于执行全局系统指令: **xcall jps**

```
[root@linux1 ~]# xcall jps
-----linux1-----
5937 QuorumPeerMain
10197 Kafka
10792 Jps
-----linux2-----
6884 Kafka
4165 QuorumPeerMain
7385 Jps
-----linux3-----
3507 QuorumPeerMain
5764 Kafka
6261 Jps
```

**xstart, xstop** 用于执行全局软件启动, 关闭指令: **xstart zk,kafka**

```
[root@linux1 kafka]# xstart kafka
=====
start kafka
=====
root@linux1
root@linux2
root@linux3
root@linux1
=====
5937 QuorumPeerMain
10197 Kafka
10285 Jps
=====
root@linux2
=====
6077 Jps
6884 Kafka
4165 QuorumPeerMain
=====
root@linux3
=====
3507 QuorumPeerMain
5764 Kafka
5854 Jps
```

**xsync** 用于执行文件同步指令: **xsync kafka**

## 2.6 SSH 无密登录配置

虚拟机之间想要互相访问, 可以通过 **ssh** 命令进行操作, 但是需要输入密码, 比较麻烦, 所以咱们的 3 台虚拟机之间已经配置了无密登录

### 2.6.1 ssh 连接另外一台电脑

```
ssh linux2
```

### 2.6.2 无密登录

➤ 生成公钥和私钥:

```
ssh-keygen -t rsa
```

然后敲 (三个回车), 就会生成两个文件 **id\_rsa** (私钥)、**id\_rsa.pub** (公钥)

➤ 将公钥拷贝到要免密登录的目标机器上

```
ssh-copy-id linux1
ssh-copy-id linux2
ssh-copy-id linux3
```

注意:

还需要在 **linux2** 上采用 **root** 账号, 配置一下无密登录到 **linux1**、**linux2**、**linux3**

还需要在 linux3 上采用 root 账号，配置一下无密登录到 linux1、linux2、linux3



## 第3章 软件环境测试

在课程中，需要使用很多的大数据软件，为了更好的学习效果和教学体验，课程中需要的一些基础软件已经事先在虚拟机中安装好了。在使用前，需要对已经安装的软件进行测试。

### 3.1 Java

Java 软件是整个大数据软件的语言基础。软件在 linux 系统的的 `/opt/module/software` 中。

```
总用量 2077848
-rw-r--r--. 1 root root 67938106 2月 4 17:34 apache-flume-1.9.0-bin.tar.gz
-rw-r--r--. 1 root root 312850286 2月 4 17:35 apache-hive-3.1.2-bin.tar.gz
-rw-r--r--. 1 root root 9311744 2月 4 22:23 apache-zookeeper-3.5.7-bin.tar.gz
-rw-r--r--. 1 root root 319112561 3月 6 22:42 elasticsearch-7.8.0-linux-x86_64.tar.gz
-rw-r--r--. 1 root root 304712931 6月 5 06:50 flink-1.13.1-bin-scala_2.12.tgz
-rw-r--r--. 1 root root 338075860 2月 4 17:35 hadoop-3.1.3.tar.gz
-rw-r--r--. 1 root root 133559228 2月 4 17:36 hbase-2.0.5-bin.tar.gz
-rw-r--r--. 1 root root 195013152 2月 4 17:36 jdk-8u212-linux-x64.tar.gz
-rw-r--r--. 1 root root 70159813 2月 4 17:36 kafka_2.11-2.4.1.tgz
-rw-r--r--. 1 root root 71542357 5月 7 21:52 kafka_2.12-2.8.0.tgz
-rw-r--r--. 1 root root 80959303 6月 2 23:51 kafka-eagle-bin-2.0.5.tar.gz
-rwxr-xr-x. 2 root root 279 2月 4 17:34 mysql
-rw-r--r--. 1 root root 224453229 2月 5 08:47 spark-3.0.0-bin-hadoop3.2.tgz
root@linux1 software]#
```

软件安装目录在 `/opt/module/java8` 中

```
[root@linux1 module]# cd java8
[root@linux1 java8]# ll
总用量 25984
drwxr-xr-x. 2 10 143 4096 4月 2 2019 bin
-r--r--r--. 1 10 143 3244 4月 2 2019 COPYRIGHT
drwxr-xr-x. 3 10 143 132 4月 2 2019 include
-rw-r--r--. 1 10 143 5213825 3月 14 2019 javafx-src.zip
drwxr-xr-x. 5 10 143 185 4月 2 2019 jre
drwxr-xr-x. 5 10 143 245 4月 2 2019 lib
-r--r--r--. 1 10 143 44 4月 2 2019 LICENSE
drwxr-xr-x. 4 10 143 47 4月 2 2019 man
-r--r--r--. 1 10 143 159 4月 2 2019 README.html
-rw-r--r--. 1 10 143 424 4月 2 2019 release
-rw-r--r--. 1 10 143 21104792 4月 2 2019 src.zip
-rw-r--r--. 1 10 143 112748 3月 14 2019 THIRDPARTYLICENSEREADME-JAVAFX.txt
-r--r--r--. 1 10 143 149725 4月 2 2019 THIRDPARTYLICENSEREADME.txt
```

环境变量配置在 `/etc/profile.d/my_env.sh` 中

```
#JAVA_HOME
export JAVA_HOME=/opt/module/java8
export PATH=$PATH:$JAVA_HOME/bin
```

命令行中输入指令：`java -version`

```
[root@linux1 java8]# java -version
openjdk version "1.8.0_161"
OpenJDK Runtime Environment (build 1.8.0_161-b14)
OpenJDK 64-Bit Server VM (build 25.161-b14, mixed mode)
[root@linux1 java8]#
```

### 3.2 Hadoop

Hadoop 框架为大数据的基础框架，在当前环境中必须安装。软件在 linux 系统的的 `/opt/module/software` 中

```
[root@linux1 software]# ll
总用量 2077848
-rw-r--r--. 1 root root 67938106 2月 4 17:34 apache-flume-1.9.0-bin.tar.gz
-rw-r--r--. 1 root root 312850286 2月 4 17:35 apache-hive-3.1.2-bin.tar.gz
-rw-r--r--. 1 root root 9311744 2月 4 22:23 apache-zookeeper-3.5.7-bin.tar.gz
-rw-r--r--. 1 root root 319112561 3月 6 22:42 elasticsearch-7.8.0-linux-x86_64.tar.gz
-rw-r--r--. 1 root root 304712931 6月 5 06:50 flink-1.13.1-bin-scala_2.12.tgz
-rw-r--r--. 1 root root 338075860 2月 4 17:35 hadoop-3.1.3.tar.gz
-rw-r--r--. 1 root root 133559228 2月 4 17:36 hbase-2.0.5-bin.tar.gz
-rw-r--r--. 1 root root 195013152 2月 4 17:36 jdk-8u212-linux-x64.tar.gz
-rw-r--r--. 1 root root 70159813 2月 4 17:36 kafka_2.11-2.4.1.tgz
-rw-r--r--. 1 root root 71542357 5月 7 21:52 kafka_2.12-2.8.0.tgz
-rw-r--r--. 1 root root 80959303 6月 2 23:51 kafka-eagle-bin-2.0.5.tar.gz
drwxr-xr-x. 2 root root 279 2月 4 17:34 mysql
-rw-r--r--. 1 root root 224453229 2月 5 08:47 spark-3.0.0-bin-hadoop3.2.tgz
```

软件安装目录在 `/opt/module/hadoop3` 中

```
[root@linux1 module]# cd hadoop3/
[root@linux1 hadoop3]# ll
总用量 180
drwxr-xr-x. 2 evaima evaima 183 9月 12 2019 bin
drwxr-xr-x. 4 root root 37 2月 4 19:28 data
drwxr-xr-x. 3 evaima evaima 20 9月 12 2019 etc
drwxr-xr-x. 2 evaima evaima 106 9月 12 2019 include
drwxr-xr-x. 3 evaima evaima 20 9月 12 2019 lib
drwxr-xr-x. 4 evaima evaima 288 9月 12 2019 libexec
-rw-rw-r--. 1 evaima evaima 147145 9月 4 2019 LICENSE.txt
drwxr-xr-x. 3 root root 4096 6月 6 17:59 logs
-rw-rw-r--. 1 evaima evaima 21867 9月 4 2019 NOTICE.txt
-rw-rw-r--. 1 evaima evaima 1366 9月 4 2019 README.txt
drwxr-xr-x. 3 evaima evaima 4096 9月 12 2019 sbin
drwxr-xr-x. 4 evaima evaima 31 9月 12 2019 share
```

环境变量配置在 `/etc/profile.d/my_env.sh` 中

```
#HADOOP_HOME
export HADOOP_HOME=/opt/module/hadoop3
export PATH=$PATH:$HADOOP_HOME/bin
export PATH=$PATH:$HADOOP_HOME/sbin
```

使用群起脚本命令启动 hadoop 软件：`xstart hadoop`

```
=====
3345 JobHistoryServer      root@linux1      =====
2898 DataNode
3479 NodeManager
3671 Jps
2716 NameNode
=====
2819 NodeManager          root@linux2      =====
3222 Jps
2375 DataNode
2655 ResourceManager
=====
2368 DataNode              root@linux3      =====
2849 Jps
2503 SecondaryNameNode
2647 NodeManager
```

在浏览器中输入网址查看 UI 页面：<http://linux1:9870>

**Hadoop** Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities ▾

### Overview 'linux1:8020' (active)

Started:	Mon Jun 07 16:35:57 +0800 2021
Version:	3.1.3, rba631c436b806728f8ec2f54ab1e289526c90579
Compiled:	Thu Sep 12 10:47:00 +0800 2019 by ztang from branch-3.1.3
Cluster ID:	CID-b05f6537-e421-4b34-a56c-ff1cc03a7279
Block Pool ID:	BP-658550041-192.168.1.101-1612438342047

使用群起脚本命令关闭 hadoop 软件：`xstop hadoop`

### 3.3 Hive

Hive 软件为基于 Hadoop 软件的 SQL 操作工具软件，所以使用前，必须保证 hadoop 可以正确启动

```
=====
3345 JobHistoryServer      root@linux1      =====
2898 DataNode
3479 NodeManager
3671 Jps
2716 NameNode
=====
2819 NodeManager          root@linux2      =====
3222 Jps
2375 DataNode
2655 ResourceManager
=====
2368 DataNode              root@linux3      =====
2849 Jps
2503 SecondaryNameNode
2647 NodeManager
```

软件在 linux 系统的 `/opt/module/software` 中

```
[root@linux1 software]# ll
总用量 2077848
-rw-r--r--. 1 root root 67938106 2月 4 17:34 apache-flume-1.9.0-bin.tar.gz
-rw-r--r--. 1 root root 312850286 2月 4 17:35 apache-hive-3.1.2-bin.tar.gz
-rw-r--r--. 1 root root 9311744 2月 4 22:23 apache-zookeeper-3.5.7-bin.tar.gz
-rw-r--r--. 1 root root 319112561 2月 6 22:42 elasticsearch-7.8.0-linux-x86_64.tar.gz
-rw-r--r--. 1 root root 304712931 6月 5 06:50 flink-1.13.1-bin-scala_2.12.tgz
-rw-r--r--. 1 root root 338075860 2月 4 17:35 hadoop-3.1.3.tar.gz
-rw-r--r--. 1 root root 133559228 2月 4 17:36 hbase-2.0.5-bin.tar.gz
-rw-r--r--. 1 root root 195013152 2月 4 17:36 jdk-8u212-linux-x64.tar.gz
-rw-r--r--. 1 root root 70159813 2月 4 17:36 kafka_2.11-2.4.1.tgz
-rw-r--r--. 1 root root 71542357 5月 7 21:52 kafka_2.12-2.8.0.tgz
-rw-r--r--. 1 root root 80959303 6月 2 23:51 kafka-eagle-bin-2.0.5.tar.gz
drwxr-xr-x. 2 root root 279 2月 4 17:34 mysql
-rw-r--r--. 1 root root 224453229 2月 5 08:47 spark-3.0.0-bin-hadoop3.2.tgz
```

软件安装目录在 `/opt/module/hive3` 中

```
[root@linux1 module]# cd hive3
[root@linux1 hive3]# ll
总用量 56
drwxr-xr-x. 3 root root 180 2月 4 22:09 bin
drwxr-xr-x. 2 root root 4096 2月 4 21:48 conf
-rw-r--r--. 1 root root 115 2月 22 16:59 emp.txt
drwxr-xr-x. 4 root root 34 2月 4 21:41 examples
drwxr-xr-x. 7 root root 68 2月 4 21:41 hcatalog
drwxr-xr-x. 2 root root 44 2月 4 21:41 jdbc
drwxr-xr-x. 4 root root 12288 2月 4 21:47 lib
-rw-r--r--. 1 root root 20798 8月 23 2019 LICENSE
drwxr-xr-x. 2 root root 50 2月 4 22:09 logs
-rw-r--r--. 1 root root 230 8月 23 2019 NOTICE
-rw-r--r--. 1 root root 2469 8月 23 2019 RELEASE_NOTES.txt
drwxr-xr-x. 4 root root 35 2月 4 21:41 scripts
```

环境变量配置在 `/etc/profile.d/my_env.sh` 中

```
#HIVE_HOME
export HIVE_HOME=/opt/module/hive3
export PATH=$PATH:$HIVE_HOME/bin
```

使用命令行启动 Hive: `hiverservices.sh start`

```
[root@linux1 hive3]# hiverservices.sh start
[root@linux1 hive3]# jps
3345 JobHistoryServer
4289 Jps
2898 DataNode
3794 RunJar
3479 NodeManager
4121 MapreduceDependencyClasspathTool
2716 NameNode
```

启动后，执行命令: `bin/hive` 进入 Hive 中查询数据 : `show tables;`

```
[root@linux1 hive3]# bin/hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/opt/module/hbase/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/module/hadoop3/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
SLF4J: Class path contains multiple SLF4J bindings
SLF4J: Found binding in [jar:file:/opt/module/hive3/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/module/hadoop3/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Hive Session ID = 599c63c6-3864-4b57-ba55-beb08b4131de

Logging initialized using configuration in jar:file:/opt/module/hive3/lib/hive-common-3.1.2.jar!/hive-log4j2.properties Async: true
Hive Session ID = 0fedd5ad-52bd-497d-9e4d-63883ae98a12
Hive-on-MR is deprecated in Hive 2 and may not be available in the future versions. Consider using a different execution engine (i.e. spark, tez) or using Hive 1.X releases.
hive>
```

使用命令行关闭 Hive: `hiverservices.sh stop`

## 3.4 Zookeeper

软件在 linux 系统的 `/opt/module/software` 中

```
[root@linux1 software]# ll
总用量 2077848
-rw-r--r--. 1 root root 67938106 2月 4 17:34 apache-flume-1.9.0-bin.tar.gz
-rw-r--r--. 1 root root 312850286 2月 4 17:35 apache-hive-3.1.2-bin.tar.gz
-rw-r--r--. 1 root root 9311744 2月 4 22:23 apache-zookeeper-3.5.7-bin.tar.gz
-rw-r--r--. 1 root root 319112561 2月 6 22:42 elasticsearch-7.8.0-linux-x86_64.tar.gz
-rw-r--r--. 1 root root 304712931 6月 5 06:50 flink-1.13.1-bin-scala_2.12.tgz
-rw-r--r--. 1 root root 338075860 2月 4 17:35 hadoop-3.1.3.tar.gz
-rw-r--r--. 1 root root 133559228 2月 4 17:36 hbase-2.0.5-bin.tar.gz
-rw-r--r--. 1 root root 195013152 2月 4 17:36 jdk-8u212-linux-x64.tar.gz
-rw-r--r--. 1 root root 70159813 2月 4 17:36 kafka_2.11-2.4.1.tgz
-rw-r--r--. 1 root root 71542357 5月 7 21:52 kafka_2.12-2.8.0.tgz
-rw-r--r--. 1 root root 80959303 6月 2 23:51 kafka-eagle-bin-2.0.5.tar.gz
drwxr-xr-x. 2 root root 279 2月 4 17:34 mysql
-rw-r--r--. 1 root root 224453229 2月 5 08:47 spark-3.0.0-bin-hadoop3.2.tgz
```

软件安装目录在 `/opt/module/zookeeper` 中

```
[root@linux1 zookeeper]# ll
总用量 32
drwxr-xr-x. 2 502 games 232 2月 10 2020 bin
drwxr-xr-x. 2 502 games 98 6月 6 18:18 conf
drwxr-xr-x. 5 502 games 4096 2月 10 2020 docs
drwxr-xr-x. 2 root root 4096 2月 4 22:24 lib
-rw-r--r--. 1 502 games 11358 9月 13 2018 LICENSE.txt
drwxr-xr-x. 2 root root 46 2月 4 22:30 logs
-rw-r--r--. 1 root root 0 2月 4 22:25 myid
-rw-r--r--. 1 502 games 432 2月 10 2020 NOTICE.txt
-rw-r--r--. 1 502 games 1560 2月 7 2020 README.md
-rw-r--r--. 1 502 games 1347 2月 7 2020 README_packaging.txt
drwxr-xr-x. 3 root root 35 6月 7 00:42 zkData
```

使用群起脚本命令启动 zookeeper 软件：**xstart zk**

```
[root@linux1 zookeeper]# xstart zk
===== start zk =====
===== root@linux1 =====
ZooKeeper JMX enabled by default
Using config: /opt/module/zookeeper/bin/../conf/zoo.cfg
Starting zookeeper ... STARTED
===== root@linux2 =====
ZooKeeper JMX enabled by default
Using config: /opt/module/zookeeper/bin/../conf/zoo.cfg
Starting zookeeper ... STARTED
===== root@linux3 =====
ZooKeeper JMX enabled by default
Using config: /opt/module/zookeeper/bin/../conf/zoo.cfg
Starting zookeeper ... STARTED
===== root@linux1 =====
5937 QuorumPeerMain
6039 Jps
===== root@linux2 =====
4165 QuorumPeerMain
4266 Jps
===== root@linux3 =====
3507 QuorumPeerMain
3607 Jps
```

使用群起脚本命令关闭 zookeeper 软件：**xstop zk**

### 3.5 HBase

HBase 软件为基于 Hadoop 软件的海量数据存储软件，所以使用前，必须保证 hadoop 可以正确启动。并且 HBase 软件基于 zookeeper 进行集群管理，所以也必须保证 zookeeper 正常启动。

软件在 linux 系统的 **/opt/module/software** 中

```
[root@linux1 software]# ll
总用量 2077848
-rw-r--r--. 1 root root 67938106 2月 4 17:34 apache-flume-1.9.0-bin.tar.gz
-rw-r--r--. 1 root root 312850286 2月 4 17:35 apache-hive-3.1.2-bin.tar.gz
-rw-r--r--. 1 root root 9311744 2月 4 22:23 apache-zookeeper-3.5.7-bin.tar.gz
-rw-r--r--. 1 root root 319112561 3月 6 22:42 elasticsearch-7.8.0-linux-x86_64.tar.gz
-rw-r--r--. 1 root root 304712931 6月 5 06:50 flink-1.13.1-bin-scala_2.12.tgz
-rw-r--r--. 1 root root 338075860 2月 4 17:35 hadoop-3.1.3.tar.gz
-rw-r--r--. 1 root root 133559228 2月 4 17:36 hbase-2.0.5-bin.tar.gz
-rw-r--r--. 1 root root 195013152 2月 4 17:36 jdk-8u212-linux-x64.tar.gz
-rw-r--r--. 1 root root 70159813 2月 4 17:36 kafka_2.11-2.4.1.tgz
-rw-r--r--. 1 root root 71542357 5月 7 21:52 kafka_2.12-2.8.0.tgz
-rw-r--r--. 1 root root 80959303 6月 2 23:51 kafka-eagle-bin-2.0.5.tar.gz
drwxr-xr-x. 2 root root 279 2月 4 17:34 mysql
-rw-r--r--. 1 root root 224453229 2月 5 08:47 spark-3.0.0-bin-hadoop3.2.tgz
```

软件安装目录在 **/opt/module/hbase** 中

```
[root@linux1 hbase]# ll
总用量 1936
drwxr-xr-x. 4 evaima evaima 4096 3月 18 2019 bin
-rw-r--r--. 1 evaima evaima 846815 3月 18 2019 CHANGES.md
drwxr-xr-x. 2 evaima evaima 200 2月 20 16:12 conf
drwxr-xr-x. 7 evaima evaima 80 6月 18 2020 hbase-webapps
-rw-r--r--. 1 evaima evaima 262 3月 18 2019 LEGAL
drwxr-xr-x. 4 root root 8192 2月 20 16:02 lib
-rw-r--r--. 1 evaima evaima 11358 6月 18 2020 LICENSE.txt
drwxr-xr-x. 2 root root 4096 2月 22 08:36 logs
-rw-r--r--. 1 evaima evaima 616712 6月 18 2020 NOTICE.txt
-rw-r--r--. 1 evaima evaima 1477 3月 18 2019 README.txt
-rw-r--r--. 1 evaima evaima 470589 3月 18 2019 RELEASNOTES.md
-rw-r--r--. 1 root root 22 2月 22 16:33 splits.txt
```

环境变量配置在 **/etc/profile.d/my\_env.sh** 中

```
#HBASE_HOME
export HBASE_HOME=/opt/module/hbase
export PATH=$PATH:$HBASE_HOME/bin
```

使用群起脚本命令启动 hbase 软件：**xstart hbase**



```
[root@linux1 hbase]# xstart hbase
=====
start hbase
=====
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/opt/module/hbase/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/module/hadoop3/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
running master, logging to /opt/module/hbase/logs/hbase-root-master-linux1.out
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/opt/module/hbase/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/module/hadoop3/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
linux1: running regionserver, logging to /opt/module/hbase/logs/hbase-root-regionserver-linux1.out
linux3: running regionserver, logging to /opt/module/hbase/bin/../logs/hbase-root-regionserver-linux3.out
linux2: running regionserver, logging to /opt/module/hbase/bin/../logs/hbase-root-regionserver-linux2.out
linux3: running master, logging to /opt/module/hbase/bin/../logs/hbase-root-master-linux3.out
```

```
=====
root@linux1
=====
5937 QuorumPeerMain
6498 DataNode
6309 NameNode
8039 Jps
7066 NodeManager
7610 HMaster
7835 HRegionServer
6926 JobHistoryServer
=====
root@linux2
=====
4165 QuorumPeerMain
4662 ResourceManager
4393 DataNode
5562 Jps
4844 NodeManager
5470 HRegionServer
=====
root@linux3
=====
3507 QuorumPeerMain
4362 HRegionServer
4474 HMaster
3996 NodeManager
3727 DataNode
3871 SecondaryNameNode
4575 Jps
```

在浏览器中输入网址查看 UI 页面：<http://linux1:16010>

<div> <div>APACHE HBASE</div> <div> <a href="#">Home</a> <a href="#">Table Details</a> <a href="#">Procedures &amp; Locks</a> <a href="#">Process Metrics</a> <a href="#">Local Logs</a> <a href="#">Log Level</a> <a href="#">Debug Dump</a> <a href="#">Metrics Dump</a> <a href="#">HBase Configuration</a> </div> </div>						
Master linux1						
Region Servers						
<div> <div>Base Stats</div> <div>Memory</div> <div>Requests</div> <div>Storefiles</div> <div>Compactions</div> </div>						
ServerName	Start time	Last contact	Version	Requests Per Second	Num. Regions	
linux1,16020,1623056530540	Mon Jun 07 17:02:10 CST 2021	1 s	2.0.5	0	16	
linux2,16020,1623056530104	Mon Jun 07 17:02:10 CST 2021	1 s	2.0.5	0	15	
linux3,16020,1623056530054	Mon Jun 07 17:02:10 CST 2021	1 s	2.0.5	0	16	
Total:3				0	47	

使用群起脚本命令关闭 hbase 软件：**xstop hbase**

## 3.6 MySQL

mysql 已经作为系统服务，系统启动后可以直接访问。

linux1的配置

说明

名称: linux1

连接

主机: 192.168.1.101

端口: 3306

连接类型: 内建

登录信息

用户: root

密码: 123123

数据库:

帮助

关闭

## 3.7 Kafka

Kafka 软件基于 zookeeper 进行集群管理，所以也必须保证 zookeeper 正常启动。

```
=====
5937 QuorumPeerMain      root@linux1      =====
9724 Jps
=====
6418 Jps                  root@linux2      =====
4165 QuorumPeerMain
=====
3507 QuorumPeerMain      root@linux3      =====
5294 Jps
```

软件在 linux 系统的的/opt/module/software 中

```
[root@linux1 software]# ll
总用量 2077848
-rw-r--r--. 1 root root 67938106 2月 4 17:34 apache-flume-1.9.0-bin.tar.gz
-rw-r--r--. 1 root root 312850286 2月 4 17:35 apache-hive-3.1.2-bin.tar.gz
-rw-r--r--. 1 root root 9311744 2月 4 22:23 apache-zookeeper-3.5.7-bin.tar.gz
-rw-r--r--. 1 root root 319112561 3月 6 22:42 elasticsearch-7.8.0-linux-x86_64.tar.gz
-rw-r--r--. 1 root root 304712931 6月 5 06:50 flink-1.13.1-bin-scala_2.12.tgz
-rw-r--r--. 1 root root 338075860 2月 4 17:35 hadoop-3.1.3.tar.gz
-rw-r--r--. 1 root root 133559228 2月 4 17:36 hbase-2.0.5-bin.tar.gz
-rw-r--r--. 1 root root 195013152 2月 4 17:36 jdk-8u212-linux-x64.tar.gz
-rw-r--r--. 1 root root 70159813 2月 4 17:36 kafka-2.11-2.4.1.tgz
-rw-r--r--. 1 root root 71542357 5月 7 21:52 kafka-2.12-2.8.0.tgz
-rw-r--r--. 1 root root 80959303 6月 2 23:51 kafka-eagle-bin-2.0.5.tar.gz
drwxr-xr-x. 2 root root 279 2月 4 17:34 mysql
-rw-r--r--. 1 root root 224453229 2月 5 08:47 spark-3.0.0-bin-hadoop3.2.tgz
```

软件安装目录在/opt/module/kafka 中

```
[root@linux1 module]# cd kafka
[root@linux1 kafka]# ll
总用量 44
drwxr-xr-x. 3 root root 4096 4月 14 22:31 bin
drwxr-xr-x. 3 root root 4096 4月 14 22:31 config
drwxr-xr-x. 2 root root 8192 6月 6 18:33 libs
-rw-r--r--. 1 root root 14515 4月 14 22:28 LICENSE
drwxr-xr-x. 2 root root 236 4月 14 22:31 licenses
drwxr-xr-x. 20 root root 4096 6月 7 00:42 logs
-rw-r--r--. 1 root root 953 4月 14 22:28 NOTICE
drwxr-xr-x. 2 root root 44 4月 14 22:31 site-docs
```

环境变量配置在/etc/profile.d/my\_env.sh 中

```
#KAFKA_HOME
export KAFKA_HOME=/opt/module/kafka
export PATH=$PATH:$KAFKA_HOME/bin
```

使用群起脚本命令启动 kafka 软件：**xstart kafka**

```
[root@linux1 kafka]# xstart kafka
=====
start kafka
=====
root@linux1
=====
root@linux2
=====
root@linux3
=====
root@linux1
=====
5937 QuorumPeerMain
10197 Kafka
10285 Jps
=====
root@linux2
=====
6077 Jps
6884 Kafka
4165 QuorumPeerMain
=====
root@linux3
=====
3507 QuorumPeerMain
5764 Kafka
5854 Jps
```

使用群起脚本命令关闭 kafka 软件：**xstop kafka**