

步骤 1. 创建 VMmachines 文件夹

创建文件夹 VMmachines, 例如可以在 E 盘下创建该文件夹, 如下图 1 所示。



图 1

步骤 2. 拷贝 3 个虚拟机压缩文件

将 node01.zip、node02.zip、node03.zip 拷贝至 VMmachines 文件夹中, 如下图 2 所示。

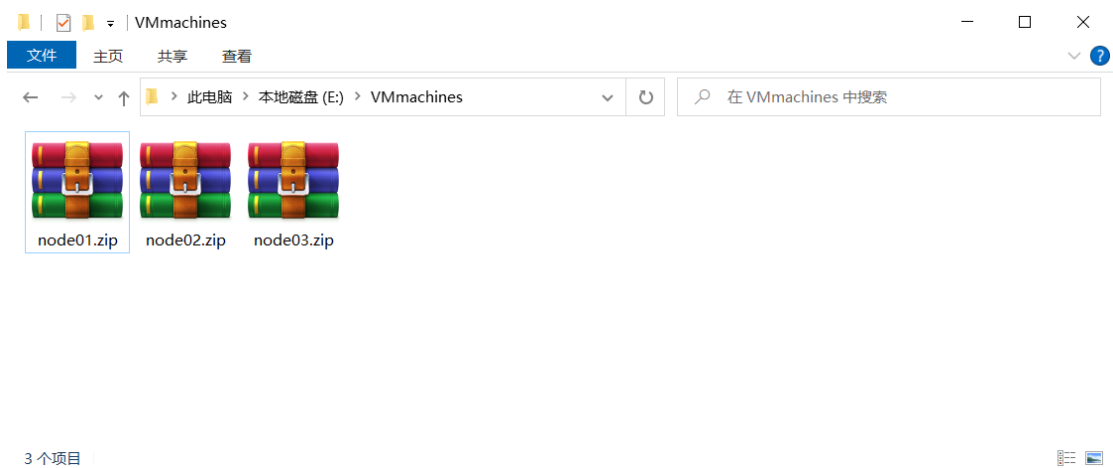


图 2

步骤 3. 分别解压 3 个虚拟机压缩文件

右键点击 node01.zip，选择“解压到当前文件夹”，如图 3 所示，解压后如图 4 所示。

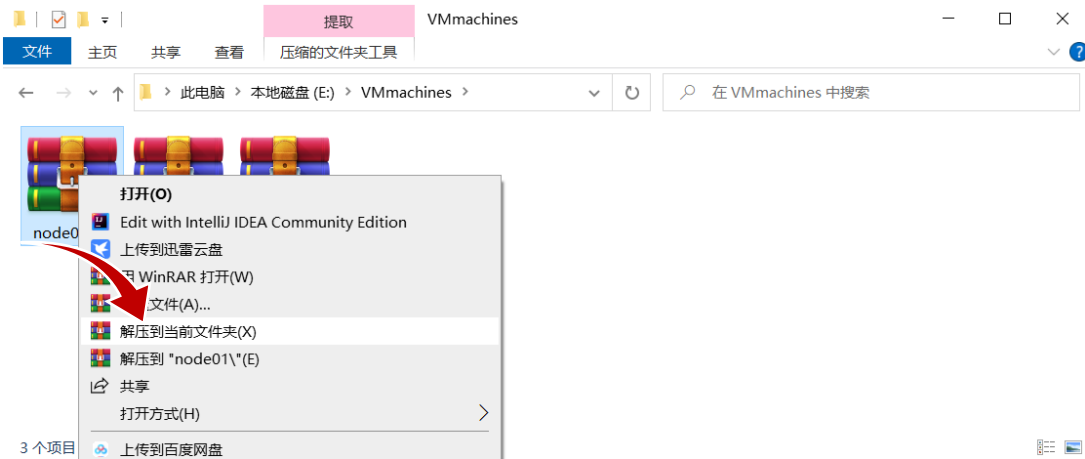


图 3

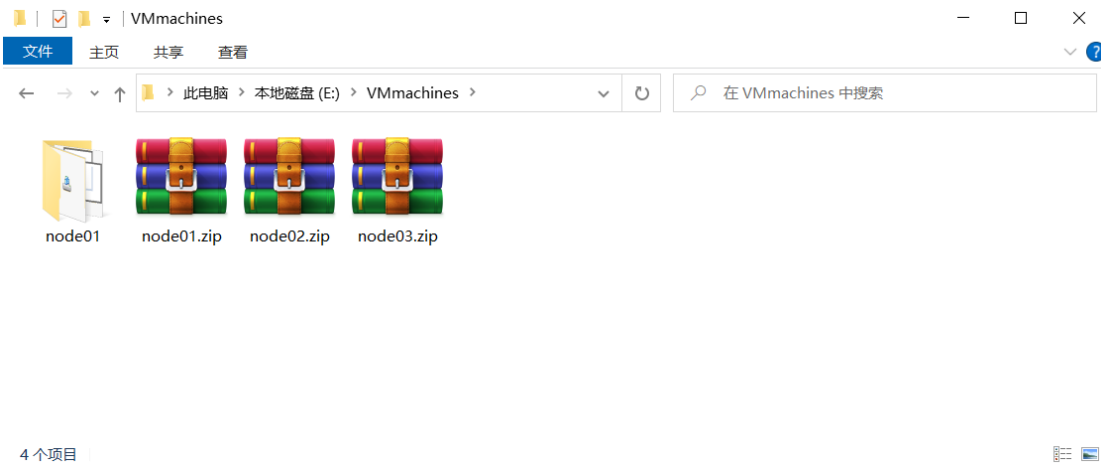


图 4

按照上述操作，分别对 node02.zip、node03.zip 进行解压，解压后如图 5 所示。

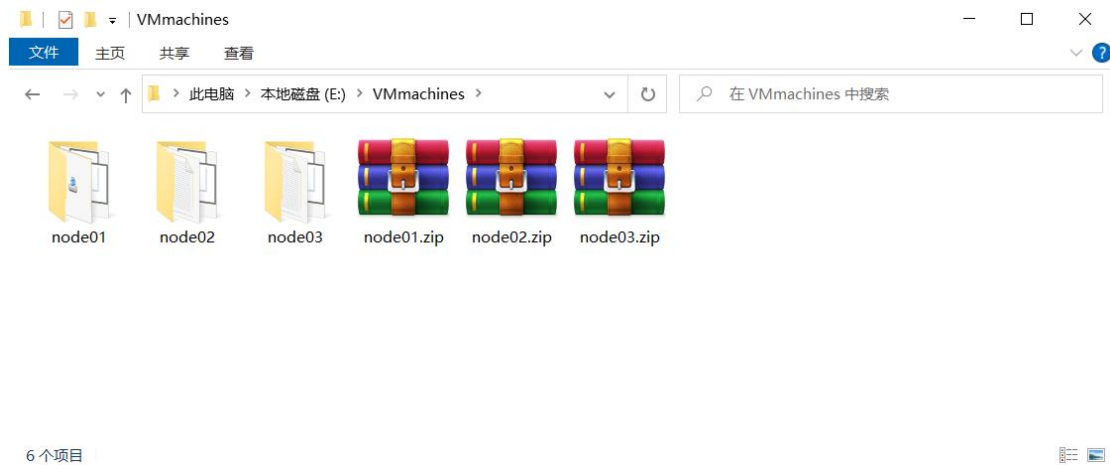


图 5

步骤 4. 开启虚拟机

打开 VMWareWorkstation，在主页中选择“打开虚拟机”，如图 6 所示。

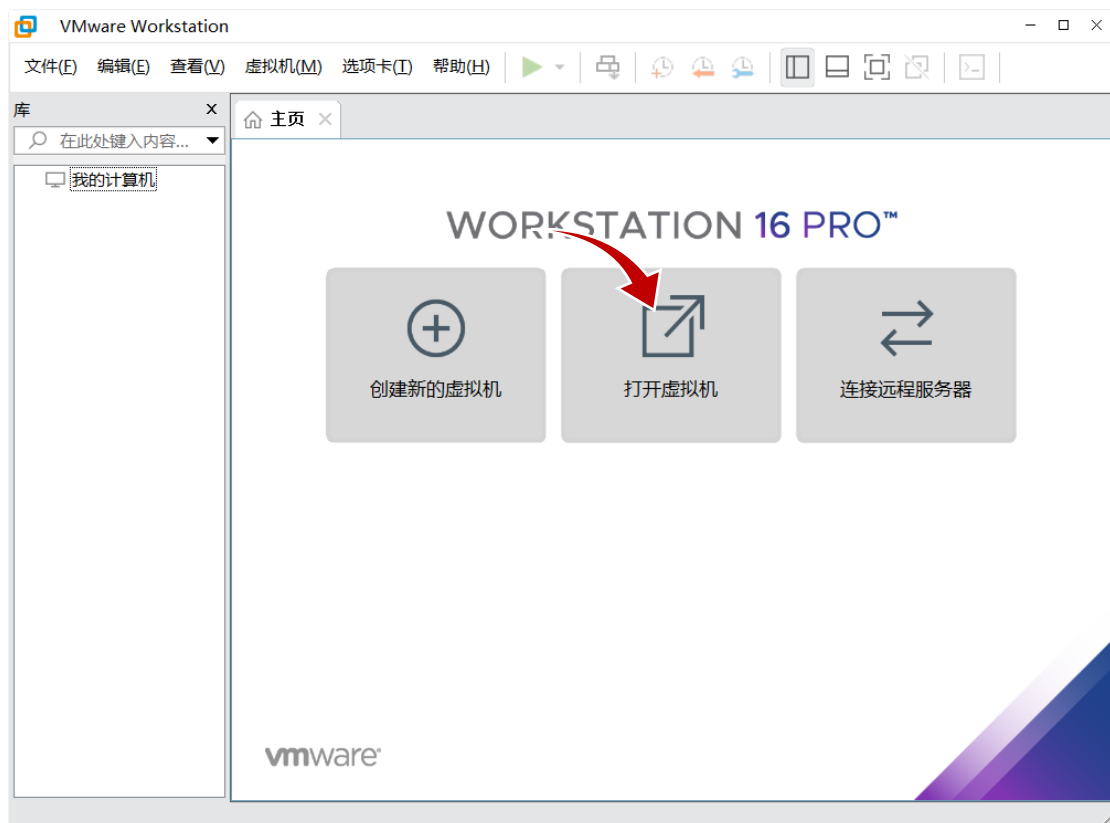


图 6

选择步骤 3 中解压出来的 node01 文件夹中的 node01.vmx 文件，点击“打

开”，如下图 7 所示。

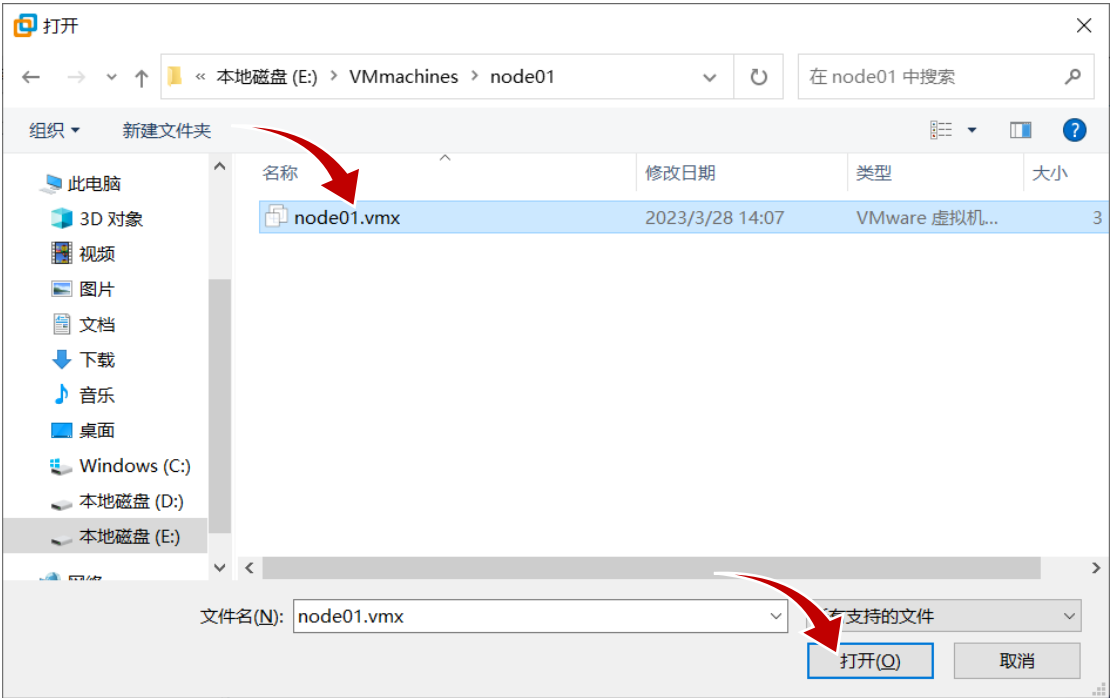


图 7

重复上述步骤，在 VMWareWorkstation 的主页中分别打开 node02.vmx 和 node03.vmx，如下图 8-9 所示。

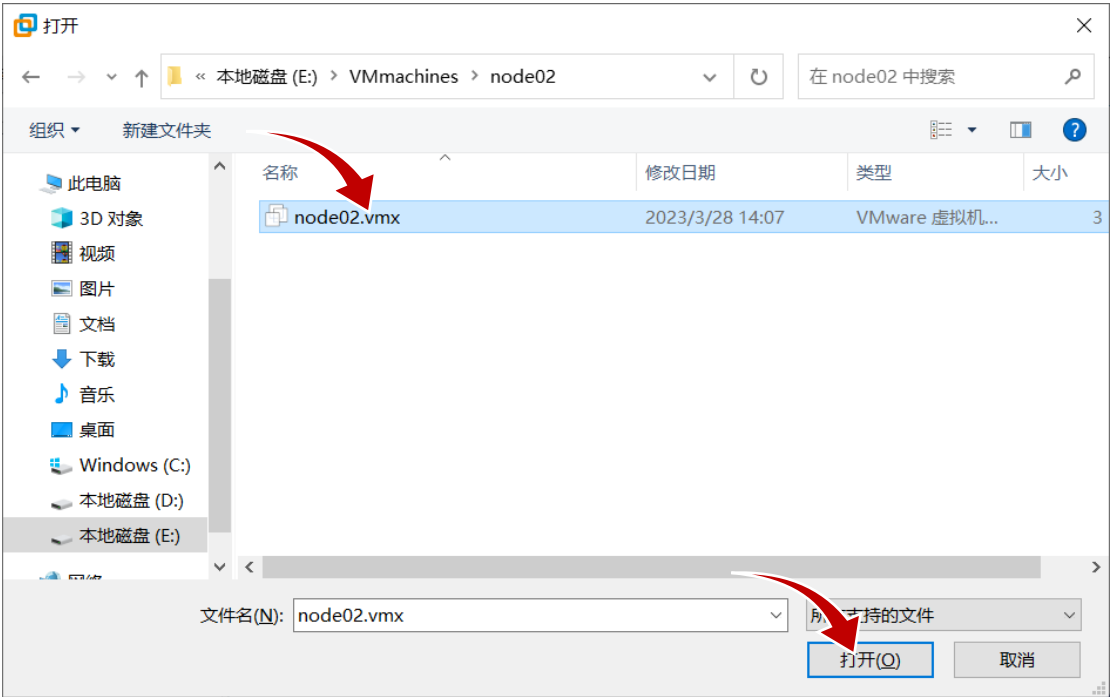


图 8

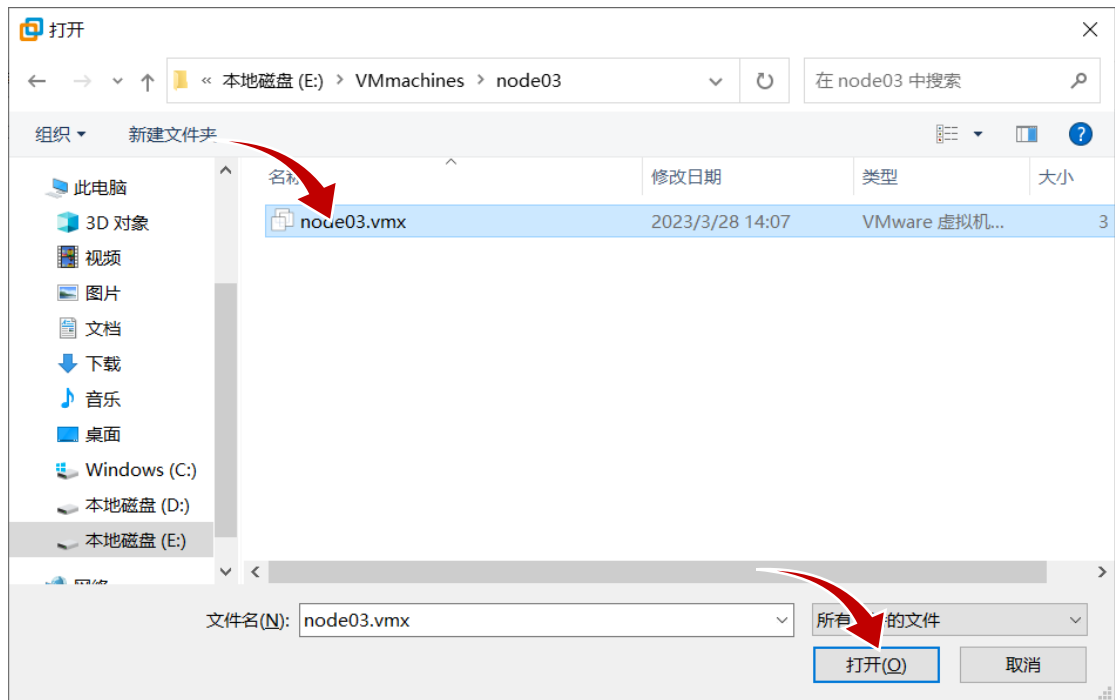


图 9

在 VMWareWorkstation 中选择 node01 虚拟机，点击“开启此虚拟机”，如图 10 所示。



图 10

开启 3 台虚拟机后，如下图 11 所示。

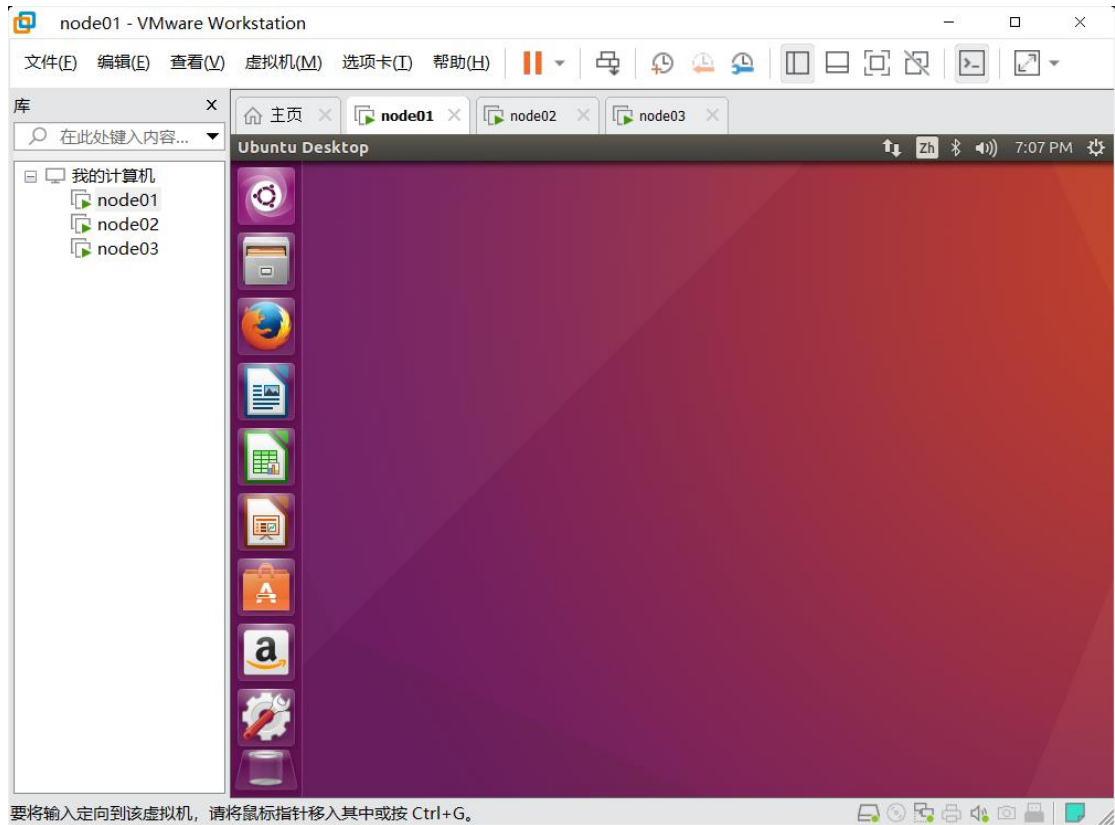


图 11

步骤 5. 开启 Hadoop

在 node01 桌面空白处，鼠标右键点击“Open Terminal”

在打开的终端中键入 `start-all.sh` 命令后按回车，如下图 12 所示。

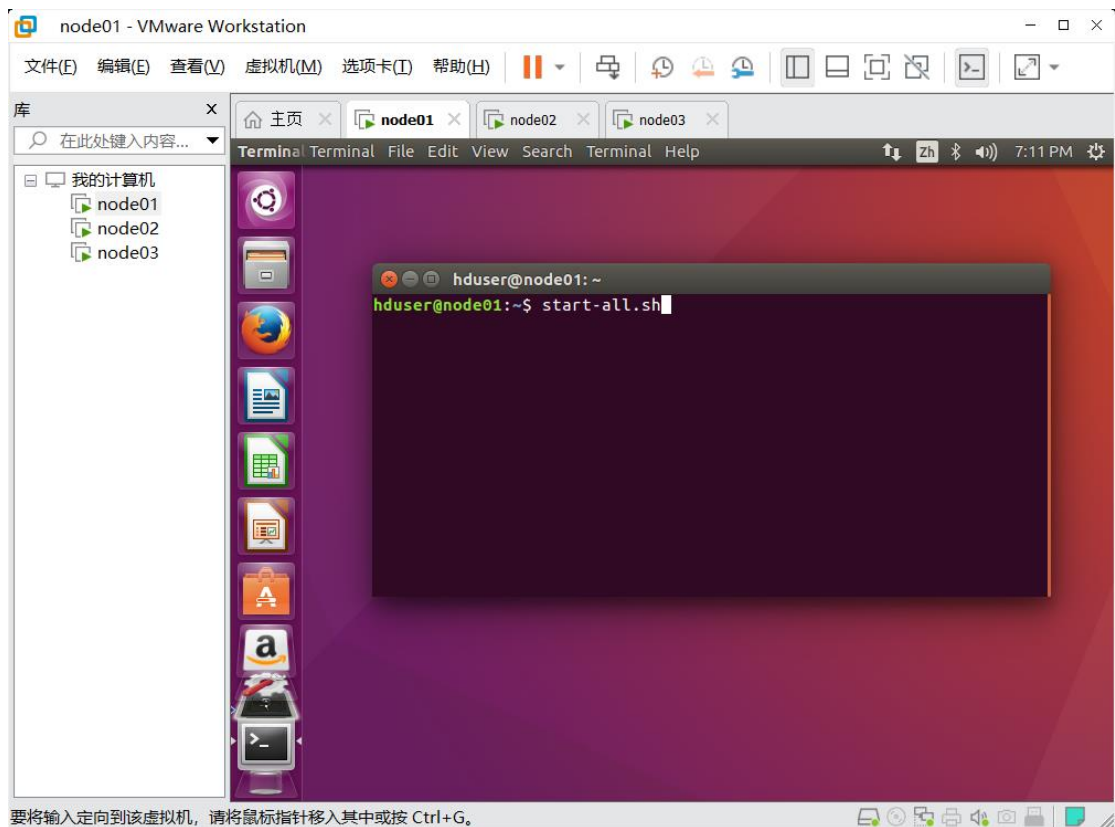


图 12

验证 Hadoop 集群是否开启成功，分别在 node01、node02、node03 打开终端，键入 jps 命令后按回车，查看进程，如下图 13-15 所示。

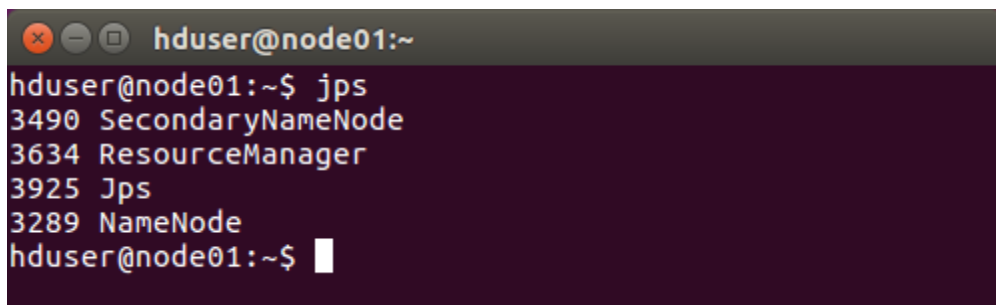
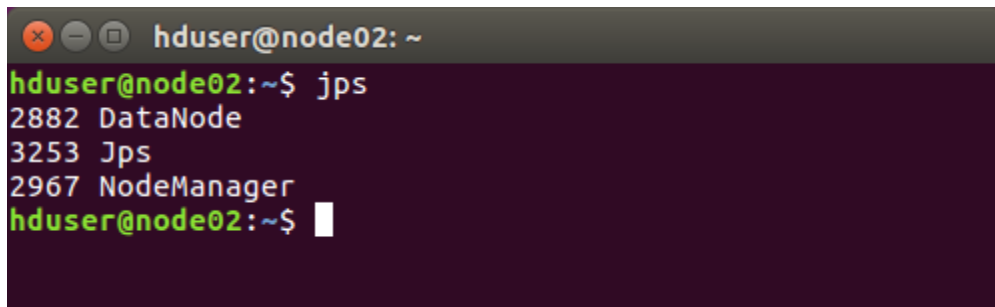
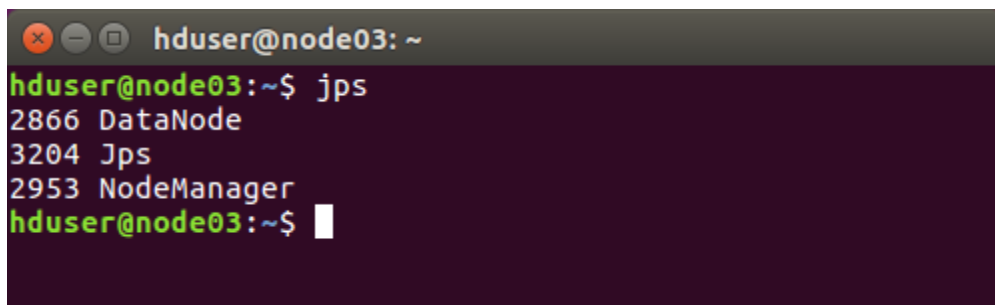


图 13

A terminal window titled 'hduser@node02: ~' with a dark purple background. The prompt is 'hduser@node02:~\$'. The command 'jps' has been executed, showing three lines of output: '2882 DataNode', '3253 Jps', and '2967 NodeManager'. The prompt is now 'hduser@node02:~\$' with a white cursor.

```
hduser@node02: ~
hduser@node02:~$ jps
2882 DataNode
3253 Jps
2967 NodeManager
hduser@node02:~$
```

图 14

A terminal window titled 'hduser@node03: ~' with a dark purple background. The prompt is 'hduser@node03:~\$'. The command 'jps' has been executed, showing three lines of output: '2866 DataNode', '3204 Jps', and '2953 NodeManager'. The prompt is now 'hduser@node03:~\$' with a white cursor.

```
hduser@node03: ~
hduser@node03:~$ jps
2866 DataNode
3204 Jps
2953 NodeManager
hduser@node03:~$
```

图 15

步骤 6. 开启 Spark

在桌面空白处，鼠标右键点击“Open Terminal”。

键入 `cd spark-2.2.0` 命令后按回车，进入 `spark-2.2.0` 文件夹。

键入 `sbin/start-all.sh` 命令后按回车，开启 spark，如下图 16 所示。

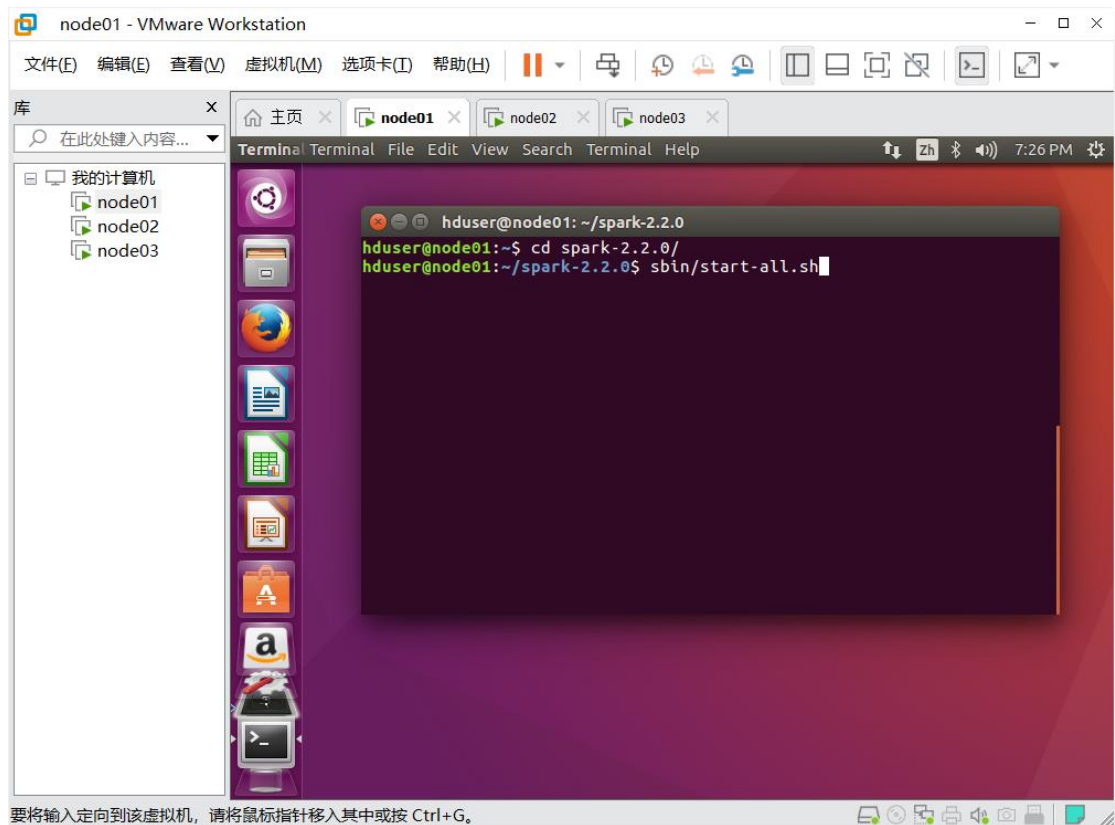


图 16

验证 Spark 集群是否开启成功, 分别在 node01、node02、node03 打开终端, 键入 jps 命令后按回车, 查看进程, 如下图 17-19 所示。

```
hduser@node01:~/spark-2.2.0
hduser@node01:~/spark-2.2.0$ sbin/start-all.sh
starting org.apache.spark.deploy.master.Master, logging to /home/hduser/spark-2.2.0/logs/spark-hduser-org.apache.spark.deploy.master.Master-1-node01.out
192.168.146.131: starting org.apache.spark.deploy.worker.Worker, logging to /home/hduser/spark-2.2.0/logs/spark-hduser-org.apache.spark.deploy.worker.Worker-1-node03.out
192.168.146.130: starting org.apache.spark.deploy.worker.Worker, logging to /home/hduser/spark-2.2.0/logs/spark-hduser-org.apache.spark.deploy.worker.Worker-1-node02.out
hduser@node01:~/spark-2.2.0$ jps
3490 SecondaryNameNode
3634 ResourceManager
5701 Master
5767 Jps
3289 NameNode
hduser@node01:~/spark-2.2.0$
```

图 17

```
hduser@node02: ~  
hduser@node02:~$ jps  
2882 DataNode  
2967 NodeManager  
4439 Worker  
4524 Jps  
hduser@node02:~$
```

图 18

```
hduser@node03: ~  
hduser@node03:~$ jps  
2866 DataNode  
2953 NodeManager  
4058 Worker  
4143 Jps  
hduser@node03:~$
```

图 19

步骤 7. 开启 Spark Shell

在 node01 打开终端，通过键入 `cd spark-2.2.0` 指令进入 `spark-2.2.0` 目录，然后键入 `bin/spark-shell` 命令后按回车，如下图 20 所示。

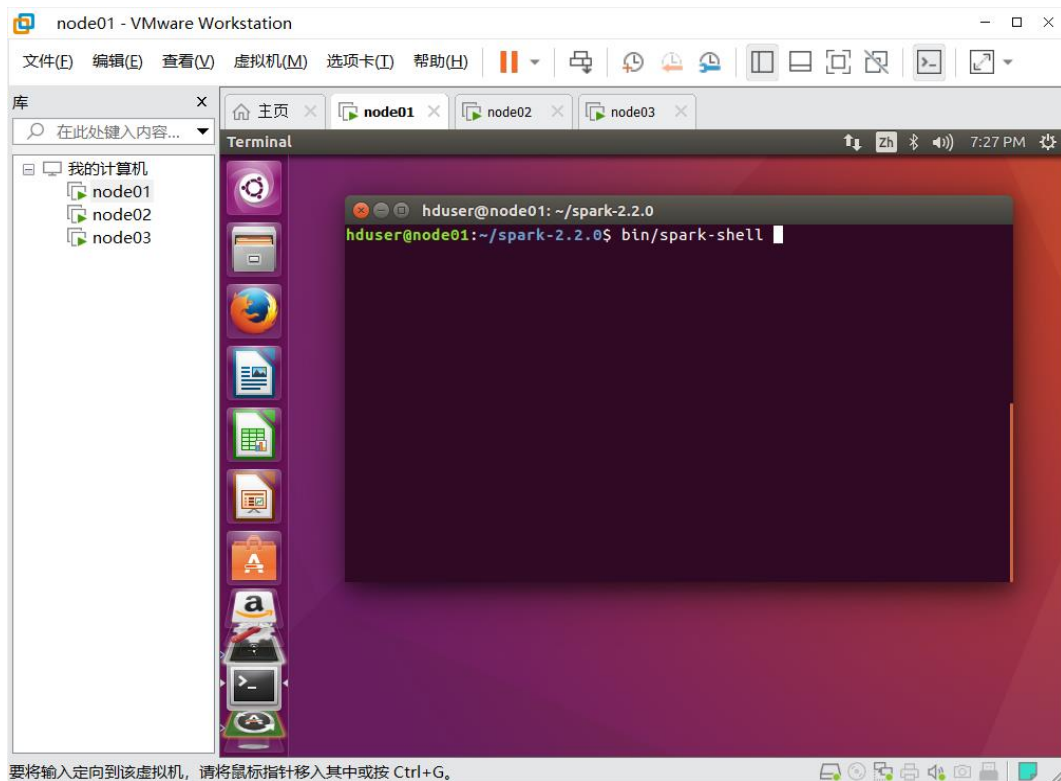
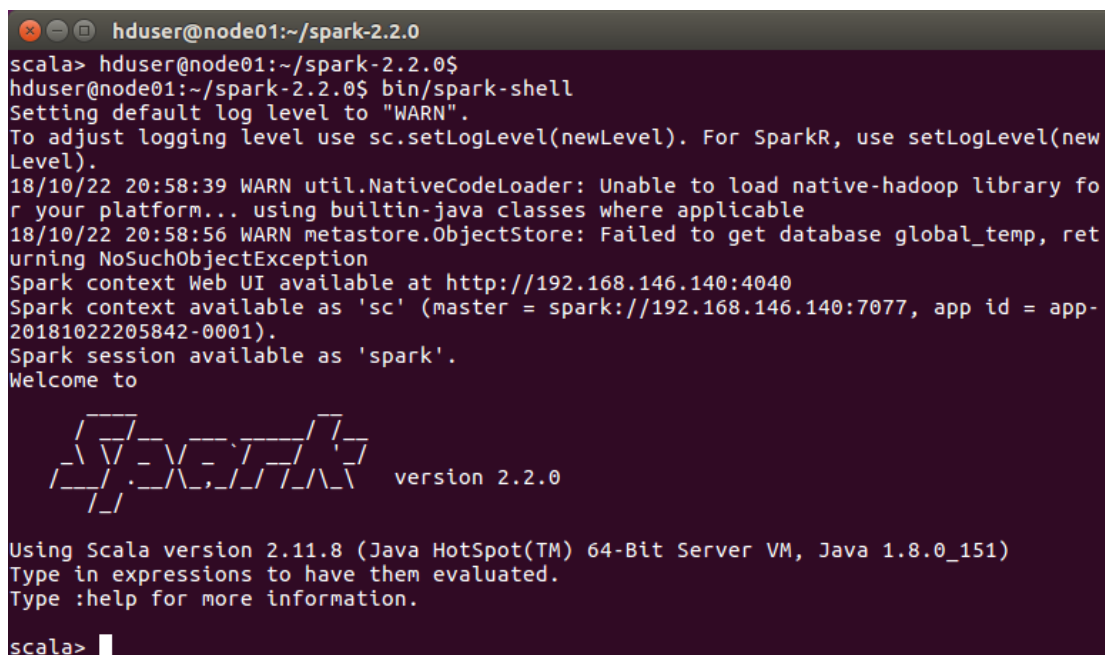


图 20

开启成功后如下图 21 所示。



```
hduser@node01:~/spark-2.2.0
scala> hduser@node01:~/spark-2.2.0$
hduser@node01:~/spark-2.2.0$ bin/spark-shell
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
18/10/22 20:58:39 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/10/22 20:58:56 WARN metastore.ObjectStore: Failed to get database global_temp, returning NoSuchObjectException
Spark context Web UI available at http://192.168.146.140:4040
Spark context available as 'sc' (master = spark://192.168.146.140:7077, app id = app-20181022205842-0001).
Spark session available as 'spark'.
Welcome to

  ____      __
 / ___ |    /  \
| |  \| |  / ____\
| |__| | / /___  \
|  __  |/ ___|  /
| |  | | \___ \|
|_|  |_|  \___/

version 2.2.0

Using Scala version 2.11.8 (Java HotSpot(TM) 64-Bit Server VM, Java 1.8.0_151)
Type in expressions to have them evaluated.
Type :help for more information.
scala> |
```

图 21

步骤 8. 关闭集群

关闭 Spark 集群：在 spark-2.2.0 目录下键入 sbin/stop-all.sh 命令后按回车

关闭 Hadoop 集群：在任意位置键入 stop-all.sh 命令后按回车

关闭集群后，在终端键入 jps 指令后按回车，如下图 22 所示。

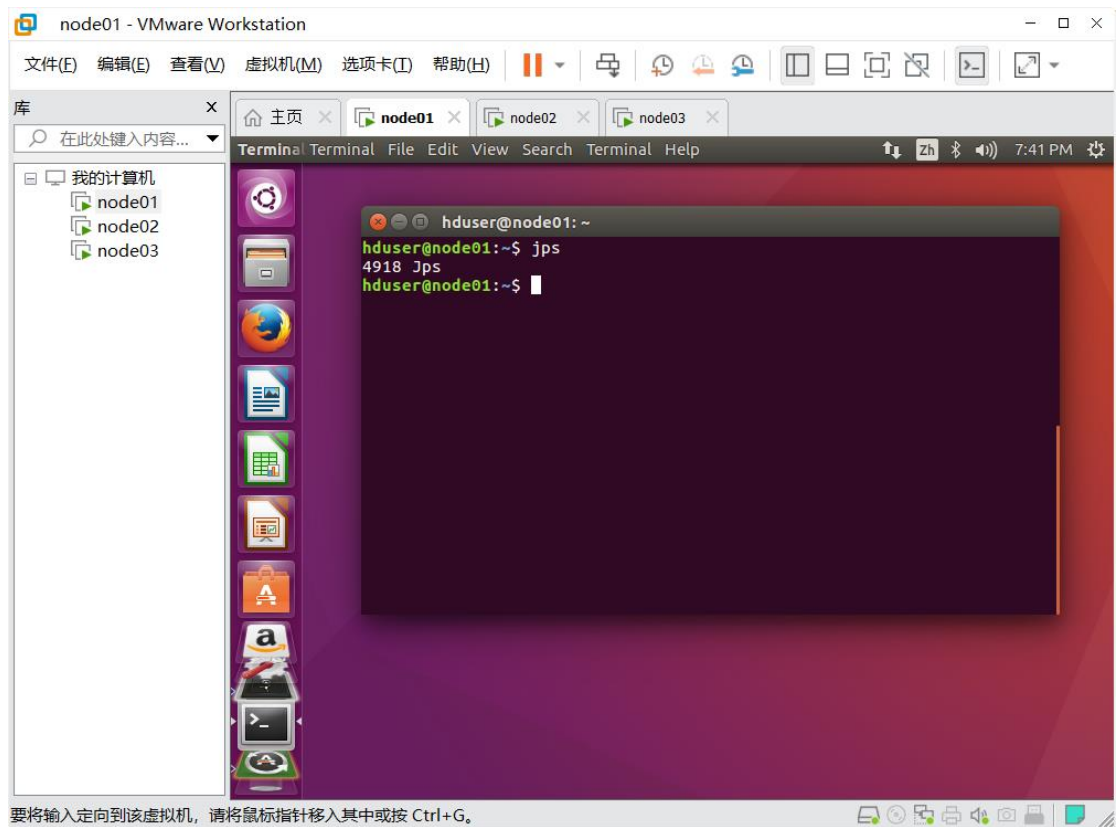


图 22