# 实验三 配置HDFS机架感知

**·实验要求：**

启动hdfs集群，查看配置前集群拓扑，再将两台从服务器的rack分别配置成 /c1/d1和/c1/d2。实现配置的方式有很多种，可以根据自己的习惯来实现（shell脚本）。

**·实验步骤：**

实验背景：三台Linux主机，分别为：master、slave1、slave2，都已经安装配置好hadoop集群。

1. 启动集群
   1. 启动集群

首先启动集群，查看集群的rack信息。

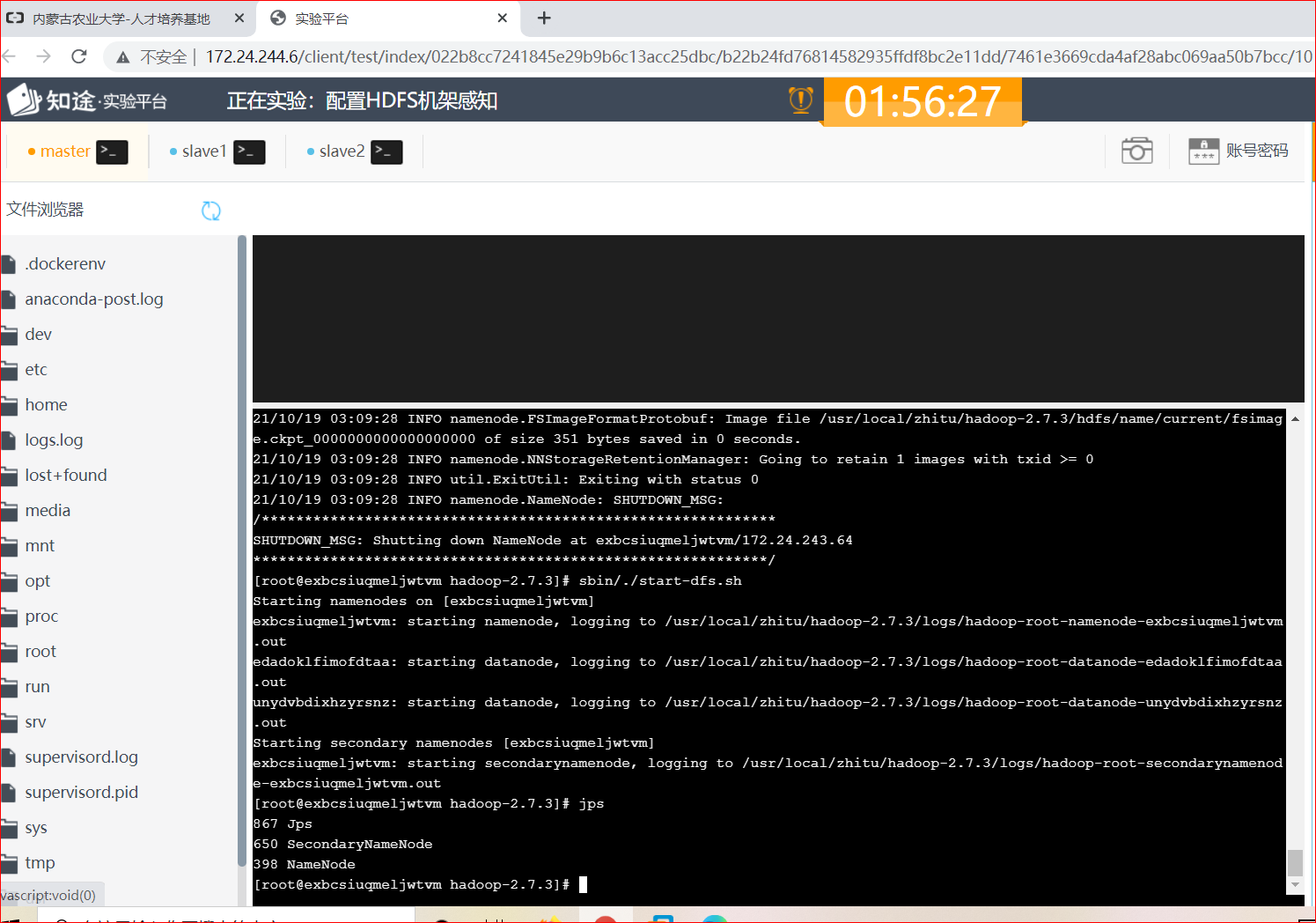
在master服务器上，进入到/usr/local/zhitu/hadoop-2.7.3目录下，使用命令启动hdfs集群。

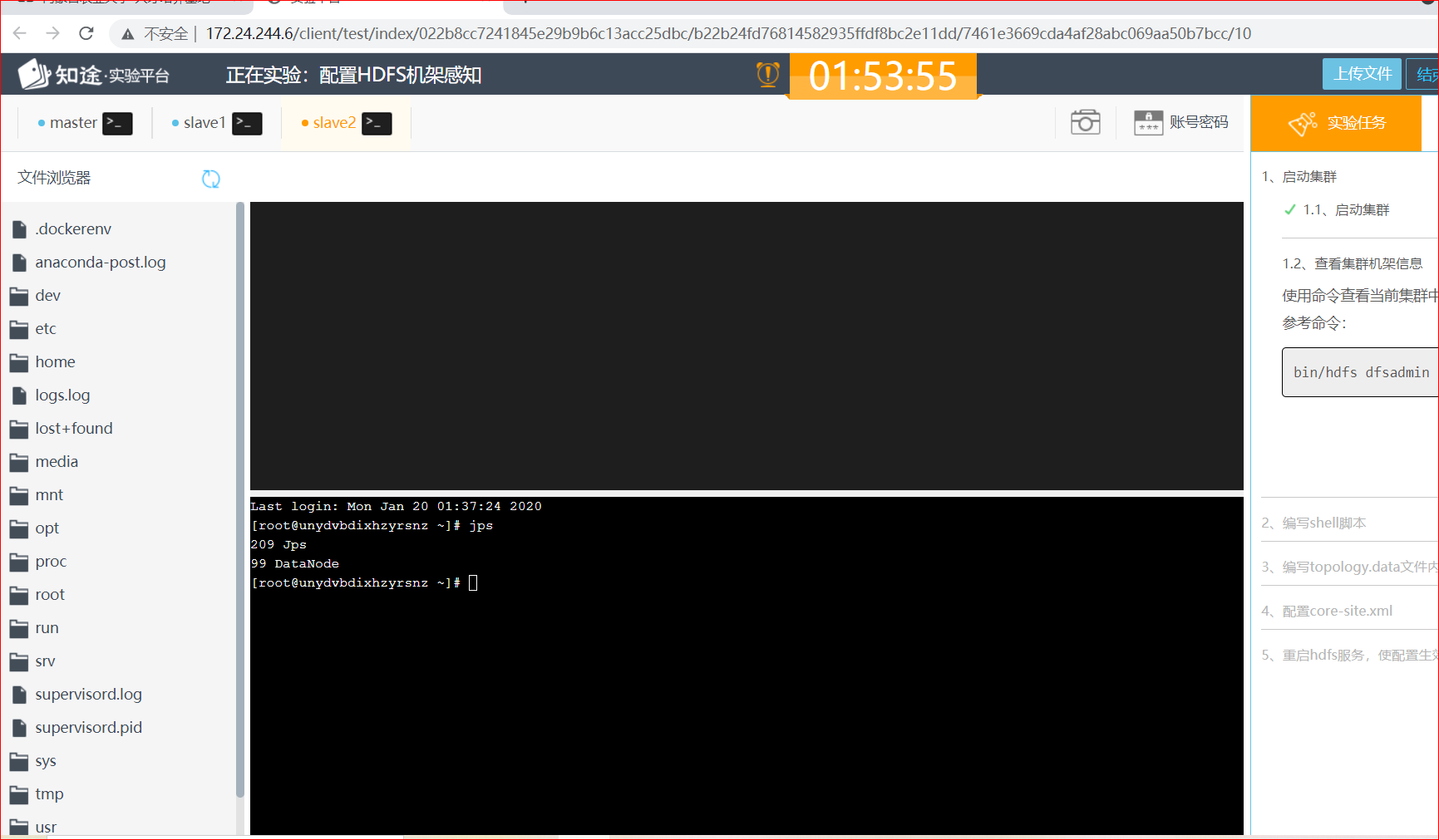
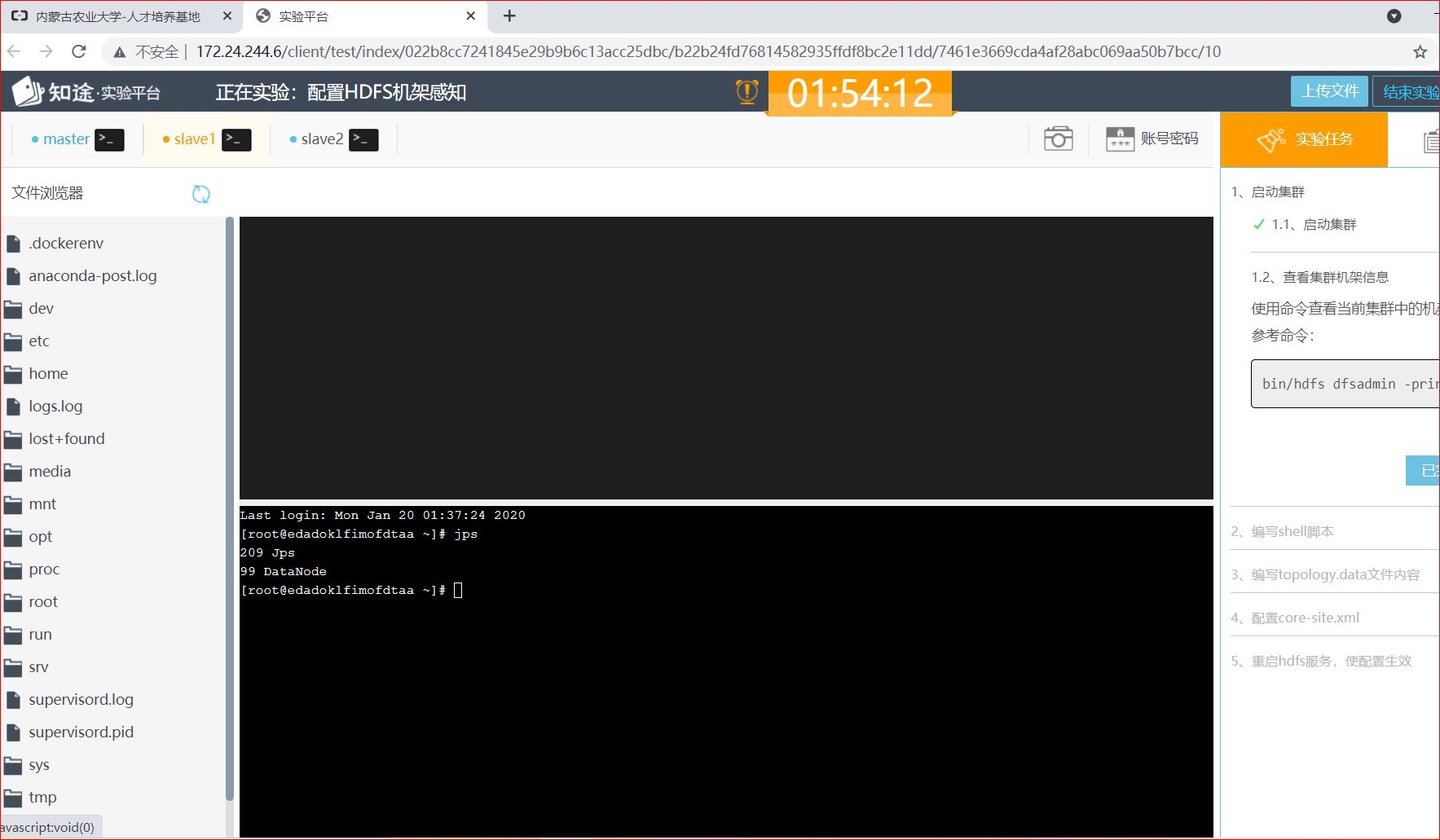
|  |
| --- |
| cd /usr/local/zhitu/hadoop-2.7.3  bin/hdfs namenode -format  sbin/./start-dfs.sh |

1

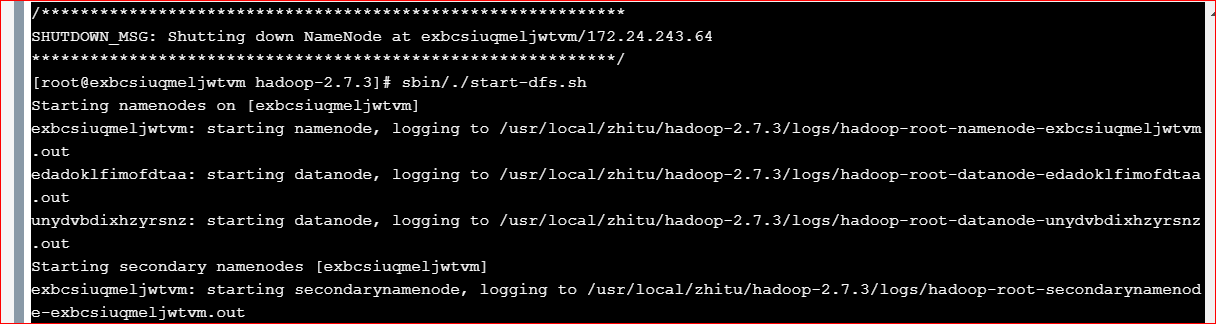
2

查看各台服务器的服务是否都启动成功：



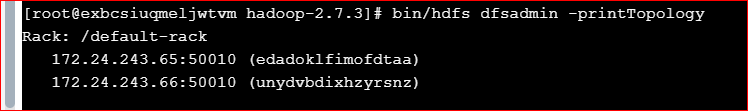


4



* 1. 查看集群机架信息

|  |
| --- |
| bin/hdfs dfsadmin -printTopology |

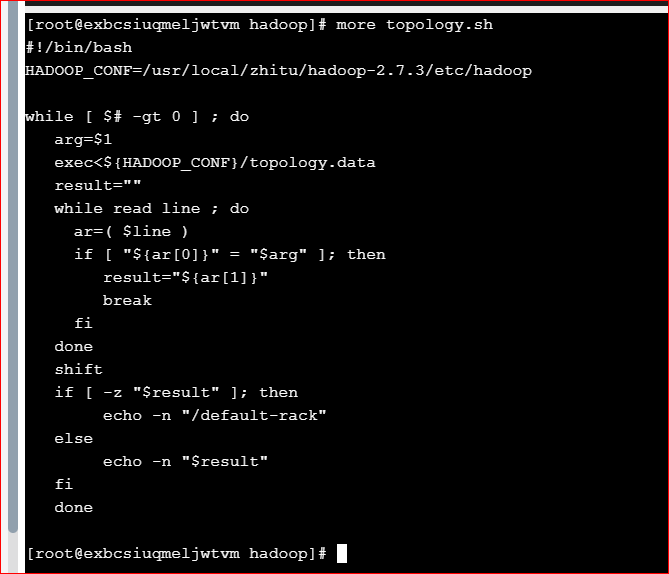
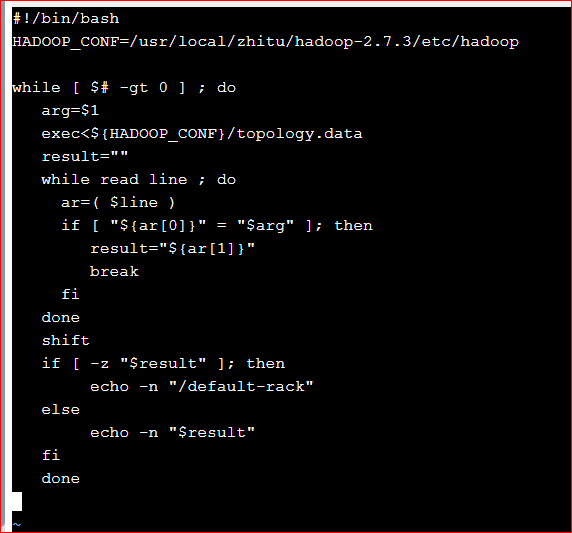


1. 编写shell脚本
   1. 在/usr/local/zhitu/hadoop-2.7.3/etc/hadoop目录下新建topology.sh  该脚本接收一个服务器的ip地址作为参数，返回该服务器的rack信息。（也可以用 vi etc/hadoop/topology.sh）

|  |
| --- |
| #!/bin/bash  HADOOP\_CONF=/usr/local/zhitu/hadoop-2.7.3/etc/hadoop    while [ $# -gt 0 ] ; do //参数大于0往下执行  arg=$1 //$1表示b，$0是a，$2是c  exec<${HADOOP\_CONF}/topology.data //输入流读取配置文件，ip地址对应的信息写入  result=""  while read line ; do  ar=( $line ) //按空格切分  if [ "${ar[0]}" = "$arg" ]; then //判断第一个参数符合IP地址，则配置第二个  result="${ar[1]}"  break  Fi //if的结束  done  shift  if [ -z "$result" ]; then //判断是否为空  echo -n "/default-rack"  else  echo -n "$result"  fi  done |

进入hadoop目录下

8

9

编写好topology.sh 脚本之后需要给脚本赋予执行的权限。

|  |
| --- |
| chmod +x /usr/local/zhitu/hadoop-2.7.3/etc/hadoop/topology.sh |

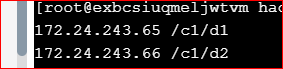
12

1. 编写topology.data文件内容
   1. 编写topology.data文件内容

上一步中使用了文件

/usr/local/zhitu/hadoop-2.7.3/etc/hadoop/topology.data，新建此文件，在文件中添加ip与rack的对应信息。

13



测试topolpgy.xml文件是否正确：

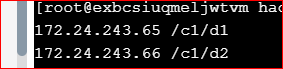
17-1

17-0

1. 配置core-site.xml
   1. 配置core-site.xml文件

进入hadoop-2.7.3目录下的etc/hadoop目录下，修改core-site.xml文件，配置集群拓扑脚本文件。

|  |
| --- |
| <property>  <name>topology.script.file.name</name>  <value>/usr/local/zhitu/hadoop-2.7.3/etc/hadoop/topology.sh</value>  </property> |

15

5、重启hdfs服务，使配置生效

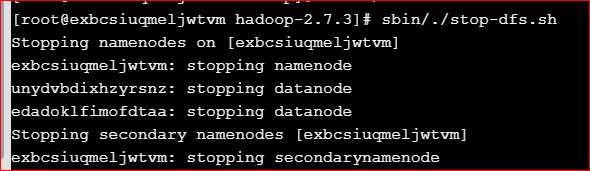
5.1、重启hdfs服务，使配置生效

  要是配置生效，需要重启hdfs服务。重启hdfs服务，查看配置是否生效。

|  |
| --- |
| cd /usr/local/zhitu/hadoop-2.7.3  sbin/./stop-dfs.sh  sbin/./start-dfs.sh  bin/hdfs dfsadmin -printTopology |

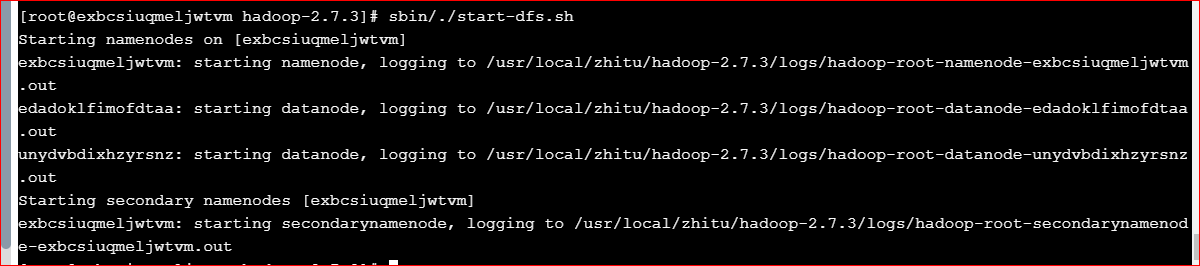
先关闭所有服务：

17

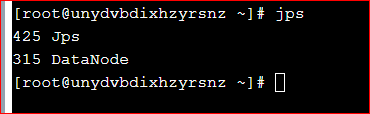
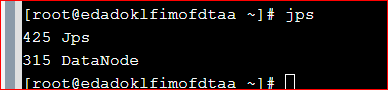
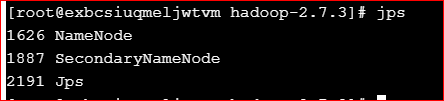


启动所有服务：

19



使用jps命令查看master、slave1、slave2所有服务：



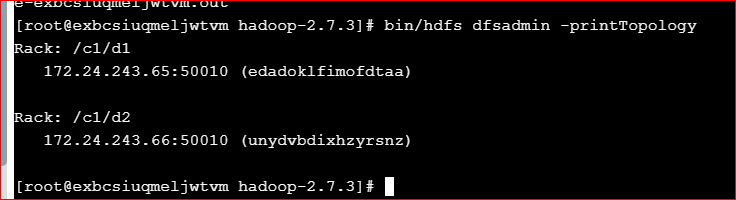
Slave2

Slave1

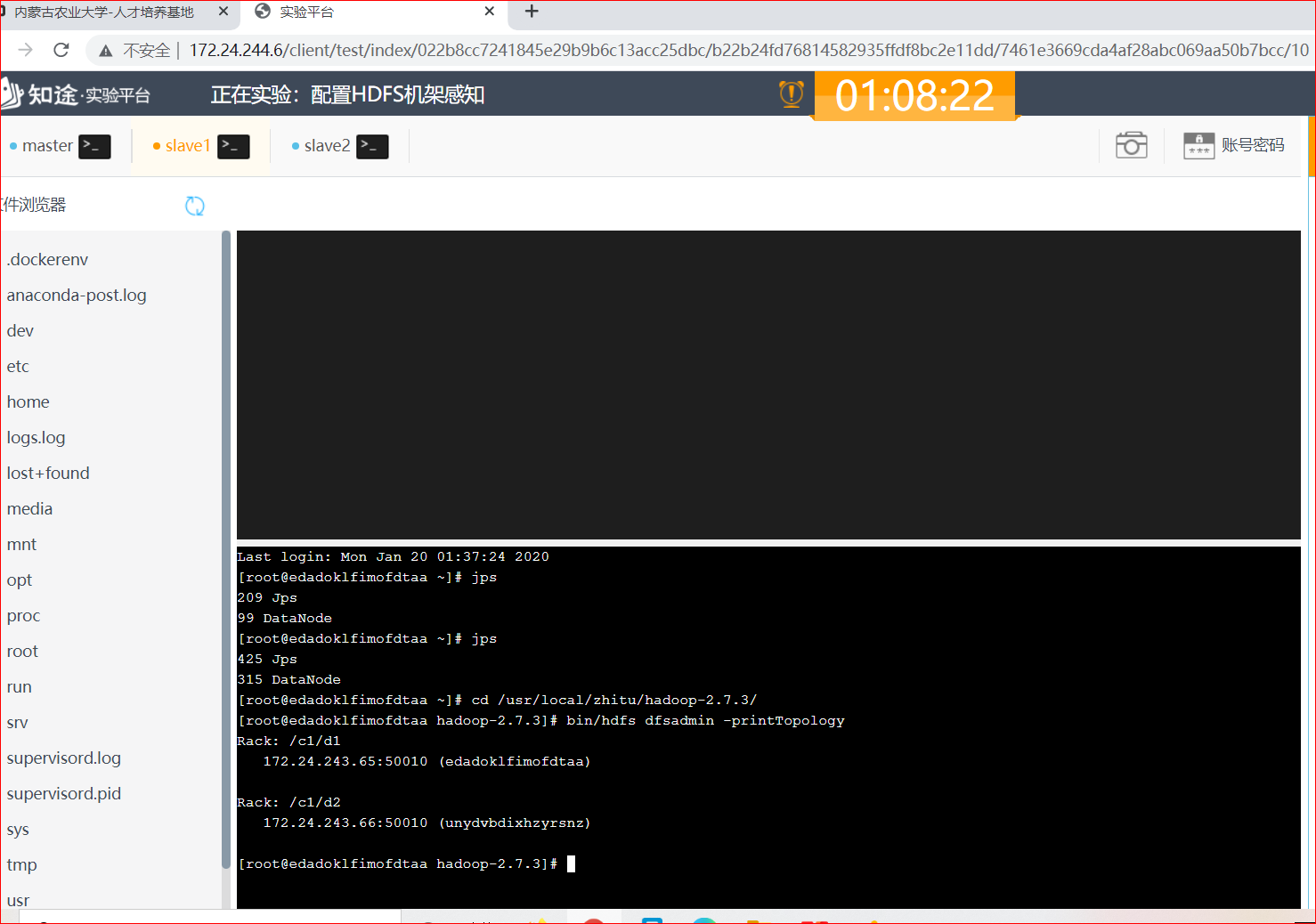
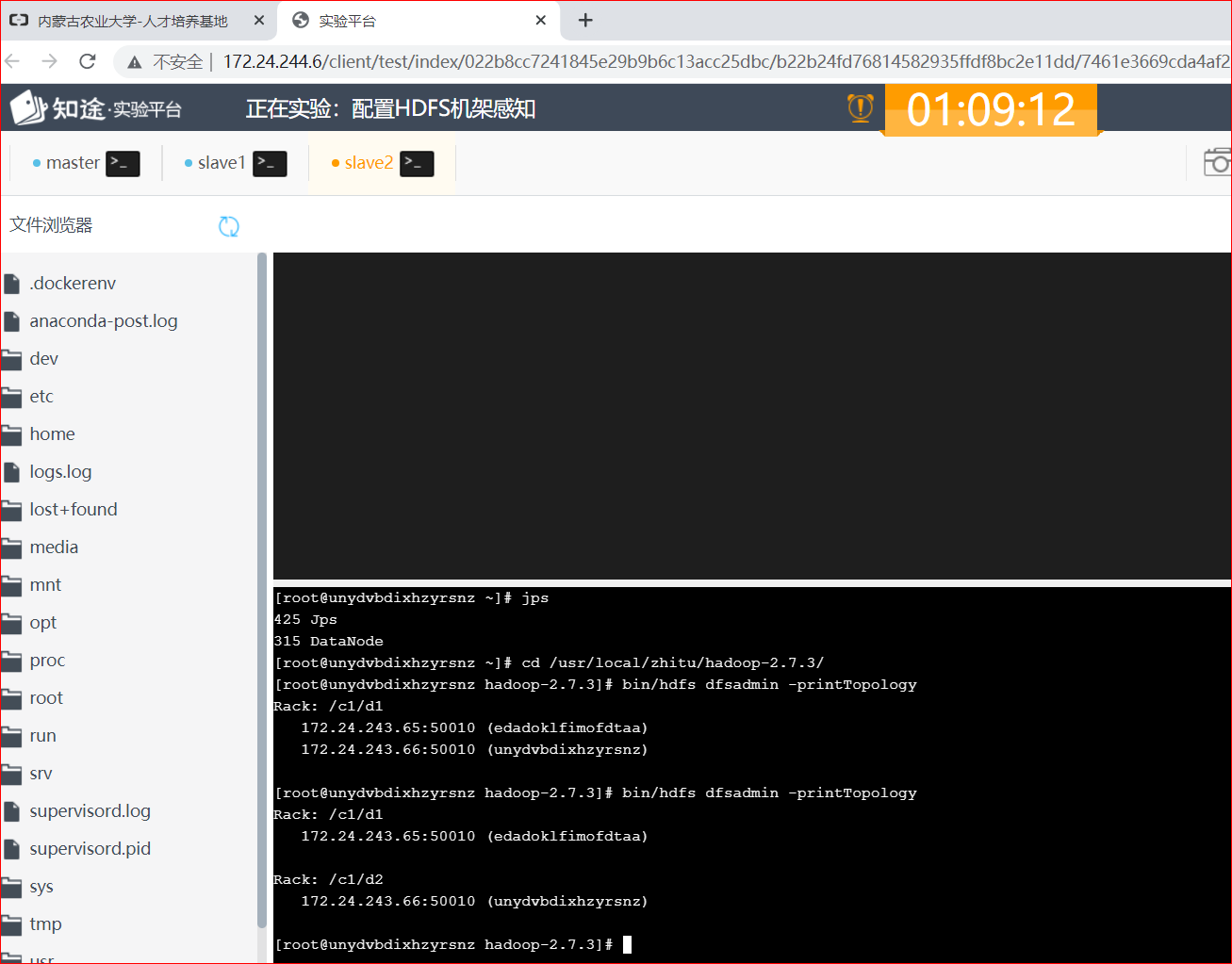
master

再次查看master上的rack信息：

21

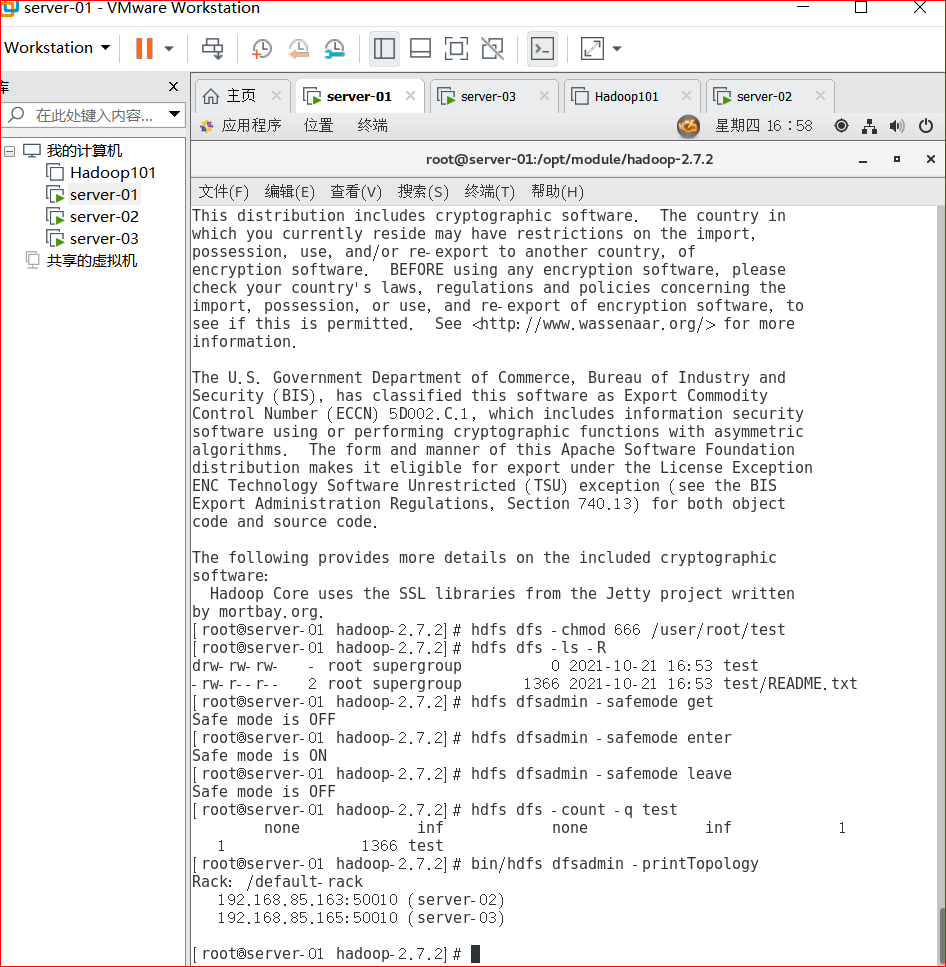


查看slave1和slave2当前机架信息信息：

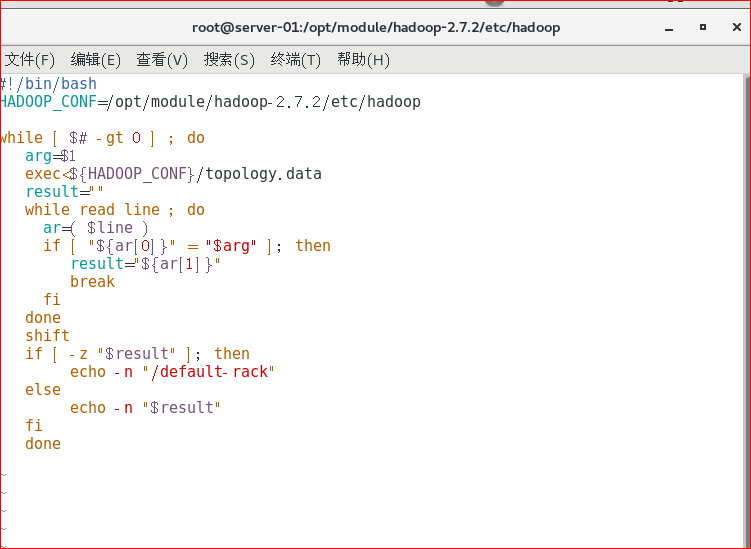


·自己环境运行结果：

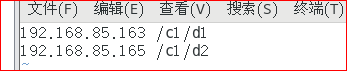
查看server-01机架信息：



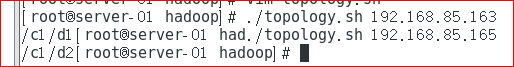
写入topology.sh的内容：



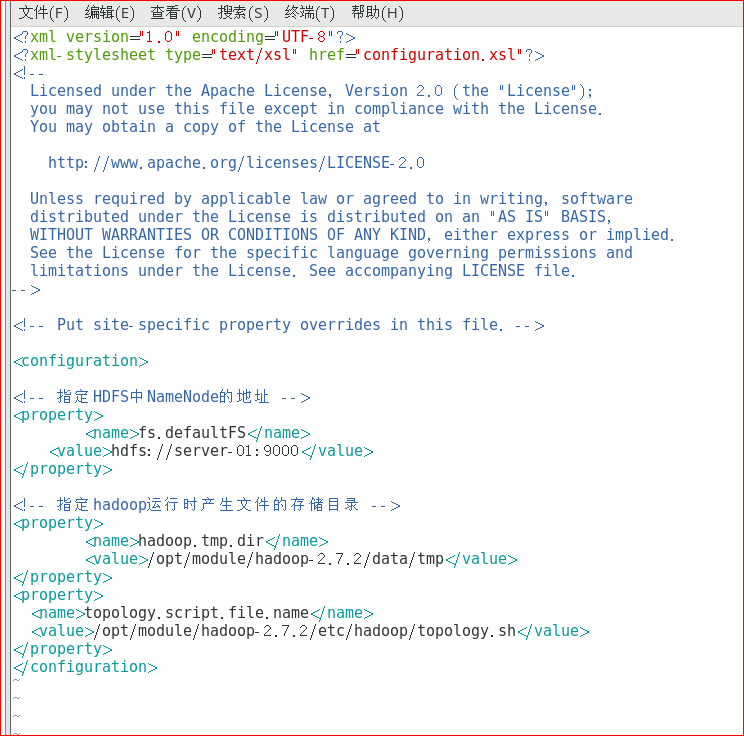
写入topology.data的内容：



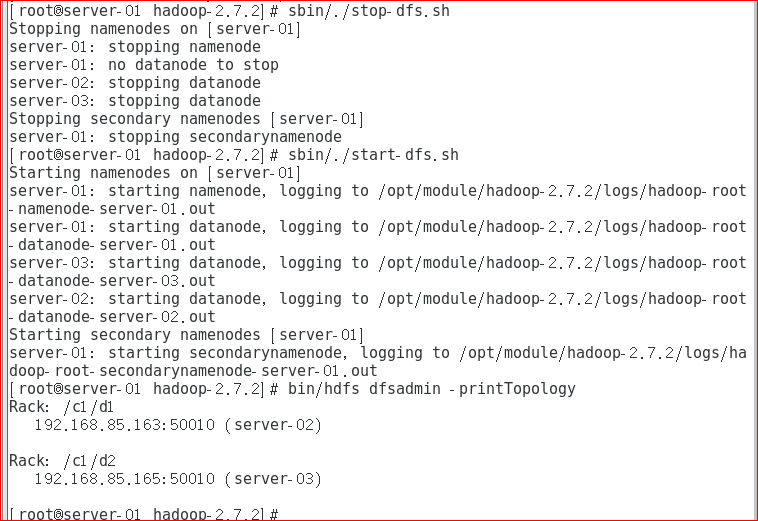
检测topology文件是否写入正确：



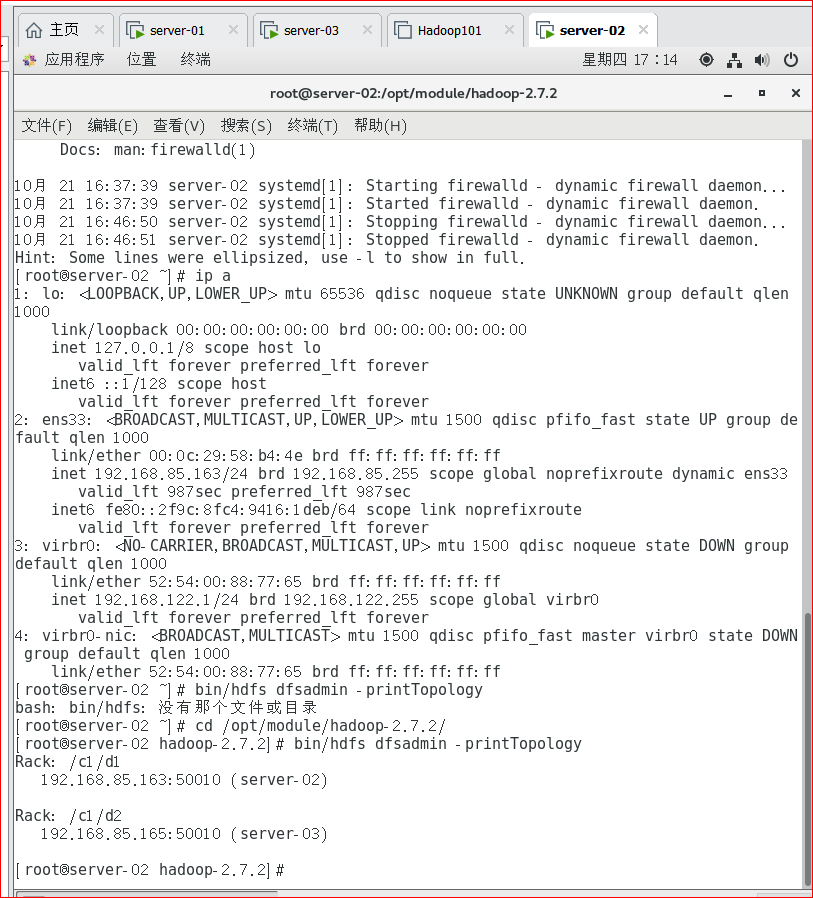
core-site.xml文件增加内容：



停止、启动所有服务，并查看server-01的机架信息：



Server-02机架信息:



Server-03机架信息:



**·实验心得**

通过本次实验学习了如何通过脚本配置HDFS机架感知，学会了如何启动和停止hdfs集群，如何查看当前集群中的机架信息，了解了shell脚本中代码表示的含义，还有如何给脚本赋予执行权限。