

实验三 Hadoop 2.x 部署与编程 (分布式)

3.1 实验目的

- 掌握免密钥登录的方法
- 学习 Hadoop 2.x 部署，理解单机集中式部署、单机伪分布式部署和分布式部署三种部署方式之间的区别
- 学会通过查找系统日志中的错误来解决系统部署中遇到的问题
- 通过系统部署理解 Hadoop 2.x 的架构，以及 Hadoop 1.x 和 Hadoop 2.x 之间的差异
- 学习在分布部署方式下运行 MapReduce 程序

3.2 实验任务

- 实现本机以及多主机之间的免密钥登录
- 完成 Hadoop 2.x 的分布式部署，并运行示例程序
- 在分布式部署方式下调试、提交并运行 MapReduce WordCount 程序

3.3 实验环境

- 操作系统：Ubuntu 18.04 或 Ubuntu 20.04
- JDK 版本：1.8
- Hadoop 版本：2.10.1

3.4 实验步骤

注意：本教程以名为 dase-local 或 dase-dis 的用户为例开展实验，读者在实验过程中需替换为自己的用户名（云主机上的默认用户名为 ubuntu）。

3.4.1 本机免密钥登录

(1) ssh 免密钥登录本机的安装和配置

a) 生成本机公钥

```
1 cd ~  
2 mkdir .ssh #可能提示该文件已存在，不影响其他操作  
3 cd ~/.ssh/
```

4 ssh-keygen -t rsa # 生成公钥私钥对，对于系统提示，按回车即可

通过上述操作生成的密钥文件所在位置:~/.ssh/, 用户公钥的位置:“~/.ssh/id_rsa.pub”,
用户私钥的位置:“~/.ssh/id_rsa”。

b) 将本机公钥发送给本机需要免密 ssh 访问的节点（即本机）

1 cat id_rsa.pub >> authorized_keys # 查看公钥内容是否追加写到授权文件

c) 执行以下命令来修改授权文件及其所在文件夹的权限

1 chmod 600 ~/.ssh/authorized_keys
2 chmod 700 -R ~/.ssh

(2) 验证本机免密钥登录是否配置成功

执行以下命令，若没有出现需要输入密码的情况，则代表配置成功。

1 ssh localhost

3.4.2 多台机器之间免密钥登录

假设现有四台机器，要使得四台机器能够通过“ssh”命令实现彼此之间免密钥登录，执行如下操作：

(1) 在每台机器上创建一个同名的用户“dase-dis”，此后所有操作都在此用户名下，软件也安装在此用户目录下

(2) 在每台机器上的“dase-dis”用户下实现本机免密钥登录（参考上文“本机免密钥登录”相关内容）

(3) 分别修改四台机器的主机名

a) 在第 1 台主机上执行以下命令

1 sudo hostnamectl set-hostname ecnu01

b) 在第 2 台主机上执行以下命令

1 sudo hostnamectl set-hostname ecnu02

c) 在第 3 台主机上执行以下命令

1 sudo hostnamectl set-hostname ecnu03

d) 在第 4 台主机上执行以下命令

```
1 sudo hostnamectl set-hostname ecnu04
```

- (4) 在每台机器上修改/etc/hosts 文件，添加主机名“ecnu01”、“encu02”、“ecnu03”、“ecnu04”和对应的 IP 地址之间的映射，使得每台机器都可以通过主机名登录其他机器。具体操作如下：

- hosts 配置文件路径：/etc/hosts，该配置文件的作用为将一些常用的网址域名与其对应的 IP 地址关联起来。当用户在浏览器中输入一个需要登录的网址时，系统会首先从 hosts 文件中寻找对应的 IP 地址，一旦找到，系统就会立即打开对应网页，如果没有找到，则系统会将网址提交 DNS 域名解析服务器进行 IP 地址的解析
- 执行以下命令，打开 hosts 文件

```
1 sudo vim /etc/hosts
```

- 打开文件后，按“i”键进入编辑模式，添加映射信息，映射信息如以下列表所示。映射信息添加完成后，按“Esc”键退出编辑模式，并输入“:wq!”保存并退出

```
1 #IP地址 主机名
2 10.24.15.11 ecnu01
3 10.24.15.12 ecnu02
4 10.24.15.13 ecnu03
5 10.24.15.14 ecnu04
6 # 注意，以上10.24.15.11、10.24.15.12、10.24.15.13、10.24.15.14  
四个IP地址仅作示范，请分别替换为四台机器实际的IP地址（云主机的内网IP）
```

- (5) 将主机名为“encu02”、“ecnu03”、“ecnu04”机器上的公钥发送给主机名为“ecnu01”机器上的授权文件。即分别切换到主机名为“ecnu02”、“ecnu03”、“ecnu04”的机器上执行命令（共执行三次）

```
1 cat ~/.ssh/id_rsa.pub | ssh dase-dis@ecnu01 'cat - >> ~/.ssh/authorized_keys'
```

- (6) 将主机名为“ecnu01”机器上的授权文件（包含“ecnu02”、“ecnu03”、“ecnu04”的公钥）发送到主机名为“ecnu02”、“ecnu03”、“ecnu04”的机器上。即切换到主机名为“ecnu01”的机器上执行命令，覆盖原有的授权文件

```
1 scp ~/.ssh/authorized_keys dase-dis@ecnu02:/home/dase-dis/.ssh/authorized_keys
2 scp ~/.ssh/authorized_keys dase-dis@ecnu03:/home/dase-dis/.ssh/authorized_keys
3 scp ~/.ssh/authorized_keys dase-dis@ecnu04:/home/dase-dis/.ssh/authorized_keys
```

- (7) 验证四台机器是否彼此之间都可以免密钥登录。在每台机器上执行以下命令，若没有出现需要输入密码的情况，则代表配置成功

```
1 ssh dase-dis@ecnu01
2 exit
3 ssh dase-dis@ecnu02
4 exit
5 ssh dase-dis@ecnu03
6 exit
7 ssh dase-dis@ecnu04
8 exit
```

3.4.3 分布式部署

HDFS

(1) 准备工作

- 准备 4 台机器，其中包括 1 个主节点（主机名：ecnu01）、2 个从节点（主机名：ecnu02 和 ecnu03）、1 个客户端（主机名：ecnu04）
- 4 台机器均已创建用户 dase-dis
- 4 台机器之间实现免密钥登录
- 登录 dase-dis 用户
- 检查 IP 与主机名映射是否正确，若 IP 地址发生改变，参考第3.4.2节进行修改

(2) 修改 HDFS 配置

注意：修改之前务必停止正在运行的 HDFS 服务。

在主节点执行以下操作：

- 修改 slaves 文件（文件路径：~/hadoop-2.10.1/etc/hadoop/slaves）

```
1 ecnu02
2 ecnu03
```

- 修改 core-site.xml（文件路径：~/hadoop-2.10.1/etc/hadoop/core-site.xml）

```
1 <configuration>
2   <property>
3     <name>hadoop.tmp.dir</name>
4     <value>/home/dase-dis/hadoop-2.10.1/tmp</value>
5   </property>
6   <!-- 填写主节点主机名-->
7   <property>
8     <name>fs.defaultFS</name>
9     <value>hdfs://ecnu01:9000</value>
```

```
10  </property>
11 </configuration>
```

- 修改 hdfs-site.xml (文件路径: `~/hadoop-2.10.1/etc/hadoop/hdfs-site.xml`)

```
1 <configuration>
2   <property>
3     <name>dfs.replication</name>
4     <value>2</value>
5   </property>
6   <property>
7     <name>dfs.namenode.name.dir</name>
8     <value>file:/home/dase-dis/hadoop-2.10.1/tmp/dfs/name</value>
9   </property>
10  <property>
11    <name>dfs.datanode.data.dir</name>
12    <value>file:/home/dase-dis/hadoop-2.10.1/tmp/dfs/data</value>
13  </property>
14 </configuration>
```

- 修改 hadoop-env.sh 文件 (文件路径: `~/hadoop-2.10.1/etc/hadoop/hadoop-env.sh`)

找到 `export JAVA_HOME=$JAVA_HOME` 这行，将此行修改为：

`export JAVA_HOME=/usr/local/jdk1.8` 即可。

- 将整个 hadoop-2.10.1 文件夹拷贝至其余机器

```
1 scp -r /home/dase-dis/hadoop-2.10.1 dase-dis@ecnu02:/home/dase-dis/
2 scp -r /home/dase-dis/hadoop-2.10.1 dase-dis@ecnu03:/home/dase-dis/
3 scp -r /home/dase-dis/hadoop-2.10.1 dase-dis@ecnu04:/home/dase-dis/
```

(3) 启动 HDFS 服务

在主节点执行如下命令：

- 格式化 NameNode

注意：仅在第一次启动 HDFS 时才需要格式化 NameNode，如果是重启 HDFS，那么跳过“格式化 NameNode”这一步，直接执行下一步“启动 HDFS 服务”即可。此外，在进行 NameNode 格式化之前，如果 `~/hadoop-2.10.1/tmp/` 文件夹已存在，那么需要删除该文件夹后再执行以下格式化命令。

```
1 ~/hadoop-2.10.1/bin/hdfs namenode -format
```

- 启动 HDFS 服务

```
1 ~/hadoop-2.10.1/sbin/start-dfs.sh
```

(4) 查看 HDFS 服务信息

- 使用 jps 查看进程，验证是否成功启动服务

主节点正常启动结果如图3.1所示，从节点正常启动结果如图3.2所示。如启动失败，可参照下面的查看进程日志来排查原因。

```
dase-dis@ecnu01:~$ jps
25536 SecondaryNameNode
25169 NameNode
28058 Jps
```

图 3.1 主节点 jps 运行结果

```
dase-dis@ecnu02:~$ jps
6355 Jps
4325 DataNode
```

图 3.2 从节点 jps 运行结果

- 查看 NameNode、DataNode、SecondaryNameNode 进程日志

本实验的进程日志记录在 `~/hadoop-2.10.1/logs/` 路径下，后缀为.log 的文件中。

- NameNode 进程日志

默认位置：`~/hadoop-2.10.1/logs/hadoop-*-namenode-*.`log

- DataNode 进程日志

默认位置：`~/hadoop-2.10.1/logs/hadoop-*-datanode-*.`log

- SecondaryNameNode 进程日志

默认位置：`~/hadoop-2.10.1/logs/hadoop-dase-local-secondarynamenode-*.`log

- 访问 HDFS 网页：<http://ecnu01:50070>，如图3.3所示

(5) 常用的 HDFS Shell 命令

- 目录操作（在客户端执行以下操作）

注意：第一次使用 HDFS 时，需要首先在 HDFS 中创建用户目录，如下方代码所示。

HDFS 中的目录操作包括：新建目录、查看目录内容、删除目录等。示例如下：

```
1 cd ~/hadoop-2.10.1
2 ./bin/dfs dfs -mkdir -p /user/dase-dis #在 HDFS 中为当前 dase-dis
    用户创建一个用户根目录 hdfs:///user/dase-dis
```

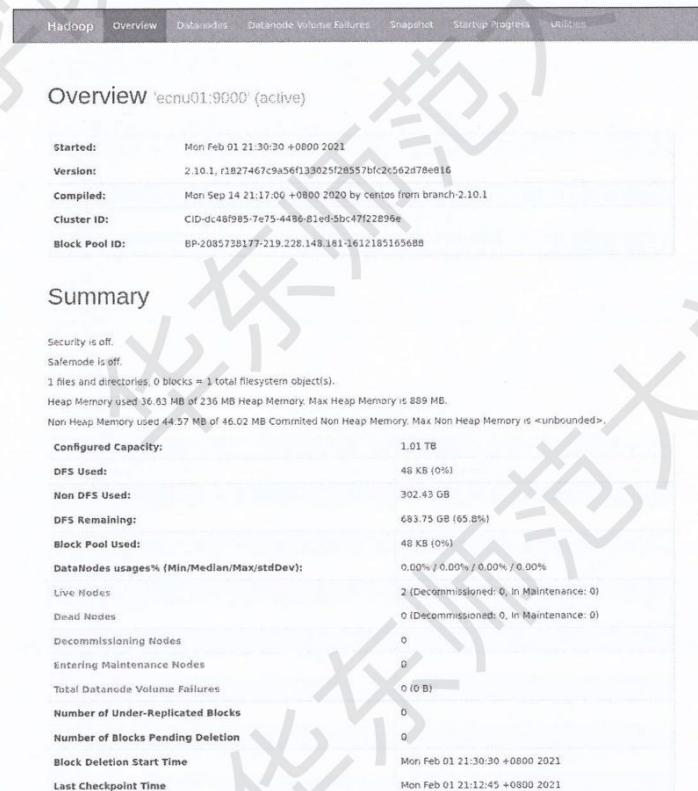


图 3.3 HDFS 信息

```

3 ./bin/hdfs dfs -ls . #显示 hdfs:///user/dase-dis 下的文件
4 ./bin/hdfs dfs -ls /user/dase-dis #显示 hdfs:///user/dase-dis 下的文件
5 ./bin/hdfs dfs -mkdir input #新建 hdfs:///user/dase-dis/input 目录
6 ./bin/hdfs dfs -rm -r input #删除 hdfs:///user/dase-dis/input 目录
7 ./bin/hdfs dfs -mkdir /input #新建 hdfs:///input 目录
8 ./bin/hdfs dfs -rm -r /input #删除 hdfs:///input 目录

```

- 文件操作（在客户端执行以下操作）

HDFS 中的文件操作包括：上传文件、下载文件、拷贝文件等。示例如下：

- 从本地文件系统向 HDFS 中上传文件

```

1 cd ~/hadoop-2.10.1
2 ./bin/hdfs dfs -mkdir input #新建 hdfs:///user/dase-dis/input 文件夹
3 ./bin/hdfs dfs -put README.txt input/ #把本地的 README.txt 上传到
   hdfs:///user/dase-dis/input
4 ./bin/hdfs dfs -cat input/README.txt #查看上传的 README.txt

```

- 将之前下载保存至 ~/input/ 的文件 pd.train 上传至 HDFS

```

1 mkdir ~/input
2 scp dase-local@localhost:~/input/pd.train ~/input/ #
   将 dase-local 用户的 “~/input/pd.train” 文件拷贝至当前用户
3 ./bin/hdfs dfs -put ~/input/pd.train input/
   #将本地文件系统的 “~/input/pd.train” 上传至
   hdfs:///user/dase-dis/input/ 目录下

```

- 把 HDFS 中的文件下载到本地文件系统中

```

1 ~/hadoop-2.10.1/bin/hdfs dfs -get input/README.txt ~/Downloads #把
   hdfs:///user/dase-dis/input/README.txt 下载到本地 ~/Downloads
2 cat ~/Downloads/README.txt #查看下载的 README.txt

```

- 把文件从 HDFS 中的一个目录拷贝至 HDFS 中的另外一个目录

```

1 ~/hadoop-2.10.1/bin/hdfs dfs -cp input/README.txt . #把
   hdfs:///user/dase-dis/input/README.txt 拷贝至 hdfs:///user/dase-dis
   目录下
2 ~/hadoop-2.10.1/bin/hdfs dfs -cat README.txt #查看拷贝的 README.txt

```

(6) 停止 HDFS 服务

- 在主节点执行以下停止命令

```
~/hadoop-2.10.1/sbin/stop-dfs.sh
```

- 查看进程，验证是否成功停止服务

使用 jps 在主节点和从节点查看进程，不再出现 NameNode、DataNode、SecondaryNameNode，则表示服务停止。

MapReduce

(1) 修改 MapReduce 配置

以下操作均在主节点进行。

- 新建 mapred-site.xml(文件位置：`~/hadoop-2.10.1/etc/hadoop/`)

```
1 cp ~/hadoop-2.10.1/etc/hadoop/mapred-site.xml.template
2   ~/hadoop-2.10.1/etc/hadoop/mapred-site.xml
3   #通过复制模板文件新建mapred-site.xml
```

- 修改 mapred-site.xml

```
1 <configuration>
2   <property>
3     <name>mapreduce.framework.name</name>
4     <value>yarn</value>
5   </property>
6   <property>
7     <name>mapreduce.jobhistory.address</name>
8     <value>ecnu01:10020</value>
9   </property>
10  <property>
11    <name>mapreduce.jobhistory.webapp.address</name>
12    <value>ecnu01:19888</value>
13  </property>
14  <!-- 指定作业运行中产生的临时文件的暂存目录位置-->
15  <property>
16    <name>yarn.app.mapreduce.am.staging-dir</name>
17    <value>/user</value>
18  </property>
19 </configuration>
```

- 修改 yarn-site.xml(文件位置：`~/hadoop-2.10.1/etc/hadoop/`)

```
1 <configuration>
2   <property>
3     <name>yarn.resourcemanager.hostname</name>
4     <value>ecnu01</value>
5   </property>
```

```

6   <property>
7     <name>yarn.nodemanager.aux-services</name>
8     <value>mapreduce_shuffle</value>
9   </property>
10  <property>
11    <name>yarn.log-aggregation-enable</name>
12    <value>true</value>
13  </property>
14  <property>
15    <name>yarn.log-aggregation.retain-seconds</name>
16    <value>302400</value>
17  </property>
18  <property>
19    <name>yarn.nodemanager.vmem-pmem-ratio</name>
20    <value>2.1</value>
21  </property>
22 </configuration>

```

- 将配置文件拷贝到从节点和客户端

```

1 scp -r ~/hadoop-2.10.1/etc/hadoop dase-dis@ecnu02:~/hadoop-2.10.1/etc/
2 scp -r ~/hadoop-2.10.1/etc/hadoop dase-dis@ecnu03:~/hadoop-2.10.1/etc/
3 scp -r ~/hadoop-2.10.1/etc/hadoop dase-dis@ecnu04:~/hadoop-2.10.1/etc/

```

(2) 启动 Yarn 服务

在主节点执行以下命令：

- 启动 Yarn 服务

```

1 ~/hadoop-2.10.1/sbin/start-yarn.sh #启动Yarn
2 ~/hadoop-2.10.1/sbin/mr-jobhistory-daemon.sh start historyserver
  #开启历史服务器，才能在 web 中查看作业日志

```

- 启动 HDFS 服务

```
1 ~/hadoop-2.10.1/sbin/start-dfs.sh
```

(3) 查看 Yarn 服务信息

- 使用 jps 查看进程，验证是否成功启动服务

分别在主节点和从节点执行 jps 命令，主节点根据配置应出现图3.4所示进程，从节点应出现如图3.5所示进程。

如果启动后缺少某些进程，说明启动过程中出现错误：参考下方查看进程日志来排查原因，根据日志中的出错信息，查找相关解决方法。

```
dase-dis@ecnu01:~$ jps
12292 JobHistoryServer
15813 NameNode
11737 ResourceManager
16667 Jps
16173 SecondaryNameNode
```

图 3.4 主节点进程

```
dase-dis@ecnu02:~$ jps
18273 Jps
16356 DataNode
11503 NodeManager
```

图 3.5 从节点进程

- 查看 ResourceManager、NodeManager 进程日志
本实验的进程日志记录在 `~/hadoop-2.10.1/logs/` 路径下，后缀为.log 的相关文件中。
 - ResourceManager 进程日志
默认位置：`~/hadoop-2.10.1/logs/yarn-*-resourcemanager-*.``log`
 - NodeManager 进程日志
默认位置：`~/hadoop-2.10.1/logs/yarn-*-nodemanager-*.``log`

- 访问 Yarn 的 Web 界面
通过 `http://ecnu01:8088`，查看 Yarn 的 cluster 信息，如图3.6所示。

(4) 运行 MapReduce 应用程序

- 在客户端提交 jar 包并查看运行结果，运行结果如图3.7所示。

```
1 cd ~/hadoop-2.10.1
2 ./bin/hdfs dfs -mkdir -p input/grep #新建输入文件夹
3 ./bin/hdfs dfs -put etc/hadoop/* input/grep #上传输入文件
4 ./bin/hdfs dfs -rm -r output/grep #删除输出路径
5 ./bin/yarn jar share/hadoop/mapreduce/hadoop-mapreduce-examples-2.10.1.jar
   grep input/grep output/grep 'dfs[a-z.]+'
6 ./bin/hdfs dfs -cat output/grep/p* #查看输出文件
```



图 3.6 Cluster 界面

```

6      dfs.audit.logger
4      dfs.class
3      dfs.logger
3      dfs.server.namenode.
2      dfs.audit.log.maxbackupindex
2      dfs.period
2      dfs.audit.log.maxfilesize
1      dfs.replication
1      dfs.log
1      dfs.file
1      dfs.datanode.name.dir
1      dfs.servers
1      dfsadmin
1      dfsmetrics.log
1      dfs.namenode.name.dir

```

图 3.7 grep 示例运行结果

- 查看运行过程中的进程

在客户端提交 WordCount 示例作业，并且查看系统执行该作业过程中启动的进程。

主节点中进程如图3.8所示，从节点中进程如图3.9所示，客户端进程如图3.10所示。

```

1 cd ~/hadoop-2.10.1
2 ./bin/hdfs dfs -put ~/input/pd.train input #上传输入文件
3 ./bin/hdfs dfs -rm -r output/wordcount #删除输出路径
4 ./bin/yarn jar share/hadoop/mapreduce/hadoop-mapreduce-examples-2.10.1.jar
   wordcount input/pd.train output/wordcount

```

(5) 查看 MapReduce 程序运行信息

- 访问 Hadoop Map/Reduce Administration 界面

通过 <http://ecnu01:8088/>，可以看到本次启动 Yarn 后提交的所有应用程序的相关信息，如图3.11所示。

- 查看 MapReduce 应用程序日志

- Task 日志

默认位置: `~/hadoop-2.10.1/logs/userlogs/<jobid>/<attempt-id>`

```
8672 NameNode  
7333 JobHistoryServer  
9029 SecondaryNameNode  
15800 Jps  
6878 ResourceManager
```

图 3.8 主节点中进程

```
dase-dis@ecnu03:~$ jps  
433574 YarnChild  
433572 YarnChild  
433589 YarnChild  
433171 MRAppMaster  
433505 YarnChild  
399121 NodeManager  
399549 DataNode  
433515 YarnChild  
433691 Jps  
433609 YarnChild
```

图 3.9 从节点中进程

9382 RunJan
9465 Jps

图 3.10 客户端中进程

图 3.11 All Applications

- 查看 MapReduce 应用程序历史记录

通过 <http://ecnu01:19888>，即可看到程序的历史运行记录，如图3.12所示。

Retired Jobs											Search:		
Submit Time	Start Time	Finish Time	Job ID	Name	User	Queue	State	Maps Total	Maps Completed	Reduces Total	Reduces Completed	Elapsed Time	
2021-02-02 11:32:03 CST	2021-02-02 11:32:07 CST	2021-02-02 11:37:12 CST	job_1612236145191_0003	word count	dase-dis	default	SUCCEEDED	17	17	1	1	00hrs, 05mins, 05sec	
2021-02-02 11:31:29 CST	2021-02-02 11:31:37 CST	2021-02-02 11:31:38 CST	job_1612236145191_0002	grep sort	dase-dis	default	SUCCEEDED	1	1	1	1	00hrs, 00mins, 01sec	
2021-02-02 11:31:00 CST	2021-02-02 11:31:05 CST	2021-02-02 11:31:26 CST	job_1612236145191_0001	grep search	dase-dis	default	SUCCEEDED	29	29	1	1	00hrs, 00mins, 21sec	

图 3.12 JobHistory

3.4.4 运行 MapReduce 应用程序

在分布式部署的 Hadoop 上运行 WordCount 应用程序。以下操作在各节点均以 dase-dis 用户身份进行：

(1) 将之前打好的 jar 包拷贝至 `/home/dase-dis/hadoop-2.10.1/myApp` 目录下。

(2) 上传输入文件

输入文件 `pd.train` 在上一节中已上传至 `hdfs://ecnu01:9000/user/dase-dis/input/`。

(3) 通过提交 jar 包运行应用程序

在客户端节点执行以下命令：

```

1 cd ~/hadoop-2.10.1/
2 ./bin/hadoop jar ./myApp/WordCount.jar input/pd.train output
3 ./bin/hdfs dfs -tail output/part-r-00000 #查看输出结果

```

该程序实现了统计 HDFS “`/user/dase-dis/input/`” 目录下的 “`pd.train`” 文件单词数量的操作。运行结果存储在 HDFS 的 “`/user/dase-dis/output/`” 目录下的 “`part-r-00000`” 文件中。

(4) 停止服务

在主节点执行以下命令：

```

1 ~/hadoop-2.10.1/sbin/stop-yarn.sh
2 ~/hadoop-2.10.1/sbin/mr-jobhistory-daemon.sh stop historyserver
3 ~/hadoop-2.10.1/sbin/stop-dfs.sh

```

正常停止后，执行 `jps` 命令，各节点不会出现 ResourceManger、NodeManager、JobHistoryServer、NameNode、SecondaryNameNode、DataNode 进程。

3.5 思考题

- 1 根据分布式部署的 HDFS 的 Web UI，能看到 HDFS 架构中的哪些角色？每个角色对应所在主机的 IP 地址是多少？
- 2 根据分布式部署的 Yarn 的 Web UI，能看到 Yarn 架构中的哪些角色？每个角色对应所在主机的 IP 地址是多少？
- 3 能否从 Hadoop Map/Reduce Administration 的 Web UI 中获取某个 MapReduce 应用程序的总执行时间？请简要说明。
- 4 在近两次实验中，单机伪分布式部署方式和分布式部署方式下执行 WordCount 作业的时间分别是多少？
- 5 在本实验中，在分布式部署方式下尝试增加或减少 slaves 文件中的节点（注意：需要同时修改所有机器上的 slaves 文件），观察执行 WordCount 作业的时间变化？
- 6 编写一个利用 Combine 机制实现的 WordCount 程序，使用实验中的 pd.train 数据集，在分布式部署方式下分别运行本实验中的 WordCount 程序和使用 Combine 机制的 WordCount 程序，观察二者的时间差并分析原因。