

# Hadoop配置



## SSH远程登录配置

```
1 # vi /etc/ssh/sshd_conf
2 PermitEmptyPasswords yes
3 PasswordAuthentication yes
4 PermitRootLogin yes
5 # 注释掉有个重复的 no
6
7 [root] systemctl restart sshd
```

## 创建单独为hadoop工作的hadoop用户

```
1 [root] useradd hadoop
2 [root] passwd hadoop
```

## 安装java

```
1 yum install java-11-openjdk java-11-openjdk-devel
2
3 # 在系统环境中配置java环境变量
4 [root] vi /etc/profile
5 export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-11.0.14.1.1-1.el7_9.x86_64
6 export PATH=$PATH:$JAVA_HOME/bin
7
8 [root] source /etc/profile
```

## 规划服务器角色

1	bigdata100	namenode	datanode	(主节点、数据节点、namenode)
2	bigdata99	Secondarynamenode	datanode	(备份主节点、数据节点)
3	bigdata98	dn1	datanode	(从节点、数据节点)

## 域名解析

```
1 [root] vi /etc/hosts
2 #!!!!!!!!!!!! 用192的域名，别用172的
3 172.xxx.xxx.xxx    namenode
4 172.xxx.xxx.xxx    Secondarynamenode
5 172.xxx.xxx.xxx    dn1
```

## 分发密钥

```
1 # 生成密钥
2 [root] ssh-keygen -t rsa
3 [root] ssh-copy-id Secondarynamenode
4 [root] ssh-copy-id dn1
5 [root] ssh-copy-id namenode # 给自己一个密钥，后面用到
```

## 配置Hadoop

```
1 # 安装wget和用wget下载hadoop
2 [root] yum install wget
3 [root] su hadoop
4 [hadoop] cd
5 [hadoop] wget https://archive.apache.org/dist/hadoop/common/hadoop-
3.3.2/hadoop-3.3.2.tar.gz
6 # 解压hadoop到~
7 [hadoop] tar -xvf hadoop-3.3.2/hadoop-3.3.2.tar.gz
8 # 创建data文件夹存数据
9 [hadoop ~] mkdir data
10
11
12 # hadoop命令存放在/home/hadoop/hadoop3.3.2/bin下
13 # 把路径加入到hadoop环境变量中，以后不用写路径调命令
14 # 用户的环境变量修改.bashrc即可
15 [hadoop] cd
16 [hadoop] vi .bashrc
17 export HADOOP_HOME=/home/hadoop/hadoop-3.3.2
18 export PATH=$PATH:$HADOOP_HOME/bin:$HADOOP_HOME/sbin
19
20 [hadoop] source .bashrc
21 # 测试下
22 [hadoop ~] hadoop version
23 # 出现下面内容即成功
24 #Hadoop 3.3.2
25 #Source code repository git@github.com:apache/hadoop.git -r
#0bcb014209e219273cb6fd4152df7df713cbac61
26 #Compiled by chao on 2022-02-21T18:39Z
27 #Compiled with protoc 3.7.1
28 #From source with checksum 4b40fff8bb27201ba07b6fa5651217fb
29 #This command was run using /home/hadoop/hadoop-
#3.3.2/share/hadoop/common/hadoop-common-3.3.2.jar
```

## 编辑hadoop配置文件

```
1 # 位置 /home/hadoop/hadoop-3.3.2/etc/hadoop
2 [hadoop] cd /home/hadoop/hadoop-3.3.2/etc/hadoop
3 [hadoop] vi core-site.xml
4 # 修改添加core-site.xml内容， 确保data文件夹已创建
5 <configuration>
6     <property>
7         <name>fs.defaultFS</name>
8         <value>hdfs://namenode:9000</value>
9     </property>
10    <property>
```

```
11         <name>hadoop.tmp.dir</name>
12         <value>file:/home/hadoop/data/tmp</value>
13     </property>
14 </configuration>
15
16
17 [hadoop] vi hdfs-sit.xml
18 # 修改添加hdfs-sit.xml内容， 确保data文件夹已创建
19 <configuration>
20     <property>
21         <name>dfs.namenode.secondary.http-address</name>
22         <value>Secondarynamenode:50090</value>
23     </property>
24     <property>
25         <name>dfs.replication</name>
26         <value>3</value>
27     </property>
28     <property>
29         <name>dfs.namenode.name.dir</name>
30         <value>file:/home/hadoop/data/name</value>
31     </property>
32     <property>
33         <name>dfs.datanode.data.dir</name>
34         <value>file:/home/hadoop/data/data</value>
35     </property>
36 </configuration>
37
38 # 修改添加mapred-site.xml内容
39     <property>
40         <name>mapreduce.framework.name</name>
41         <value>yarn</value>
42     </property>
43     <property>
44         <name>mapreduce.jobhistory.address</name>
45         <value>namenode:10020</value>
46     </property>
47     <property>
48         <name>mapreduce.jobhistory.webapp.address</name>
49         <value>namenode:19888</value>
50     </property>
51 # 修改添加yarn-site.xml内容
52     <property>
53         <name>yarn.resourcemanager.hostname</name>
54         <value>namenode</value>
55     </property>
56     <property>
57         <name>yarn.nodemanager.aux-services</name>
58         <value>mapreduce_shuffle</value>
59     </property>
60 # 修改添加workers内容
61 # 删除localhost
62     namenode
63     Secondarynamenode
64     dn1
65 # 修改hadoop-env.sh
66 export HADOOP_HOME=/home/hadoop/hadoop-3.3.2
67 export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-11.0.14.1.1-1.e17_9.x86_64
68
```

## 分发hadoop

```
1 [hadoop] scp -r hadoop-3.3.2 hadoop@Secordarynamenode:~/
2 [hadoop] scp -r hadoop-3.3.2 hadoop@dn1:~/
```

## 给本机一个ssh

```
1 [hadoop] ssh-copy-id namenode
```

## 启动

```
1 hdfs namenode -format
2
3 start-dfs.sh
4 start-yarn.sh
5 jps
```

## 文件操作

```
1 hadoop fs -put [from] [target]
2 hadoop fs -get
3 hadoop fs -ls /
```

## Web查看

```
1 [master ip]:9870
```

つ\_●\_●つ ➡ 出错多看logs! ! ! ! !

配Hadoop配麻了! 😞



