# Hadoop-2.7.2 分布式安装指南

--李琦

## 1.机器信息

五台centos 64位机器

## 2.集群规划

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Server Name** | **Hadoop Cluster** | **Zookeeper Ensemble** | **HBase Cluster** | **Ip** |
| **Hadoop01** | Name node & Resource manager |  | Master | 192.168.12.60 |
| **Hadoop02** | Secondary name node |  |  | 192.168.12.61 |
| **Hadoop03** | Data node & Node manager | √ | Region server | 192.168.12.62 |
| **Hadoop04** | Data node & Node manager | √ | Region server | 192.168.12.63 |
| **Hadoop05** | Data node & Node manager | √ | Region server | 192.168.12.64 |

## 3. hadoop 集群

#### 3.1core-site.xml

<configuration>

<property>

<name>hadoop.tmp.dir</name>

<value>/usr/app/hadoop-2.7.2</value>

<description>A base for other temporary directories.</description>

</property>

<property>

<name>fs.defaultFS</name>

<value>hdfs://hadoop01:9000</value>

</property>

<property>

<name>io.file.buffer.size</name>

<value>131072</value>

</property>

</configuration>

### 3.2hdfs-site.xml

<configuration>

<property>

<name>dfs.namenode.name.dir</name>

<value>file:/root/hadoopdata/namenode</value>

</property>

<property>

<name>dfs.datanode.data.dir</name>

<value>file:/root/hadoopdata/datanode</value>

</property>

<property>

<name>dfs.namenode.secondary.http-address</name>

<value>hadoop02:9001</value>

</property>

<property>

<name>io.file.buffer.size</name>

<value>131072</value>

</property>

<property>

<name>dfs.namenode.handler.count</name>

<value>100</value>

</property>

</configuration>

### 3.3mapred-site.xml

<configuration>

<property>

<name>mapreduce.framework.name</name>

<value>yarn</value>

</property>

<property>

<name>mapreduce.jobhistory.address</name>

<value>hadoop01:10020</value>

</property>

<property>

<name>mapreduce.jobhistory.address</name>

<value>hadoop01:10020</value>

</property>

</configuration>

### 3.4yarn-site.xml

<property>

<name>yarn.nodemanager.aux-services</name>

<value>mapreduce\_shuffle</value>

</property>

<property>

<name>yarn.nodemanager.aux-services.mapreduce.shuffle.class</name>

<value>org.apache.hadoop.mapred.ShuffleHandler</value>

</property>

<property>

<name>yarn.resourcemanager.resource-tracker.address</name>

<value>hadoop01:8035</value>

</property>

<property>

<name>yarn.resourcemanager.scheduler.address</name>

<value>hadoop01:8030</value>

</property>

<property>

<name>yarn.resourcemanager.address</name>

<value>hadoop01:8032</value>

</property>

<property>

<name>yarn.acl.enable</name>

<value>false</value>

</property>

<property>

<name>yarn.admin.acl</name>

<value>\*</value>

</property>

<property>

<name>yarn.log-aggregation-enable</name>

<value>false</value>

</property>

<property>

<name>yarn.resourcemanager.webapp.address</name>

<value>hadoop01:8088</value>

</property>

<property>

<name>yarn.resourcemanager.hostname</name>

<value>hadoop01</value>

</property>

<property>

<name>yarn.nodemanager.aux-services</name>

<value>mapreduce\_shuffle</value>

</property>

### 3.5slaves

hadoop03

hadoop04

hadoop05

### 3.6hadoop-env.sh

修改java\_home 改成绝对路径

export JAVA\_HOME=/usr/app/jdk1.7.0\_51

## 4修改host文件

命令 vim /etc/hosts

192.168.12.60 hadoop01

192.168.12.61 hadoop02

192.168.12.62 hadoop03

192.168.12.63 hadoop04

192.168.12.64 hadoop05

### 5.修改环境变量

vim /etc/profile

编辑如下：

HADOOP\_PREFIX=/usr/app/hadoop-2.7.2

HADOOP\_MAPRED\_PREFIX=$HADOOP\_PREFIX

HADOOP\_HDFS\_PREFIX=$HADOOP\_PREFIX

HADOOP\_YARN\_PREFIX=$HADOOP\_PREFIX

PATH=$PATH:$HADOOP\_PREFIX/bin

PATH=$PATH:$HADOOP\_PREFIX/sbin

## 6配置ssh 免密码登录

命令 cd ~/.ssh

发现 没有问价 那么就自己创建一个.ssh 文件

利用ls –al 来查看文件

再执行 命令 cd ~/.ssh

执行 ssh-keygen -t rsa （四个回车）

命令ssh-copy-id localhost 和其他机器

## 7 集群配置

将配置好的hadoop-2.7.2文件发送到各个机器上 并修改各个机器的配置文件

### 8启动集群

命令 start-dfs.sh 显示除了本机之外的权限不够 Permission denied

登录其他机器 修改调用文件的权限 例如 chmod 777 文件名

## 8总结

遇到的问题 我在配置的时候 ssh 时候一直配置不成功 当时使用的命令是scp ~/.ssh/id\_rsa.pub [root@192.168.0.2:/root/.ssh](mailto:root@192.168.0.2:/root/.ssh)

后来使用ssh-copy-id 主机名 还有一个问题就是启动hdfs 时候 目标节点显示权限不够 (很困扰)困在这里很长时间。

修改权限之后就可以运行了。