# Hadoop2.6.0伪分布式设置

准备工作：把JDK和Hadoop安装包上传到linux系统（oracle用户的根目录）

## Jdk安装

oracle用户替换为你的用户

1. 在oracle用户的根目录，Jdk解压，（hadoop用户操作）

tar -zxvf jdk-8u121-linux-x64.tar.gz

解压完成后，在oracle用户的根目录有一个jdk1.8.0\_121目录

1. 配置环境变量，需要修改/etc/profile文件（root用户操作）

切到root用户，输入su命令

vi /etc/profile

进去编辑器后，输入i，进入vi编辑器的插入模式

在profile文件最后添加

|  |
| --- |
| JAVA\_HOME=/home/hadoop/jdk1.8.0\_121  export PATH=$PATH:$JAVA\_HOME/bin |

编辑完成后，按下esc退出插入模式

输入：，这时在左下角有一个冒号的标识

q 退出不保存

wq保存退出

q! 强制退出

1. 把修改的环境变量生效（hadoop用户操作）

执行source /etc/profile

## Hadoop安装

1. 在hadoop用户的根目录，解压（hadoop用户操作）

tar -zxvf hadoop-2.6.0.tar.gz

解压完成在oracle用户的根目录下有一个hadoop-2.6.0目录

1. 修改配置文件hadoop-2.6.0/etc/hadoop/hadoop-env.sh（hadoop用户操作）

export JAVA\_HOME=/home/hadoop/jdk1.8.0\_121

1. 修改配置文件hadoop-2.6.0/etc/hadoop/core-site.xml，添加（hadoop用户操作）

|  |
| --- |
| <property>  <name>fs.defaultFS</name>  <value>hdfs://oracle:9000</value>  </property> |

1. 修改配置文件hadoop-2.6.0/etc/hadoop/hdfs-site.xml，添加（hadoop用户操作）

|  |
| --- |
| <property>  <name>dfs.replication</name>  <value>1</value>  </property>  把HDFS的储存信息放到一个安全地方  在hdfs-site.xml文件里设置  <property>  <name>dfs.name.dir</name>  <value>/home/hadoop/appData/hadoop/name</value>  </property>  <property>  <name>dfs.data.dir</name>  <value>/home/hadoop/appData/hadoop/data</value>  </property> |

1. 修改修改配置文件hadoop-2.6.0/etc/hadoop/mapred-site.xml（hadoop用户操作）

这个文件没有，需要复制一份

cp etc/hadoop/mapred-site.xml.template etc/hadoop/mapred-site.xml

添加

|  |
| --- |
| <property>  <name>mapreduce.framework.name</name>  <value>yarn</value>  </property> |

1. 修改配置文件hadoop-2.6.0/etc/hadoop/yarn-site.xml，添加（hadoop用户操作）

|  |
| --- |
| <property>  <name>yarn.nodemanager.aux-services</name>  <value>mapreduce\_shuffle</value>  </property> |

1. 修改/etc/hosts文件（root用户操作）,添加：ip主机名称

192.168.199.100 oracle

1. 格式化HDFS，在hadoop解压目录下，执行如下命令：（hadoop用户操作）

bin/hdfs namenode -format

有必要吐槽一句 横线是英文的 上次写中文的鼓捣了一天 醉了

注意：格式化只能操作一次，如果因为某种原因，集群不能用，需要再次格式化，需要把上一次格式化的信息删除，在/tmp目录里执行rm–rf \*

1. 启动集群，在hadoop解压目录下，执行如下命令：（hadoop用户操作）

启动集群：sbin/start-all.sh需要输入四次当前用户的密码(通过配置ssh互信解决)

启动后，在命令行输入jps有以下输出

|  |
| --- |
| [oracle@oracle hadoop-2.6.0]$ jps  32033 Jps  31718 SecondaryNameNode  31528 DataNode  31852 ResourceManager  31437 NameNode  31949 NodeManager |

关闭集群：sbin/stop-all.sh需要输入四次当前用户的密码(通过配置ssh互信解决)

## SSH互信配置（hadoop用户操作）

rsa加密方法，公钥和私钥

1. 生成公钥和私钥

在命令行执行ssh-keygen，然后回车，然后会提示输入内容，什么都不用写，一路回车

在oracle用户根目录下，有一个.ssh目录

id\_rsa 私钥

id\_rsa.pub 公钥

known\_hosts 通过SSH链接到本主机，都会在这里有记录

1. 把公钥给信任的主机(本机)（相当于分布式中把公钥给从节点）

在命令行输入ssh-copy-id 主机名称

ssh-copy-id oracle

复制的过程中需要输入信任主机的密码

1. 验证，在命令行输入：ssh信任主机名称

ssh oracle

如果没有提示输入密码，则配置成功