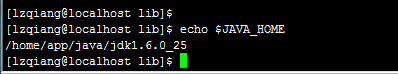
1. 配置java环境变量

Echo $JAVA\_HOME



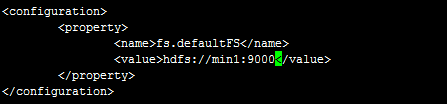
配置主机映射

vim /etc/hosts

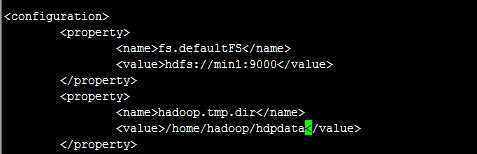
2编辑 /etc/hadoop 下的hadoop-env.sh



1. 配置 core-site.xml

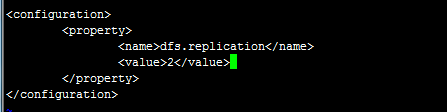


默认文件系统：使用hdfs文件系统 就是namenode地址



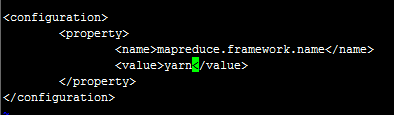
Hadoop进程产生的临时文件（数据目录）

1. 配置 hdfs-site.xml



副本的个数 集群中文件都会存2份 默认是3

1. 配置mapred-site.xml



默认值是：local 在本地跑 mapreduce程序

Mapreduce运行环境是yarn

修改名称



1. 修改yarn-site.xml

<!-- 指定YARN的老大（ResourceManager）的地址 -->

<property>

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

<value>min1</value>

</property>

<!-- reducer获取数据的方式 -->

<property>

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

<value>mapreduce\_shuffle</value>

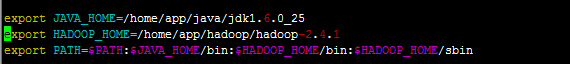
</property>

1. 将配置好的 分发到其他机器上



1. 配置hadoop环境变量

Vim /etc/profile

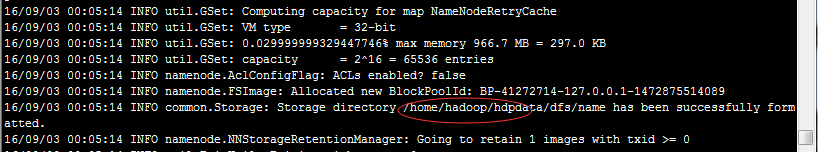


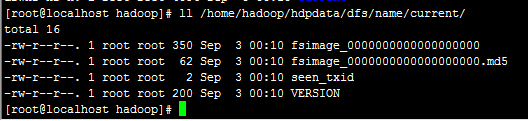
刷新环境变量



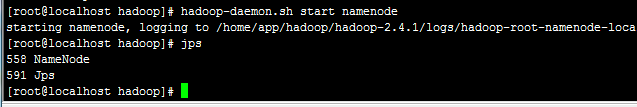
1. hdfs需要格式化





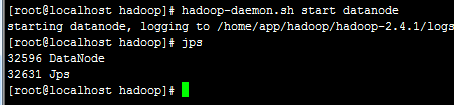


1. 启动namenode

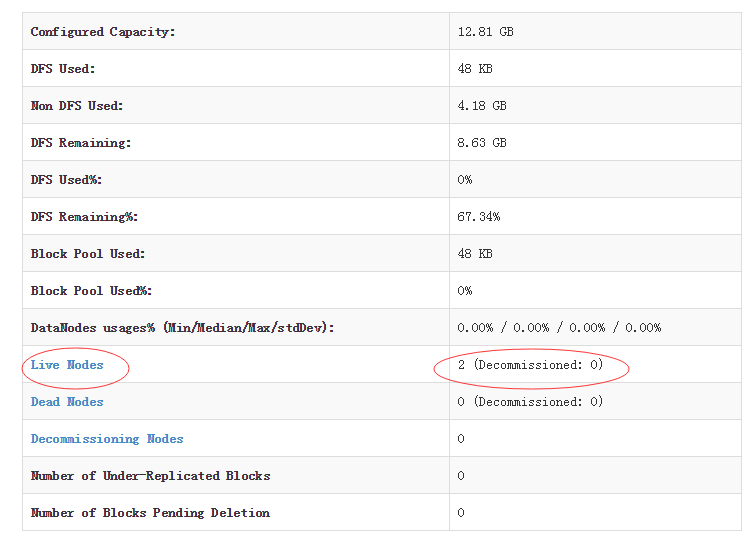


浏览器访问 <http://192.168.88.194:50070/>

在min2中启动 datanode



启动两台datanode



4.配置ssh免登陆

#生成ssh免登陆密钥

#进入到我的home目录

cd ~/.ssh

ssh-keygen -t rsa （四个回车）

执行完这个命令后，会生成两个文件id\_rsa（私钥）、id\_rsa.pub（公钥）

将公钥拷贝到要免密登陆的目标机器上

ssh-copy-id localhost





一次性启动多台

要在slaves中添加 要启动的机器地址