**Hadoop完全分布式搭建手册（图片28-）**

**需要去官网提前准备：**

**JDK、Hadoop**

**JDK：[www.orcle.com](http://www.orcle.com)**

**Hadoop：[www.apache.org](http://www.apache.org)**

**Linux:**

**lsof**

**Ps -ef| grep zookeeper**

**整体步骤：**

**确定好总共几个节点**

**第一台虚拟机**

1. **创建 指定用户hadoop**

**涉及到权限问题**

**将文件使用工具xftp 5传到指定目录下**

**ll 查看该目录下的文件信息（包含权限和归属）**

**更改文件权限root模式下:**

**如果文件归属于其他用户，可以将文件强制归属hadoop用户**

**chown -R hadoop:hadoop hadoop-2.7.3/**

**提升权限**

**切换到root权限下 vi /etc/sudoers**

**如果root权限下无法修改此文件**

**修改文件权限：chmod 777 /etc/sudoers**

**在root ALL```(需要补图)**

**下添加hadoop all....**

**修改完后需要修改回去：chmod 400 /etc/sudoers**

1. **在指定用户hadoop下解压JDK和Hadoop**
2. **在指定用户hadoop下JDK和Hadoop的路径和环境配置**

**vi ~/.bashrc 是当前用户下配置**

**vi /etc/profile是对这台虚拟机中所有用户进行配置**

**这里需要编辑完后source一下**

**免密设置（疑惑）**

**一．**

**是第一台虚拟机**

**su hadoop**

**ssh-keygen –t rsa**

**一直回车创建后密钥对**

**在最后的时候，克隆两台机子，再将**

**ssh-copy-id -i ~/.ssh/id\_rsa.pub hadoop@node1**

**ssh-copy-id -i ~/.ssh/id\_rsa.pub hadoop@node2**

**ssh-copy-id -i ~/.ssh/id\_rsa.pub hadoop@node3**

**chmod 0600 ~/.ssh/authorized\_keys**

1. **还是其他配置完成后，将第一台虚拟机克隆两台后**

**再来每一台机子**

**su hadoop**

**ssh-keygen –t rsa**

**再在一台机子上**

**ssh-copy-id -i ~/.ssh/id\_rsa.pub hadoop@node1**

**ssh-copy-id -i ~/.ssh/id\_rsa.pub hadoop@node2**

**ssh-copy-id -i ~/.ssh/id\_rsa.pub hadoop@node3**

**chmod 0600 ~/.ssh/authorized\_keys**

1. **配置hadoop**

**4个常规文件**

**注意yarn-env.sh下JDK路径、hadoop-env.sh java路径**

**（一定都需要？还是自己hadoop包有问题，待验证）**

**systemctl status firewalld.service**

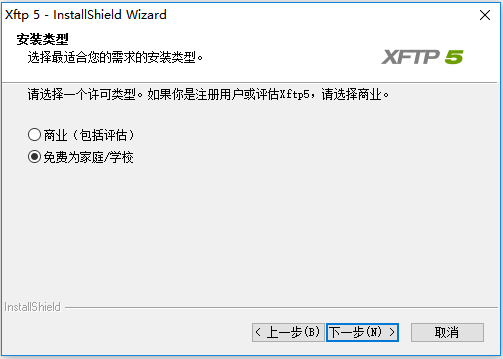
**systemctl stop firewalld.service**

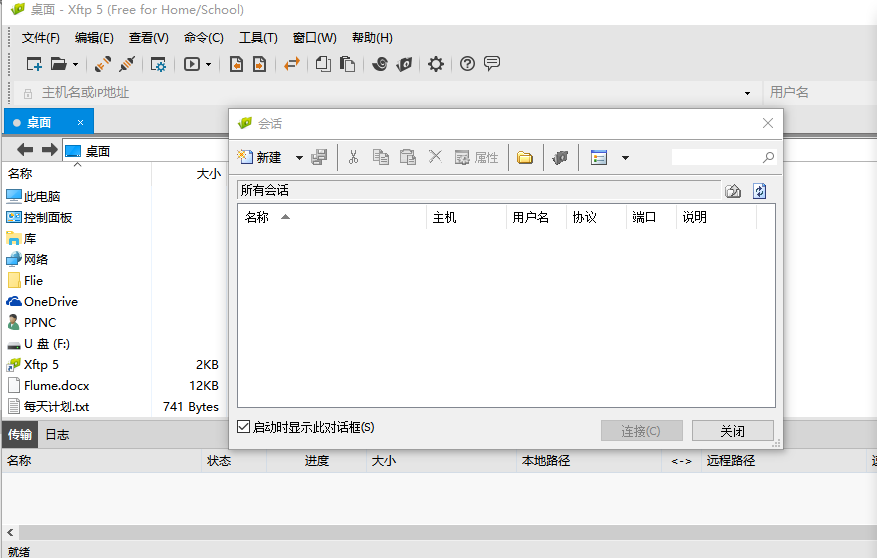
**systemctl disable firewalld.service 禁用防火墙服务(如果当前没有关闭，是不会关闭的，需要stop)**

**下载好后是在本地电脑上的，并没有在虚拟机中**

**这时需要使用软件将本地文件传入虚拟机中**

**这里推荐:xftp 5**

****

****

**CentOS7安装指南 里面已经设置好网络配置**

**点击文件下的图中图标即可创建新的会话**

**30.1**

****

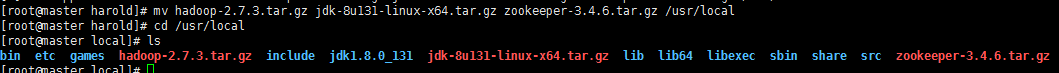
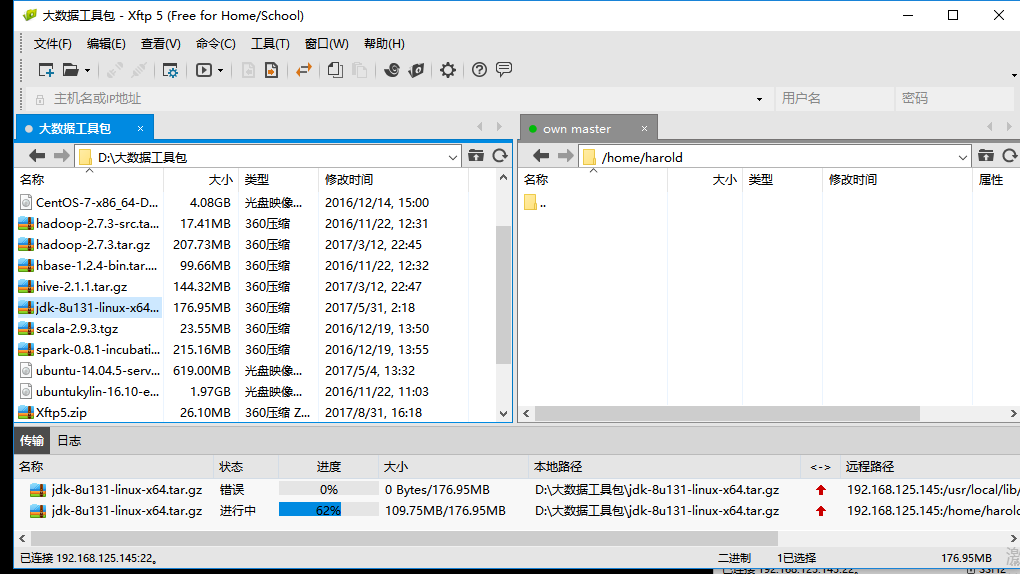
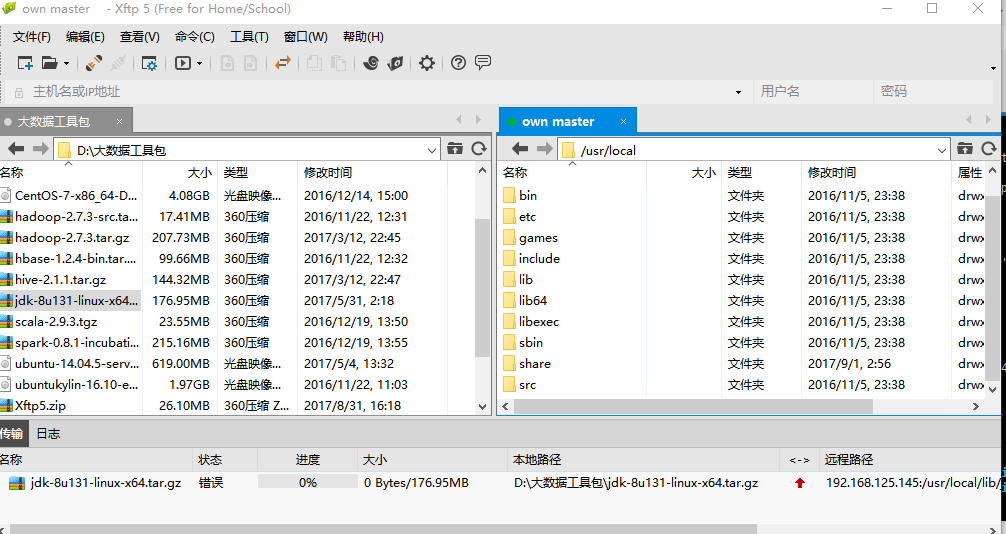
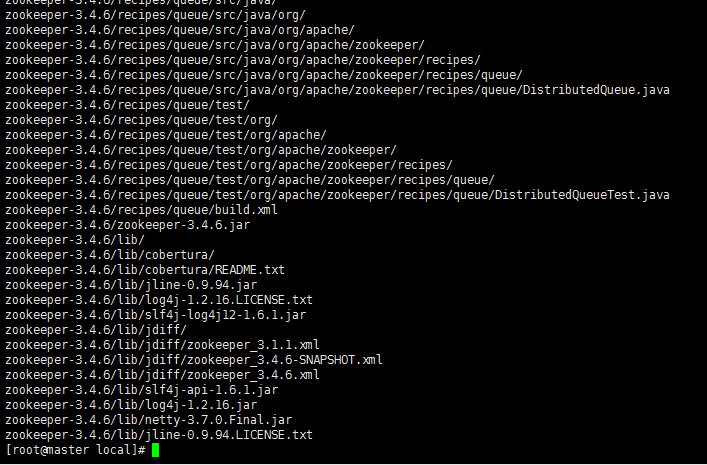
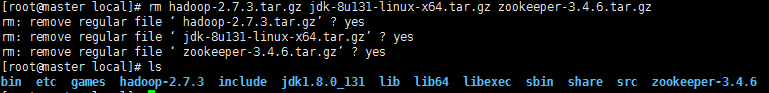
**将需要传送文件的IP和账号密码填入**

**这里协议需要改为SFTP**

**FTP和SFTP相同点：都是文件传输协议，两者中的FTP指的都是**

**FTP：**

**SFTP：**

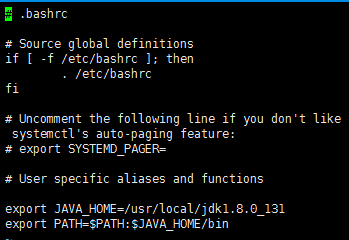
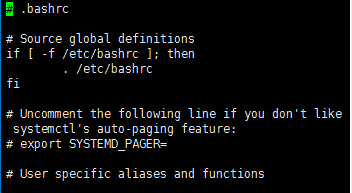
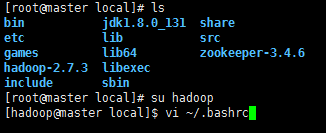
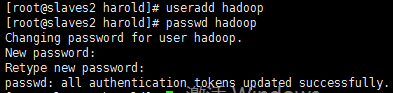
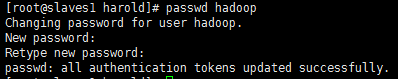
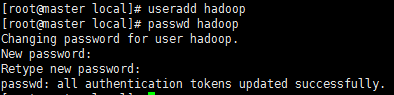
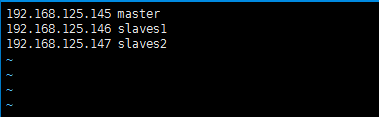
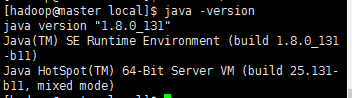
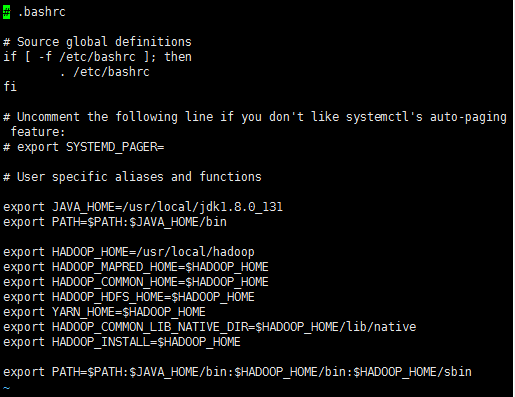
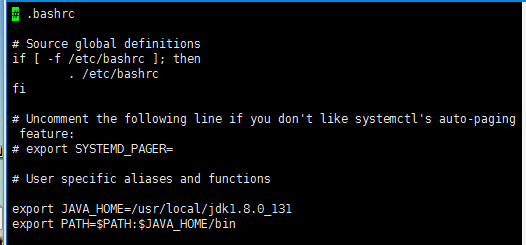
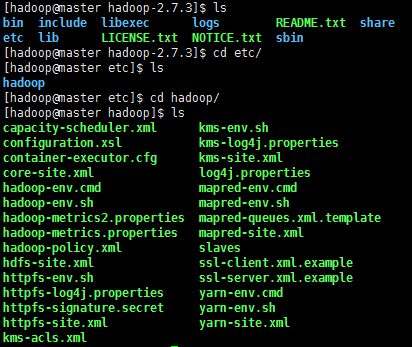
**3638**

**40.0**

**40.1**

**40.2**

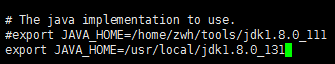
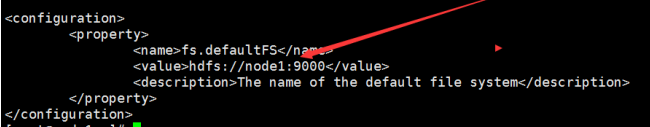
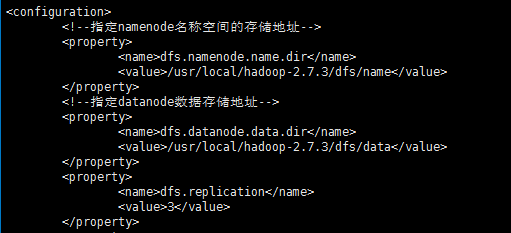
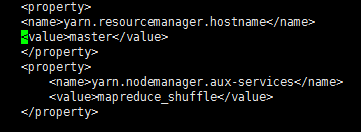
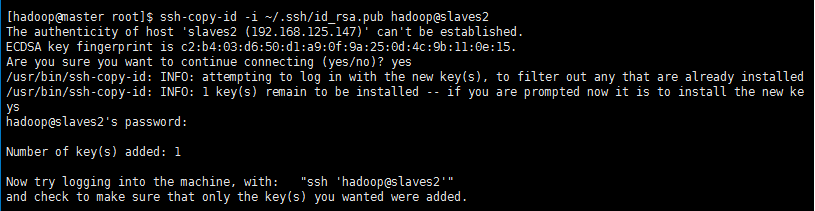
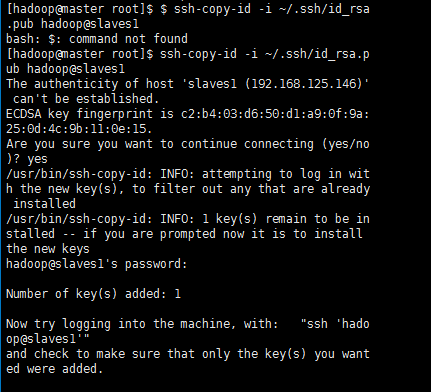
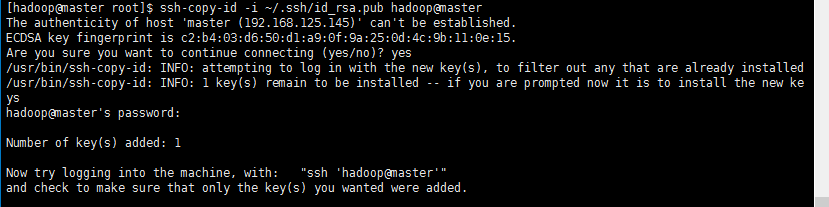
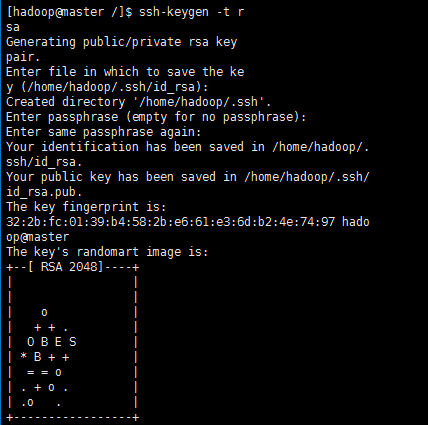
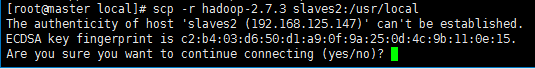
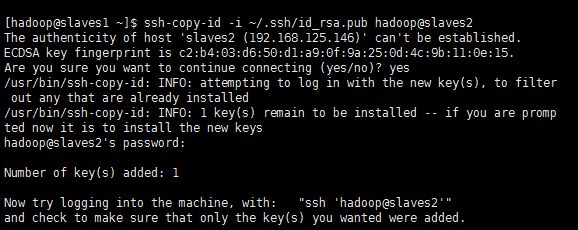
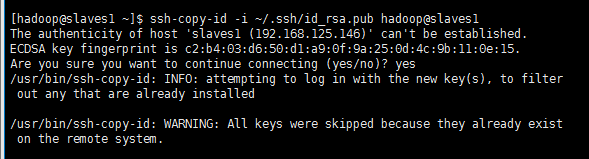
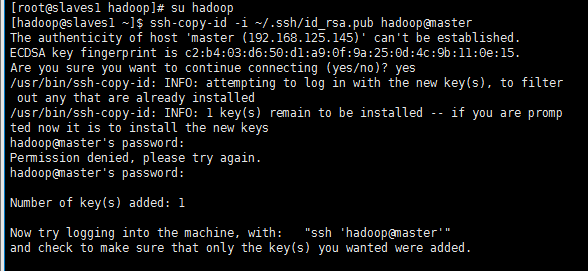
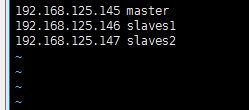
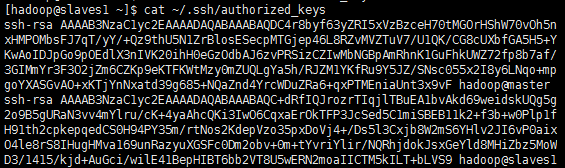
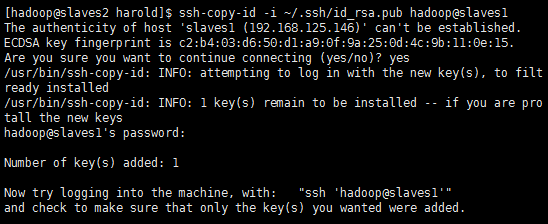
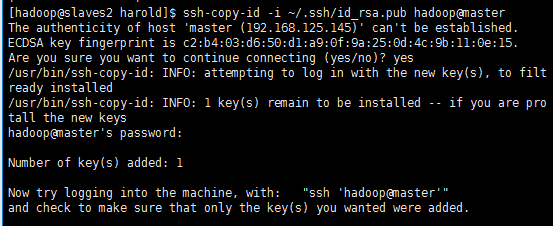
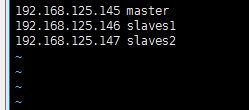
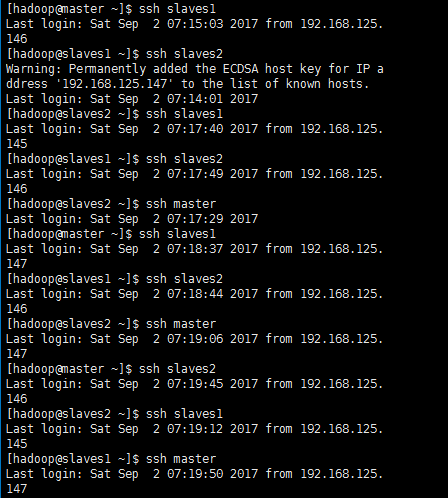
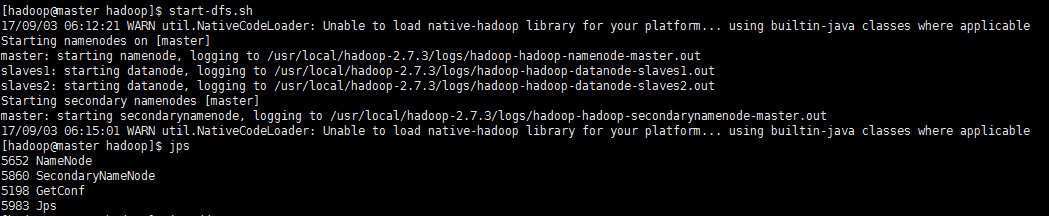
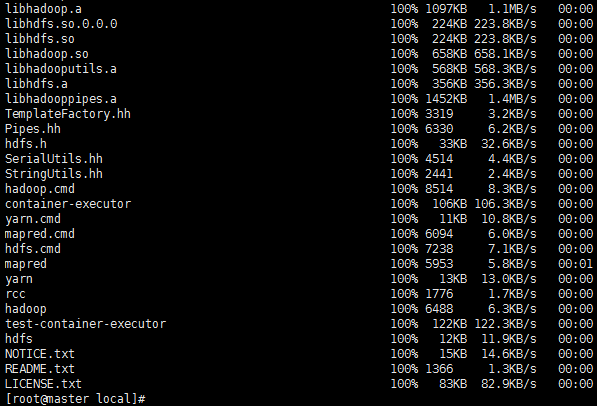
**40.3**

**41495154**

**为了用java开发hadoop程序，需要修改hadoop-env.sh中的JAVA\_HOME，指向系统的java路径**

**export JAVA\_HOME=/usr/local/jdk1.7.0\_71**

**还有yarn-env.sh**

**576063656671.07577818386**

**87**

**第一台master全部配好后**

**只需要克隆2台slaves**

1. **vi /etc/hostname**

**2.vi /etc/sysconfig/network-scripts/ifcfg-ens33**

**namenode出错 解决方案：**

**namenode 默认在/tmp下建立临时文件，但关机后，/tmp下文档自动删除，再次启动Master造成文件不匹配，所以namenode启动失败。  
在core-site.xml中指定临时文件位置，然后重新格式化，终极解决！  
<property>  
<name>hadoop.tmp.dir</name>  
<value>/usr/grid/hadoop1.7.0\_17/hadoop\_${user.name}</value>  
<property>  
value中的路径只要不是/tmp 就行**

**Hadoop 50070无法访问问题解决**

1. **检查防火墙（尤其是重启过的虚拟机，防火墙会自动开启，需要关闭这项服务）**
2. **Root下：vi /etc/selinux/config**

**将SELINUX=enforcing改为SELINUX=disabled**

1. **在hadoop-env.sh中设置java绝对路径**
2. **查看$HADOOP\_HOME/etc/hadoop下的core-site.xml和hdfs-site.xml是否配置好**