

环境搭建

一、 配置主机.....	2
二、 关闭防火墙.....	2
三、 安装 ssh, Xshell, FileZilla.....	2
四、 JDK 安装.....	2
五、 ssh 免密登录.....	3
六、 搭建伪分布式 Hadoop.....	3
七、 安装 Scala.....	7
八、 安装 Zookeeper.....	8
九、 安装 Hbase.....	10
十、 安装 spark.....	12
十一、 配置本机环境.....	14
文件资源.....	17
参考文献：	18

一、 配置主机

1. 配置 hosts 文件

```
vi /etc/hosts
```

添加以下内容并保存：(删除有 localhost 的行)

```
192.168.114.135 master
```

注：搭建伪分布式只需设置一个

二、 关闭防火墙

1. 执行命令: service iptables stop

验证: service iptables status

三、 安装 ssh, Xshell, FileZilla

1. 安装 ssh

安装 ssh: apt-get install openssh-server

启动 ssh: /etc/init.d/ssh start

检测 SSH 进程是否已经开启: ps -e| grep sshd

2. 查看服务器 ip: ifconfig

3. 安装使用 Xshell

打开 Xshell——新建会话

输入用户名和密码，连接成功

打开 filezilla 新建站点并连接

上传所需安装包

四、 JDK 安装

(安装包放在/home/lxy/software 下， 软件放在/home/lxy/install 下)

1. 解压 jdk:

```
sudo tar -zxvf jdk-8u251-linux-x64.tar.gz -C ../install
```

```
sudo mv jdk-8u251 jdk
```

2. 配置环境变量: sudo vi /etc/profile

3. 添加以下内容并保存

```
export JAVA_HOME=/home/lxy/install/jdk
```

```

export JRE_HOME=$JAVA_HOME/jre
export CLASSPATH=.:$JAVA_HOME/lib:$JRE_HOME/lib
export PATH=$PATH:$JAVA_HOME/bin:$JRE_HOME/bin
done
unset i
fi

export JAVA_HOME=/home/lxy/install/jdk
export JRE_HOME=$JAVA_HOME/jre
export CLASSPATH=.::$JAVA_HOME/lib:$JRE_HOME/lib
export PATH=$PATH:$JAVA_HOME/bin:$JRE_HOME/bin

```

4. 使环境配置生效: source /etc/profile
5. 查看是否配置成功: java -version

```

lxy@master:~/install/jdk$ java -version
java version "1.8.0_251"
Java(TM) SE Runtime Environment (build 1.8.0_251-b08)
Java HotSpot(TM) 64-Bit Server VM (build 25.251-b08, mixed mode)

```

五、 ssh 免密登录

1. 设置免密登录: ssh-keygen -t rsa (一直回车即可)
2. 导入 authorized_keys: cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
3. 检测: ssh localhost

六、 搭建伪分布式 Hadoop

1. 配置 Hadoop 环境

- 1) 解压:

```

tar -zvxf hadoop-2.7.7.tar.gz -C ./install
sudo mv hadoop-2.7.7 hadoop

```

- 2) 配置环境: vi /etc/profile 增加内容如下:

```

export HADOOP_HOME=/home/lxy/install/hadoop
export PATH=.:$HADOOP_HOME/bin:$HADOOP_HOME/sbin:$PATH

```

```

done
unset i
fi

export JAVA_HOME=/home/lxy/install/jdk
export JRE_HOME=$JAVA_HOME/jre
export CLASSPATH=.::$JAVA_HOME/lib:$JRE_HOME/lib
export PATH=$PATH:$JAVA_HOME/bin:$JRE_HOME/bin

export HADOOP_HOME=/home/lxy/install/hadoop
export PATH=.:$HADOOP_HOME/bin:$HADOOP_HOME/sbin:$PATH

```

- 3) 使配置生效: source /etc/profile

4) 验证: hadoop

```
lxy@master:~/install$ hadoop
Usage: hadoop [--config confdir] [COMMAND | CLASSNAME]
  CLASSNAME          run the class named CLASSNAME
  or
  where COMMAND is one of:
    fs                  run a generic filesystem user client
    version             print the version
    jar <jar>           run a jar file
                        note: please use "yarn jar" to launch
                        YARN applications, not this command.
    checknative [-a|-h]  check native hadoop and compression libraries availability
    distcp <srcurl> <desturl> copy file or directories recursively
    archive -archiveName NAME -p <parent path> <src>* <dest> create a hadoop archive
    classpath           prints the class path needed to get the
    credential          interact with credential providers
    daemonlog           Hadoop jar and the required libraries
    trace               get/set the log level for each daemon
                        view and modify Hadoop tracing settings
```

2. 修改 conf 目录下的配置文件 hadoop-env.sh、core-site.xml、hdfs-site.xml、mapred-site.xml

1) 进入文件目录

```
lxy@master:~/install/hadoop$ ls
bin etc include lib libexec LICENSE.txt NOTICE.txt README.txt sbin share
lxy@master:~/install/hadoop$ cd etc/hadoop/
lxy@master:~/install/hadoop/etc/hadoop$ ls
capacity-scheduler.xml      httpfs-env.sh      mapred-env.sh
configuration.xsl            httpfs-log4j.properties  mapred-queues.xml.template
container-executor.cfg       httpfs-signature.secret  mapred-site.xml.template
core-site.xml                httpfs-site.xml     slaves
hadoop-env.cmd              kms-acls.xml      ssl-client.xml.example
hadoop-env.sh                kms-env.sh        ssl-server.xml.example
hadoop-metrics2.properties  kms-log4j.properties  yarn-env.cmd
hadoop-metrics.properties   kms-site.xml      yarn-env.sh
hadoop-policy.xml           log4j.properties   yarn-site.xml
hdfs-site.xml                mapred-env.cmd
```

2) sudo vi hadoop-env.sh

```
export JAVA_HOME=/home/lxy/install/jdk
```

```
# The java implementation to use.
export JAVA_HOME=/home/lxy/install/jdk
```

3) sudo vi core-site.xml (黄色强调部分改为自己主机名)

```
<property>
  <name>fs.defaultFS</name>
  <value>hdfs://master:8020</value>
  <!--value 填的是默认文件系统的 URL 地址，格式是
  "hdfs://host:port",host 为 ip 或者主机名, port 默认为 8020-->
```

```
</property>
```

```
<property>
```

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

```
  <value>/home/lxy/install/hadoop/tmp</value>
```

```
  <!--sda-->
```

```
</property>
```

```
<!--如果需要用到 spark 的话, 需要在 core-site.xml 中加入如下:【否则会出错】-->
```

```
<property>
```

```
  <name>hadoop.proxyuser.root.hosts</name>
```

```
  <value>*</value>
```

```

</property>
<property>
    <name>hadoop.proxyuser.root.groups</name>
    <value>*</value>
</property>

<configuration>

<property>
    <name>fs.defaultFS</name>
    <value>hdfs://master:8020</value>
    <!--value填的是默认文件系统的URL地址，格式是“hdfs://host:port”，host为ip或者主机名，port默认为8020-->
</property>
<property>
    <name>hadoop.tmp.dir</name>
    <value>/home/lxy/install/hadoop/tmp</value>
    <!--sda-->
</property>

<!--如果需要用到spark的话，需要在core-site.xml中加入如下：【否则会出错】-->

<property>
    <name>hadoop.proxyuser.root.hosts</name>
    <value>*</value>
</property>
<property>
    <name>hadoop.proxyuser.root.groups</name>
    <value>*</value>
</property>

</configuration>

```

hadoop 目录下创建 777 权限 tmp: mkdir -m 777 tmp

4) sudo vi hdfs-site.xml

```

<configuration>

<property>
    <name>dfs.replication</name>
    <value>1</value>
    <!--默认为 3， 搭建伪分布式时只有一个节点、改为 1 即可
-->
</property>
</configuration>

<configuration>

<property>
    <name>dfs.replication</name>
    <value>1</value>
    <!--默认为3， 搭建伪分布式时只有一个节点、改为1即可-->
</property>

</configuration>

```

```

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this file. -->

<configuration>
<property>
    <name>dfs.namenode.http-address</name>
    <value>master:50070</value>
</property>
<property>
    <name>dfs.namenode.secondary.http-address</name>
    <value>master:50090</value>
</property>
<property>
    <name>dfs.replication</name>
    <value>1</value>
</property>
<property>
    <name>dfs.permissions</name>
    <value>false</value>
</property>
<property>
    <name>dfs.namenode.name.dir</name>
    <value>/usr/local/hadoop/name</value>
</property>
<property>
    <name>dfs.datanode.data.dir</name>
    <value>/usr/local/hadoop/data</value>
</property>
</configuration>

```

5) 先: cp mapred-site.xml.template mapred-site.xml

再: sudo vi mapred-site.xml

<configuration>

<property>

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

<value>yarn</value>

<!--value 默认为 local, 设置为 yarn 使其运行在 YARN 框架上-->

</property>

</configuration>

<configuration>

<property>

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

<value>yarn</value>

<!--value 默认为 local, 设置为 yarn 使其运行在 YARN 框架上-->

</property>

</configuration>

~

6) sudo vi yarn-site.xml

<configuration>

<property>

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

<value>master</value>

<!--设置资源管理器的主机, 设置为主机名或者 IP-->

</property>

```

<property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
    <!-- 设置节点管理器的辅助服务器，默认为空，设置为 mapreduce_shuffle-->
</property>
</configuration>

<configuration>

<property>
    <name>yarn.resourcemanager.hostname</name>
    <value>master</value>
    <!-- 设置资源管理器的主机，设置为主机名或者IP-->
</property>
<property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
    <!-- 设置节点管理器的辅助服务器，默认为空，设置为mapreduce_shuffle-->
</property>

</configuration>

```

```

hdfs-site.xml          mapred-env.cmd
lxy1@master:~/hadoop/hadoop-2.7.7/etc/hadoop$ vim /core-site.xml
lxy1@master:~/hadoop/hadoop-2.7.7/etc/hadoop$ sudo vim /core-site.xml
[sudo] password for lxy1:
lxy1@master:~/hadoop/hadoop-2.7.7/etc/hadoop$ sudo vim /hdfs-site.xml
lxy1@master:~/hadoop/hadoop-2.7.7/etc/hadoop$ sudo vim /mapred-site.xml
lxy1@master:~/hadoop/hadoop-2.7.7/etc/hadoop$ sudo vim /yarn-site.xml

```

3. 初始化: hadoop namenode -format

4. 启动:start-all.sh

5. 查看: jps

```

lxy1@master:~/install/hadoop$ jps
4049 DataNode
4241 SecondaryNameNode
3894 NameNode
4843 Jps
4397 ResourceManager
4702 NodeManager

```

七、安装 Scala

1. 使用管理员权限解压 scala

```
sudo tar -zvxf scala-2.11.8.tgz -C ../install
```

在 install 目录:

```
sudo mv scala-2.11.8 scala
```

2. 配置环境变量

```
vi /etc/profile
```

添加以下内容并保存:

```
export SCALA_HOME=/home/lxy/install/scala
```

```
export PATH=$PATH:$SCALA_HOME/bin
```

3. 使配置生效&查验，如下：

```
source /etc/profile
```

```
[lxy@master:~/install$ sudo vi /etc/profile
lxy@master:~/install$ source /etc/profile
lxy@master:~/install$ scala -version
Scala code runner version 2.11.8 -- Copyright 2002-2016, LAMP/EPFL
lxy@master:~/install$ ^C
lxy@master:~/install$ ]
```

八、安装Zookeeper

1. 解压zookeeper：

```
sudo tar -zxvf zookeeper-3.4.14.tar.gz -C ./install
```

```
sudo mv zookeeper-3.4.14 zookeeper
```

2. 创建用于存储数据和日志的文件夹

```
cd /home/lxy/install/zookeeper
```

```
sudo mkdir -m 777 zkData
```

3. 创建配置文件

```
export JAVA_HOME=/home/lxy/install/jdk
export JRE_HOME=$JAVA_HOME/jre
export CLASSPATH=.:${JAVA_HOME}/lib:${JRE_HOME}/lib
export PATH=$PATH:${JAVA_HOME}/bin:${JRE_HOME}/bin

export HADOOP_HOME=/home/lxy/install/hadoop
export PATH=.:${HADOOP_HOME}/bin:${HADOOP_HOME}/sbin:$PATH

export ZOOKEEPER_HOME=/home/lxy/install/zookeeper
export PATH=$PATH:${ZOOKEEPER_HOME}/bin
export ZOO_LOG_DIR=/home/lxy/install/zookeeper/log

export HBASE_HOME=/home/lxy/install/hbase
export PATH=$PATH:${HBASE_HOME}/bin

export SCALA_HOME=/home/lxy/install/scala
export PATH=$PATH:${SCALA_HOME}/bin

export SPARK_HOME=/home/lxy/install/spark
export PATH=$PATH:${SPARK_HOME}/bin:${SPARK_HOME}/sbin
```

```
cd conf
```

```
sudo mv zoo_sample.cfg zoo.cfg
```

```
sudo vi zoo.cfg
```

```
dataDir=/home/lxy/install/zookeeper/zkData
```

```
# the directory where the snapshot is stored.
# do not use /tmp for storage, /tmp here is just
# example sakes.
dataDir=/home/lxy/install/zookeeper/zkData
# the port at which the clients will connect
clientPort=2181
```

4. 环境变量配置

```
export ZOOKEEPER_HOME=/home/lxy/install/zookeeper
```

```
export PATH=$PATH:$ZOOKEEPER_HOME/bin
export ZOO_LOG_DIR=/home/lxy/install/zookeeper/log

export JAVA_HOME=/home/lxy/install/jdk
export JRE_HOME=$JAVA_HOME/jre
export CLASSPATH=.:$JAVA_HOME/lib:$JRE_HOME/lib
export PATH=$PATH:$JAVA_HOME/bin:$JRE_HOME/bin

export HADOOP_HOME=/home/lxy/install/hadoop
export PATH=.:$HADOOP_HOME/bin:$HADOOP_HOME/sbin:$PATH

export ZOOKEEPER_HOME=/home/lxy/install/zookeeper
export PATH=$PATH:$ZOOKEEPER_HOME/bin
export ZOO_LOG_DIR=/home/lxy/install/zookeeper/log

export HBASE_HOME=/home/lxy/install/hbase
export PATH=$PATH:$HBASE_HOME/bin

export SCALA_HOME=/home/lxy/install/scala
export PATH=$PATH:$SCALA_HOME/bin

export SPARK_HOME=/home/lxy/install/spark
export PATH=$PATH:$SPARK_HOME/bin:$SPARK_HOME/sbin
```

5. 启动 zookeeper: (摘自尚硅谷文档)

(1) 启动 Zookeeper

```
6. [atguigu@hadoop102 zookeeper-3.4.10]$ bin/zkServer.sh start
```

(2) 查看进程是否启动

```
7. [atguigu@hadoop102 zookeeper-3.4.10]$ jps
8. 4020 Jps
9. 4001 QuorumPeerMain
```

(3) 查看状态:

```
10. [atguigu@hadoop102 zookeeper-3.4.10]$ bin/zkServer.sh
    status
11. ZooKeeper JMX enabled by default
12. Using config:
    /opt/module/zookeeper-3.4.10/bin/../conf/zoo.cfg
13. Mode: standalone
```

(4) 启动客户端:

```
14. [atguigu@hadoop102 zookeeper-3.4.10]$ bin/zkCli.sh
```

(5) 退出客户端:

```
15. [zk: localhost:2181(CONNECTED) 0] quit
```

(6) 停止 Zookeeper

```
16. [atguigu@hadoop102 zookeeper-3.4.10]$ bin/zkServer.sh
    stop
```

```
g session
2020-06-17 17:21:39,891 [myid:] - INFO  [main-SendThread(localhost:2181);ClientCnxn
Thread@1299] - Session establishment complete on server localhost/0:0:0:0:0:0:1:2
sessionid = 0x10000ac1c430000, negotiated timeout = 30000
[zk: localhost:2181(CONNECTED) 0]
WATCHER::

WatchedEvent state:SyncConnected type:None path:null

[zk: localhost:2181(CONNECTED) 0] ls /
[zookeeper]
[zk: localhost:2181(CONNECTED) 1] █
```

九、 安装 Hbase

1. 解压

```
sudo tar -zxvf hbase-1.3.6-bin.tar.gz -C ./install
mv hbase-1.3.6 hbas
```

2. sudo vi hbase-env.sh

```
export JAVA_HOME=/home/lxy/install/jdk
export HBASE_PID_DIR=/home/lxy/install/hbase/pids
export HBASE_MANAGES_ZK=false
```

3. sudo vi hbase-site.xml

```
<configuration>
```

```
<property>
    <name>hbase.rootdir</name>
    <value>hdfs://master:9000/hbase</value>
</property>
<property>
    <name>hbase.cluster.distributed</name>
    <value>true</value>
</property>
<!-- 0.98 后的新变动，之前版本没有.port，默认端口为 60000 --&gt;
&lt;property&gt;
    &lt;name&gt;hbase.master.port&lt;/name&gt;
    &lt;value&gt;16000&lt;/value&gt;
&lt;/property&gt;
&lt;property&gt;
    &lt;name&gt;hbase.zookeeper.quorum&lt;/name&gt;
    &lt;value&gt;master:2181&lt;/value&gt;
&lt;/property&gt;
&lt;property&gt;
    &lt;name&gt;hbase.zookeeper.property.dataDir&lt;/name&gt;
    &lt;value&gt;/home/lxy/install/zookeeper/zkData&lt;/value&gt;
&lt;/property&gt;&lt;/configuration&gt;</pre>
```

```
<configuration>

<property>
    <name>hbase.rootdir</name>
    <value>hdfs://master:9000/hbase</value>
</property>
<property>
    <name>hbase.cluster.distributed</name>
    <value>true</value>
</property>
<!-- 0.98后的新变动，之前版本没有.port，默认端口为60000 --&gt;
&lt;property&gt;
    &lt;name&gt;hbase.master.port&lt;/name&gt;
    &lt;value&gt;16000&lt;/value&gt;
&lt;/property&gt;
&lt;property&gt;
    &lt;name&gt;hbase.zookeeper.quorum&lt;/name&gt;
    &lt;value&gt;master:2181&lt;/value&gt;
&lt;/property&gt;
&lt;property&gt;
    &lt;name&gt;hbase.zookeeper.property.dataDir&lt;/name&gt;
    &lt;value&gt;/home/lxy/install/zookeeper/zkData&lt;/value&gt;
&lt;/property&gt;

&lt;/configuration&gt;</pre>
```

4. sudo vi regionservers

master

```
lxy@master:~/install/hbase/conf$ sudo vi regionservers
[sudo] password for lxy:
                         █
master█
```

5. 将 hadoop 的配置文件 core-site.xml 复制到 hbase 的配置文件目录中
cp /home/lxy/install/hadoop/etc/hadoop/core-site.xml conf/

6. 配置环境

```
export HBASE_HOME=/home/lxy/install/hbase
export PATH=$PATH:$HBASE_HOME/bin
```

```
export JAVA_HOME=/home/lxy/install/jdk
export JRE_HOME=$JAVA_HOME/jre
export CLASSPATH=.:${JAVA_HOME}/lib:${JRE_HOME}/lib
export PATH=$PATH:${JAVA_HOME}/bin:${JRE_HOME}/bin

export HADOOP_HOME=/home/lxy/install/hadoop
export PATH=.:${HADOOP_HOME}/bin:${HADOOP_HOME}/sbin:$PATH

export ZOOKEEPER_HOME=/home/lxy/install/zookeeper
export PATH=$PATH:${ZOOKEEPER_HOME}/bin
export ZOO_LOG_DIR=/home/lxy/install/zookeeper/log

export HBASE_HOME=/home/lxy/install/hbase
export PATH=$PATH:${HBASE_HOME}/bin

export SCALA_HOME=/home/lxy/install/scala
export PATH=$PATH:${SCALA_HOME}/bin

export SPARK_HOME=/home/lxy/install/spark
export PATH=$PATH:${SPARK_HOME}/bin:${SPARK_HOME}/sbin
```

十、安装 spark

1. 解压 spark

```
tar -xvf spark-2.2.0-bin-hadoop2.7.tgz -C ./install
```

2. 改名

```
mv spark-2.2.0-bin-hadoop2.7/ spark
```

3. 配置文件

进入到 spark 的配置目录下 (conf)

修改 slaves.template、spark-env.sh.template、spark-defaults.conf.template 等文件的文件名

```
mv slaves.template slaves
```

```
mv spark-env.sh.template spark-env.sh
```

```
mv spark-defaults.conf.template spark-defaults.conf
```

```
lxy@master:~/install/spark/conf$ ll
total 40
drwxr-xr-x  2 lxy lxy 4096 Jun 24 15:13 .
drwxr-xr-x 12 lxy lxy 4096 Jul  1 2017 ..
-rw-r--r--  1 lxy lxy  996 Jul  1 2017 docker.properties.template
-rw-r--r--  1 lxy lxy 1105 Jul  1 2017 fairscheduler.xml.template
-rw-r--r--  1 lxy lxy 2025 Jul  1 2017 log4j.properties.template
-rw-r--r--  1 lxy lxy 7313 Jul  1 2017 metrics.properties.template
-rw-r--r--  1 lxy lxy  865 Jul  1 2017 slaves
-rw-r--r--  1 lxy lxy 1292 Jul  1 2017 spark-defaults.conf
-rwxr-xr-x  1 lxy lxy 3699 Jul  1 2017 spark-env.sh*
```

4. 修改 spark-env.sh 文件

```
vi spark-env.sh
```

添加一下内容并保存

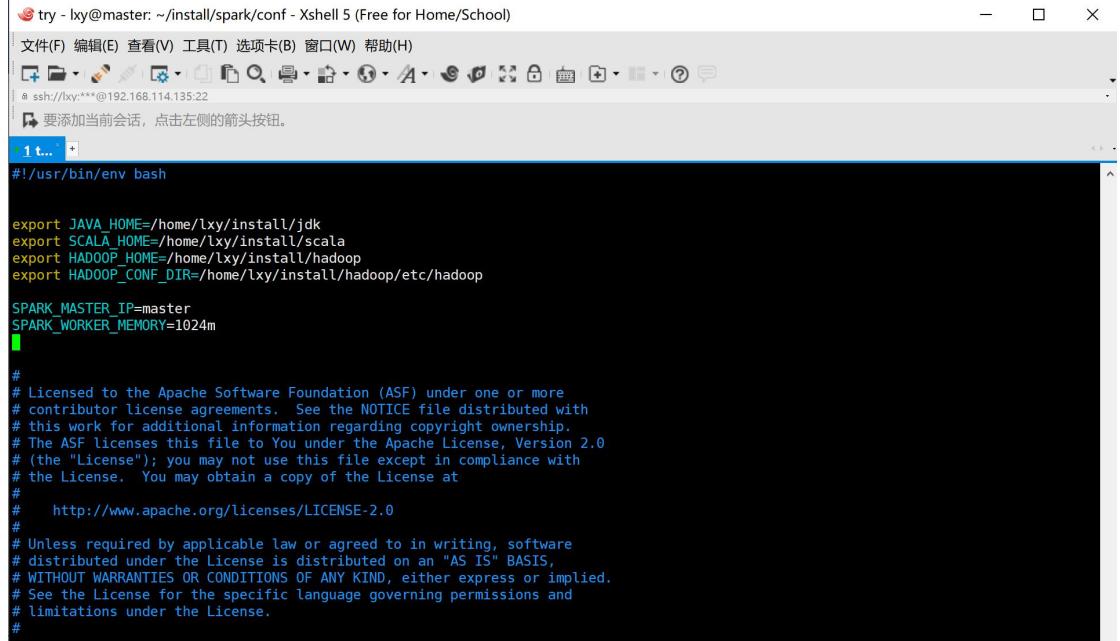
```
export JAVA_HOME=/home/lxy/install/jdk
```

```
export SCALA_HOME=/home/lxy/install/scala
```

```
export HADOOP_HOME=/home/lxy/install/hadoop
```

```
export HADOOP_CONF_DIR=/home/lxy/install/hadoop/etc/Hadoop
```

```
SPARK_MASTER_IP=master  
SPARK_WORKER_MEMORY=1024m  
~~~
```



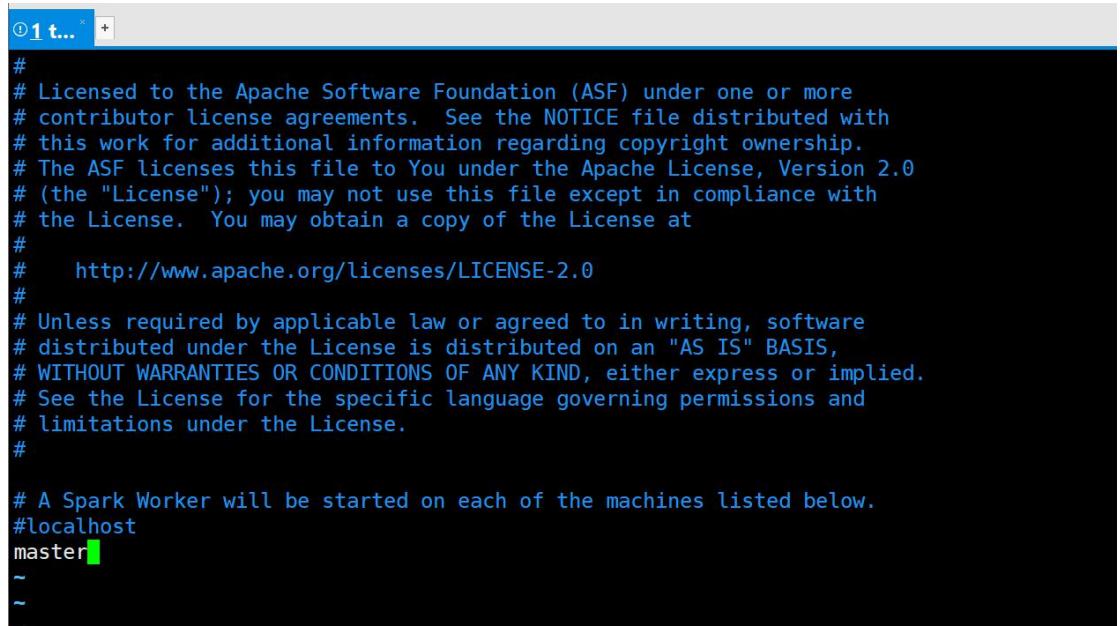
```
try - lxy@master: ~/install/spark/conf - Xshell 5 (Free for Home/School)  
文件(F) 编辑(E) 查看(V) 工具(T) 选项卡(B) 窗口(W) 帮助(H)  
ssh://lxy:**@192.168.114.135:22  
要添加当前会话，点击左侧的箭头按钮。  
① t...  
#!/usr/bin/env bash  
  
export JAVA_HOME=/home/lxy/install/jdk  
export SCALA_HOME=/home/lxy/install/scala  
export HADOOP_HOME=/home/lxy/install/hadoop  
export HADOOP_CONF_DIR=/home/lxy/install/hadoop/etc/hadoop  
  
SPARK_MASTER_IP=master  
SPARK_WORKER_MEMORY=1024m  
  
#  
# Licensed to the Apache Software Foundation (ASF) under one or more  
# contributor license agreements. See the NOTICE file distributed with  
# this work for additional information regarding copyright ownership.  
# The ASF licenses this file to You under the Apache License, Version 2.0  
# (the "License"); you may not use this file except in compliance with  
# the License. You may obtain a copy of the License at  
#  
# http://www.apache.org/licenses/LICENSE-2.0  
#  
# Unless required by applicable law or agreed to in writing, software  
# distributed under the License is distributed on an "AS IS" BASIS,  
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.  
# See the License for the specific language governing permissions and  
# limitations under the License.
```

5. 修改 slaves 文件

```
vi slaves
```

添加以下内容

master



```
① t...  
#  
# Licensed to the Apache Software Foundation (ASF) under one or more  
# contributor license agreements. See the NOTICE file distributed with  
# this work for additional information regarding copyright ownership.  
# The ASF licenses this file to You under the Apache License, Version 2.0  
# (the "License"); you may not use this file except in compliance with  
# the License. You may obtain a copy of the License at  
#  
# http://www.apache.org/licenses/LICENSE-2.0  
#  
# Unless required by applicable law or agreed to in writing, software  
# distributed under the License is distributed on an "AS IS" BASIS,  
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.  
# See the License for the specific language governing permissions and  
# limitations under the License.  
  
# A Spark Worker will be started on each of the machines listed below.  
#localhost  
master  
-  
-  
-
```

6. 配置环境变量

```
vi /etc/profile
```

添加以下内容

```

export SPARK_HOME=/home/lxy/install/spark
export PATH=$PATH:$SPARK_HOME/bin

export JAVA_HOME=/home/lxy/install/jdk
export JRE_HOME=$JAVA_HOME/jre
export CLASSPATH=.:$JAVA_HOME/lib:$JRE_HOME/lib
export PATH=$PATH:$JAVA_HOME/bin:$JRE_HOME/bin

export HADOOP_HOME=/home/lxy/install/hadoop
export PATH=.:$HADOOP_HOME/bin:$HADOOP_HOME/sbin:$PATH

export ZOOKEEPER_HOME=/home/lxy/install/zookeeper
export PATH=$PATH:$ZOOKEEPER_HOME/bin
export ZOO_LOG_DIR=/home/lxy/install/zookeeper/log

export HBASE_HOME=/home/lxy/install/hbase
export PATH=$PATH:$HBASE_HOME/bin

export SCALA_HOME=/home/lxy/install/scala
export PATH=$PATH:$SCALA_HOME/bin

export SPARK_HOME=/home/lxy/install/spark
export PATH=$PATH:$SPARK_HOME/bin

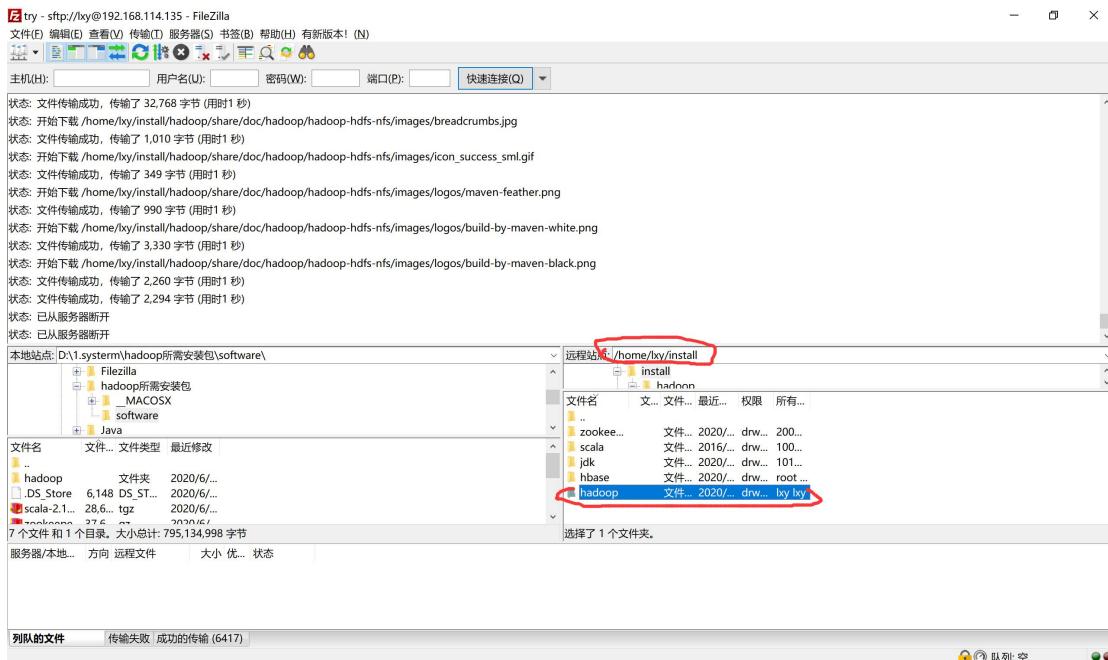
```

7. 使环境配置生效

source /etc/profile

十一、 配置本机环境

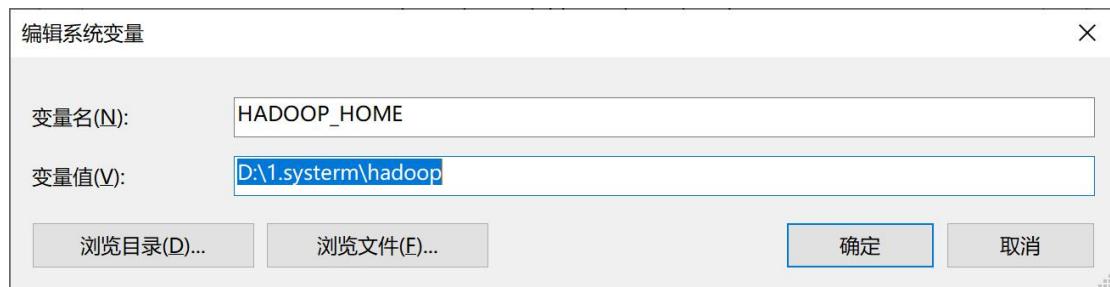
1. Java 环境&下载安装 eclipse
2. Hadoop 环境, 注: 从 Linux 系统 down 的 Hadoop 包 bin 文件夹中缺少一些文件, 可从网上下载

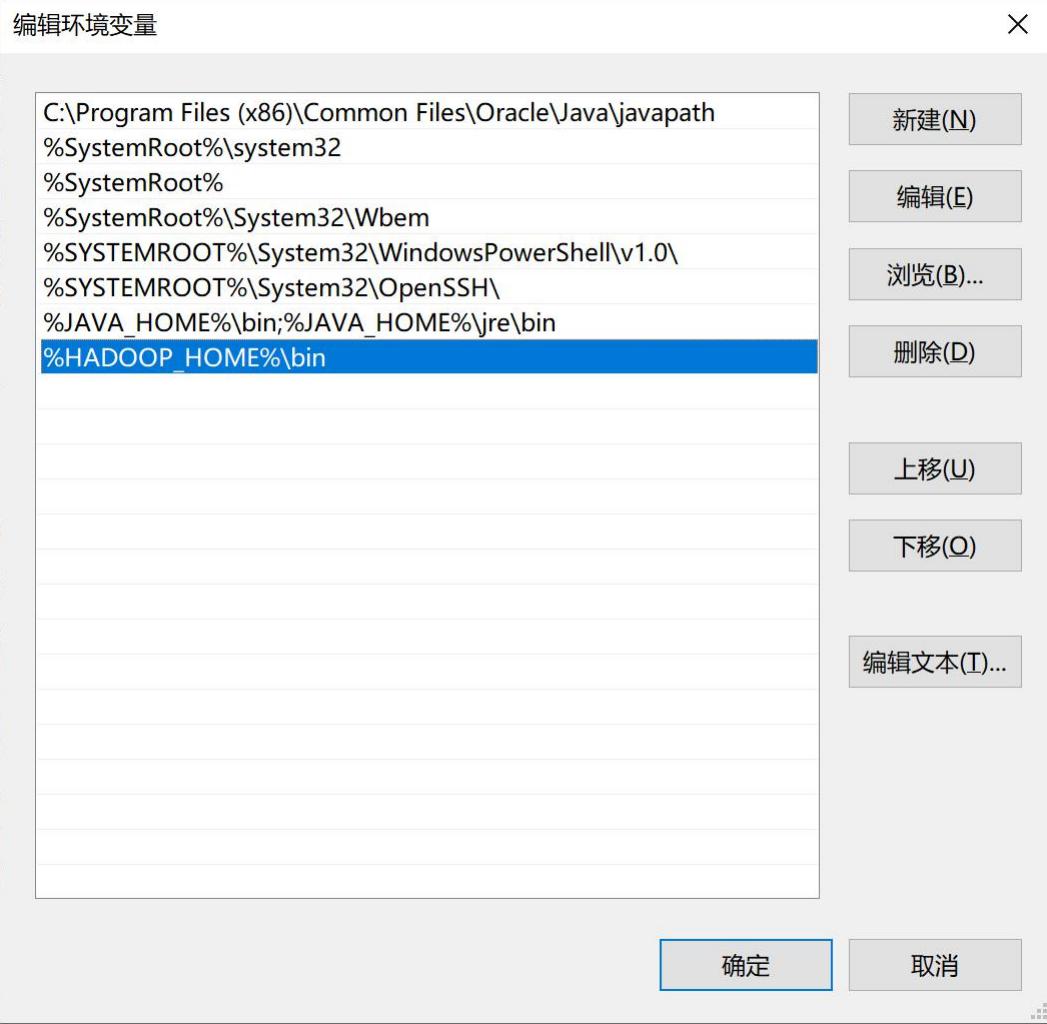


电脑 > Data (D:) > 1.systerm > hadoop

搜索 "hadoop"

名称	修改日期	类型	大小
bin	2020/6/18 10:40	文件夹	
etc	2020/6/18 10:40	文件夹	
include	2020/6/18 10:40	文件夹	
lib	2020/6/18 10:40	文件夹	
libexec	2020/6/18 10:40	文件夹	
logs	2020/6/18 10:40	文件夹	
sbin	2020/6/18 10:40	文件夹	
share	2020/6/18 10:40	文件夹	
tmp	2020/6/18 10:40	文件夹	
LICENSE	2020/6/18 10:40	文本文档	85 KB
NOTICE	2020/6/18 10:40	文本文档	15 KB
README	2020/6/18 10:40	文本文档	2 KB





3. 修改本机 hosts (本地运行 eclipse 必要操作)

文件资源

所有安装包和尚硅谷大数据文档

链接：[HTTPS://PAN.BAIDU.COM/s/1G-Jt2MDYgApXN8kwGLYETQ](https://pan.baidu.com/s/1G-Jt2MDYgApXN8kwGLYETQ)

提取码：H3NI

参考文献：

1. ubuntu16+Hadoop2.7.3 环境搭建（伪分布式）
原文链接: https://blog.csdn.net/sinat_34022298/article/details/72621887
2. Ubuntu16.04 Xshell 远程控制、Filezilla 远程传输文件
原文链接: https://blog.csdn.net/weixin_43789195/article/details/106126816
3. Ubuntu16.04 安装 Hadoop 单机和伪分布式环境超详细
原文链接: <https://blog.csdn.net/kh896424665/article/details/78765175>
4. Ubuntu16.04 下搭建 Hadoop2.7.4 单机伪分布式环境
原文链接: <https://www.jianshu.com/p/e4c709cf8daa/>
5. Cenos7 下 hadoop2.7.7 伪分布式模式安装
原文链接: <https://www.cnblogs.com/chenruhai/p/12464201.html>
6. 尚硅谷视频
7. 各类文档