

8-hdfs配置-单节点

-----成都尚学堂-mr-zeng-----

a) windows安装 xshell xftp

b) linux节点启动运行-配置好【网络和域名映射】

通过ping 域名； 命令检查是否配置成功

c) linux安装jdk (hadoop基于java开发的)

d) 安装hadoop (集成了核心功能：hdfs、mapreduce、yarn...)

1. 下载hadoop2稳定版本

下载地址：<http://archive.apache.org/dist/> --hadoop --core --hadoop-2.7.2 -- hadoop-2.7.2.tar.gz

2. 使用xftp将hadoop安装包放到linux指定目录下，并在linux解压 安装到usr下

```
tar -zxvf /root/ hadoop-2.7.2.tar.gz -C /usr/hadoop
```

3. 配置hadoop集群 环境--》参考官方文档（官方的有很多坑-大家参考的时候要慎重）



-----以下是单节点配置-----

节点node1：同时配置了：namenode，secondaryNamenode，datanode

a) 配置hadoop所使用的jdk

etc/hadoop/hadoop-env.sh（必须修改,环境变量的java_home在这里可能读取不到！）

```
vi /usr/hadoop/hadoop-2.7.2/etc/hadoop/hadoop-env.sh
```

修改export JAVA_HOME= /usr/java/jdk1.7.0_79

b) 配置 etc/hadoop/core-site.xml

```
vi core-site.xml
```

```
<property>
  <name>fs.defaultFS</name>
  <value>hdfs://namenode的域名:9000</value>
</property>
<property>
  <name>hadoop.tmp.dir</name>
  <value>/node1/hadoop-2.7.2/tmp</value>
</property>
```

（官方文档没有）

注：

fs.defaultFS：配置namenode的地址9000：是rpc端口默认8020

Hadoop.tmp.dir：配置hadoop的临时文件的路径（存储namenode，secondarNamenode，datanode的数据）

c) 配置etc/hadoop/hdfs-site.xml:

```
<property>
  <name>dfs.namenode.secondary.http-address</name>
  <value>node2:50090</value>
</property>
```

注：

dfs.namenode.secondary.http-address：secondarNamnode运行的地址和端口

d) 配置slaves 用以配置(datanode和nodemanager的地址)（-目前单节点只需要配置本机）

```
etc/hadoop/slaves
```

按行分割配置各地址

NODE1

NODE2

...

4) -----为了方便使用hadoop把命令写入系统环境-----

1)配置hadoop环境命令到系统(配置环境变量)，方便以后执行

```
vi /etc/profile
```

追加 (如下包含了上面已经配置的java环境)

```
export JAVA_HOME=/usr/java/jdk1.7.0_79
```

```
export HADOOP_HOME=/usr/hadoop/hadoop-2.7.2
```

```
export PATH=$PATH:$JAVA_HOME/bin:$HADOOP_HOME/bin:$HADOOP_HOME/sbin
```

2) 刷新配置文件 (使环境变量生效)

```
source /etc/profile
```

3) 验证命令是否生效

```
hadoop version
```

运行hadoop的hdfs

1) 格式化hdfs文件系统- (类似于硬盘格式化-这里-会把namenode需要的元数据基础信息写到对应的文件夹 (hadoop.tmp.dir))

ps----》只需要进行一次-格式化会删除当前hdfs所有的数据

```
hdfs namenode -format
```

2) 启动hadoop的hdfs (文件系统)

切换到 home/hadoop-2.7.2/sbin/ 下

```
./start-dfs.sh
```

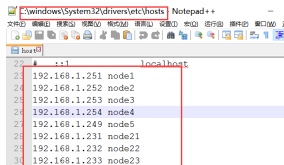
3) 防火墙关闭(为了方便其他节点的系统访问hdfs)

```
service iptables stop 关闭防火墙
```

```
service iptables status 查看当前状态
```

如果没有关闭,则需要对防火墙进行设置---》建议关闭,这样可以在其他客户端访问当前hdfs系统

4) window 里修改hosts-加入所有节点的ip和域名的映射



5) 客户端系统 (window, mac, linux等) 网页进入hadoop的hdfs文件系统管理目录

A) ip访问: <http://namenode的ip:50070/> 注意浏览器拦截问题 建议使用google

B) 域名访问: <http://namenode的域名:50070/>

6) 测试hdfs文件系统正常: 在hdfs上传文件一个文件

```
hadoop fs -put /root/jdk-7u79-linux-x64.tar.gz hdfs://node1:9000/jdk
```

7) 查看文件是否上传成功-再次重复 (4) 访问网页进入hadoop的hdfs文件系统管理目录



能成功下载说明hdfs搭建成功

-----启动服务器的小问题 (不用处理-这是个警告-处理起来特别麻烦!) -----

Unable to load native-hadoop library for your platform... using builtin-java classes where applicable解决方法:

手动编译hadoop <http://blog.csdn.net/Joy58061678/article/details/45746847>

<http://blog.csdn.net/wulantian/article/details/38111999/>

常见问题

A) window无法ip访问 hdfs服务器-》请检查ip是否正确，检查服务器的防火墙是否关闭

登录到hdfs服务器 service iptables stop

B) window无法域名访问 hdfs服务器-》请检查该域名在本window中是否配置对应的域名映射

修改文件加入对应域名映射： C:\windows\System32\drivers\etc\hosts

C) 启动报错java.net.UnknownHostException: xxxxxxxx 表示不知道域名，可能原因，本机的域名没有配置

vi /etc/sysconfig/network

或本机连接到其他机器的域名没有配置

vi /etc/hosts

B) 如果更改过配置文件（影响了namenode，datanode等），那么必须把临时文件夹删除，重新格式化，否则对应的节点可能启动不了！！

E) 如果运行错误了-请根据/usr/hadoop/hadoop-2.7.2/logs 里对应的错误日志记录文件分析解决！！