

在Fedora上部署Hadoop2.2.0伪分布式平台

如果你想配置完全分布式平台请参见本博客
[《Hadoop2.2.0完全分布式集群平台安装与设置》](#)

首先，你得在电脑上面安装好jdk7，如何安装，这里就不说了，网上一大堆教程！然后安装好ssh，如何安装请参见本博客[《Linux平台下安装SSH》](#)、并设置好无密码登录（[《Ubuntu和CentOS如何配置SSH使得无密码登陆》](#)）。好了，上面的前提条件部署好之后，下面将进入Hadoop2.2.0的部署。运行下面命令，将最新版的hadoop下载下来：

```
[wyp@wyp hadoop]$ wget W  
http://mirror.bit.edu.cn/apache/hadoop/common/hadoop-2.2.0/hadoop-2.2.0.tar.gz
```

当然，你也可以用下载的软件到上面的地址去下载。上面的命令是一行，由于此处太长了，所以强制弄成两行。假设下载好的hadoop存放在/home/wyp/Downloads/hadoop目录中，由于下载下来的hadoop是压缩好的，请将它解压，运行下面的命令：

```
[wyp@wyp hadoop]$ tar -xvf hadoop-2.2.0.tar.gz
```

解压之后，你将会看到如下目录结构：

```
[wyp@wyp hadoop]$ ls -l  
total 56  
drwxr-xr-x. 2 wyp wyp 4096 Oct 7 14:38 bin  
drwxr-xr-x. 3 wyp wyp 4096 Oct 7 14:38 etc  
drwxr-xr-x. 2 wyp wyp 4096 Oct 7 14:38 include  
drwxr-xr-x. 3 wyp wyp 4096 Oct 7 14:38 lib  
drwxr-xr-x. 2 wyp wyp 4096 Oct 7 14:38 libexec  
-rw-r--r--. 1 wyp wyp 15164 Oct 7 14:46 LICENSE.txt  
drwxrwxr-x. 3 wyp wyp 4096 Oct 28 14:38 logs  
-rw-r--r--. 1 wyp wyp 101 Oct 7 14:46 NOTICE.txt  
-rw-r--r--. 1 wyp wyp 1366 Oct 7 14:46 README.txt  
drwxr-xr-x. 2 wyp wyp 4096 Oct 28 12:37 sbin  
drwxr-xr-x. 4 wyp wyp 4096 Oct 7 14:38 share
```

下面是配置Hadoop的步骤：首先设置好Hadoop环境变量：

```
[wyp@wyp hadoop]$ sudo vim /etc/profile
```

在/etc/profile文件的末尾加上以下配置

```
export HADOOP_DEV_HOME=/home/wyp/Downloads/hadoop
export PATH=$PATH:$HADOOP_DEV_HOME/bin
export PATH=$PATH:$HADOOP_DEV_HOME/sbin
export HADOOP_MAPRED_HOME=${HADOOP_DEV_HOME}
export HADOOP_COMMON_HOME=${HADOOP_DEV_HOME}
export HADOOP_HDFS_HOME=${HADOOP_DEV_HOME}
export YARN_HOME=${HADOOP_DEV_HOME}
export HADOOP_CONF_DIR=${HADOOP_DEV_HOME}/etc/hadoop
```

然后按:wq保存。为了让刚刚的设置生效，运行下面的命令

```
[wyp@wyp hadoop]$ sudo source /etc/profile
```

接下来修改Hadoop的hadoop-env.sh配置文件，设置jdk所在的路径：

```
[wyp@wyp hadoop]$ vim etc/hadoop/hadoop-env.sh
```

在里面找到JAVA_HOME，并将它的值设置为你电脑jdk所在的绝对路径

```
# The java implementation to use.
export JAVA_HOME=/home/wyp/Downloads/jdk1.7.0_45
```

设置好之后请保存退出。接下来请配置好一下几个文件（都在hadoop目录下的etc/hadoop目录下）：

```
-----core-site.xml
<property>
  <name>fs.default.name</name>
  <value>hdfs://localhost:8020</value>
```

```
<final>true</final>
</property>
```

----- yarn-site.xml

```
<property>
  <name>yarn.nodemanager.aux-services</name>
  <value>mapreduce_shuffle</value>
</property>
```

```
<property>
  <name>yarn.nodemanager.aux-services.mapreduce.shuffle.class</name>
  <value>org.apache.hadoop.mapred.ShuffleHandler</value>
</property>
```

----- mapred-site.xml

```
<property>
  <name>mapreduce.framework.name</name>
  <value>yarn</value>
</property>
```

```
<property>
  <name>mapred.system.dir</name>
  <value>file:/opt/cloud/hadoop_space/mapred/system</value>
  <final>true</final>
</property>
```

```
<property>
  <name>mapred.local.dir</name>
  <value>file:/opt/cloud/hadoop_space/mapred/local</value>
  <final>true</final>
</property>
```

----- hdfs-site.xml

```
<property>
  <name>dfs.namenode.name.dir</name>
  <value>file:/opt/cloud/hadoop_space/dfs/name</value>
  <final>true</final>
</property>
```

```
<property>
  <name>dfs.datanode.data.dir</name>
  <value>file:/opt/cloud/hadoop_space/dfs/data</value>
  <description>Determines where on the local
  filesystem an DFS data node should store its blocks.
  If this is a comma-delimited list of directories,
  then data will be stored in all named
```

```
directories, typically on different devices.
Directories that do not exist are ignored.
</description>
<final>true</final>
</property>
<property>
  <name>dfs.replication</name>
  <value>1</value>
</property>

<property>
  <name>dfs.permissions</name>
  <value>>false</value>
</property>
```

上面的配置弄好之后，现在来进行测试，看看配置是否正确。首先格式化一下HDFS,如下命令：

```
[wyp@wyp hadoop]$ hdfs namenode -format
13/10/28 16:47:33 INFO namenode.NameNode: STARTUP_MSG:
/*****

.....此处省略好多文字.....

*****/
13/10/28 16:47:33 INFO namenode.NameNode: registered UNIX signal
handlers for [TERM, HUP, INT]
Formatting using clusterid: CID-9931f367-92d3-4693-a706-d83e120cacd6
13/10/28 16:47:34 INFO namenode.HostFileManager: read includes:
HostSet(
)
13/10/28 16:47:34 INFO namenode.HostFileManager: read excludes:
HostSet(
)

.....此处也省略好多文字.....

13/10/28 16:47:38 INFO util.ExitUtil: Exiting with status 0
13/10/28 16:47:38 INFO namenode.NameNode: SHUTDOWN_MSG:
/*****
SHUTDOWN_MSG: Shutting down NameNode at wyp/192.168.142.138
*****/
[wyp@wyp hadoop]$
```

大概出现上面的输出，好了，去启动一下你的Hadoop吧，命令如下：

```
[wyp@wyp hadoop]$ sbin/start-dfs.sh  
[wyp@wyp hadoop]$ sbin/start-yarn.sh
```

查看一下是否安装好了Hadoop，命令如下：

```
[wyp@wyp hadoop]$ jps  
9582 Main  
9684 RemoteMavenServer  
7011 DataNode  
7412 ResourceManager  
17423 Jps  
7528 NodeManager  
7222 SecondaryNameNode  
6832 NameNode  
[wyp@wyp hadoop]$
```

其中的jps是jdk自带的，如果出现NameNode、SecondaryNameNode、NodeManager、ResourceManager、DataNode这五个进程,那就恭喜你了，你的Hadoop已经安装好了！

这里附上如何关闭Hadoop各个服务

```
[wyp@wyp hadoop]$ sbin/stop-dfs.sh  
[wyp@wyp hadoop]$ sbin/stop-yarn.sh
```

本博客文章除特别声明，全部都是原创！
原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。
本文链接: [【】](#)（）