我是基于阿里云的云服务器配置的Hadoop，

配置是：1核、2G运存、40G存储；系统是Ubuntu16.04,64位。

如果在配置过程有什么问题，先自己解决，解决不了及时向我反映。

下面进入正题

安装配置Hadoop环境之前，先要知道Hadoop有三种运行模式：

  1. 单机模式（standalone）

  单机模式是Hadoop的默认模式。当首次解压Hadoop的源码包时，Hadoop无法了解硬件安装环境，便保守地选择了最小配置。在这种默认模式下所有3个XML文件均为空。当配置文件为空时，Hadoop会完全运行在本地。因为不需要与其他节点交互，单机模式就不使用HDFS，也不加载任何Hadoop的守护进程。该模式主要用于开发调试MapReduce程序的应用逻辑。

  2. 伪分布模式（Pseudo-Distributed Mode）

  伪分布模式在“单节点集群”上运行Hadoop，其中所有的守护进程都运行在同一台机器上。该模式在单机模式之上增加了代码调试功能，允许你检查内存使用情况，HDFS输入输出，以及其他的守护进程交互。

  3. 全分布模式（Fully Distributed Mode）

  Hadoop运行在一个集群上。

我安装的是单机模式

1. **创建Hadoop用户hduser**
2. $ sudo useradd -m hduser -s /bin/bash

这条命令创建了可以登陆的 hduser 用户，并使用 /bin/bash 作为 shell。

注：每一个$后面就是一条语句，写的时候不写$，#之后是注释，也不写，以后不再赘述

1. $ sudo passwd hduser

接着，这条语句是设置密码，可任意设置，简单点设置自己的名字就行，按提示输入两次密码。

1. $ sudo adduser hduser sudo

接着，这条语句是为 hadoop 用户增加管理员权限，方便部署。

1. $ sudo su hduser

输入上述语句，在输入密码后，就进入了hduser用户。

1. **安装SSH、配置SSH无密码登陆**
2. $ sudo apt-get install openssh-server

集群、单节点模式都需要用到 SSH 登陆（类似于远程登陆，你可以登录某台 Linux 主机，并且在上面运行命令），Ubuntu 默认已安装了 SSH client，此外还需要安装 SSH server，利用上面这条指令安装server

1. $ ssh localhost

此时会有SSH首次登陆提示，输入 yes 。然后按提示输入密码，这样就登陆到本机了。

1. 但这样登陆是需要每次输入密码的，若配置成SSH无密码登陆比较方便。首先退出刚才的 ssh，就回到了我们原先的终端窗口，然后利用 ssh-keygen 生成密钥，并将密钥加入到授权中：

$ exit                           # 退出刚才的 ssh localhost

$ cd ~/.ssh/                     # 若没有该目录，请先执行一次ssh localhost

$ ssh-keygen -t rsa              # 会有提示，都按回车就可以

（上面是三条语句，每一个$后面就是一条语句，写的时候不写$，每次执行完一条回车一次）

1. **安装Java环境**

（1）我安装的是Oracle的Java 8。

$ sudo add-apt-repository ppa:webupd8team/java

$ sudo apt-get update # 更新资源列表

$ sudo apt-get install oracle-java8-installer

注：

在执行这一步

$ sudo add-apt-repository ppa:webupd8team/java

很可能出现问题，

如果提示：add-apt-repository:command not found等类似的语句

解决：

先执行

$ sudo apt-get update

在执行

$ sudo apt-get install software-properties-common

最后在执行一开始的三条语句

（2）接着配置 JAVA\_HOME 环境变量，我在~/.bashrc中进行设置：

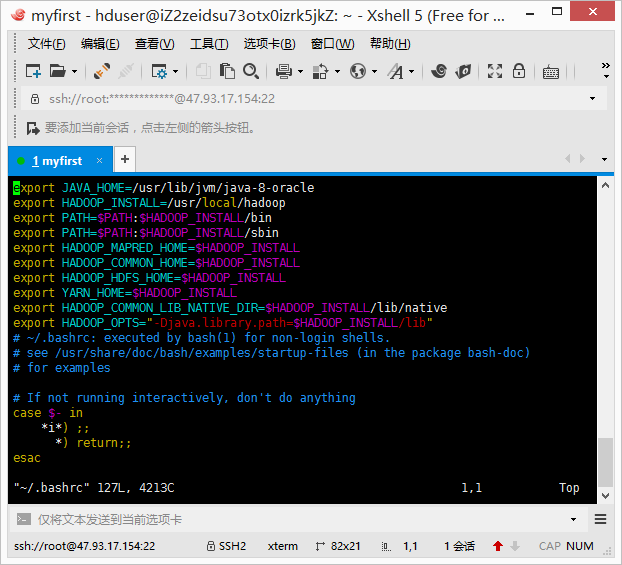
先执行$ sudo vim ~/.bashrc

注：在执行这条语句的时候，一定要在执行过sudo su hduser 和 ssh localhost之后才行，只有进入了Hadoop的用户中，改动bashrc才有效果

之后进入vim编辑界面，具有三种模式：命令模式（Command mode），插入模式（Insert mode）和底线命令模式（Last line mode）。

用户刚刚启动Vim，便进入了命令模式。

此时按i进入编辑模式，在编辑模式下你可以自由的编辑



这编辑模式下，在第一行插入这条语句

export JAVA\_HOME=/usr/lib/jvm/java-8-oracle

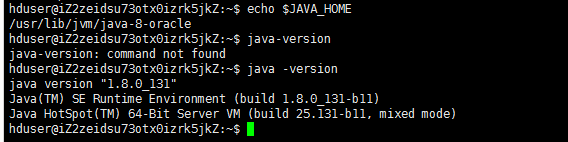
然后按ESC退出编辑模式进入命令模式，按 “：”（英文状态下），输入光标跳转到最后一行，按wq（保存退出），q（不保存直接退出）

1. 进行检测，是否已经成功安装好Java环境

$ echo $JAVA\_HOME

$ java –version

执行这两条语句，分别出现如下结果，证明成功



第一次Java –version没有成功时因为Java和-之间需要有一个空格，注意！

1. **安装 Hadoop 2.6.4与添加环境变量**
2. 下载Hadoop2.6.4

网址：<http://mirror.bit.edu.cn/apache/hadoop/common/hadoop-2.6.4/>

选择第二个180多M的压缩文件，下载好之后把它移到虚拟机或者云服务器中

利用命令

$ rz

可以把文件移动到Linux中，如果不行，百度搜一下怎么移文件

1. 把文件移动到Linux下之后，执行

$ sudo tar -zxf /hadoop-2.6.4.tar.gz -C /usr/local    # 解压到/usr/local中

$ cd /usr/local/ #打开usr/local文件夹

$ sudo mv ./hadoop-2.6.0/ ./hadoop            # 将文件夹名改为hadoop

1. 配置变量：在/usr/local/hadoop/etc/hadoop/hadoop-env.sh文件中添加Java\_Home

具体过程为

先执行语句 $ cd /usr/local/hadoop/etc/hadoop

在执行语句 $ sudo vim hadoop-env.sh

在编辑模式下，在第一行输入

export JAVA\_HOME=/usr/lib/jvm/java-8-oracle

之后保存退出

1. 在~/.bashrc文件里，添加环境变量

export JAVA\_HOME=/usr/lib/jvm/java-8-oracle

export HADOOP\_INSTALL=/usr/local/hadoop

export PATH=$PATH:$HADOOP\_INSTALL/bin

export PATH=$PATH:$HADOOP\_INSTALL/sbin

export HADOOP\_MAPRED\_HOME=$HADOOP\_INSTALL

export HADOOP\_COMMON\_HOME=$HADOOP\_INSTALL

export HADOOP\_HDFS\_HOME=$HADOOP\_INSTALL

export YARN\_HOME=$HADOOP\_INSTALL

export HADOOP\_COMMON\_LIB\_NATIVE\_DIR=$HADOOP\_INSTALL/lib/native

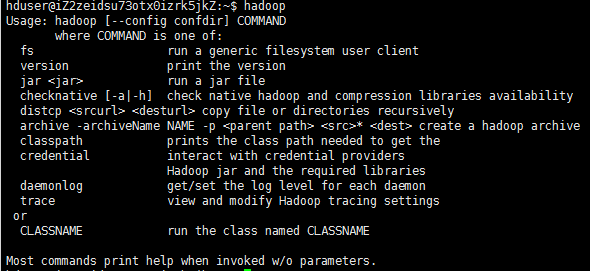
export HADOOP\_OPTS="-Djava.library.path=$HADOOP\_INSTALL/lib"

最后：wq保存退出，然后执行语句：

$ source ~/.bashrc #刷新更新使之生效

1. **单例模式配置完成，测试**
2. 在执行了sudo su hduser 和 ssh localhost之后，执行下面的语句

$ hadoop 会出现如下的显示



执行$ cd /usr/local/hadoop 进入hadoop的安装目录，

执行$ ./bin/hadoop jar ./share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.4.jar(不要漏掉最前面的一个点)

会出现如下：

