# 第二章Hadoop环境设置

## 实验一：搭建Hadoop单机模式环境

|  |
| --- |
| 【实验名称】 搭建Hadoop单机模式环境 |
| 【实验目的】 |
| 掌握Hadoop单机模式环境的搭建   1. 安装Java环境 2. 搭建Hadoop单机模式 |
| 【实验原理】 |
| 安装Java环境和搭建Hadoop单机模式环境 |
| 【实验环境】 |
| VMware® Workstation 14 Pro、Ubuntu 16.04系统  提示：本实验中Linux系统用户名为hadoop，密码为123456。 |
| 【实验步骤】   1. 启动Linux系统    1. 打开VMware   1.1  1.2、在虚拟机软件中打开虚拟机。    1.3、选择虚拟机    1.4、启动虚拟机    因为虚拟机安装文件夹移动过，第一次启动会有如下提示，选择“我已复制该  虚拟机(P)”选项。随后Linux系统启动。  1  1.5、输入系统密码,本系统密码为123456。    Linux系统启动完成。如下：  2  1.6、修改主机名，用个人姓名命名主机名，是实验是否自主完成个人标识。  如”张三“，主机名则为“zhangsan“。  （1）在桌面右键，打开终端。  3  （2）编辑文件  4  在文件中输入名字拼音，保存退出。  5   1. 重启系统。可看到主机名已改为成功。   7  在/etc/hosts 中添加本主机，否则后面代码运行出错。  添加内容“127.0.0.1 jiangyanping”，保存退出。  8   1. 安装java环境   2.1、下载 JDK1.8（本实验已将 Java 安装包上传到 Ubuntu 系统中，无需重新  下载）  2.2、解压安装包：  进入/home/hadoop/software\_tool/目录，所有实验用到的安装包都存放在本目录。  9  到usr/local查看是否有jdk1.8.0\_181目录  10  2.3、配置JDK环境变量，修改家目录下的配置文件.bashrc。  11  然后在末尾添加如下内容，保存退出。  12  2.4、刷新配置：使配置立即生效  13  2.5、检验是否安装成功  14  出现如下版本号信息为成功安装  15   1. hadoop 安装   3.1、下载安装 Hadoop（本实验已下载好安装包，这里提供下载的方法，供参  考。）  3.2、解压安装包：解压安装包至/usr/local目录下。  先切换至安装包目录/home/hadoop/software\_tool/  解压到/usr/local目录  D:\201706401102蒋延平\截图\16.PNG16  解压完成后，切换到目标目录。  17  3.3、配置 hadoop 的 java 工作环境  修改配置文件 hadoop-env.sh。该配置文件在 hadoop 的安装目录下的 etc/hadoop  目录下。  18  找到export JAVA\_HOME，并修改为：  19  3.4、把 hadoop 加入到环境变量中  配置hadoop环境变量，修改家目录下的配置文件.bashrc。  20  然后在末尾添加如下内容，保存退出。  21  刷新配置，使配置立即生效 并检验 hadoop 是否安装成功。  22  以上单机模式已安装好。  3.5、单机模式已安装好，运行 demo 程序。  先修改 hadoop-2.7.3 权限  23  在/usr/local/hadoop-2.7.3 目录下创建 input 目录， 把 etc/hadoop/下的 xml 配置文件都拷贝到 input 目录中  24  运行 demo 程序  26  查看程序的输出结果  27  至此单机模式安装和测试完成。   |  | | --- | | **实验二：搭建Hadoop伪分布式模式环境**  【实验名称】 搭建Hadoop伪分布式模式环境 | | 【实验目的】 | | 掌握 Hadoop 伪分布式环境搭建 | | 【实验原理】 | | 本实验在单机版 hadoop 基础上安装伪分布式 | | 【实验环境】 | | 已经部署好的 Hadoop 单机模式环境 |   一、配置无密登录  28  二、修改配置文件  2.1、切换到/usr/local/hadoop-2.7.3/etc/hadoop/目录,需要修改如下配置  （1）etc/hadoop/core-site.xml:  D:\201706401102蒋延平\截图\30.PNG30  68  （2）etc/hadoop/hdfs-site.xml:  33  （3）etc/hadoop/mapred-site.xml:  34  35  （4）etc/hadoop/yarn-site.xml:  36  37  三、启动hadoop，并执行demo  格式化的过程是创建初始目录和文件系统结构的过程。格式化只需进行一次，  下次启动不再需要格式化。执行以下命令：  38  39  3.2、启动 HDFS ,用jps命令验证  40  41  3.3、启动 yarn  56  用$ jps命令验证，正确启动将多出以下两个进程：  57  3.4、在 HDFS 上创建文件夹  43  建立完成可通过hdfs dfs -lsr 查看建立好的文件夹。  44  3.5、把/etc/hadoop目录上传到HDFS，保存到input目录，先建立用户目录hadoop  45  通过 hdfs dfs -ls 查看上传后新建立的 input 目录。  48  查看 input 目录里的文件。  49  3.6、执行 demo 代码  50  3.7、查看程序结果  59  查看具体内容：  D:\201706401102蒋延平\截图\60.PNG60  3.8、把在 HDFS 上的程序执行结果下载到本地 ，并查看。  61  62  四、通过web查看hadoop。  4.1、 浏览器访问Hadoop文件系统  浏览器访问Hadoop文件系统默认端口号为50070，打开Ubuntu自带的火狐浏览器，输入以下网址获得Hadoop Web UI的服务，可以查看 NameNode 和 Datanode 信  息，还可以在线查看 HDFS 中的文件。网址为：http://localhost:50070  64  4.2、 浏览器查看集群所有应用程序  访问集群中的所有应用程序的默认端口号为8088。使用以下URL访问该服务。网  址为：http://localhost:8088  五、停止所有进程  使用以下命令停止所有进程：  66  用命令jps查看：关闭了Hadoop所有进程  67  至此Hadoop伪分布式搭建完成。 |