**武汉工商学院**

**实验报告**

系：物联网系 专业：物联网工程 班级：20物联本1 课程名称：云计算技术

|  |  |  |  |
| --- | --- | --- | --- |
| 实验名称 | 实验一 Hadoop部署 | | |
| 学号 | 20423030 | 姓名 | 谢康 |
| 指导老师 | 王勇 |  |  |
| 实验目的 | 1. 理解Hadoop体系结构。 2. 熟练掌握Hadoop平台的部署方法与配置。 | | |
| **一、实验环境**  操作系统:Ubuntu-16.04、Hadoop版本:2.6.0、JDK版本:1.7  **二、实验内容及步骤**  **1、准备**  1.1创建hadoop用户  C:\Users\ADMINI~1\AppData\Local\Temp\1540133994(1).png  1.2安装SSH,设置SSH无密码登陆  C:\Users\ADMINI~1\AppData\Local\Temp\1540134279(1).png  输入完 $ ssh-keygen -t rsa　语句以后，需要连续敲击三次回车，如下图：  C:\Users\ADMINI~1\AppData\Local\Temp\1540134458(1).png  其中，第一次回车是让KEY存于默认位置，以方便后续的命令输入。第二次和第三次是确定passphrase，相关性不大。两次回车输入完毕以后，如果出现类似于下图所示的输出，即成功：  https://images2017.cnblogs.com/blog/1139343/201709/1139343-20170928110536840-777008126.jpg  之后再输入：  C:\Users\ADMINI~1\AppData\Local\Temp\1540134760(1).png  https://images2017.cnblogs.com/blog/1139343/201709/1139343-20170928110719637-31277406.png  **2、安装jdk1.7**  首先在[oracle官网](http://www.oracle.com/technetwork/java/javase/downloads/index.html%20" \t "_blank)下载jdk1.7 http://www.oracle.com/technetwork/java/javase/downloads/index.html 接下来进行安装与环境变量配置，根据个人电脑系统选择对应版本，选的是jdk-7u80-linux-x64.tar.gz  C:\Users\ADMINI~1\AppData\Local\Temp\1540135071(1).png  （注：其中如果权限不够，无法在相关目录下创建jvm文件夹，那么可以使用 $ sudo -i  语句进入root账户来创建文件夹。  另外推荐使用vim来编辑环境变量，即最后一句使用指令  C:\Users\ADMINI~1\AppData\Local\Temp\1540135426(1).png  如果没有vim,可以使用：  C:\Users\ADMINI~1\AppData\Local\Temp\1540135395(1).png  来进行下载。）  https://images2017.cnblogs.com/blog/1139343/201709/1139343-20170928111356262-1434564476.png  在.bashrc文件添加如下指令：  C:\Users\ADMINI~1\AppData\Local\Temp\1540135538(1).png  在文件修改完毕以后，输入代码：  C:\Users\ADMINI~1\AppData\Local\Temp\1540135606(1).png  如果出现如下图所示的内容，即为安装成功    **3、安装hadoop-2.6.0**  先下载hadoop-2.6.0.tar.gz，链接如下: http://mirrors.hust.edu.cn/apache/hadoop/common/  3.1 hadoop安装准备：  C:\Users\ADMINI~1\AppData\Local\Temp\1540135799(1).png  3.2给hadoop配置环境变量，将下面代码添加到.bashrc文件:  C:\Users\ADMINI~1\AppData\Local\Temp\1540135854(1).png  https://images2017.cnblogs.com/blog/1139343/201709/1139343-20170928113521434-2017853701.png  3.3同样，执行source ~./bashrc使设置生效，并查看hadoop是否安装成功    **4、伪分布式配置**  （1、Hadoop 可以在单节点上以伪分布式的方式运行，Hadoop 进程以分离的 Java 进程来运行，节点既作为 NameNode 也作为 DataNode，同时，读取的是 HDFS 中的文件。  2、Hadoop 的配置文件位于 /usr/local/hadoop/etc/hadoop/ 中，伪分布式需要修改2个配置文件 core-site.xml 和 hdfs-site.xml 。Hadoop的配置文件是 xml 格式，每个配置以声明 property 的 name 和 value 的方式来实现。）  4.1 首先将jdk1.7的路径添（export JAVA\_HOME=/usr/lib/jvm/java ）加到hadoop-env.sh文件    4.2 接下来修改core-site.xml文件：    4.3、接下来修改配置文件 hdfs-site.xml    （Hadoop 的运行方式是由配置文件决定的（运行 Hadoop 时会读取配置文件），因此如果需要从伪分布式模式切换回非分布式模式，需要删除 core-site.xml 中的配置项。此外，伪分布式虽然只需要配置 fs.defaultFS 和 dfs.replication 就可以运行（可参考官方教程），不过若没有配置 hadoop.tmp.dir 参数，则默认使用的临时目录为 /tmp/hadoo-hadoop，而这个目录在重启时有可能被系统清理掉，导致必须重新执行 format 才行。所以我们进行了设置，同时也指定 dfs.namenode.name.dir 和 dfs.datanode.data.dir，否则在接下来的步骤中可能会出错。）  4.4、配置完成后，执行 NameNode 的格式化  C:\Users\ADMINI~1\AppData\Local\Temp\1540136366(1).png  4.5、启动namenode和datanode进程，并查看启动结果  C:\Users\ADMINI~1\AppData\Local\Temp\1540136408(1).png  启动完成后，可以通过命令 jps 来判断是否成功启动，若成功启动则会列出如下进程: “NameNode”、”DataNode” 和 “SecondaryNameNode”    成功启动后，可以访问 Web 界面 [http://localhost:50070](http://localhost:50070/) 查看 NameNode 和 Datanode 信息，还可以在线查看 HDFS 中的文件。 | | | |