1、为什么要自己编译

企业的线上环境（生产环境）一般都是64位

hadoop官方网站上并没有支持64位的安装包

2、安装编译环境

<https://github.com/apache/hadoop.git/tags/release-2.6.1>    下载源代码： svn checkout

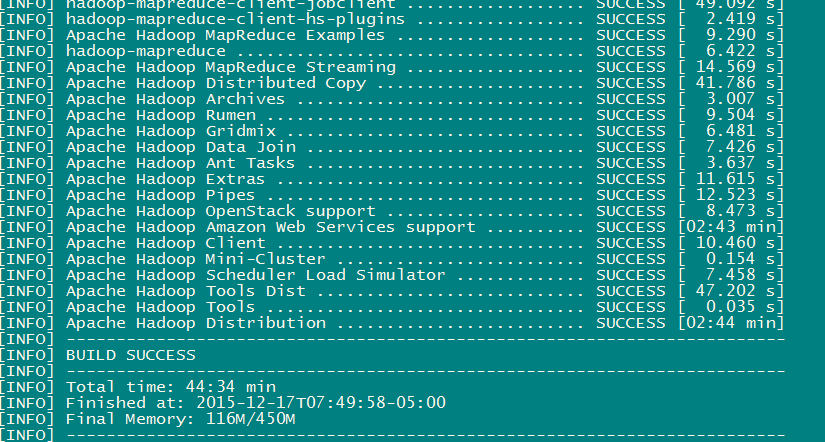


    进入项目根目录，进行编译：

    mvn package -Pdist,native -DskipTests -Dtar

    耐心等待1小时左右

    在hadoop项目的hadoop-dist/target就有了安装包



3、创建安装目录并解压文件

mkdir -p /export/servers

mkdir -p /export/software

mv hadoop-2.6.2.tar.gz /export/software/

cd /export/software/

tar -zxvf hadoop-2.6.2.tar.gz -C /export/servers/

cd /export/servers/

ln -s hadoop-2.6.2 hadoop

4、配置环境变量

            vi /etc/profile

#set hadoop env

export HADOOP\_HOME=/export/servers/hadoop

export PATH=${HADOOP\_HOME}/bin:${HADOOP\_HOME}/sbin:$PATH

5、修改Hadoop配置文件

    第一个：hadoop-env.sh

    第二个：core-site.xml

    第三个：hdfs-site.xml

    第四个：mapred-site.xml

    第五个：yarn-site.xml

6、格式化nameonde

    hadoop namenode -format

7、启动HDFS服务

手动一台一台地启动

在相应服务器上启动hdfs的相关进程：

启动namenode进程—— sbin/hadoop-daemon.sh start namenode

启动datanode进程——sbin/hadoop-daemon.sh start datanode

然后，验证hdfs的服务是否能正常提供：

bin/hdfs dfsadmin -report  查看hdfs集群的统计信息

Shell脚本批量启动方式：

在任意一台服务器上执行命令：

启动hdfs服务：sbin/start-dfs.sh

启动yarn服务：sbin/start-yarn.sh

或者：直接启动hdfs+yarn服务：sbin/start-all.sh

 yarn-daemon.sh start nodemanager

<http://www.ibm.com/developerworks/cn/data/library/bd-yarn-intro/>

<http://hadoop01:50070/dfshealth.html#tab-overview>  HDFS管理界面

<http://hadoop01:8088/cluster>  任务管理界面

           hadoop dfsadmin -safemode leave