五 Hadoop 源码编译

5.1 前期准备工作

1）CentOS 联网

配置 CentOS 能连接外网。Linux 虚拟机 ping www.baidu.com 是畅通的

注意：采用 root 角色编译，减少文件夹权限出现问题

2）jar 包准备(hadoop 源码、JDK8、maven、ant 、protobuf)

（1）hadoop-2.7.2-src.tar.gz

（2）jdk-8u144-linux-x64.tar.gz

（3）apache-ant-1.9.9-bin.tar.gz

（4）apache-maven-3.0.5-bin.tar.gz

（5）protobuf-2.5.0.tar.gz

5.2 jar 包安装

0）注意：所有操作必须在 root 用户下完成

1）JDK 解压、配置环境变量 JAVA_HOME 和 PATH，验证 java-version(如下都需要验证是否配置成功)

[root@hadoop101 software] # tar -zxf jdk-8u144-linux-x64.tar.gz -C /opt/module/

[root@hadoop101 software]# vi /etc/profile

```
#JAVA_HOME
export JAVA_HOME=/opt/module/jdk1.8.0_144
export PATH=$PATH:$JAVA_HOME/bin
```

[root@hadoop101 software]#source /etc/profile

验证命令：java -version

2）Maven 解压、配置 MAVEN_HOME 和 PATH。

[root@hadoop101 software]# tar -zxvf apache-maven-3.0.5-bin.tar.gz -C /opt/module/

[root@hadoop101 apache-maven-3.0.5]# vi conf/settings.xml

```
<mirrors>
    <!-- mirror
     | Specifies a repository mirror site to use instead of a given repository. The repository that
     | this mirror serves has an ID that matches the mirrorOf element of this mirror. IDs are used
     | for inheritance and direct lookup purposes, and must be unique across the set of
```

```
mirrors.
        |
          <mirror>
            <id>mirrorId</id>
            <mirrorOf>repositoryId</mirrorOf>
            <name>Human Readable Name for this Mirror.</name>
            <url>http://my.repository.com/repo/path</url>
          </mirror>
        -->
            <mirror>
                        <id>nexus-aliyun</id>
                        <mirrorOf>central</mirrorOf>
                        <name>Nexus aliyun</name>
                        <url>http://maven.aliyun.com/nexus/content/groups/public</url>
            </mirror>
    </mirrors>
```

[root@hadoop101 apache-maven-3.0.5]# vi /etc/profile

```
#MAVEN_HOME
export MAVEN_HOME=/opt/module/apache-maven-3.0.5
export PATH=$PATH:$MAVEN_HOME/bin
```

[root@hadoop101 software]#source /etc/profile

验证命令：mvn -version

3）ant 解压、配置 ANT _HOME 和 PATH。

[root@hadoop101 software]# tar -zxvf apache-ant-1.9.9-bin.tar.gz -C /opt/module/

[root@hadoop101 apache-ant-1.9.9]# vi /etc/profile

```
#ANT_HOME
export ANT_HOME=/opt/module/apache-ant-1.9.9
export PATH=$PATH:$ANT_HOME/bin
```

[root@hadoop101 software]#source /etc/profile

验证命令：ant -version

4）安装 glibc-headers 和 g++ 命令如下:

[root@hadoop101 apache-ant-1.9.9]# yum install glibc-headers

[root@hadoop101 apache-ant-1.9.9]# yum install gcc-c++

5）安装 make 和 cmake

[root@hadoop101 apache-ant-1.9.9]# yum install make

[root@hadoop101 apache-ant-1.9.9]# yum install cmake

6）解压 protobuf ，进入到解压后 protobuf 主目录，/opt/module/protobuf-2.5.0

然后相继执行命令：

[root@hadoop101 software]# tar -zxvf protobuf-2.5.0.tar.gz -C /opt/module/

[root@hadoop101 opt]# cd /opt/module/protobuf-2.5.0/

[root@hadoop101 protobuf-2.5.0]#./configure

[root@hadoop101 protobuf-2.5.0]# make

[root@hadoop101 protobuf-2.5.0]# make check

[root@hadoop101 protobuf-2.5.0]# make install

[root@hadoop101 protobuf-2.5.0]# ldconfig

[root@hadoop101 hadoop-dist]# vi /etc/profile

```
#LD_LIBRARY_PATH
export LD_LIBRARY_PATH=/opt/module/protobuf-2.5.0
export PATH=$PATH:$LD_LIBRARY_PATH
```

[root@hadoop101 software]#source /etc/profile

验证命令：protoc --version

7）安装 openssl 库

[root@hadoop101 software]#yum install openssl-devel

8）安装 ncurses-devel 库：

[root@hadoop101 software]#yum install ncurses-devel

到此，编译工具安装基本完成。

5.3 编译源码

1）解压源码到/opt/目录

[root@hadoop101 software]# tar -zxvf hadoop-2.7.2-src.tar.gz -C /opt/

2）进入到 hadoop 源码主目录

[root@hadoop101 hadoop-2.7.2-src]# pwd

/opt/hadoop-2.7.2-src

3）通过 maven 执行编译命令

[root@hadoop101 hadoop-2.7.2-src]#mvn package -Pdist,native -DskipTests -Dtar

等待时间 30 分钟左右，最终成功是全部 SUCCESS。

```
[INFO] Apache Hadoop Common ........................... SUCCESS [3:35.094s]
[INFO] Apache Hadoop NFS .............................. SUCCESS [5.004s]
[INFO] Apache Hadoop KMS .............................. SUCCESS [54.027s]
[INFO] Apache Hadoop Common Project ................... SUCCESS [0.022s]
[INFO] Apache Hadoop HDFS ............................. SUCCESS [3:58.444s]
[INFO] Apache Hadoop HttpFS ........................... SUCCESS [1:02.562s]
[INFO] Apache Hadoop HDFS BookKeeper Journal .......... SUCCESS [33.138s]
[INFO] Apache Hadoop HDFS-NFS ......................... SUCCESS [3.993s]
[INFO] Apache Hadoop HDFS Project ..................... SUCCESS [0.022s]
[INFO] hadoop-yarn .................................... SUCCESS [0.037s]
[INFO] hadoop-yarn-api ................................ SUCCESS [1:26.119s]
[INFO] hadoop-yarn-common ............................. SUCCESS [1:20.025s]
[INFO] hadoop-yarn-server ............................. SUCCESS [0.168s]
[INFO] hadoop-yarn-server-common ...................... SUCCESS [9.107s]
[INFO] hadoop-yarn-server-nodemanager ................. SUCCESS [19.867s]
[INFO] hadoop-yarn-server-web-proxy ................... SUCCESS [3.397s]
[INFO] hadoop-yarn-server-applicationhistoryservice ... SUCCESS [7.432s]
[INFO] hadoop-yarn-server-resourcemanager ............. SUCCESS [17.078s]
[INFO] hadoop-yarn-server-tests ....................... SUCCESS [3.998s]
[INFO] hadoop-yarn-client ............................. SUCCESS [5.962s]
[INFO] hadoop-yarn-server-sharedcachemanager .......... SUCCESS [2.803s]
[INFO] hadoop-yarn-applications ....................... SUCCESS [0.024s]
[INFO] hadoop-yarn-applications-distributedshell ...... SUCCESS [1.841s]
[INFO] hadoop-yarn-applications-unmanaged-am-launcher .... SUCCESS [1.876s]
```

4）成功的 64 位 hadoop 包在/opt/hadoop-2.7.2-src/hadoop-dist/target 下。

[root@hadoop101 target]# pwd

/opt/hadoop-2.7.2-src/hadoop-dist/target

5.4 常见的问题及解决方案

1）MAVEN install 时候 JVM 内存溢出

处理方式：在环境配置文件和 maven 的执行文件均可调整 MAVEN_OPT 的 heap 大小。

（详情查阅 MAVEN 编译 JVM 调优问题，如：

http://outofmemory.cn/code-snippet/12652/maven-outofmemoryerror-method）

2）编译期间 maven 报错。可能网络阻塞问题导致依赖库下载不完整导致，多次执行命令（一次通过比较难）：

[root@hadoop101 hadoop-2.7.2-src]#mvn package -Pdist,native -DskipTests -Dtar

3）报 ant、protobuf 等错误，插件下载未完整或者插件版本问题，最开始链接有较多特殊情况，同时推荐

2.7.0 版本的问题汇总帖子 http://www.tuicool.com/articles/IBn63qf