Hadoop源码编译

准备工作

(1) CentOS联网

配置CentOS能连接外网。Linux虚拟机ping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 (build工具, 打包用的)
- (4) apache-maven-3.0.5-bin.tar.gz
- (5) protobuf-2.5.0.tar.gz (序列化的框架)

编译工具安装

(1) 安装JDK

验证命令: java -version

<!-- mirror

[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

(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>

```
| 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.
  <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) 配置ANT

```
[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++

```
[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) 安装protobuf

```
[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

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

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

(4) 编译源码过程中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-outofmemoryerrormethod)

(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,nativeN -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