Spark 编译与部署（中）

--Hadoop 编译安装

目 录

Spark 编译与部署（中）

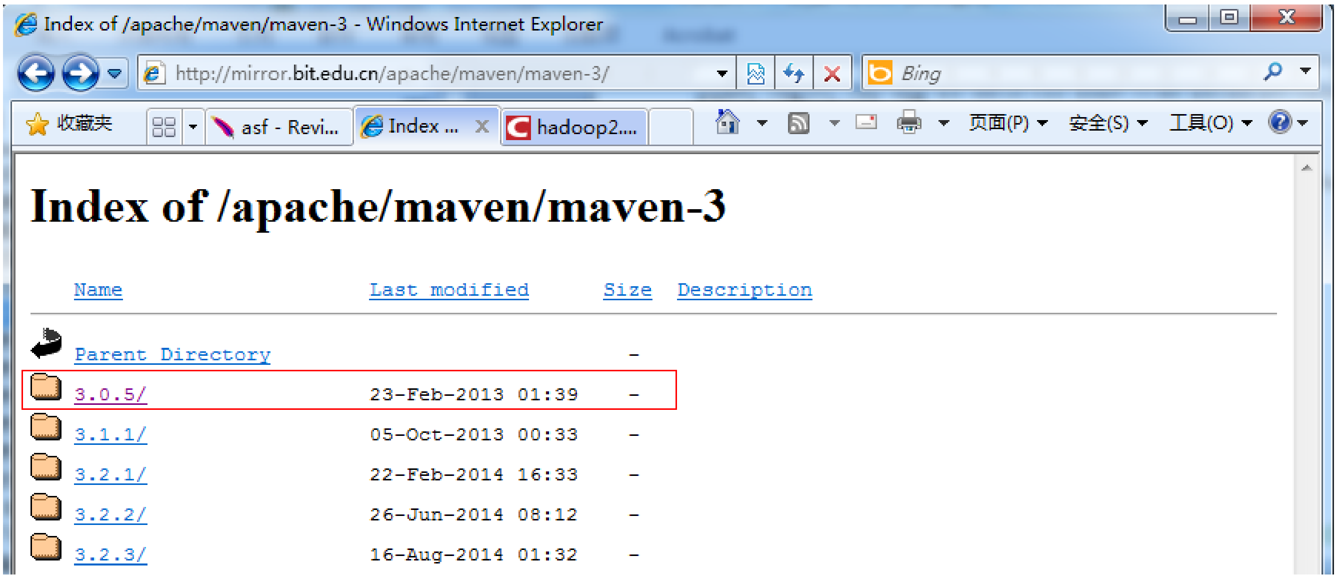
1 编译 Hadooop

1.1 搭建环境

1.1.1 安装并设置 maven

1. 下载 maven 安装包，建议安装 3.0 以上版本，本次安装选择的是 maven3.0.5 的二进制包，下载地址如下

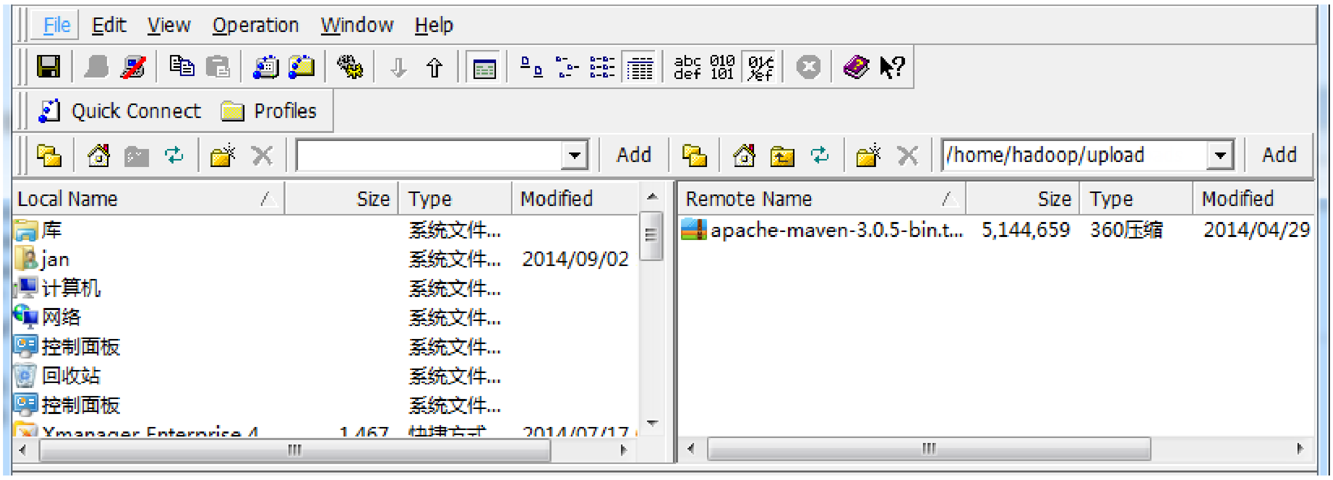
http://mirror.bit.edu.cn/apache/maven/maven-3/


Index of /apache/maven/maven-3 - Windows Internet Explorer
http://mirror.bit.edu.cn/apache/maven/maven-3/
Bing
收藏夹
asf - Revi...
Index ...
hadoop2....
页面(P)
安全(S)
工具(O)

Index of /apache/maven/maven-3

```
     Name                      Last modified      Size  Description
   Parent Directory                                -
   3.0.5/                      23-Feb-2013 01:39   -
   3.1.1/                      05-Oct-2013 00:33   -
   3.2.1/                      22-Feb-2014 16:33   -
   3.2.2/                      26-Jun-2014 08:12   -
   3.2.3/                      16-Aug-2014 01:32   -
```

2. 使用 ssh 工具把 maven 包上传到/home/hadoop/upload 目录

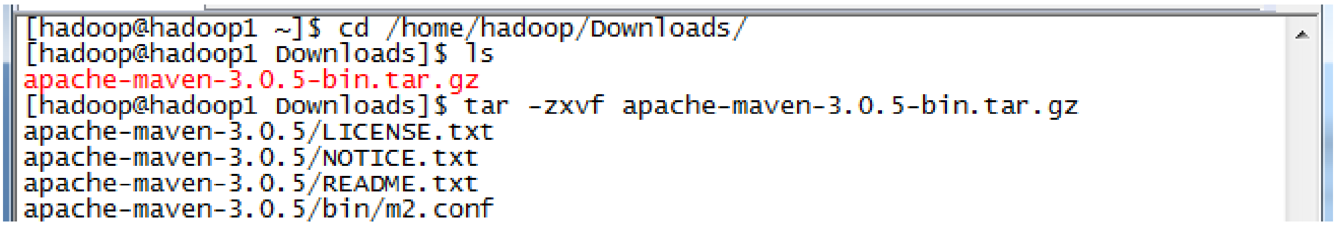


Local Name	Size	Type	Modified
库		系统文件...	
jan		系统文件...	2014/09/02
计算机		系统文件...	
网络		系统文件...	
控制面板		系统文件...	
回收站		系统文件...	
控制面板		系统文件...	
Xmanager Enterprise 4	1,467	快捷方式	2014/07/17

Remote Name	Size	Type	Modified
apache-maven-3.0.5-bin.t...	5,144,659	360压缩	2014/04/29

3. 解压缩 apache-maven-3.0.5-bin.tar.gz 包

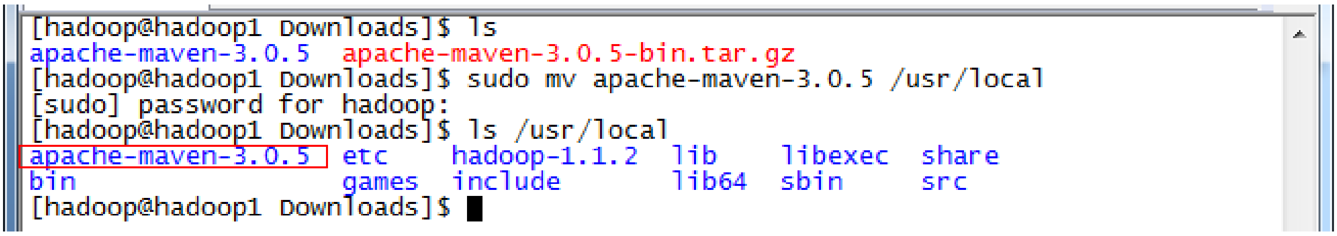
$tar -zxvf apache-maven-3.0.5-bin.tar.gz



```
[hadoop@hadoop1 ~]$ cd /home/hadoop/Downloads/
[hadoop@hadoop1 Downloads]$ ls
apache-maven-3.0.5-bin.tar.gz
[hadoop@hadoop1 Downloads]$ tar -zxvf apache-maven-3.0.5-bin.tar.gz
apache-maven-3.0.5/LICENSE.txt
apache-maven-3.0.5/NOTICE.txt
apache-maven-3.0.5/README.txt
apache-maven-3.0.5/bin/m2.conf
```

4. 把 apache-maven-3.0.5 目录移到/usr/local 目录下

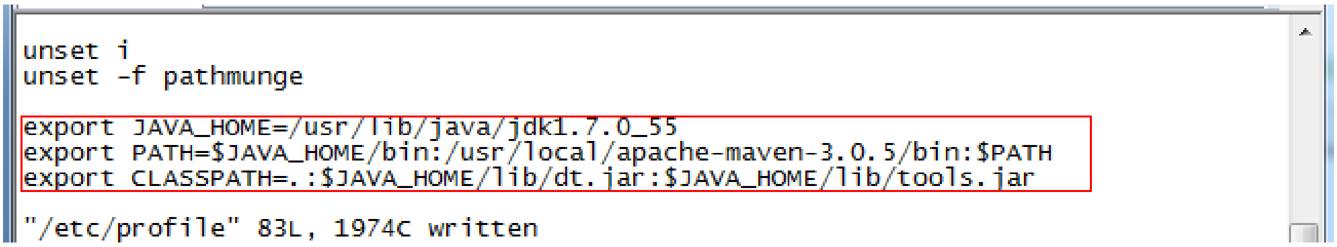
$sudo mv apache-maven-3.0.5 /usr/local



```
[hadoop@hadoop1 Downloads]$ ls
apache-maven-3.0.5  apache-maven-3.0.5-bin.tar.gz
[hadoop@hadoop1 Downloads]$ sudo mv apache-maven-3.0.5 /usr/local
[sudo] password for hadoop:
[hadoop@hadoop1 Downloads]$ ls /usr/local
apache-maven-3.0.5  etc    hadoop-1.1.2  lib    libexec  share
bin                 games  include       lib64  sbin     src
[hadoop@hadoop1 Downloads]$ 
```

5. 在/etc/profile 配置文件中加入如下设置

export PATH=$JAVA_HOME/bin:/usr/local/apache-maven-3.0.5/bin:$PATH

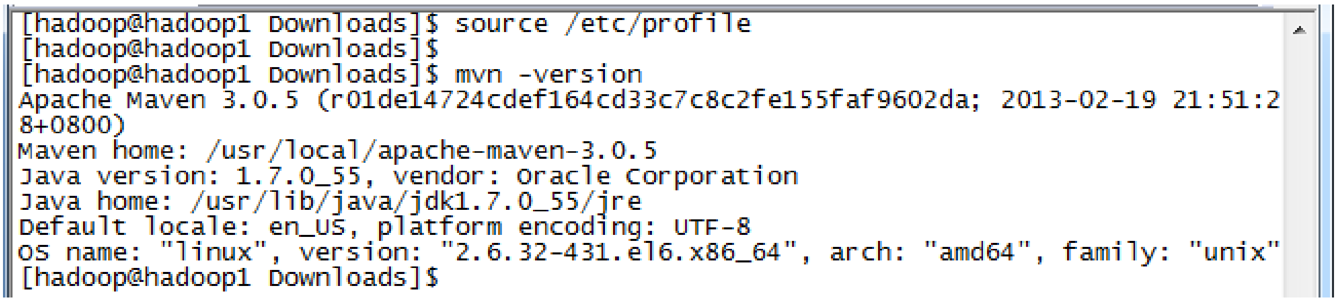


```

unset i
unset -f pathmunge

export JAVA_HOME=/usr/lib/java/jdk1.7.0_55
export PATH=$JAVA_HOME/bin:/usr/local/apache-maven-3.0.5/bin:$PATH
export CLASSPATH=.:$JAVA_HOME/lib/dt.jar:$JAVA_HOME/lib/tools.jar

"/etc/profile" 83L, 1974C written
```

6. 编辑/etc/profile 文件并验证配置是否成功:

$source /etc/profile

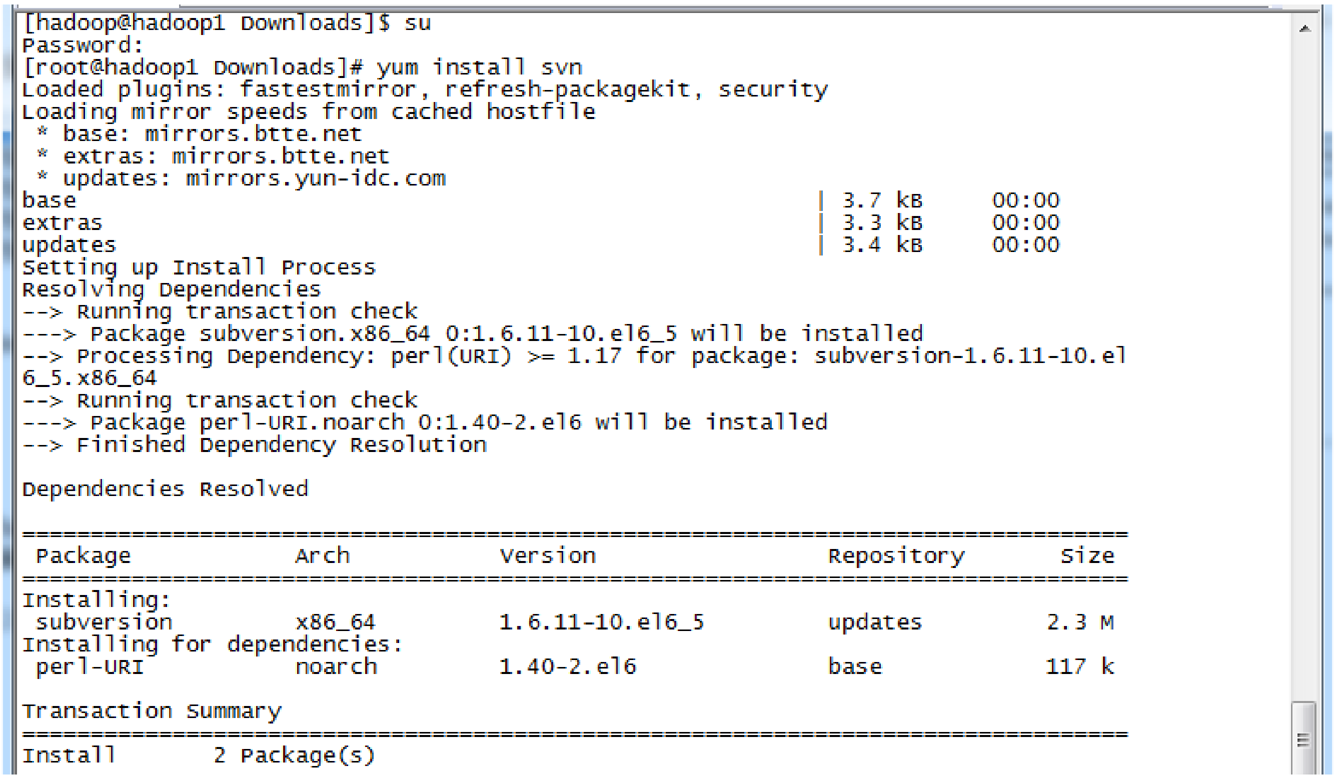
$mvn -version



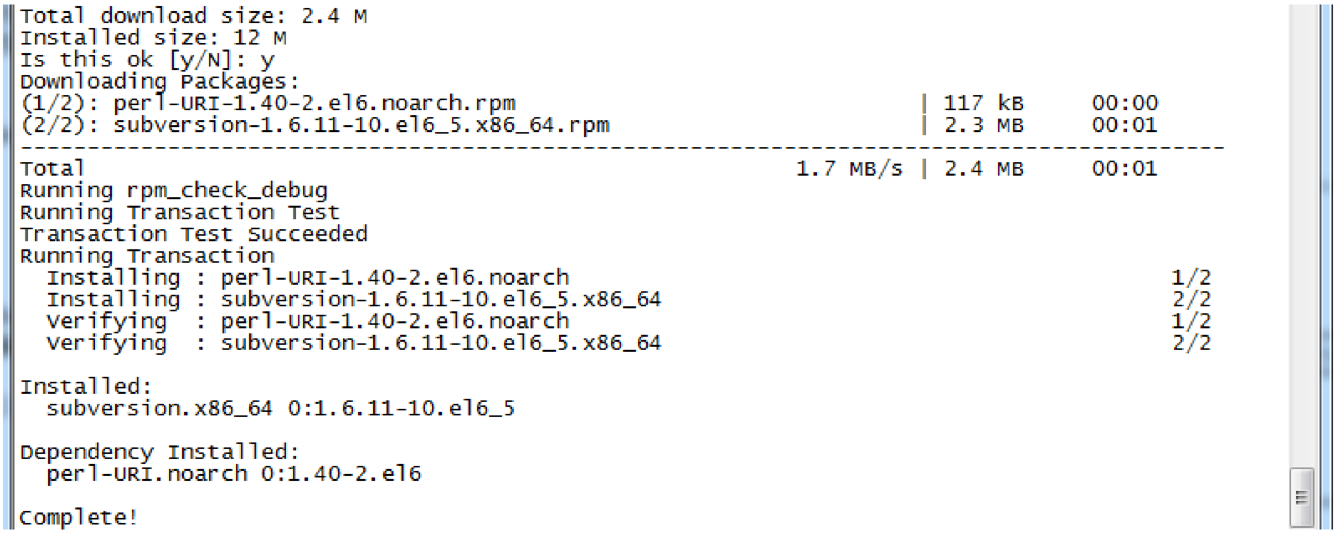
```
[hadoop@hadoop1 Downloads]$ source /etc/profile
[hadoop@hadoop1 Downloads]$
[hadoop@hadoop1 Downloads]$ mvn -version
Apache Maven 3.0.5 (r01de14724cdef164cd33c7c8c2fe155faf9602da; 2013-02-19 21:51:2
8+0800)
Maven home: /usr/local/apache-maven-3.0.5
Java version: 1.7.0_55, vendor: Oracle Corporation
Java home: /usr/lib/java/jdk1.7.0_55/jre
Default locale: en_US, platform encoding: UTF-8
OS name: "linux", version: "2.6.32-431.el6.x86_64", arch: "amd64", family: "unix"
[hadoop@hadoop1 Downloads]$
```

1.1.2 以 root 用户使用 yum 安装 svn

#yum install svn



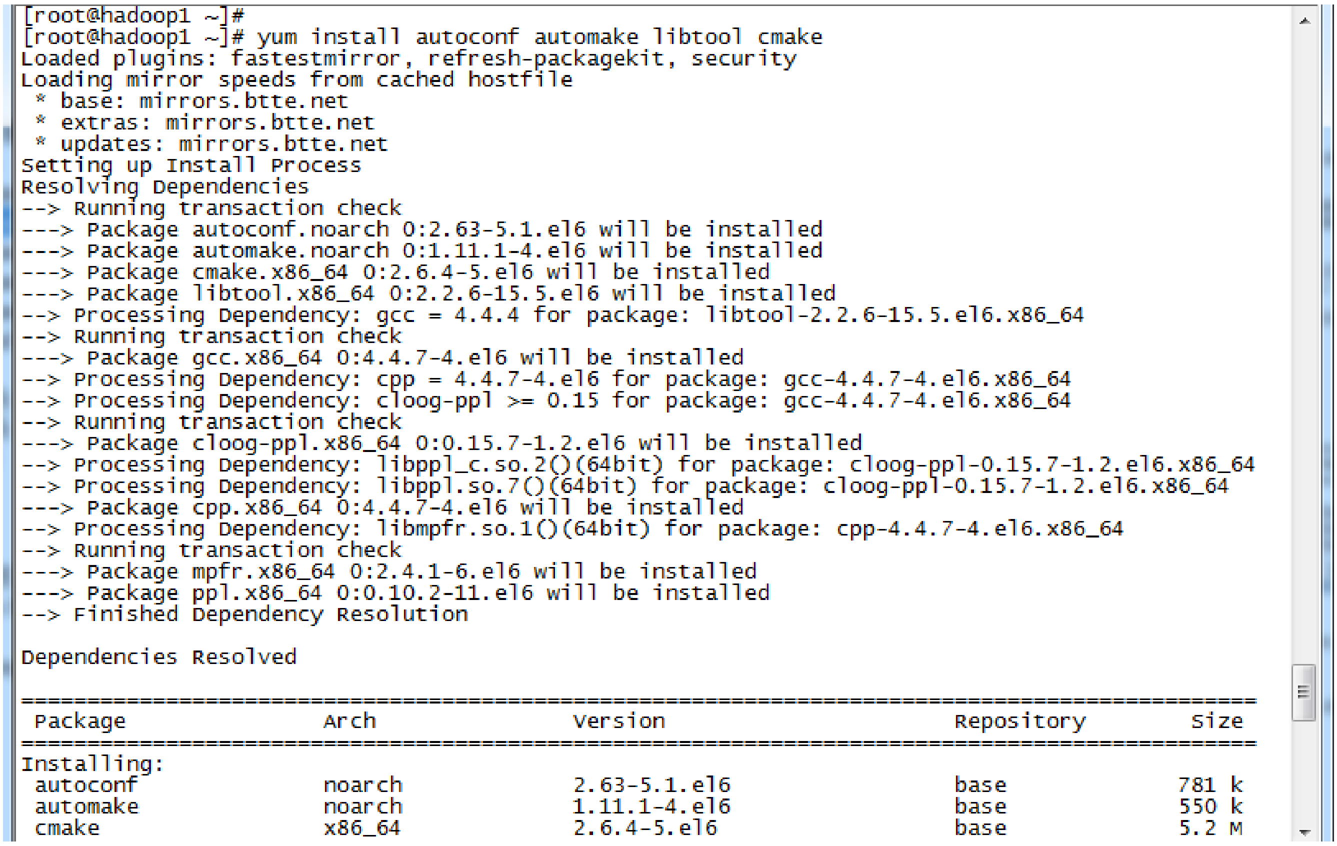
```
[hadoop@hadoop1 Downloads]$ su
Password:
[root@hadoop1 Downloads]# yum install svn
Loaded plugins: fastestmirror, refresh-packagekit, security
Loading mirror speeds from cached hostfile
 * base: mirrors.btte.net
 * extras: mirrors.btte.net
 * updates: mirrors.yun-idc.com
base                                                     | 3.7 kB     00:00
extras                                                   | 3.3 kB     00:00
updates                                                  | 3.4 kB     00:00
Setting up Install Process
Resolving Dependencies
--> Running transaction check
---> Package subversion.x86_64 0:1.6.11-10.el6_5 will be installed
--> Processing Dependency: perl(URI) >= 1.17 for package: subversion-1.6.11-10.el
6_5.x86_64
--> Running transaction check
---> Package perl-URI.noarch 0:1.40-2.el6 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

================================================================================
 Package              Arch            Version                Repository    Size
================================================================================
Installing:
 subversion           x86_64          1.6.11-10.el6_5        updates      2.3 M
Installing for dependencies:
 perl-URI             noarch          1.40-2.el6             base         117 k

Transaction Summary
================================================================================
Install       2 Package(s)
```



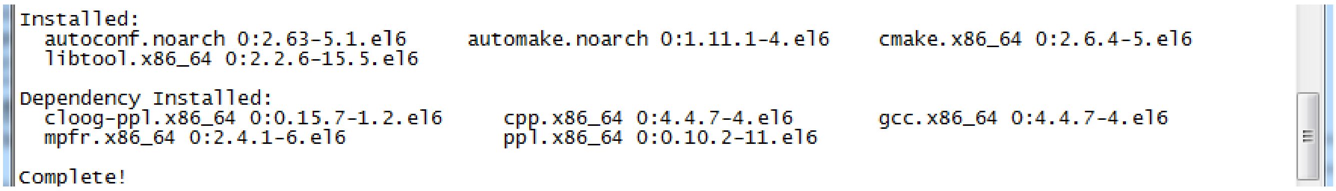
```
Total download size: 2.4 M
Installed size: 12 M
Is this ok [y/N]: y
Downloading Packages:
(1/2): perl-URI-1.40-2.el6.noarch.rpm                    | 117 kB     00:00
(2/2): subversion-1.6.11-10.el6_5.x86_64.rpm             | 2.3 MB     00:01
--------------------------------------------------------------------------------
Total                                           1.7 MB/s | 2.4 MB     00:01
Running rpm_check_debug
Running Transaction Test
Transaction Test Succeeded
Running Transaction
  Installing : perl-URI-1.40-2.el6.noarch                                   1/2
  Installing : subversion-1.6.11-10.el6_5.x86_64                            2/2
  Verifying  : perl-URI-1.40-2.el6.noarch                                   1/2
  Verifying  : subversion-1.6.11-10.el6_5.x86_64                            2/2

Installed:
  subversion.x86_64 0:1.6.11-10.el6_5

Dependency Installed:
  perl-URI.noarch 0:1.40-2.el6

Complete!
```

1.1.3 以 root 用户使用 yum 安装 autoconf automake libtool cmake

#yum install autoconf automake libtool cmake



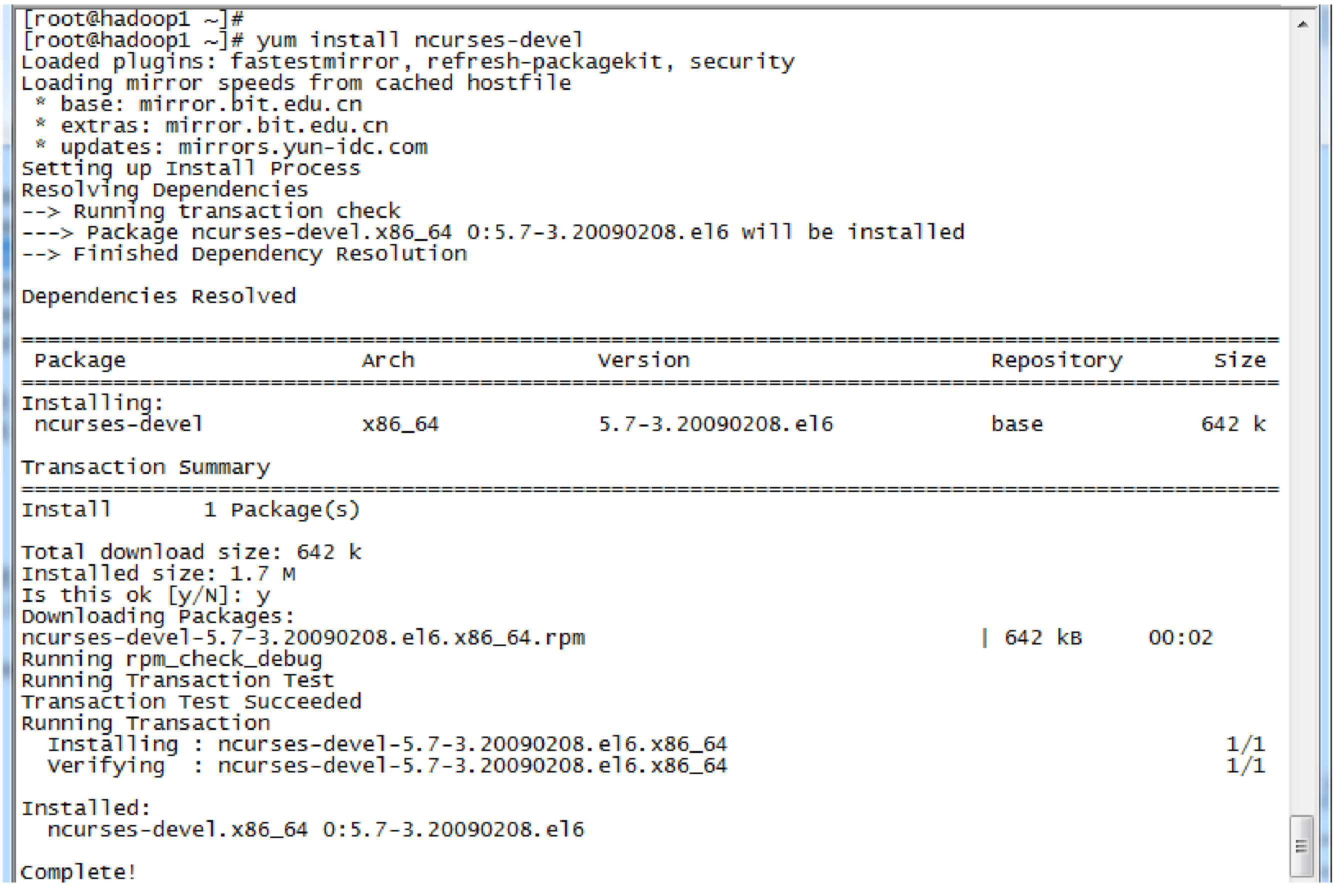
```
[root@hadoop1 ~]#
[root@hadoop1 ~]# yum install autoconf automake libtool cmake
Loaded plugins: fastestmirror, refresh-packagekit, security
Loading mirror speeds from cached hostfile
 * base: mirrors.btte.net
 * extras: mirrors.btte.net
 * updates: mirrors.btte.net
Setting up Install Process
Resolving Dependencies
--> Running transaction check
---> Package autoconf.noarch 0:2.63-5.1.el6 will be installed
---> Package automake.noarch 0:1.11.1-4.el6 will be installed
---> Package cmake.x86_64 0:2.6.4-5.el6 will be installed
---> Package libtool.x86_64 0:2.2.6-15.5.el6 will be installed
--> Processing Dependency: gcc = 4.4.4 for package: libtool-2.2.6-15.5.el6.x86_64
--> Running transaction check
---> Package gcc.x86_64 0:4.4.7-4.el6 will be installed
--> Processing Dependency: cpp = 4.4.7-4.el6 for package: gcc-4.4.7-4.el6.x86_64
--> Processing Dependency: cloog-ppl >= 0.15 for package: gcc-4.4.7-4.el6.x86_64
--> Running transaction check
---> Package cloog-ppl.x86_64 0:0.15.7-1.2.el6 will be installed
--> Processing Dependency: libppl_c.so.2()(64bit) for package: cloog-ppl-0.15.7-1.2.el6.x86_64
--> Processing Dependency: libppl.so.7()(64bit) for package: cloog-ppl-0.15.7-1.2.el6.x86_64
---> Package cpp.x86_64 0:4.4.7-4.el6 will be installed
--> Processing Dependency: libmpfr.so.1()(64bit) for package: cpp-4.4.7-4.el6.x86_64
--> Running transaction check
---> Package mpfr.x86_64 0:2.4.1-6.el6 will be installed
---> Package ppl.x86_64 0:0.10.2-11.el6 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

==============================================================================================
 Package                 Arch              Version                      Repository       Size
==============================================================================================
Installing:
 autoconf                noarch            2.63-5.1.el6                 base            781 k
 automake                noarch            1.11.1-4.el6                 base            550 k
 cmake                   x86_64            2.6.4-5.el6                  base            5.2 M
```



```
Installed:
  autoconf.noarch 0:2.63-5.1.el6        automake.noarch 0:1.11.1-4.el6      cmake.x86_64 0:2.6.4-5.el6
  libtool.x86_64 0:2.2.6-15.5.el6

Dependency Installed:
  cloog-ppl.x86_64 0:0.15.7-1.2.el6        cpp.x86_64 0:4.4.7-4.el6         gcc.x86_64 0:4.4.7-4.el6
  mpfr.x86_64 0:2.4.1-6.el6                ppl.x86_64 0:0.10.2-11.el6

Complete!
```

1.1.4 以 root 用户使用 yum 安装 ncurses-devel

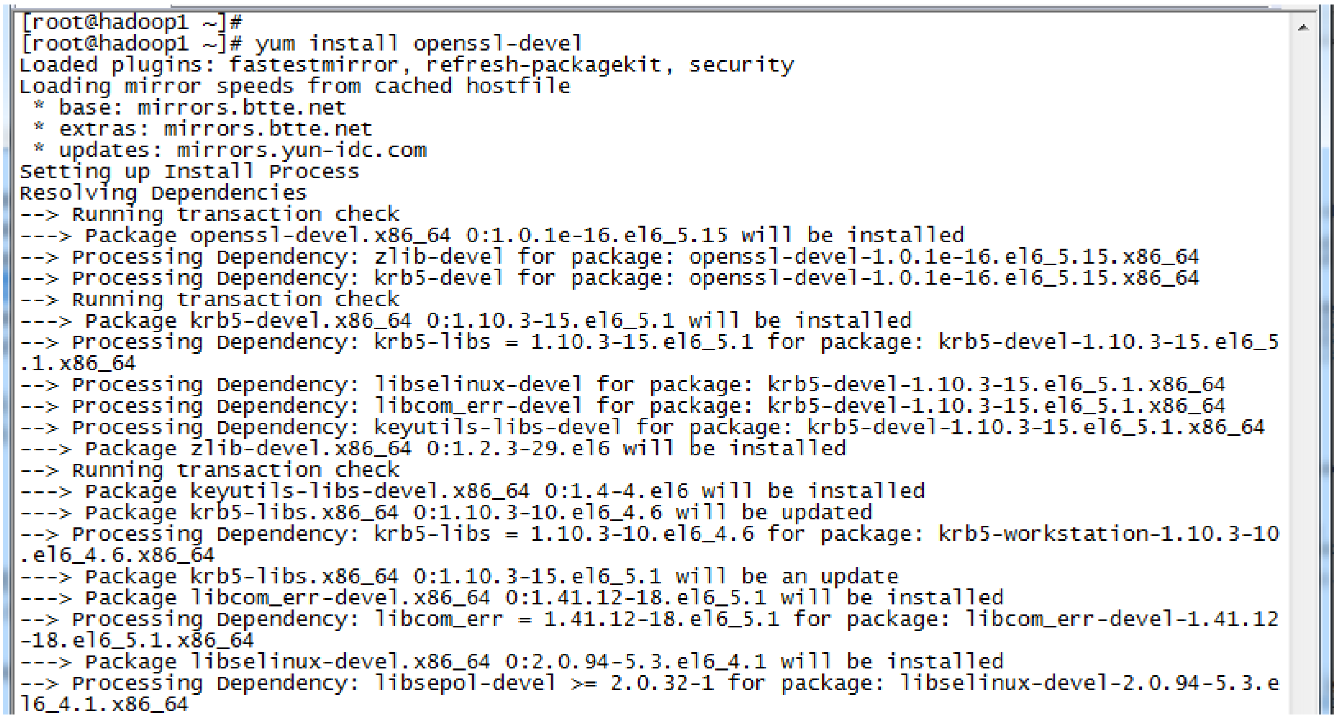
#yum install ncurses-devel



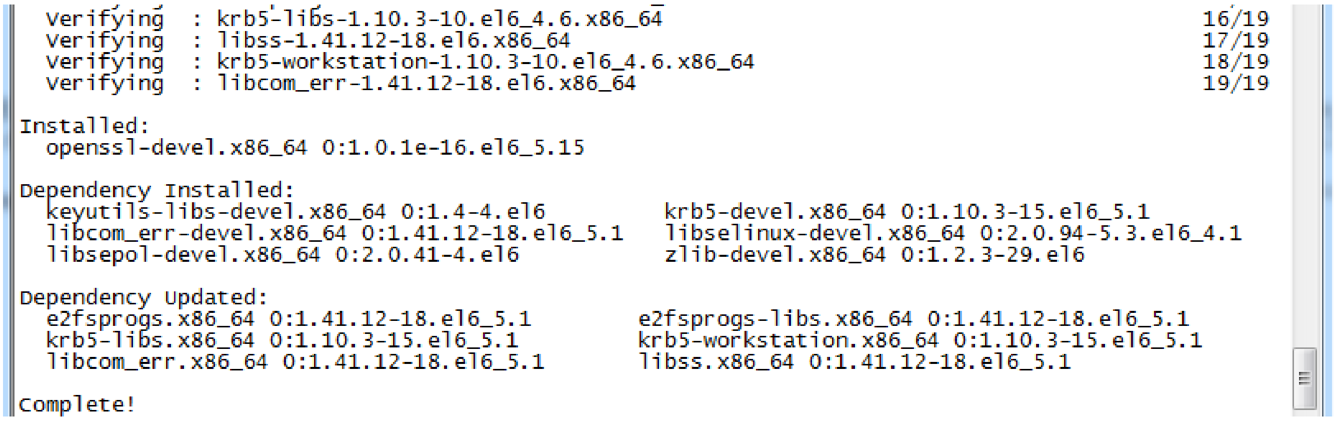
```
[root@hadoop1 ~]#
[root@hadoop1 ~]# yum install ncurses-devel
Loaded plugins: fastestmirror, refresh-packagekit, security
Loading mirror speeds from cached hostfile
 * base: mirror.bit.edu.cn
 * extras: mirror.bit.edu.cn
 * updates: mirrors.yun-idc.com
Setting up Install Process
Resolving Dependencies
--> Running transaction check
---> Package ncurses-devel.x86_64 0:5.7-3.20090208.el6 will be installed
--> Finished Dependency Resolution

Dependencies Resolved

================================================================================================
 Package                  Arch              Version                       Repository       Size
================================================================================================
Installing:
 ncurses-devel            x86_64            5.7-3.20090208.el6            base            642 k

Transaction Summary
================================================================================================
Install       1 Package(s)

Total download size: 642 k
Installed size: 1.7 M
Is this ok [y/N]: y
Downloading Packages:
ncurses-devel-5.7-3.20090208.el6.x86_64.rpm                              | 642 kB     00:02
Running rpm_check_debug
Running Transaction Test
Transaction Test Succeeded
Running Transaction
  Installing : ncurses-devel-5.7-3.20090208.el6.x86_64                                      1/1
  Verifying  : ncurses-devel-5.7-3.20090208.el6.x86_64                                      1/1

Installed:
  ncurses-devel.x86_64 0:5.7-3.20090208.el6

Complete!
```

1.1.5 以 root 用户使用 yum 安装 openssl-devel

#yum install openssl-devel



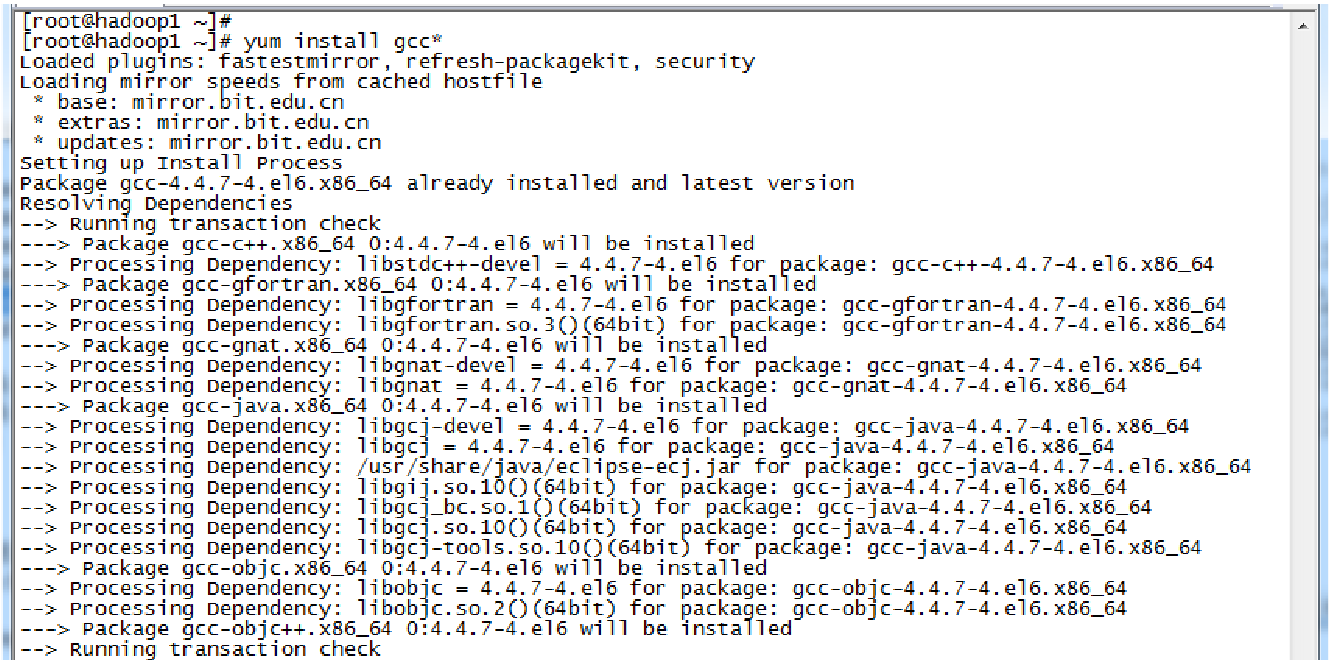
```
[root@hadoop1 ~]#
[root@hadoop1 ~]# yum install openssl-devel
Loaded plugins: fastestmirror, refresh-packagekit, security
Loading mirror speeds from cached hostfile
 * base: mirrors.btte.net
 * extras: mirrors.btte.net
 * updates: mirrors.yun-idc.com
Setting up Install Process
Resolving Dependencies
--> Running transaction check
---> Package openssl-devel.x86_64 0:1.0.1e-16.el6_5.15 will be installed
--> Processing Dependency: zlib-devel for package: openssl-devel-1.0.1e-16.el6_5.15.x86_64
--> Processing Dependency: krb5-devel for package: openssl-devel-1.0.1e-16.el6_5.15.x86_64
--> Running transaction check
---> Package krb5-devel.x86_64 0:1.10.3-15.el6_5.1 will be installed
--> Processing Dependency: krb5-libs = 1.10.3-15.el6_5.1 for package: krb5-devel-1.10.3-15.el6_5
.1.x86_64
--> Processing Dependency: libselinux-devel for package: krb5-devel-1.10.3-15.el6_5.1.x86_64
--> Processing Dependency: libcom_err-devel for package: krb5-devel-1.10.3-15.el6_5.1.x86_64
--> Processing Dependency: keyutils-libs-devel for package: krb5-devel-1.10.3-15.el6_5.1.x86_64
---> Package zlib-devel.x86_64 0:1.2.3-29.el6 will be installed
--> Running transaction check
---> Package keyutils-libs-devel.x86_64 0:1.4-4.el6 will be installed
---> Package krb5-libs.x86_64 0:1.10.3-10.el6_4.6 will be updated
--> Processing Dependency: krb5-libs = 1.10.3-10.el6_4.6 for package: krb5-workstation-1.10.3-10
.el6_4.6.x86_64
---> Package krb5-libs.x86_64 0:1.10.3-15.el6_5.1 will be an update
---> Package libcom_err-devel.x86_64 0:1.41.12-18.el6_5.1 will be installed
--> Processing Dependency: libcom_err = 1.41.12-18.el6_5.1 for package: libcom_err-devel-1.41.12
-18.el6_5.1.x86_64
---> Package libselinux-devel.x86_64 0:2.0.94-5.3.el6_4.1 will be installed
--> Processing Dependency: libsepol-devel >= 2.0.32-1 for package: libselinux-devel-2.0.94-5.3.e
l6_4.1.x86_64
```



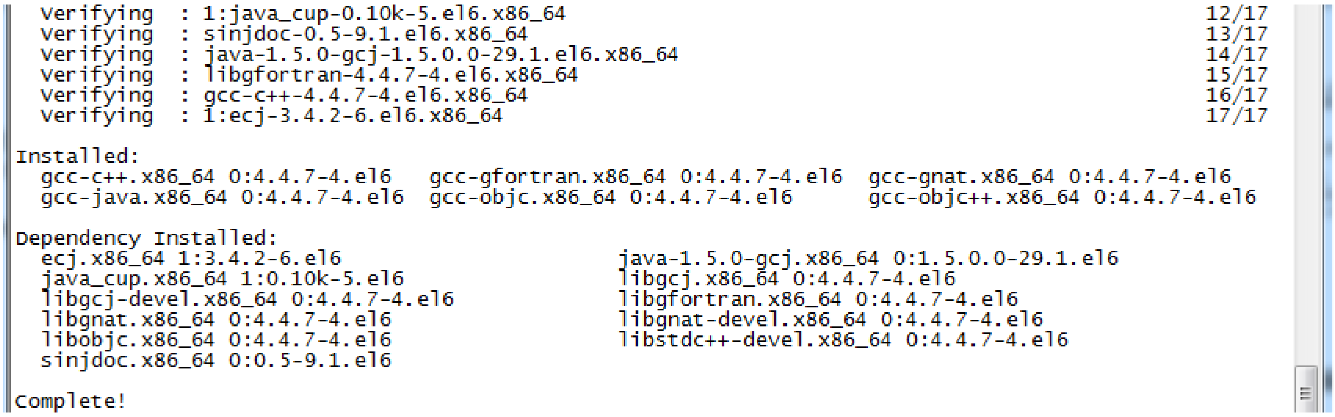
```
  Verifying  : krb5-libs-1.10.3-10.el6_4.6.x86_64                         16/19
  Verifying  : libss-1.41.12-18.el6.x86_64                                17/19
  Verifying  : krb5-workstation-1.10.3-10.el6_4.6.x86_64                  18/19
  Verifying  : libcom_err-1.41.12-18.el6.x86_64                           19/19

Installed:
  openssl-devel.x86_64 0:1.0.1e-16.el6_5.15

Dependency Installed:
  keyutils-libs-devel.x86_64 0:1.4-4.el6         krb5-devel.x86_64 0:1.10.3-15.el6_5.1
  libcom_err-devel.x86_64 0:1.41.12-18.el6_5.1   libselinux-devel.x86_64 0:2.0.94-5.3.el6_4.1
  libsepol-devel.x86_64 0:2.0.41-4.el6           zlib-devel.x86_64 0:1.2.3-29.el6

Dependency Updated:
  e2fsprogs.x86_64 0:1.41.12-18.el6_5.1        e2fsprogs-libs.x86_64 0:1.41.12-18.el6_5.1
  krb5-libs.x86_64 0:1.10.3-15.el6_5.1         krb5-workstation.x86_64 0:1.10.3-15.el6_5.1
  libcom_err.x86_64 0:1.41.12-18.el6_5.1       libss.x86_64 0:1.41.12-18.el6_5.1

Complete!
```

1.1.6 以 root 用户使用 yum 安装 gcc*

*#yum install gcc**



```
[root@hadoop1 ~]#
[root@hadoop1 ~]# yum install gcc*
Loaded plugins: fastestmirror, refresh-packagekit, security
Loading mirror speeds from cached hostfile
 * base: mirror.bit.edu.cn
 * extras: mirror.bit.edu.cn
 * updates: mirror.bit.edu.cn
Setting up Install Process
Package gcc-4.4.7-4.el6.x86_64 already installed and latest version
Resolving Dependencies
--> Running transaction check
---> Package gcc-c++.x86_64 0:4.4.7-4.el6 will be installed
--> Processing Dependency: libstdc++-devel = 4.4.7-4.el6 for package: gcc-c++-4.4.7-4.el6.x86_64
---> Package gcc-gfortran.x86_64 0:4.4.7-4.el6 will be installed
--> Processing Dependency: libgfortran = 4.4.7-4.el6 for package: gcc-gfortran-4.4.7-4.el6.x86_64
--> Processing Dependency: libgfortran.so.3()(64bit) for package: gcc-gfortran-4.4.7-4.el6.x86_64
---> Package gcc-gnat.x86_64 0:4.4.7-4.el6 will be installed
--> Processing Dependency: libgnat-devel = 4.4.7-4.el6 for package: gcc-gnat-4.4.7-4.el6.x86_64
--> Processing Dependency: libgnat = 4.4.7-4.el6 for package: gcc-gnat-4.4.7-4.el6.x86_64
---> Package gcc-java.x86_64 0:4.4.7-4.el6 will be installed
--> Processing Dependency: libgcj-devel = 4.4.7-4.el6 for package: gcc-java-4.4.7-4.el6.x86_64
--> Processing Dependency: libgcj = 4.4.7-4.el6 for package: gcc-java-4.4.7-4.el6.x86_64
--> Processing Dependency: /usr/share/java/eclipse-ecj.jar for package: gcc-java-4.4.7-4.el6.x86_64
--> Processing Dependency: libgij.so.10()(64bit) for package: gcc-java-4.4.7-4.el6.x86_64
--> Processing Dependency: libgcj_bc.so.1()(64bit) for package: gcc-java-4.4.7-4.el6.x86_64
--> Processing Dependency: libgcj.so.10()(64bit) for package: gcc-java-4.4.7-4.el6.x86_64
--> Processing Dependency: libgcj-tools.so.10()(64bit) for package: gcc-java-4.4.7-4.el6.x86_64
---> Package gcc-objc.x86_64 0:4.4.7-4.el6 will be installed
--> Processing Dependency: libobjc = 4.4.7-4.el6 for package: gcc-objc-4.4.7-4.el6.x86_64
--> Processing Dependency: libobjc.so.2()(64bit) for package: gcc-objc-4.4.7-4.el6.x86_64
---> Package gcc-objc++.x86_64 0:4.4.7-4.el6 will be installed
--> Running transaction check
```



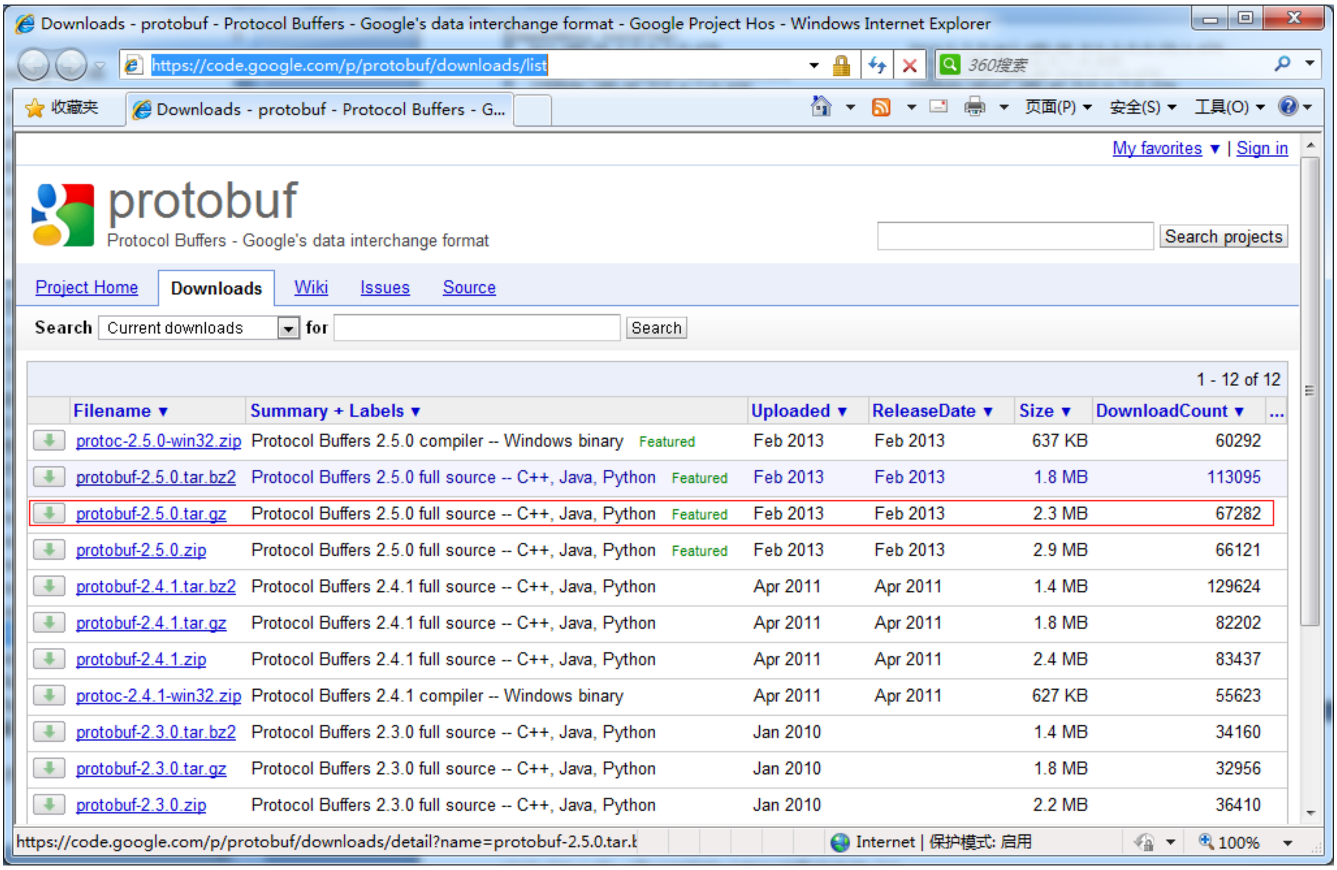
```
  Verifying  : 1:java_cup-0.10k-5.el6.x86_64                               12/17
  Verifying  : sinjdoc-0.5-9.1.el6.x86_64                                  13/17
  Verifying  : java-1.5.0-gcj-1.5.0.0-29.1.el6.x86_64                      14/17
  Verifying  : libgfortran-4.4.7-4.el6.x86_64                              15/17
  Verifying  : gcc-c++-4.4.7-4.el6.x86_64                                  16/17
  Verifying  : 1:ecj-3.4.2-6.el6.x86_64                                    17/17

Installed:
  gcc-c++.x86_64 0:4.4.7-4.el6    gcc-gfortran.x86_64 0:4.4.7-4.el6  gcc-gnat.x86_64 0:4.4.7-4.el6
  gcc-java.x86_64 0:4.4.7-4.el6   gcc-objc.x86_64 0:4.4.7-4.el6      gcc-objc++.x86_64 0:4.4.7-4.el6

Dependency Installed:
  ecj.x86_64 1:3.4.2-6.el6                     java-1.5.0-gcj.x86_64 0:1.5.0.0-29.1.el6
  java_cup.x86_64 1:0.10k-5.el6                libgcj.x86_64 0:4.4.7-4.el6
  libgcj-devel.x86_64 0:4.4.7-4.el6            libgfortran.x86_64 0:4.4.7-4.el6
  libgnat.x86_64 0:4.4.7-4.el6                 libgnat-devel.x86_64 0:4.4.7-4.el6
  libobjc.x86_64 0:4.4.7-4.el6                 libstdc++-devel.x86_64 0:4.4.7-4.el6
  sinjdoc.x86_64 0:0.5-9.1.el6

Complete!
```

1.1.7 安装并设置 protobuf

注：该程序包需要在 gcc 安装完毕后才能安装，否则提示无法找到 gcc 编译器。

1. 下载 protobuf 安装包

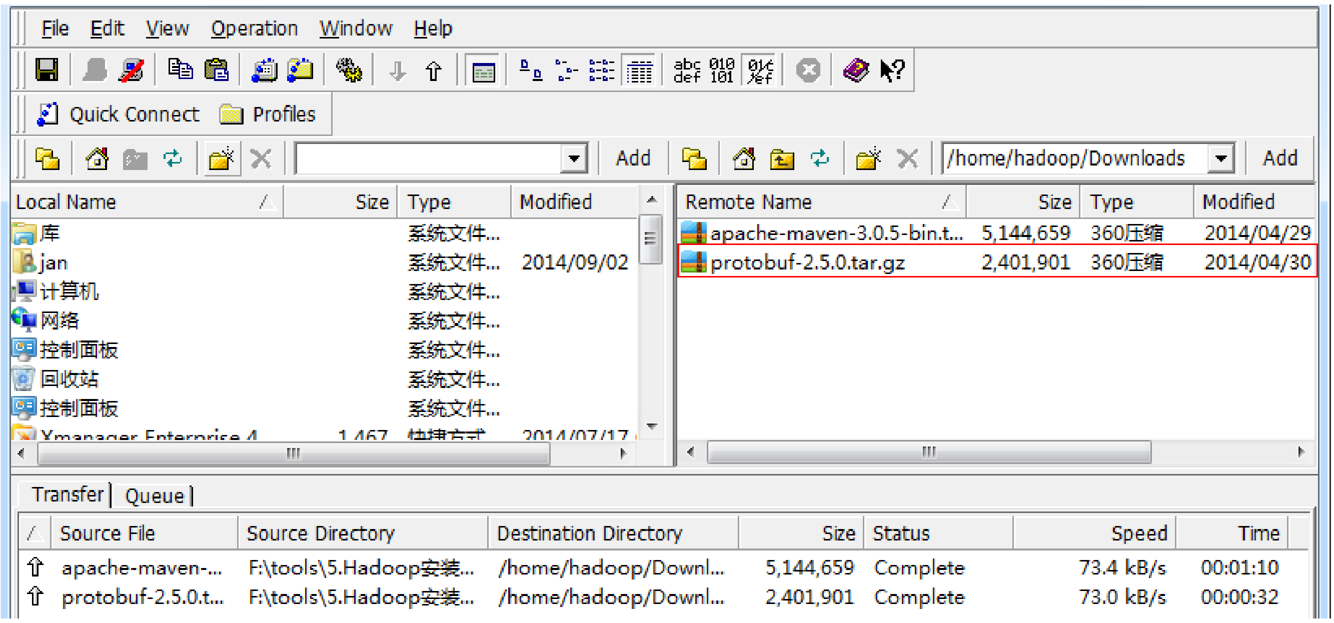
下载链接为： https://code.google.com/p/protobuf/downloads/list


Downloads - protobuf - Protocol Buffers - Google's data interchange format - Google Project Hos - Windows Internet Explorer
https://code.google.com/p/protobuf/downloads/list
My favorites ▾ | Sign in
protobuf
Protocol Buffers - Google's data interchange format
Search projects
Project Home
Downloads
Wiki
Issues
Source
Search
Current downloads
for
Search
https://code.google.com/p/protobuf/downloads/detail?name=protobuf-2.5.0.tar.
Internet | 保护模式: 启用
100%

1 - 12 of 12

Filename ▾	Summary + Labels ▾	Uploaded ▾	ReleaseDate ▾	Size ▾	DownloadCount ▾
protoc-2.5.0-win32.zip	Protocol Buffers 2.5.0 compiler -- Windows binary Featured	Feb 2013	Feb 2013	637 KB	60292
protobuf-2.5.0.tar.bz2	Protocol Buffers 2.5.0 full source -- C++, Java, Python Featured	Feb 2013	Feb 2013	1.8 MB	113095
protobuf-2.5.0.tar.gz	Protocol Buffers 2.5.0 full source -- C++, Java, Python Featured	Feb 2013	Feb 2013	2.3 MB	67282
protobuf-2.5.0.zip	Protocol Buffers 2.5.0 full source -- C++, Java, Python Featured	Feb 2013	Feb 2013	2.9 MB	66121
protobuf-2.4.1.tar.bz2	Protocol Buffers 2.4.1 full source -- C++, Java, Python	Apr 2011	Apr 2011	1.4 MB	129624
protobuf-2.4.1.tar.gz	Protocol Buffers 2.4.1 full source -- C++, Java, Python	Apr 2011	Apr 2011	1.8 MB	82202
protobuf-2.4.1.zip	Protocol Buffers 2.4.1 full source -- C++, Java, Python	Apr 2011	Apr 2011	2.4 MB	83437
protoc-2.4.1-win32.zip	Protocol Buffers 2.4.1 compiler -- Windows binary	Apr 2011	Apr 2011	627 KB	55623
protobuf-2.3.0.tar.bz2	Protocol Buffers 2.3.0 full source -- C++, Java, Python	Jan 2010		1.4 MB	34160
protobuf-2.3.0.tar.gz	Protocol Buffers 2.3.0 full source -- C++, Java, Python	Jan 2010		1.8 MB	32956
protobuf-2.3.0.zip	Protocol Buffers 2.3.0 full source -- C++, Java, Python	Jan 2010		2.2 MB	36410

2. 使用 ssh 工具把 protobuf-2.5.0.tar.gz 包上传到/home/hadoop/Downloads 目录



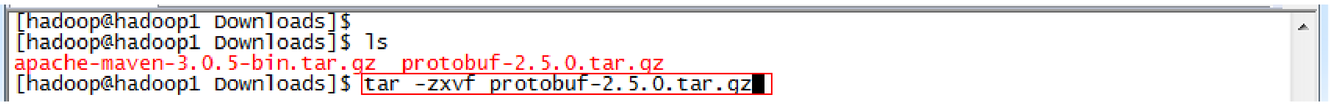
Local Name	Size	Type	Modified
库		系统文件...	
jan		系统文件...	2014/09/02
计算机		系统文件...	
网络		系统文件...	
控制面板		系统文件...	
回收站		系统文件...	
控制面板		系统文件...	
Xmanager Enterprise 4	1,467	快捷方式	2014/07/17

Remote Name	Size	Type	Modified
apache-maven-3.0.5-bin.t...	5,144,659	360压缩	2014/04/29
protobuf-2.5.0.tar.gz	2,401,901	360压缩	2014/04/30

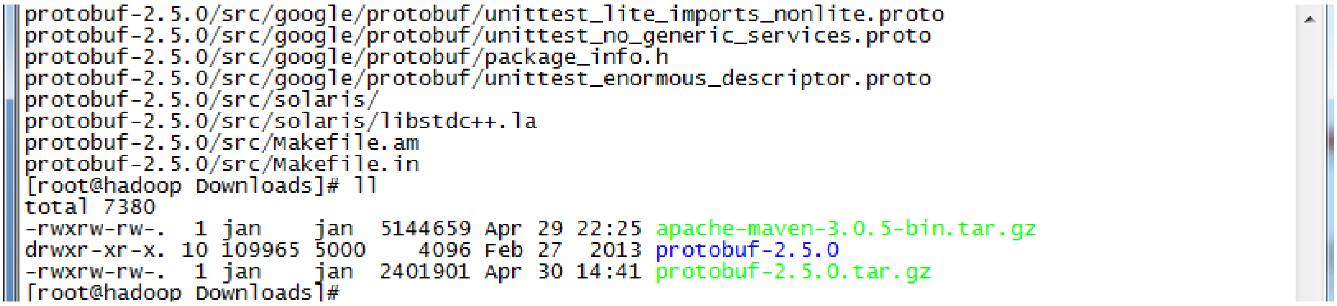
Source File	Source Directory	Destination Directory	Size	Status	Speed	Time
apache-maven-...	F:\tools\5.Hadoop安装...	/home/hadoop/Downl...	5,144,659	Complete	73.4 kB/s	00:01:10
protobuf-2.5.0.t...	F:\tools\5.Hadoop安装...	/home/hadoop/Downl...	2,401,901	Complete	73.0 kB/s	00:00:32

3. 解压安装包

$tar -zxvf protobuf-2.5.0.tar.gz



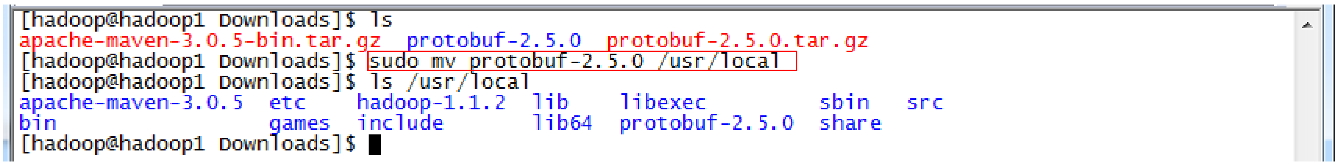
```
[hadoop@hadoop1 Downloads]$
[hadoop@hadoop1 Downloads]$ ls
apache-maven-3.0.5-bin.tar.gz  protobuf-2.5.0.tar.gz
[hadoop@hadoop1 Downloads]$ tar -zxvf protobuf-2.5.0.tar.gz
```



```
protobuf-2.5.0/src/google/protobuf/unittest_lite_imports_nonlite.proto
protobuf-2.5.0/src/google/protobuf/unittest_no_generic_services.proto
protobuf-2.5.0/src/google/protobuf/package_info.h
protobuf-2.5.0/src/google/protobuf/unittest_enormous_descriptor.proto
protobuf-2.5.0/src/solaris/
protobuf-2.5.0/src/solaris/libstdc++.la
protobuf-2.5.0/src/Makefile.am
protobuf-2.5.0/src/Makefile.in
[root@hadoop Downloads]# ll
total 7380
-rwxrw-rw-.  1 jan    jan  5144659 Apr 29 22:25 apache-maven-3.0.5-bin.tar.gz
drwxr-xr-x. 10 109965 5000    4096 Feb 27  2013 protobuf-2.5.0
-rwxrw-rw-.  1 jan    jan  2401901 Apr 30 14:41 protobuf-2.5.0.tar.gz
[root@hadoop Downloads]#
```

4. 把 protobuf-2.5.0 目录转移到/usr/local 目录下

$sudo mv protobuf-2.5.0 /usr/local



```
[hadoop@hadoop1 Downloads]$ ls
apache-maven-3.0.5-bin.tar.gz  protobuf-2.5.0  protobuf-2.5.0.tar.gz
[hadoop@hadoop1 Downloads]$ sudo mv protobuf-2.5.0 /usr/local
[hadoop@hadoop1 Downloads]$ ls /usr/local
apache-maven-3.0.5  etc    hadoop-1.1.2  lib    libexec         sbin   src
bin                 games  include       lib64  protobuf-2.5.0  share
[hadoop@hadoop1 Downloads]$
```

5. 进行目录运行命令

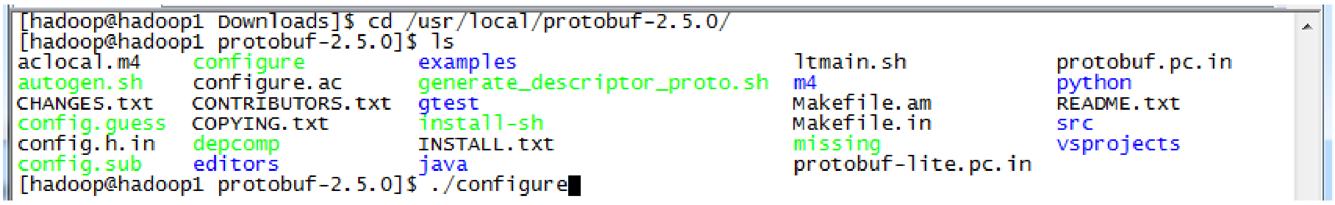
进入目录以 root 用户运行如下命令：

#./configure

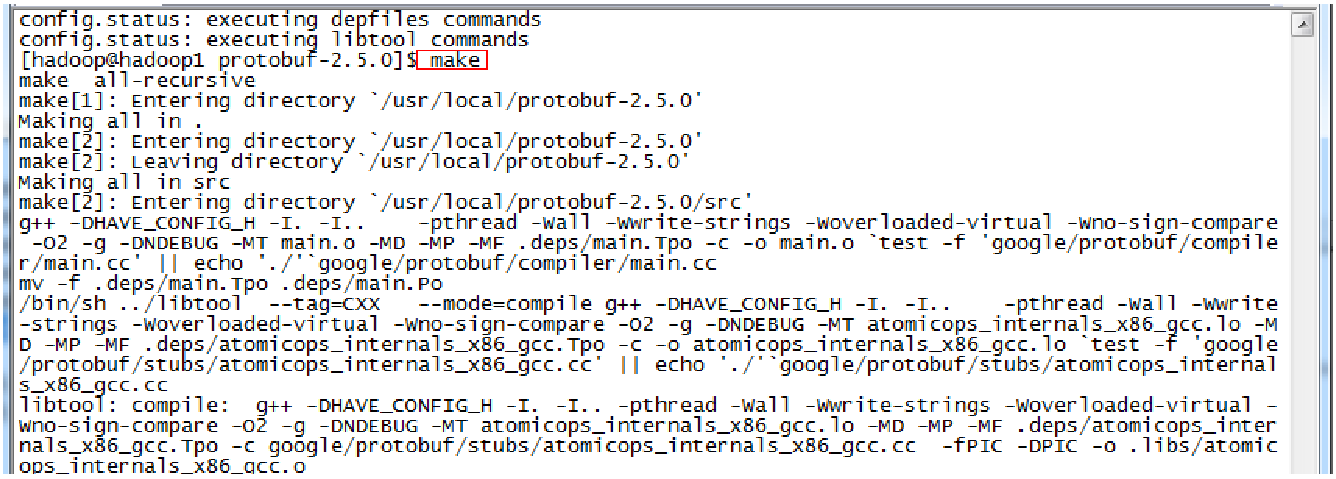
#make

#make check

#make install



```
[hadoop@hadoop1 Downloads]$ cd /usr/local/protobuf-2.5.0/
[hadoop@hadoop1 protobuf-2.5.0]$ ls
aclocal.m4    configure        examples                        ltmain.sh            protobuf.pc.in
autogen.sh    configure.ac     generate_descriptor_proto.sh    m4                   python
CHANGES.txt   CONTRIBUTORS.txt gtest                           Makefile.am          README.txt
config.guess  COPYING.txt      install-sh                      Makefile.in          src
config.h.in   depcomp          INSTALL.txt                     missing              vsprojects
config.sub    editors          java                            protobuf-lite.pc.in
[hadoop@hadoop1 protobuf-2.5.0]$ ./configure
```

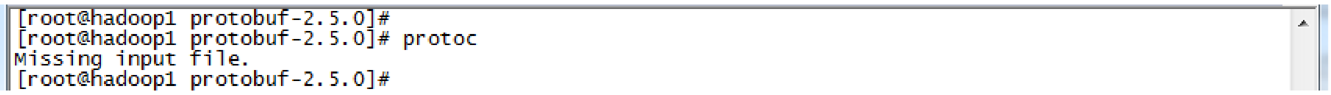


```
config.status: executing depfiles commands
config.status: executing libtool commands
[hadoop@hadoop1 protobuf-2.5.0]$ make
make  all-recursive
make[1]: Entering directory `/usr/local/protobuf-2.5.0'
Making all in .
make[2]: Entering directory `/usr/local/protobuf-2.5.0'
make[2]: Leaving directory `/usr/local/protobuf-2.5.0'
Making all in src
make[2]: Entering directory `/usr/local/protobuf-2.5.0/src'
g++ -DHAVE_CONFIG_H -I. -I..    -pthread -Wall -Wwrite-strings -Woverloaded-virtual -Wno-sign-compare
 -O2 -g -DNDEBUG -MT main.o -MD -MP -MF .deps/main.Tpo -c -o main.o `test -f 'google/protobuf/compile
r/main.cc' || echo './'`google/protobuf/compiler/main.cc
mv -f .deps/main.Tpo .deps/main.Po
/bin/sh ../libtool  --tag=CXX   --mode=compile g++ -DHAVE_CONFIG_H -I. -I..    -pthread -Wall -Wwrite
-strings -Woverloaded-virtual -Wno-sign-compare -O2 -g -DNDEBUG -MT atomicops_internals_x86_gcc.lo -M
D -MP -MF .deps/atomicops_internals_x86_gcc.Tpo -c -o atomicops_internals_x86_gcc.lo `test -f 'google
/protobuf/stubs/atomicops_internals_x86_gcc.cc' || echo './'`google/protobuf/stubs/atomicops_internal
s_x86_gcc.cc
libtool: compile:  g++ -DHAVE_CONFIG_H -I. -I.. -pthread -Wall -Wwrite-strings -Woverloaded-virtual -
Wno-sign-compare -O2 -g -DNDEBUG -MT atomicops_internals_x86_gcc.lo -MD -MP -MF .deps/atomicops_inter
nals_x86_gcc.Tpo -c google/protobuf/stubs/atomicops_internals_x86_gcc.cc  -fPIC -DPIC -o .libs/atomic
ops_internals_x86_gcc.o
```

6. 验证是否安装成功

运行成功之后，通过如下方式来验证是否安装成功

#protoc



```
[root@hadoop1 protobuf-2.5.0]#
[root@hadoop1 protobuf-2.5.0]# protoc
Missing input file.
[root@hadoop1 protobuf-2.5.0]#
```

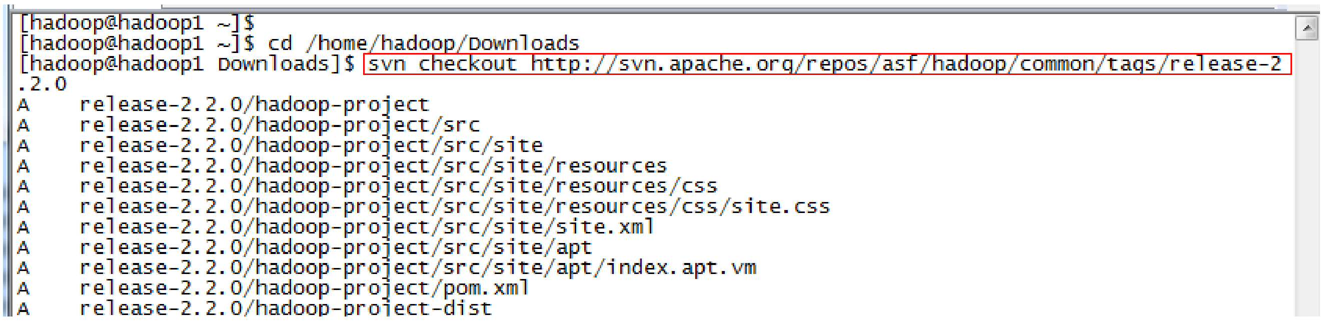
1.2 编译 Hadoop

1.2.1 下载 Hadoop 源代码 Release2.2.0

通过 SVN 获取 Hadoop2.2.0 源代码，在/home/hadoop/Downloads 目录下命令：

$svn checkout http://svn.apache.org/repos/asf/hadoop/common/tags/release-2.2.0

获取时间较长，大小约 75.3M



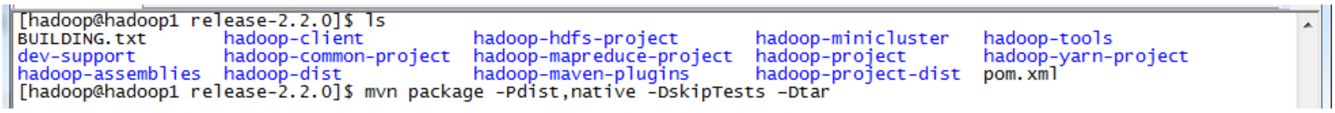
```
[hadoop@hadoop1 ~]$
[hadoop@hadoop1 ~]$ cd /home/hadoop/Downloads
[hadoop@hadoop1 Downloads]$ svn checkout http://svn.apache.org/repos/asf/hadoop/common/tags/release-2
.2.0
A    release-2.2.0/hadoop-project
A    release-2.2.0/hadoop-project/src
A    release-2.2.0/hadoop-project/src/site
A    release-2.2.0/hadoop-project/src/site/resources
A    release-2.2.0/hadoop-project/src/site/resources/css
A    release-2.2.0/hadoop-project/src/site/resources/css/site.css
A    release-2.2.0/hadoop-project/src/site/site.xml
A    release-2.2.0/hadoop-project/src/site/apt
A    release-2.2.0/hadoop-project/src/site/apt/index.apt.vm
A    release-2.2.0/hadoop-project/pom.xml
A    release-2.2.0/hadoop-project-dist
```

1.2.2 编译 Hadoop 源代码

:) 由于 hadoop2.2.0 在 svn 中 pom.xml 有点问题，会造成编译中出错，可先参考 3.2 修复该问题。在 Hadoop 源代码的根目录执行如下命令：

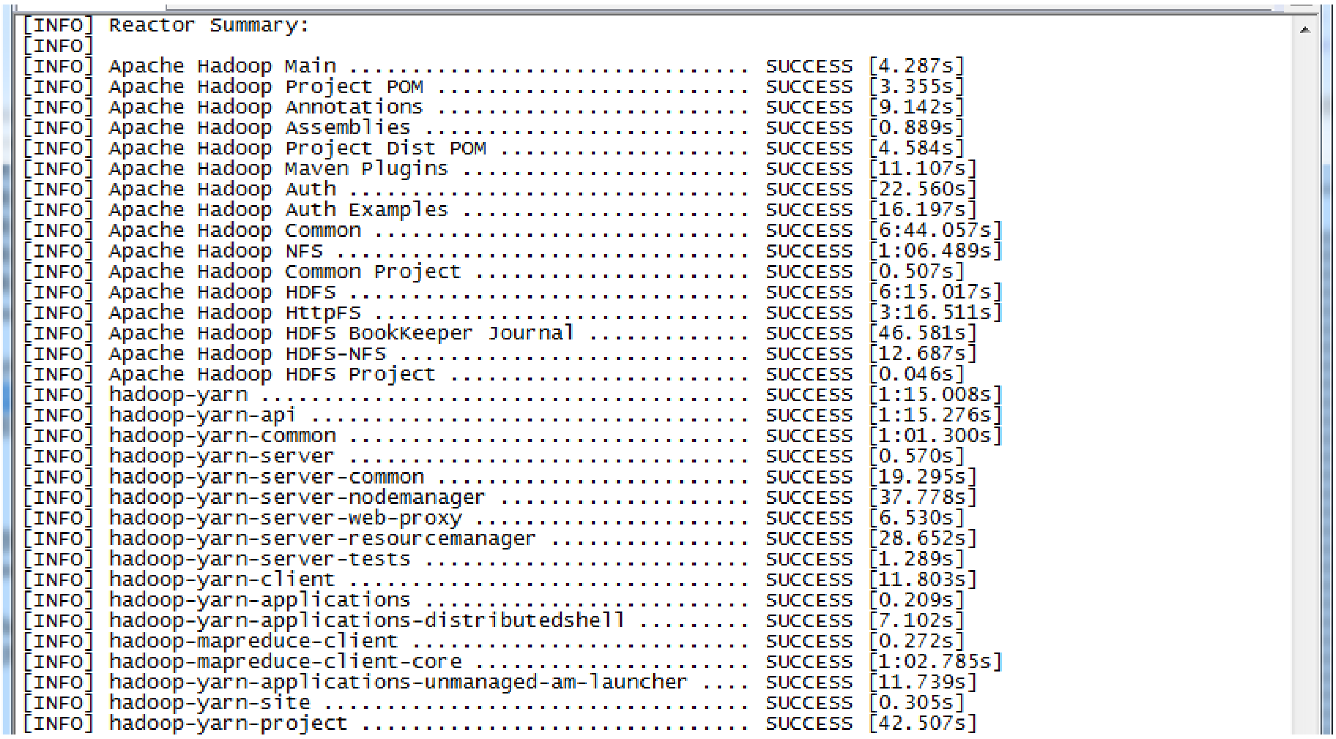
$mvn package -Pdist,native -DskipTests –Dtar

(**注意：**这行命令需要手工输入，如果复制执行会报异常！)

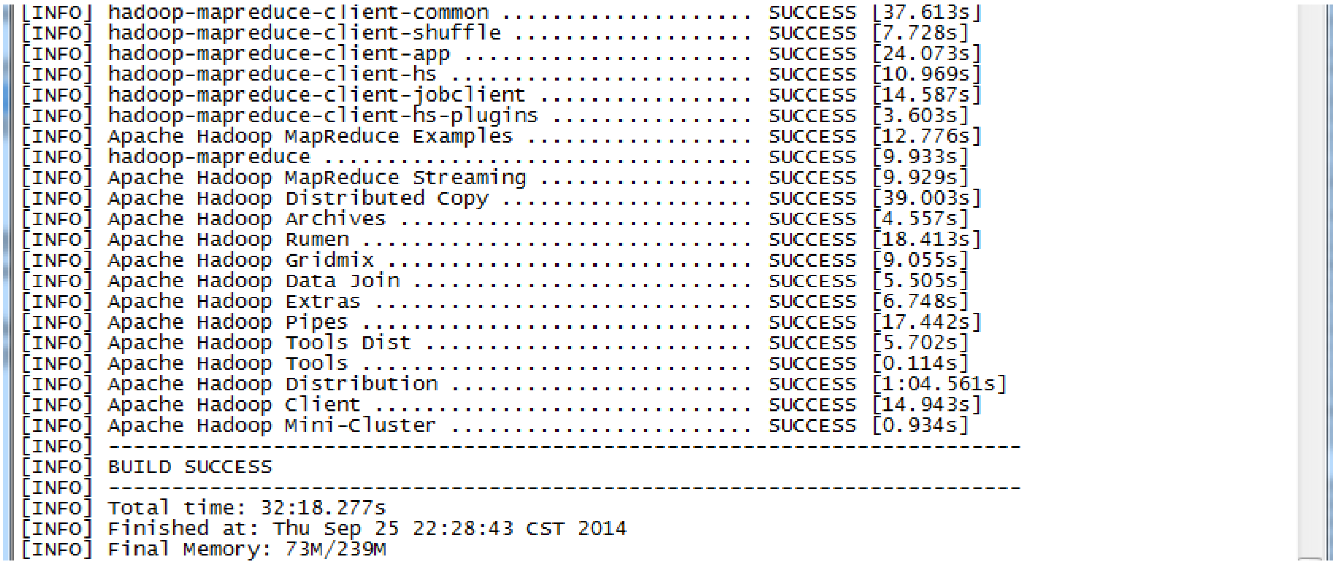


```
[hadoop@hadoop1 release-2.2.0]$ ls
BUILDING.txt       hadoop-client          hadoop-hdfs-project       hadoop-minicluster    hadoop-tools
dev-support        hadoop-common-project  hadoop-mapreduce-project  hadoop-project        hadoop-yarn-project
hadoop-assemblies  hadoop-dist            hadoop-maven-plugins      hadoop-project-dist   pom.xml
[hadoop@hadoop1 release-2.2.0]$ mvn package -Pdist,native -DskipTests -Dtar
```

耗费的时间较长，总共花费了 32 分钟，在编译过程需要联网，从网络中下载所需要的资料。



```
[INFO] Reactor Summary:
[INFO]
[INFO] Apache Hadoop Main ................................ SUCCESS [4.287s]
[INFO] Apache Hadoop Project POM ......................... SUCCESS [3.355s]
[INFO] Apache Hadoop Annotations ......................... SUCCESS [9.142s]
[INFO] Apache Hadoop Assemblies .......................... SUCCESS [0.889s]
[INFO] Apache Hadoop Project Dist POM .................... SUCCESS [4.584s]
[INFO] Apache Hadoop Maven Plugins ....................... SUCCESS [11.107s]
[INFO] Apache Hadoop Auth ................................ SUCCESS [22.560s]
[INFO] Apache Hadoop Auth Examples ....................... SUCCESS [16.197s]
[INFO] Apache Hadoop Common .............................. SUCCESS [6:44.057s]
[INFO] Apache Hadoop NFS ................................. SUCCESS [1:06.489s]
[INFO] Apache Hadoop Common Project ...................... SUCCESS [0.507s]
[INFO] Apache Hadoop HDFS ................................ SUCCESS [6:15.017s]
[INFO] Apache Hadoop HttpFS .............................. SUCCESS [3:16.511s]
[INFO] Apache Hadoop HDFS BookKeeper Journal ............. SUCCESS [46.581s]
[INFO] Apache Hadoop HDFS-NFS ............................ SUCCESS [12.687s]
[INFO] Apache Hadoop HDFS Project ........................ SUCCESS [0.046s]
[INFO] hadoop-yarn ....................................... SUCCESS [1:15.008s]
[INFO] hadoop-yarn-api ................................... SUCCESS [1:15.276s]
[INFO] hadoop-yarn-common ................................ SUCCESS [1:01.300s]
[INFO] hadoop-yarn-server ................................ SUCCESS [0.570s]
[INFO] hadoop-yarn-server-common ......................... SUCCESS [19.295s]
[INFO] hadoop-yarn-server-nodemanager .................... SUCCESS [37.778s]
[INFO] hadoop-yarn-server-web-proxy ...................... SUCCESS [6.530s]
[INFO] hadoop-yarn-server-resourcemanager ................ SUCCESS [28.652s]
[INFO] hadoop-yarn-server-tests .......................... SUCCESS [1.289s]
[INFO] hadoop-yarn-client ................................ SUCCESS [11.803s]
[INFO] hadoop-yarn-applications .......................... SUCCESS [0.209s]
[INFO] hadoop-yarn-applications-distributedshell ......... SUCCESS [7.102s]
[INFO] hadoop-mapreduce-client ........................... SUCCESS [0.272s]
[INFO] hadoop-mapreduce-client-core ...................... SUCCESS [1:02.785s]
[INFO] hadoop-yarn-applications-unmanaged-am-launcher .... SUCCESS [11.739s]
[INFO] hadoop-yarn-site .................................. SUCCESS [0.305s]
[INFO] hadoop-yarn-project ............................... SUCCESS [42.507s]
```



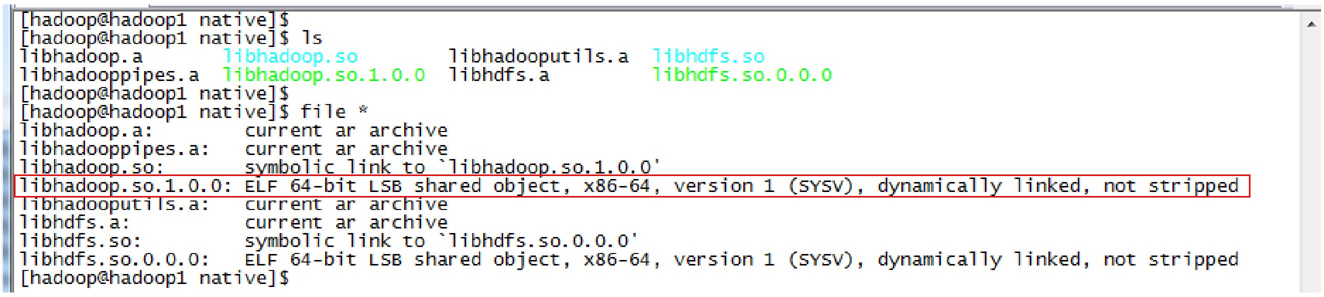
```
[INFO] hadoop-mapreduce-client-common .................... SUCCESS [37.613s]
[INFO] hadoop-mapreduce-client-shuffle ................... SUCCESS [7.728s]
[INFO] hadoop-mapreduce-client-app ....................... SUCCESS [24.073s]
[INFO] hadoop-mapreduce-client-hs ........................ SUCCESS [10.969s]
[INFO] hadoop-mapreduce-client-jobclient ................. SUCCESS [14.587s]
[INFO] hadoop-mapreduce-client-hs-plugins ................ SUCCESS [3.603s]
[INFO] Apache Hadoop MapReduce Examples .................. SUCCESS [12.776s]
[INFO] hadoop-mapreduce .................................. SUCCESS [9.933s]
[INFO] Apache Hadoop MapReduce Streaming ................. SUCCESS [9.929s]
[INFO] Apache Hadoop Distributed Copy .................... SUCCESS [39.003s]
[INFO] Apache Hadoop Archives ............................ SUCCESS [4.557s]
[INFO] Apache Hadoop Rumen ............................... SUCCESS [18.413s]
[INFO] Apache Hadoop Gridmix ............................. SUCCESS [9.055s]
[INFO] Apache Hadoop Data Join ........................... SUCCESS [5.505s]
[INFO] Apache Hadoop Extras .............................. SUCCESS [6.748s]
[INFO] Apache Hadoop Pipes ............................... SUCCESS [17.442s]
[INFO] Apache Hadoop Tools Dist .......................... SUCCESS [5.702s]
[INFO] Apache Hadoop Tools ............................... SUCCESS [0.114s]
[INFO] Apache Hadoop Distribution ........................ SUCCESS [1:04.561s]
[INFO] Apache Hadoop Client .............................. SUCCESS [14.943s]
[INFO] Apache Hadoop Mini-Cluster ........................ SUCCESS [0.934s]
[INFO] ------------------------------------------------------------------------
[INFO] BUILD SUCCESS
[INFO] ------------------------------------------------------------------------
[INFO] Total time: 32:18.277s
[INFO] Finished at: Thu Sep 25 22:28:43 CST 2014
[INFO] Final Memory: 73M/239M
```

1.2.3 验证编译是否成功

到 hadoop-dist/target/hadoop-2.2.0/lib/native 目录中查看 libhadoop.so.1.0.0 属性：

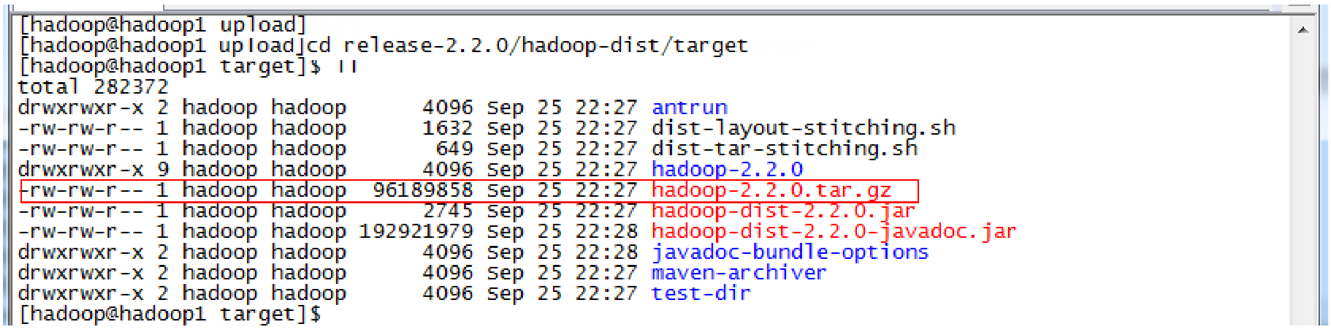
$file ./libhadoop.so.1.0.0

该文件为 ELF 64-bit LSB 则表示文件成功编译为 64 位



```
[hadoop@hadoop1 native]$
[hadoop@hadoop1 native]$ ls
libhadoop.a        libhadoop.so        libhadooputils.a  libhdfs.so
libhadooppipes.a   libhadoop.so.1.0.0  libhdfs.a         libhdfs.so.0.0.0
[hadoop@hadoop1 native]$
[hadoop@hadoop1 native]$ file *
libhadoop.a:        current ar archive
libhadooppipes.a:   current ar archive
libhadoop.so:       symbolic link to `libhadoop.so.1.0.0'
libhadoop.so.1.0.0: ELF 64-bit LSB shared object, x86-64, version 1 (SYSV), dynamically linked, not stripped
libhadooputils.a:   current ar archive
libhdfs.a:          current ar archive
libhdfs.so:         symbolic link to `libhdfs.so.0.0.0'
libhdfs.so.0.0.0:   ELF 64-bit LSB shared object, x86-64, version 1 (SYSV), dynamically linked, not stripped
[hadoop@hadoop1 native]$
```

在 hadoop-dist/target 目录中已经打包好了 hadoop-2.2.0.tar.gz ,该文件作为 Hadoop2.X 64 位安装包。



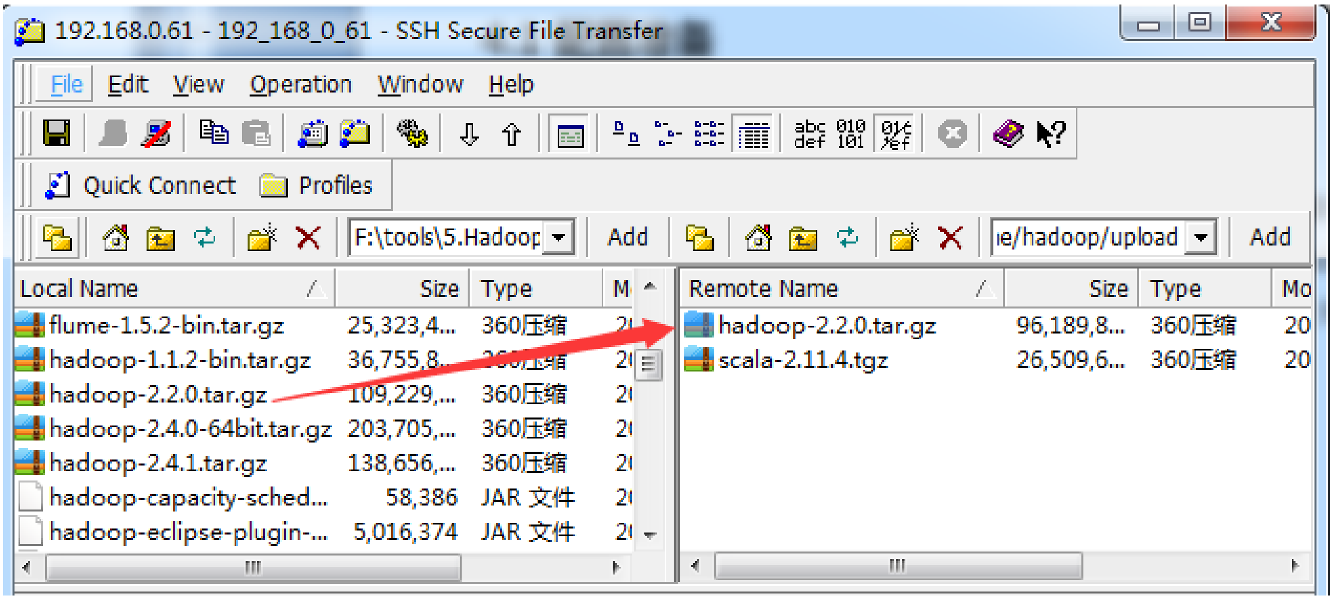
```
[hadoop@hadoop1 upload]
[hadoop@hadoop1 upload]cd release-2.2.0/hadoop-dist/target
[hadoop@hadoop1 target]$ ll
total 282372
drwxrwxr-x 2 hadoop hadoop      4096 Sep 25 22:27 antrun
-rw-rw-r-- 1 hadoop hadoop      1632 Sep 25 22:27 dist-layout-stitching.sh
-rw-rw-r-- 1 hadoop hadoop       649 Sep 25 22:27 dist-tar-stitching.sh
drwxrwxr-x 9 hadoop hadoop      4096 Sep 25 22:27 hadoop-2.2.0
-rw-rw-r-- 1 hadoop hadoop  96189858 Sep 25 22:27 hadoop-2.2.0.tar.gz
-rw-rw-r-- 1 hadoop hadoop      2745 Sep 25 22:27 hadoop-dist-2.2.0.jar
-rw-rw-r-- 1 hadoop hadoop 192921979 Sep 25 22:28 hadoop-dist-2.2.0-javadoc.jar
drwxrwxr-x 2 hadoop hadoop      4096 Sep 25 22:28 javadoc-bundle-options
drwxrwxr-x 2 hadoop hadoop      4096 Sep 25 22:27 maven-archiver
drwxrwxr-x 2 hadoop hadoop      4096 Sep 25 22:27 test-dir
[hadoop@hadoop1 target]$
```

2 安装 Hadoop

2.1 配置准备

2.1.1 上传并解压 Hadoop 安装包

1. 在 Apache 网站上提供 Hadoop2.X 安装包只支持 32 位操作系统安装，在 64 位服务器安装会出现 4.1 的错误异常。我们使用上一步骤编译好的 hadoop-2.2.0.tar.gz 文件作为安装包（也可以从网上下载 native 文件夹或者打包好的 64 位 hadoop 安装包），使用" Spark 编译与部署（上）"中 1. 3.1 介绍的工具上传到/home/hadoop/upload 目录下



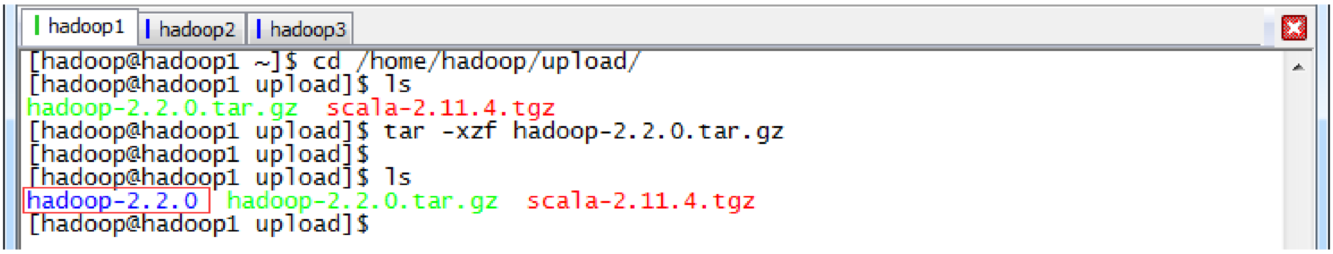
Local Name	Size	Type	M
flume-1.5.2-bin.tar.gz	25,323,4...	360压缩	2(
hadoop-1.1.2-bin.tar.gz	36,755,8...	360压缩	2(
hadoop-2.2.0.tar.gz	109,229,...	360压缩	2(
hadoop-2.4.0-64bit.tar.gz	203,705,...	360压缩	2(
hadoop-2.4.1.tar.gz	138,656,...	360压缩	2(
hadoop-capacity-sched...	58,386	JAR 文件	2(
hadoop-eclipse-plugin-...	5,016,374	JAR 文件	2(

Remote Name	Size	Type	Mo
hadoop-2.2.0.tar.gz	96,189,8...	360压缩	20
scala-2.11.4.tgz	26,509,6...	360压缩	20

2. 在主节点上解压缩

$cd /home/hadoop/upload/

$tar -xzf hadoop-2.2.0.tar.gz

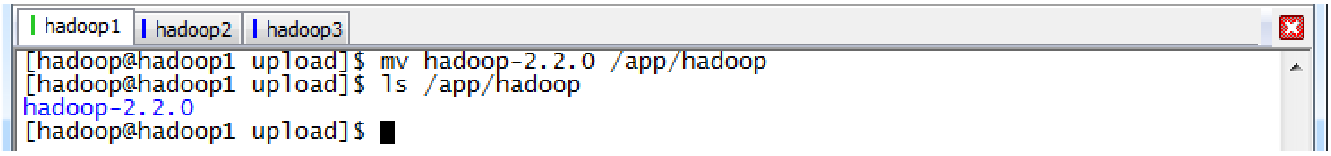


```
[hadoop@hadoop1 ~]$ cd /home/hadoop/upload/
[hadoop@hadoop1 upload]$ ls
hadoop-2.2.0.tar.gz  scala-2.11.4.tgz
[hadoop@hadoop1 upload]$ tar -xzf hadoop-2.2.0.tar.gz
[hadoop@hadoop1 upload]$
[hadoop@hadoop1 upload]$ ls
hadoop-2.2.0  hadoop-2.2.0.tar.gz  scala-2.11.4.tgz
[hadoop@hadoop1 upload]$
```

3. 把 hadoop-2.2.0 目录移到/app/hadoop 目录下

$ mv hadoop-2.2.0 /app/hadoop

$ls /app/hadoop



```
[hadoop@hadoop1 upload]$ mv hadoop-2.2.0 /app/hadoop
[hadoop@hadoop1 upload]$ ls /app/hadoop
hadoop-2.2.0
[hadoop@hadoop1 upload]$ 
```

2.1.2 在 Hadoop 目录下创建子目录

hadoop 用户在/app/hadoop/hadoop-2.2.0 目录下创建 tmp、name 和 data 目录

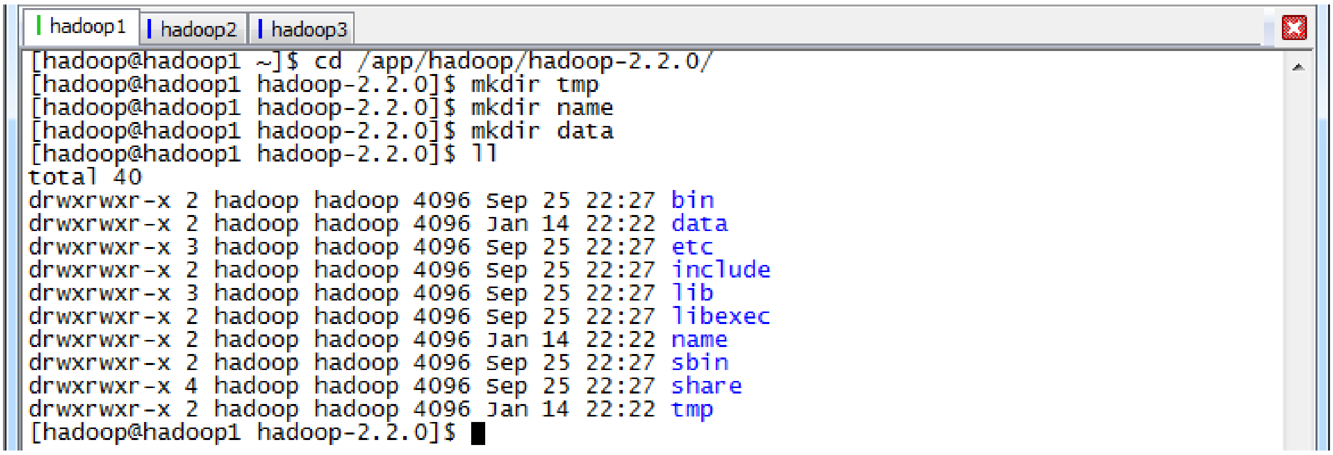
$cd /app/hadoop/hadoop-2.2.0/

$mkdir tmp

$mkdir name

$mkdir data

$ll



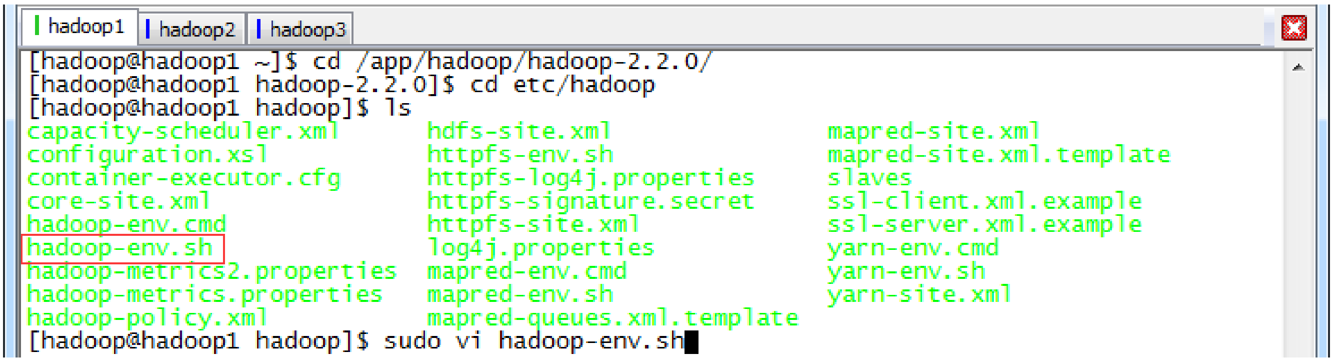
```
[hadoop@hadoop1 ~]$ cd /app/hadoop/hadoop-2.2.0/
[hadoop@hadoop1 hadoop-2.2.0]$ mkdir tmp
[hadoop@hadoop1 hadoop-2.2.0]$ mkdir name
[hadoop@hadoop1 hadoop-2.2.0]$ mkdir data
[hadoop@hadoop1 hadoop-2.2.0]$ ll
total 40
drwxrwxr-x 2 hadoop hadoop 4096 Sep 25 22:27 bin
drwxrwxr-x 2 hadoop hadoop 4096 Jan 14 22:22 data
drwxrwxr-x 3 hadoop hadoop 4096 Sep 25 22:27 etc
drwxrwxr-x 2 hadoop hadoop 4096 Sep 25 22:27 include
drwxrwxr-x 3 hadoop hadoop 4096 Sep 25 22:27 lib
drwxrwxr-x 2 hadoop hadoop 4096 Sep 25 22:27 libexec
drwxrwxr-x 2 hadoop hadoop 4096 Jan 14 22:22 name
drwxrwxr-x 2 hadoop hadoop 4096 Sep 25 22:27 sbin
drwxrwxr-x 4 hadoop hadoop 4096 Sep 25 22:27 share
drwxrwxr-x 2 hadoop hadoop 4096 Jan 14 22:22 tmp
[hadoop@hadoop1 hadoop-2.2.0]$ 
```

2.1.3 配置 hadoop-env.sh

1. 打开配置文件 hadoop-env.sh

$cd /app/hadoop/hadoop-2.2.0/etc/hadoop

$sudo vi hadoop-env.sh

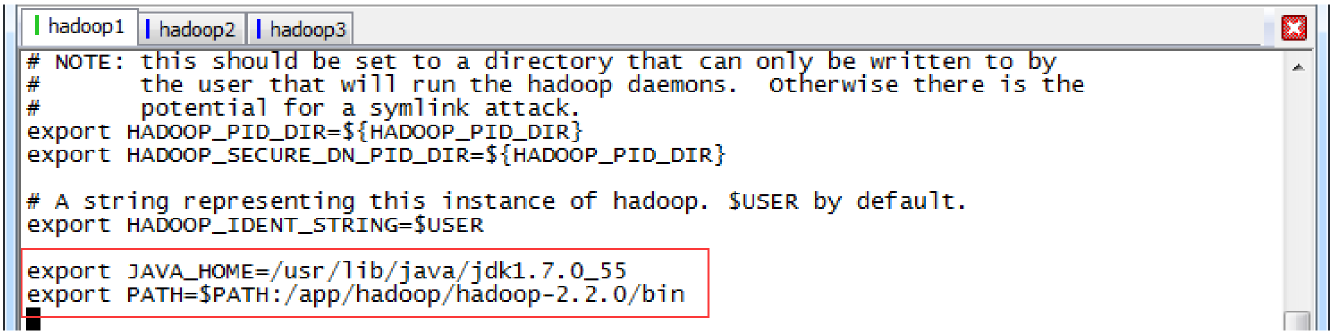


```
[hadoop@hadoop1 ~]$ cd /app/hadoop/hadoop-2.2.0/
[hadoop@hadoop1 hadoop-2.2.0]$ cd etc/hadoop
[hadoop@hadoop1 hadoop]$ ls
capacity-scheduler.xml       hdfs-site.xml              mapred-site.xml
configuration.xsl            httpfs-env.sh              mapred-site.xml.template
container-executor.cfg       httpfs-log4j.properties    slaves
core-site.xml                httpfs-signature.secret    ssl-client.xml.example
hadoop-env.cmd               httpfs-site.xml            ssl-server.xml.example
hadoop-env.sh                log4j.properties           yarn-env.cmd
hadoop-metrics2.properties   mapred-env.cmd             yarn-env.sh
hadoop-metrics.properties    mapred-env.sh              yarn-site.xml
hadoop-policy.xml            mapred-queues.xml.template
[hadoop@hadoop1 hadoop]$ sudo vi hadoop-env.sh
```

2. 加入配置内容，设置 JAVA_HOME 和 PATH 路径

export JAVA_HOME=/usr/lib/java/jdk1.7.0_55

export PATH=$PATH:/app/hadoop/hadoop-2.2.0/bin

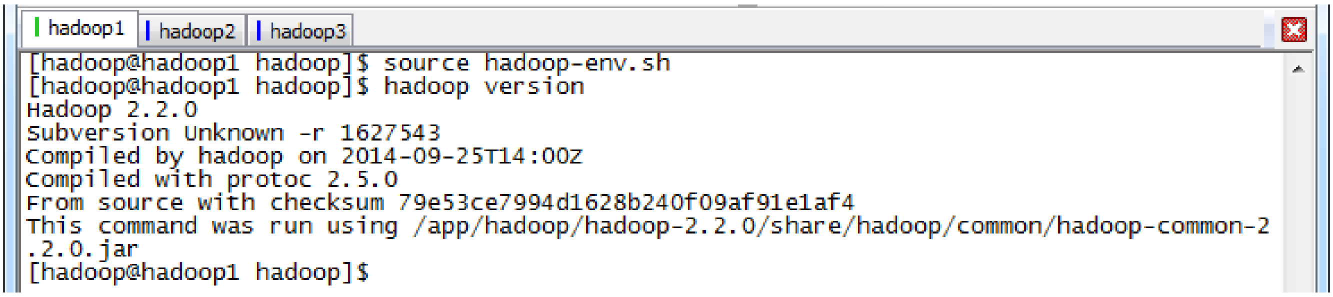


```
# NOTE: this should be set to a directory that can only be written to by
#       the user that will run the hadoop daemons.  Otherwise there is the
#       potential for a symlink attack.
export HADOOP_PID_DIR=${HADOOP_PID_DIR}
export HADOOP_SECURE_DN_PID_DIR=${HADOOP_PID_DIR}

# A string representing this instance of hadoop. $USER by default.
export HADOOP_IDENT_STRING=$USER

export JAVA_HOME=/usr/lib/java/jdk1.7.0_55
export PATH=$PATH:/app/hadoop/hadoop-2.2.0/bin
```

3. 编译配置文件 hadoop-env.sh，并确认生效

$source hadoop-env.sh

$hadoop version



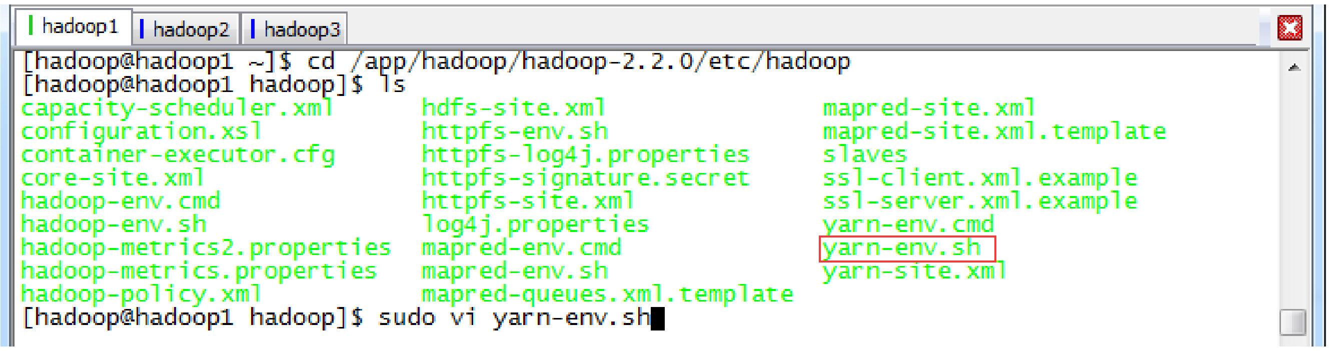
```
[hadoop@hadoop1 hadoop]$ source hadoop-env.sh
[hadoop@hadoop1 hadoop]$ hadoop version
Hadoop 2.2.0
Subversion Unknown -r 1627543
Compiled by hadoop on 2014-09-25T14:00Z
Compiled with protoc 2.5.0
From source with checksum 79e53ce7994d1628b240f09af91e1af4
This command was run using /app/hadoop/hadoop-2.2.0/share/hadoop/common/hadoop-common-2
.2.0.jar
[hadoop@hadoop1 hadoop]$
```

2.1.4 配置 yarn-env.sh

1. 在/app/hadoop/hadoop-2.2.0/etc/hadoop 打开配置文件 yarn-env.sh

$cd /app/hadoop/hadoop-2.2.0/etc/hadoop

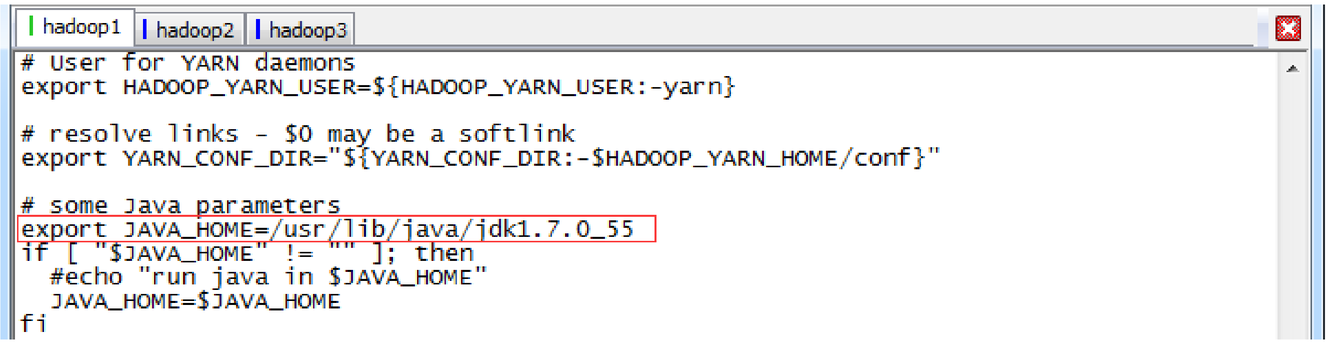
$sudo vi yarn-env.sh



```
[hadoop@hadoop1 ~]$ cd /app/hadoop/hadoop-2.2.0/etc/hadoop
[hadoop@hadoop1 hadoop]$ ls
capacity-scheduler.xml        hdfs-site.xml                  mapred-site.xml
configuration.xsl             httpfs-env.sh                  mapred-site.xml.template
container-executor.cfg        httpfs-log4j.properties        slaves
core-site.xml                 httpfs-signature.secret        ssl-client.xml.example
hadoop-env.cmd                httpfs-site.xml                ssl-server.xml.example
hadoop-env.sh                 log4j.properties               yarn-env.cmd
hadoop-metrics2.properties    mapred-env.cmd                 yarn-env.sh
hadoop-metrics.properties     mapred-env.sh                  yarn-site.xml
hadoop-policy.xml             mapred-queues.xml.template
[hadoop@hadoop1 hadoop]$ sudo vi yarn-env.sh
```

2. 加入配置内容，设置 JAVA_HOME 路径

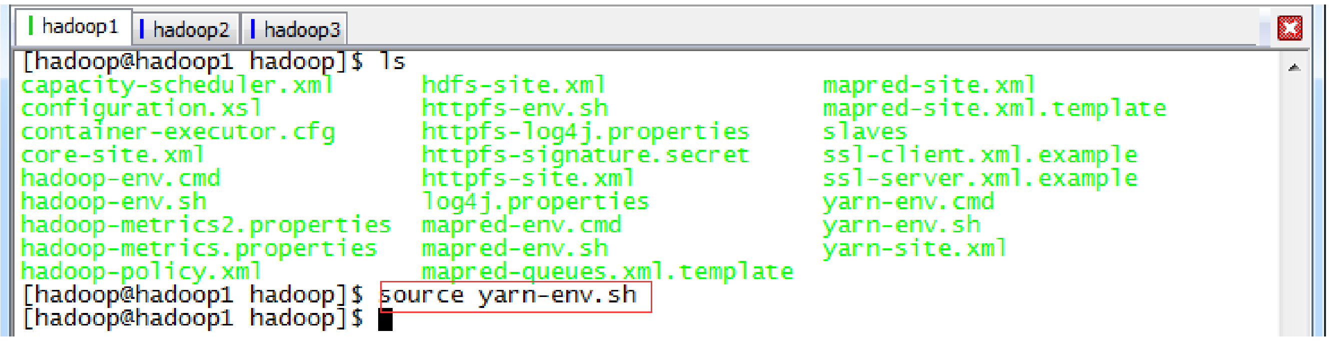
export JAVA_HOME=/usr/lib/java/jdk1.7.0_55



```
# User for YARN daemons
export HADOOP_YARN_USER=${HADOOP_YARN_USER:-yarn}

# resolve links - $0 may be a softlink
export YARN_CONF_DIR="${YARN_CONF_DIR:-$HADOOP_YARN_HOME/conf}"

# some Java parameters
export JAVA_HOME=/usr/lib/java/jdk1.7.0_55
if [ "$JAVA_HOME" != "" ]; then
  #echo "run java in $JAVA_HOME"
  JAVA_HOME=$JAVA_HOME
fi
```

3. 编译配置文件 yarn-env.sh，并确认生效

$source yarn-env.sh

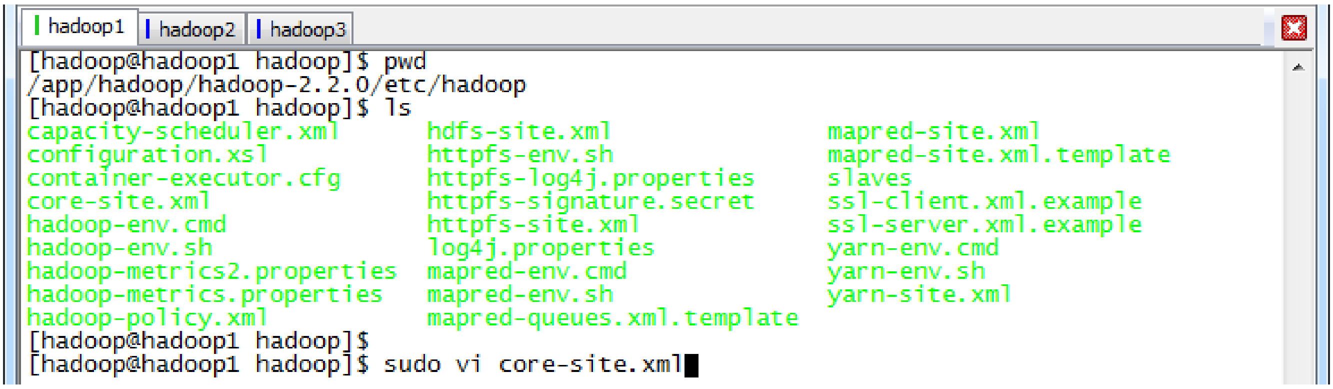


```
[hadoop@hadoop1 hadoop]$ ls
capacity-scheduler.xml        hdfs-site.xml                  mapred-site.xml
configuration.xsl             httpfs-env.sh                  mapred-site.xml.template
container-executor.cfg        httpfs-log4j.properties        slaves
core-site.xml                 httpfs-signature.secret        ssl-client.xml.example
hadoop-env.cmd                httpfs-site.xml                ssl-server.xml.example
hadoop-env.sh                 log4j.properties               yarn-env.cmd
hadoop-metrics2.properties    mapred-env.cmd                 yarn-env.sh
hadoop-metrics.properties     mapred-env.sh                  yarn-site.xml
hadoop-policy.xml             mapred-queues.xml.template
[hadoop@hadoop1 hadoop]$ source yarn-env.sh
[hadoop@hadoop1 hadoop]$ 
```

2.1.5 配置 core-site.xml

1. 使用如下命令打开 core-site.xml 配置文件

$sudo vi core-site.xml

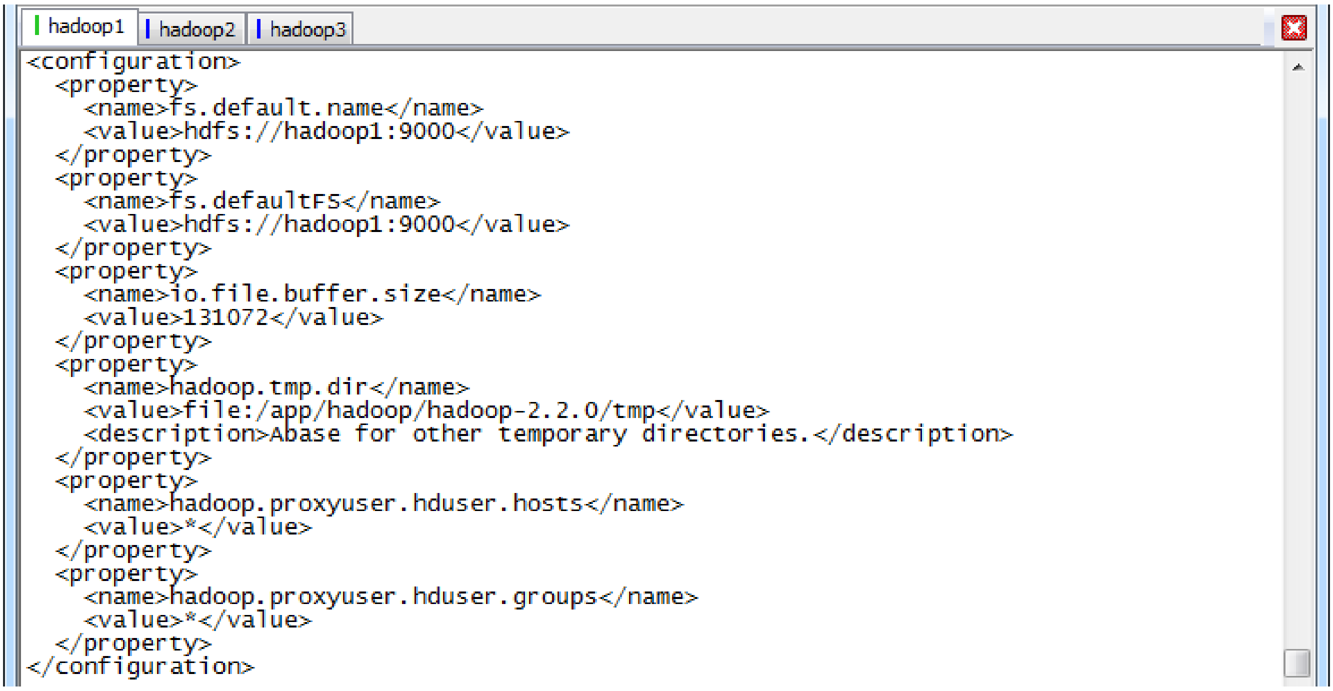


```
[hadoop@hadoop1 hadoop]$ pwd
/app/hadoop/hadoop-2.2.0/etc/hadoop
[hadoop@hadoop1 hadoop]$ ls
capacity-scheduler.xml      hdfs-site.xml               mapred-site.xml
configuration.xsl           httpfs-env.sh               mapred-site.xml.template
container-executor.cfg      httpfs-log4j.properties     slaves
core-site.xml               httpfs-signature.secret     ssl-client.xml.example
hadoop-env.cmd              httpfs-site.xml             ssl-server.xml.example
hadoop-env.sh               log4j.properties            yarn-env.cmd
hadoop-metrics2.properties  mapred-env.cmd              yarn-env.sh
hadoop-metrics.properties   mapred-env.sh               yarn-site.xml
hadoop-policy.xml           mapred-queues.xml.template
[hadoop@hadoop1 hadoop]$
[hadoop@hadoop1 hadoop]$ sudo vi core-site.xml
```

2. 在配置文件中，按照如下内容进行配置

```
<configuration>
  <property>
    <name>fs.default.name</name>
    <value>hdfs://hadoop1:9000</value>
  </property>
  <property>
    <name>fs.defaultFS</name>
    <value>hdfs://hadoop1:9000</value>
  </property>
  <property>
    <name>io.file.buffer.size</name>
    <value>131072</value>
  </property>
  <property>
    <name>hadoop.tmp.dir</name>
    <value>file:/app/hadoop/hadoop-2.2.0/tmp</value>
    <description>Abase for other temporary directories.</description>
  </property>
  <property>
    <name>hadoop.proxyuser.hduser.hosts</name>
    <value>*</value>
  </property>
  <property>
    <name>hadoop.proxyuser.hduser.groups</name>
    <value>*</value>
  </property>
```

</configuration>

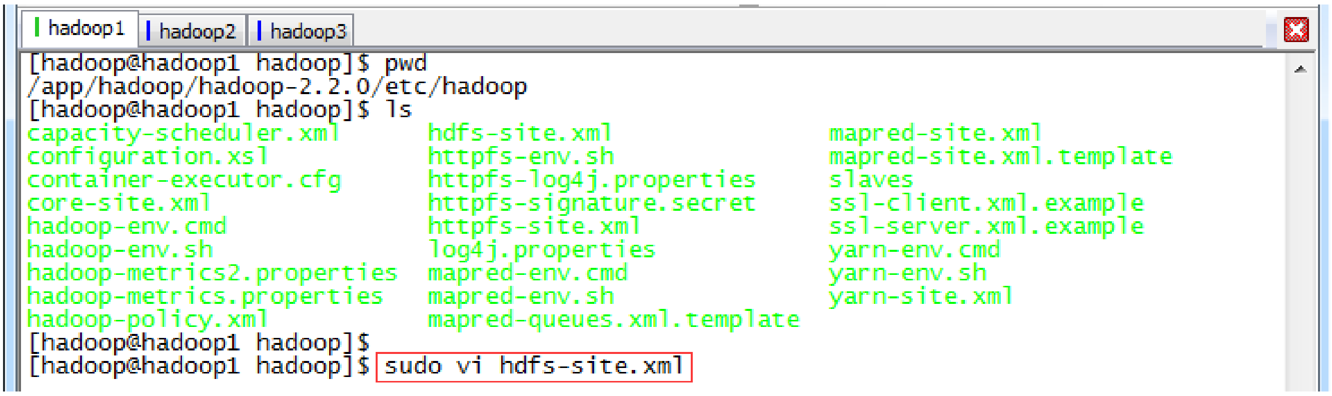


```
<configuration>
  <property>
    <name>fs.default.name</name>
    <value>hdfs://hadoop1:9000</value>
  </property>
  <property>
    <name>fs.defaultFS</name>
    <value>hdfs://hadoop1:9000</value>
  </property>
  <property>
    <name>io.file.buffer.size</name>
    <value>131072</value>
  </property>
  <property>
    <name>hadoop.tmp.dir</name>
    <value>file:/app/hadoop/hadoop-2.2.0/tmp</value>
    <description>Abase for other temporary directories.</description>
  </property>
  <property>
    <name>hadoop.proxyuser.hduser.hosts</name>
    <value>*</value>
  </property>
  <property>
    <name>hadoop.proxyuser.hduser.groups</name>
    <value>*</value>
  </property>
</configuration>
```

2.1.6 配置 hdfs-site.xml

1. 使用如下命令打开 hdfs-site.xml 配置文件

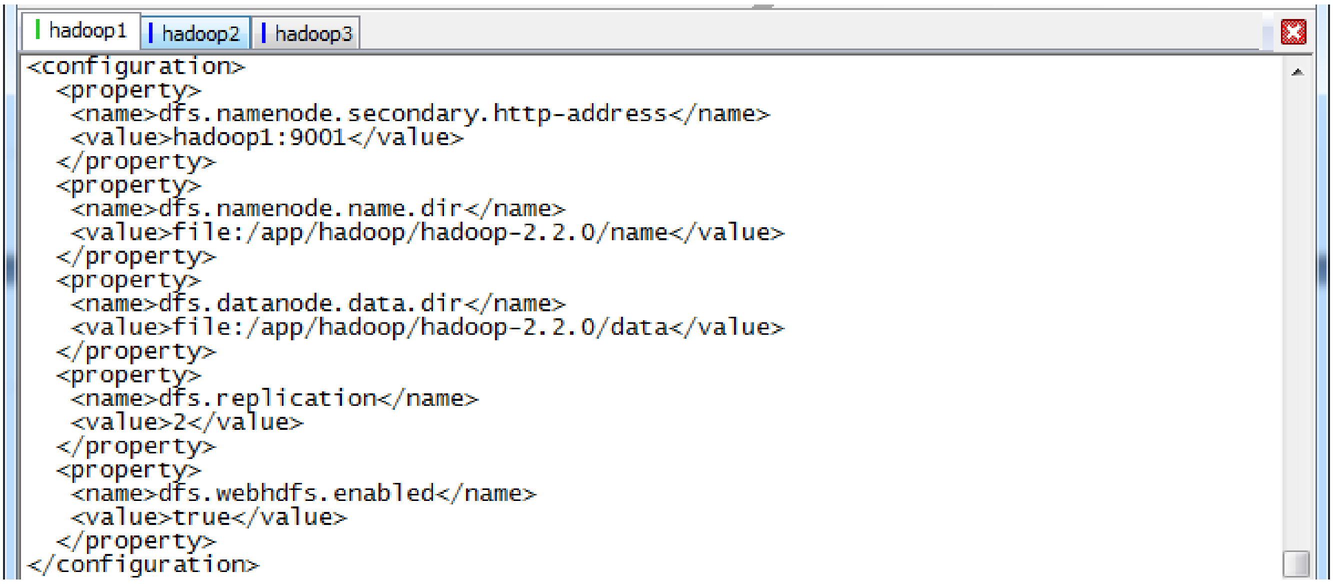
$sudo vi hdfs-site.xml



```
[hadoop@hadoop1 hadoop]$ pwd
/app/hadoop/hadoop-2.2.0/etc/hadoop
[hadoop@hadoop1 hadoop]$ ls
capacity-scheduler.xml      hdfs-site.xml               mapred-site.xml
configuration.xsl           httpfs-env.sh               mapred-site.xml.template
container-executor.cfg      httpfs-log4j.properties     slaves
core-site.xml               httpfs-signature.secret     ssl-client.xml.example
hadoop-env.cmd              httpfs-site.xml             ssl-server.xml.example
hadoop-env.sh               log4j.properties            yarn-env.cmd
hadoop-metrics2.properties  mapred-env.cmd              yarn-env.sh
hadoop-metrics.properties   mapred-env.sh               yarn-site.xml
hadoop-policy.xml           mapred-queues.xml.template
[hadoop@hadoop1 hadoop]$
[hadoop@hadoop1 hadoop]$ sudo vi hdfs-site.xml
```

2. 在配置文件中，按照如下内容进行配置

```
<configuration>
  <property>
   <name>dfs.namenode.secondary.http-address</name>
   <value>hadoop1:9001</value>
  </property>
  <property>
   <name>dfs.namenode.name.dir</name>
   <value>file:/app/hadoop/hadoop-2.2.0/name</value>
  </property>
  <property>
   <name>dfs.datanode.data.dir</name>
   <value>file:/app/hadoop/hadoop-2.2.0/data</value>
  </property>
  <property>
   <name>dfs.replication</name>
   <value>2</value>
  </property>
  <property>
   <name>dfs.webhdfs.enabled</name>
   <value>true</value>
  </property>
</configuration>
```

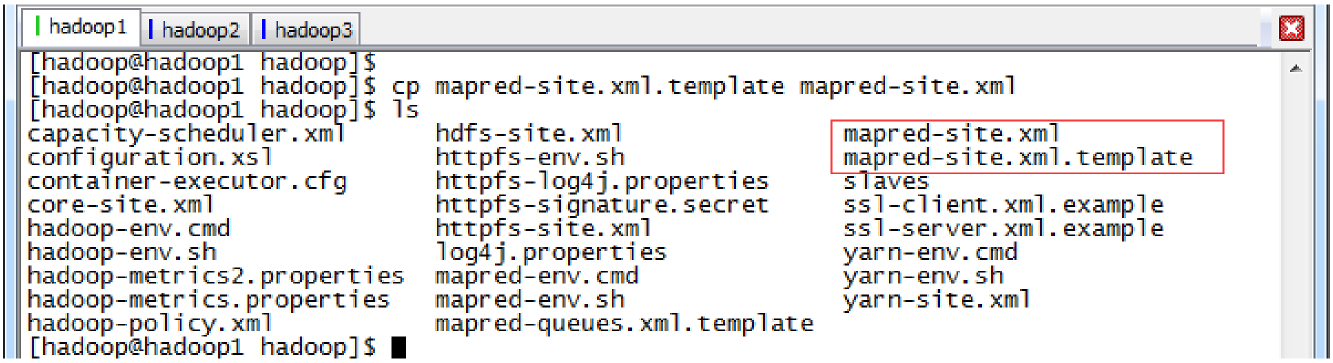


```
<configuration>
  <property>
   <name>dfs.namenode.secondary.http-address</name>
   <value>hadoop1:9001</value>
  </property>
  <property>
   <name>dfs.namenode.name.dir</name>
   <value>file:/app/hadoop/hadoop-2.2.0/name</value>
  </property>
  <property>
   <name>dfs.datanode.data.dir</name>
   <value>file:/app/hadoop/hadoop-2.2.0/data</value>
  </property>
  <property>
   <name>dfs.replication</name>
   <value>2</value>
  </property>
  <property>
   <name>dfs.webhdfs.enabled</name>
   <value>true</value>
  </property>
</configuration>
```

2.1.7 配置 mapred-site.xml

1. 默认情况下不存在 mapred-site.xml 文件，可以从模板拷贝一份

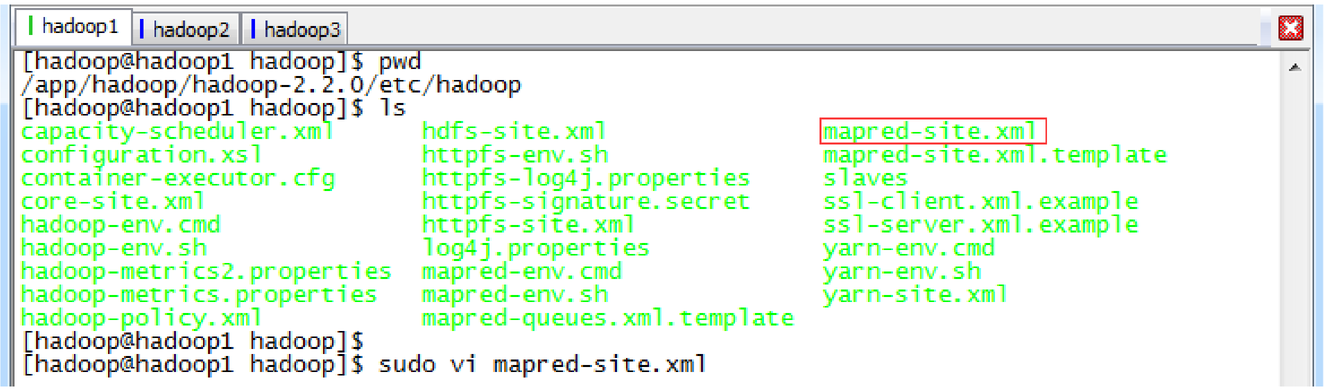
```
$cp mapred-site.xml.template mapred-site.xml
```



```
[hadoop@hadoop1 hadoop]$
[hadoop@hadoop1 hadoop]$ cp mapred-site.xml.template mapred-site.xml
[hadoop@hadoop1 hadoop]$ ls
capacity-scheduler.xml      hdfs-site.xml               mapred-site.xml
configuration.xsl           httpfs-env.sh               mapred-site.xml.template
container-executor.cfg      httpfs-log4j.properties     slaves
core-site.xml               httpfs-signature.secret     ssl-client.xml.example
hadoop-env.cmd              httpfs-site.xml             ssl-server.xml.example
hadoop-env.sh               log4j.properties            yarn-env.cmd
hadoop-metrics2.properties  mapred-env.cmd              yarn-env.sh
hadoop-metrics.properties   mapred-env.sh               yarn-site.xml
hadoop-policy.xml           mapred-queues.xml.template
[hadoop@hadoop1 hadoop]$ █
```

2. 使用如下命令打开 mapred-site.xml 配置文件

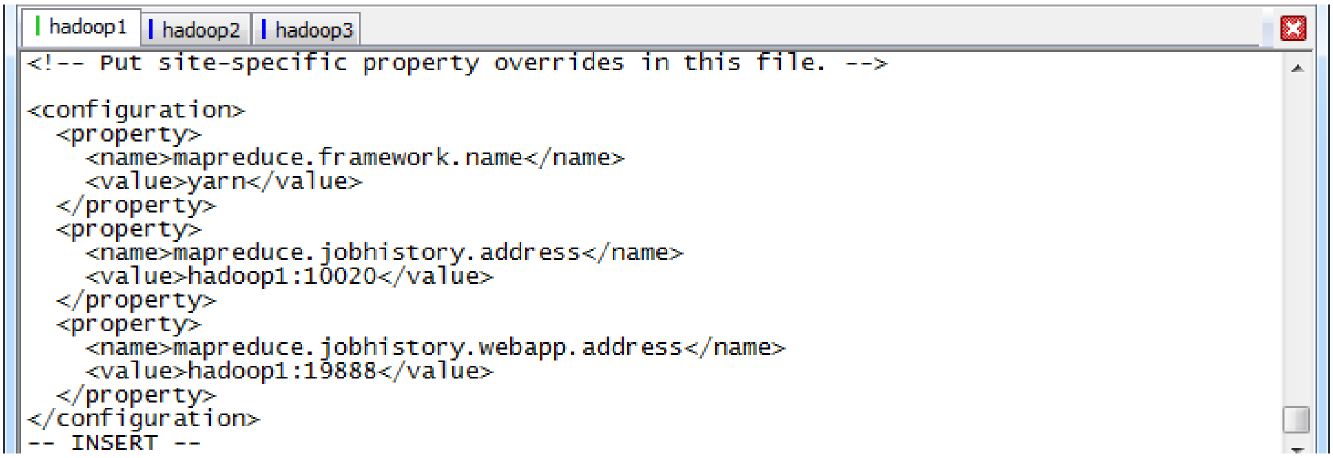
$sudo vi mapred-site.xml



```
[hadoop@hadoop1 hadoop]$ pwd
/app/hadoop/hadoop-2.2.0/etc/hadoop
[hadoop@hadoop1 hadoop]$ ls
capacity-scheduler.xml        hdfs-site.xml                mapred-site.xml
configuration.xsl             httpfs-env.sh                mapred-site.xml.template
container-executor.cfg        httpfs-log4j.properties      slaves
core-site.xml                 httpfs-signature.secret      ssl-client.xml.example
hadoop-env.cmd                httpfs-site.xml              ssl-server.xml.example
hadoop-env.sh                 log4j.properties             yarn-env.cmd
hadoop-metrics2.properties    mapred-env.cmd               yarn-env.sh
hadoop-metrics.properties     mapred-env.sh                yarn-site.xml
hadoop-policy.xml             mapred-queues.xml.template
[hadoop@hadoop1 hadoop]$
[hadoop@hadoop1 hadoop]$ sudo vi mapred-site.xml
```

3. 在配置文件中，按照如下内容进行配置

```
<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
  <property>
    <name>mapreduce.jobhistory.address</name>
    <value>hadoop1:10020</value>
  </property>
  <property>
    <name>mapreduce.jobhistory.webapp.address</name>
    <value>hadoop1:19888</value>
  </property>
</configuration>
```

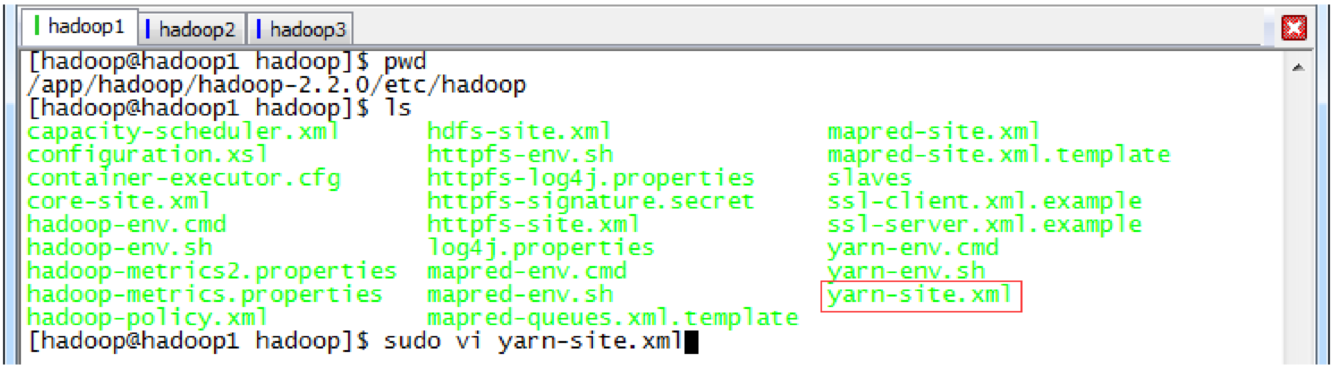


```
<!-- Put site-specific property overrides in this file. -->

<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
  <property>
    <name>mapreduce.jobhistory.address</name>
    <value>hadoop1:10020</value>
  </property>
  <property>
    <name>mapreduce.jobhistory.webapp.address</name>
    <value>hadoop1:19888</value>
  </property>
</configuration>
-- INSERT --
```

2.1.8 配置 yarn-site.xml

1. 使用如下命令打开 yarn-site.xml 配置文件

$sudo vi yarn-site.xml

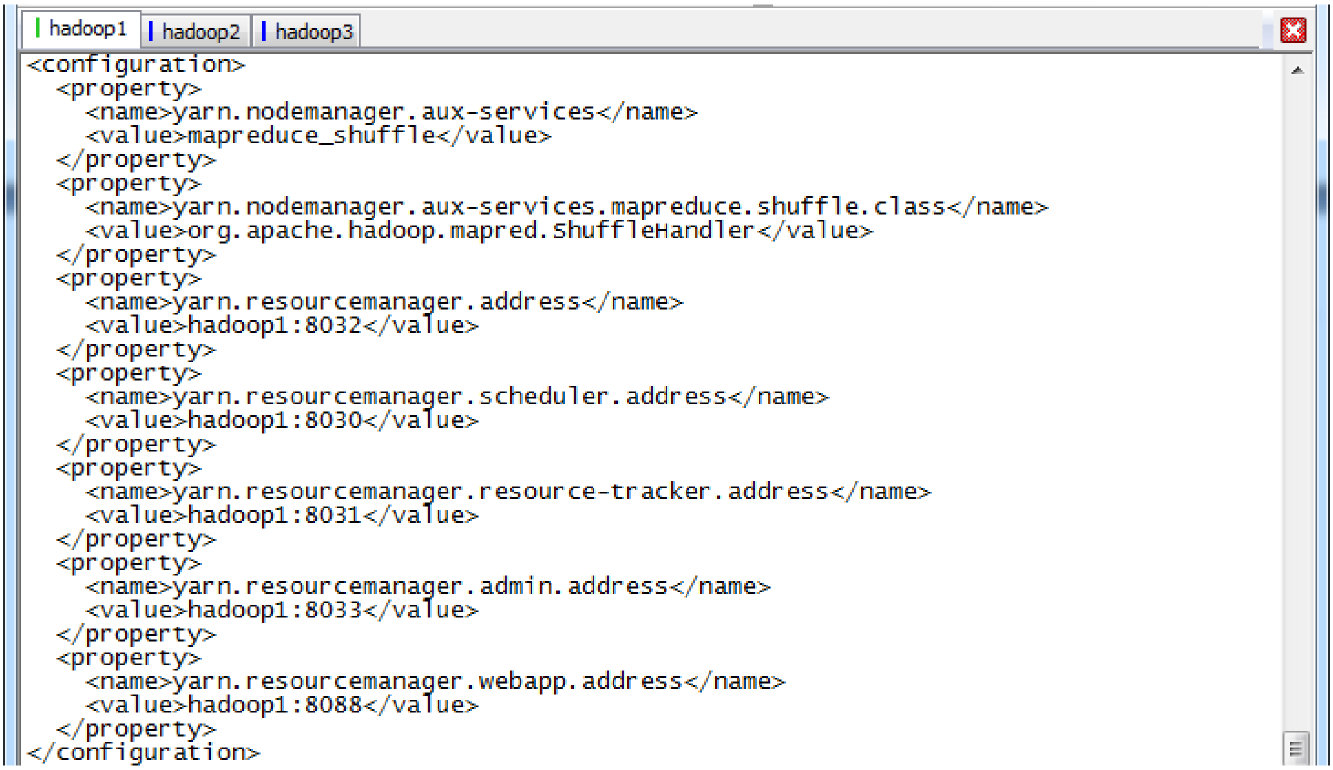


```
[hadoop@hadoop1 hadoop]$ pwd
/app/hadoop/hadoop-2.2.0/etc/hadoop
[hadoop@hadoop1 hadoop]$ ls
capacity-scheduler.xml      hdfs-site.xml              mapred-site.xml
configuration.xsl           httpfs-env.sh              mapred-site.xml.template
container-executor.cfg      httpfs-log4j.properties    slaves
core-site.xml               httpfs-signature.secret    ssl-client.xml.example
hadoop-env.cmd              httpfs-site.xml            ssl-server.xml.example
hadoop-env.sh               log4j.properties           yarn-env.cmd
hadoop-metrics2.properties  mapred-env.cmd             yarn-env.sh
hadoop-metrics.properties   mapred-env.sh              yarn-site.xml
hadoop-policy.xml           mapred-queues.xml.template
[hadoop@hadoop1 hadoop]$ sudo vi yarn-site.xml
```

2. 在配置文件中，按照如下内容进行配置

```
<configuration>
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
  <property>
    <name>yarn.nodemanager.aux-services.mapreduce.shuffle.class</name>
    <value>org.apache.hadoop.mapred.ShuffleHandler</value>
  </property>
  <property>
    <name>yarn.resourcemanager.address</name>
    <value>hadoop1:8032</value>
  </property>
  <property>
    <name>yarn.resourcemanager.scheduler.address</name>
    <value>hadoop1:8030</value>
  </property>
  <property>
    <name>yarn.resourcemanager.resource-tracker.address</name>
    <value>hadoop1:8031</value>
  </property>
  <property>
    <name>yarn.resourcemanager.admin.address</name>
```

```
      <value>hadoop1:8033</value>
   </property>
   <property>
      <name>yarn.resourcemanager.webapp.address</name>
      <value>hadoop1:8088</value>
   </property>
</configuration>
```

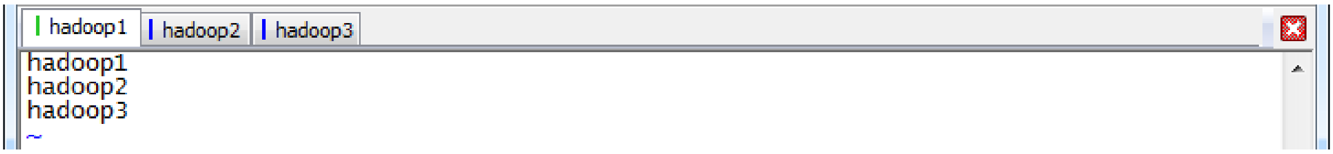


```
<configuration>
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
  <property>
    <name>yarn.nodemanager.aux-services.mapreduce.shuffle.class</name>
    <value>org.apache.hadoop.mapred.ShuffleHandler</value>
  </property>
  <property>
    <name>yarn.resourcemanager.address</name>
    <value>hadoop1:8032</value>
  </property>
  <property>
    <name>yarn.resourcemanager.scheduler.address</name>
    <value>hadoop1:8030</value>
  </property>
  <property>
    <name>yarn.resourcemanager.resource-tracker.address</name>
    <value>hadoop1:8031</value>
  </property>
  <property>
    <name>yarn.resourcemanager.admin.address</name>
    <value>hadoop1:8033</value>
  </property>
  <property>
    <name>yarn.resourcemanager.webapp.address</name>
    <value>hadoop1:8088</value>
  </property>
</configuration>
```

2.1.9 配置 Slaves 文件

使用*$sudo vi slaves*打开从节点配置文件，在文件中加入

```
hadoop1
hadoop2
hadoop3
```



```
hadoop1
hadoop2
hadoop3
~
```

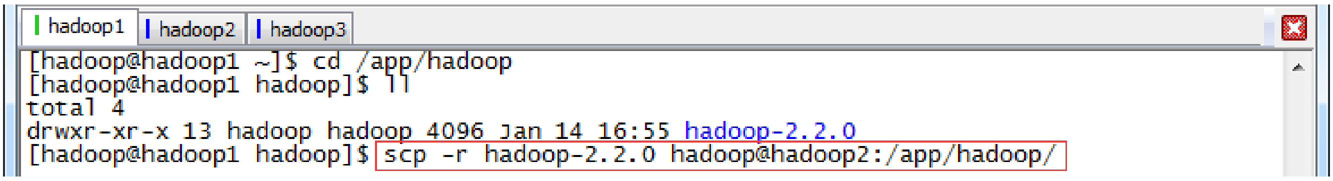
2.1.10 向各节点分发 Hadoop 程序

1. 确认 hadoop2 和 hadoop3 节点/app/hadoop 所属组和用户均为 hadoop，然后进入 hadoop1 机器/app/hadoop 目录，使用如下命令把 hadoop 文件夹复制到 hadoop2 和 hadoop3 机器

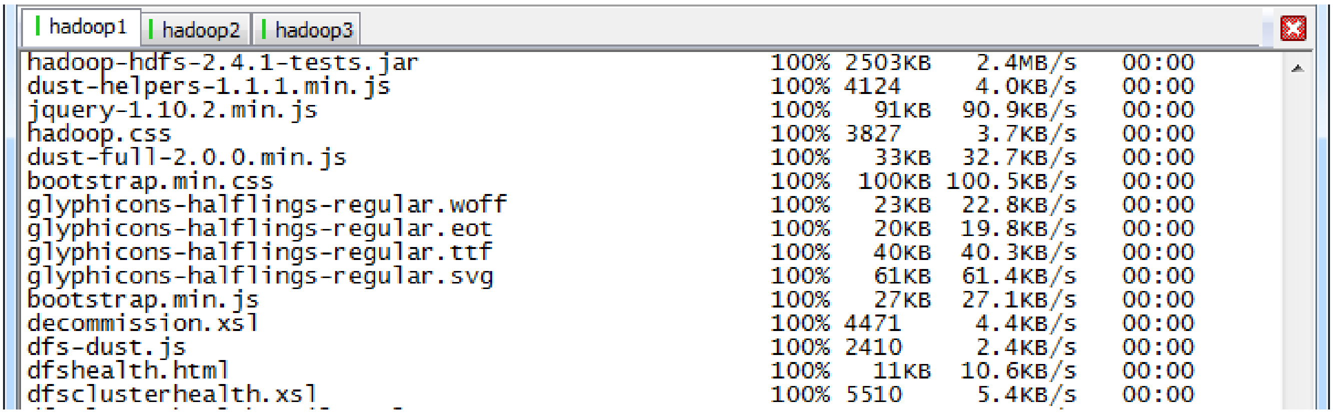
$cd /app/hadoop

$scp -r hadoop-2.2.0 hadoop@hadoop2:/app/hadoop/

$scp -r hadoop-2.2.0 hadoop@hadoop3:/app/hadoop/

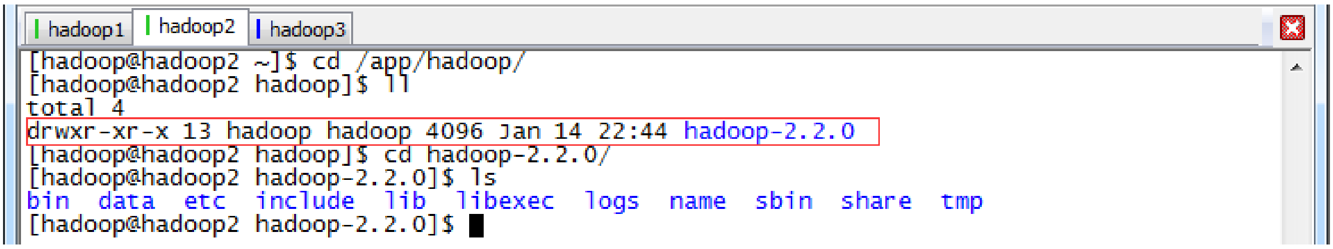


```
[hadoop@hadoop1 ~]$ cd /app/hadoop
[hadoop@hadoop1 hadoop]$ ll
total 4
drwxr-xr-x 13 hadoop hadoop 4096 Jan 14 16:55 hadoop-2.2.0
[hadoop@hadoop1 hadoop]$ scp -r hadoop-2.2.0 hadoop@hadoop2:/app/hadoop/
```



```
hadoop-hdfs-2.4.1-tests.jar                          100% 2503KB   2.4MB/s   00:00
dust-helpers-1.1.1.min.js                            100% 4124     4.0KB/s   00:00
jquery-1.10.2.min.js                                 100%   91KB  90.9KB/s   00:00
hadoop.css                                           100% 3827     3.7KB/s   00:00
dust-full-2.0.0.min.js                               100%   33KB  32.7KB/s   00:00
bootstrap.min.css                                    100%  100KB 100.5KB/s   00:00
glyphicons-halflings-regular.woff                    100%   23KB  22.8KB/s   00:00
glyphicons-halflings-regular.eot                     100%   20KB  19.8KB/s   00:00
glyphicons-halflings-regular.ttf                     100%   40KB  40.3KB/s   00:00
glyphicons-halflings-regular.svg                     100%   61KB  61.4KB/s   00:00
bootstrap.min.js                                     100%   27KB  27.1KB/s   00:00
decommission.xsl                                     100% 4471     4.4KB/s   00:00
dfs-dust.js                                          100% 2410     2.4KB/s   00:00
dfshealth.html                                       100%   11KB  10.6KB/s   00:00
dfsclusterhealth.xsl                                 100% 5510     5.4KB/s   00:00
```

2. 在从节点查看是否复制成功



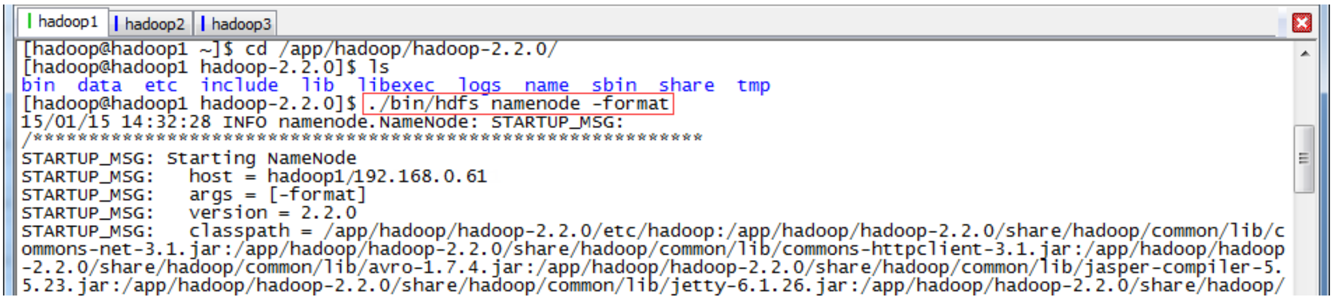
```
[hadoop@hadoop2 ~]$ cd /app/hadoop/
[hadoop@hadoop2 hadoop]$ ll
total 4
drwxr-xr-x 13 hadoop hadoop 4096 Jan 14 22:44 hadoop-2.2.0
[hadoop@hadoop2 hadoop]$ cd hadoop-2.2.0/
[hadoop@hadoop2 hadoop-2.2.0]$ ls
bin  data  etc  include  lib  libexec  logs  name  sbin  share  tmp
[hadoop@hadoop2 hadoop-2.2.0]$
```

2.2 启动部署

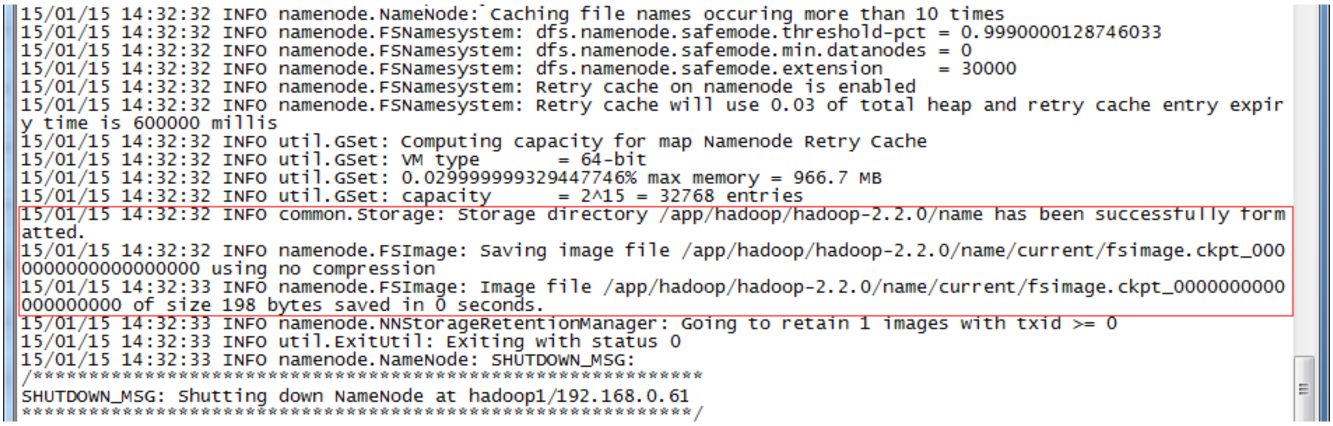
2.2.1 格式化 NameNode

$cd /app/hadoop/hadoop-2.2.0/

$./bin/hdfs namenode -format



```
[hadoop@hadoop1 ~]$ cd /app/hadoop/hadoop-2.2.0/
[hadoop@hadoop1 hadoop-2.2.0]$ ls
bin  data  etc  include  lib  libexec  logs  name  sbin  share  tmp
[hadoop@hadoop1 hadoop-2.2.0]$ ./bin/hdfs namenode -format
15/01/15 14:32:28 INFO namenode.NameNode: STARTUP_MSG:
/************************************************************
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = hadoop1/192.168.0.61
STARTUP_MSG:   args = [-format]
STARTUP_MSG:   version = 2.2.0
STARTUP_MSG:   classpath = /app/hadoop/hadoop-2.2.0/etc/hadoop:/app/hadoop/hadoop-2.2.0/share/hadoop/common/lib/c
ommons-net-3.1.jar:/app/hadoop/hadoop-2.2.0/share/hadoop/common/lib/commons-httpclient-3.1.jar:/app/hadoop/hadoop
-2.2.0/share/hadoop/common/lib/avro-1.7.4.jar:/app/hadoop/hadoop-2.2.0/share/hadoop/common/lib/jasper-compiler-5.
5.23.jar:/app/hadoop/hadoop-2.2.0/share/hadoop/common/lib/jetty-6.1.26.jar:/app/hadoop/hadoop-2.2.0/share/hadoop/
```

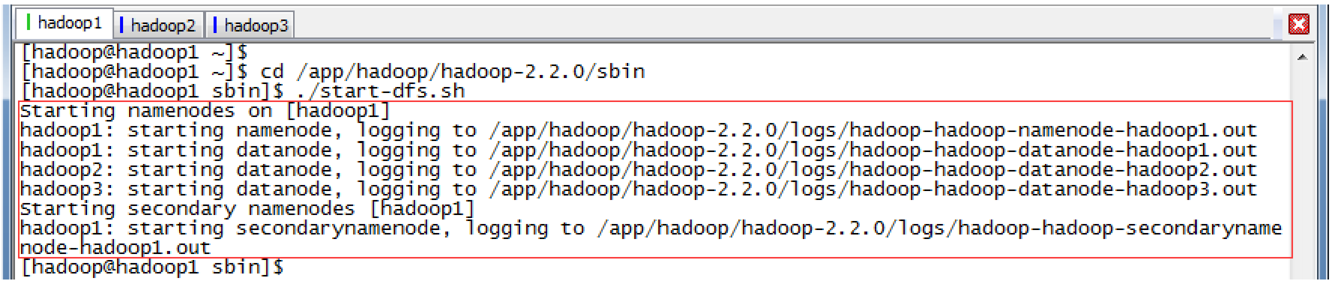


```
15/01/15 14:32:32 INFO namenode.NameNode: Caching file names occuring more than 10 times
15/01/15 14:32:32 INFO namenode.FSNamesystem: dfs.namenode.safemode.threshold-pct = 0.9990000128746033
15/01/15 14:32:32 INFO namenode.FSNamesystem: dfs.namenode.safemode.min.datanodes = 0
15/01/15 14:32:32 INFO namenode.FSNamesystem: dfs.namenode.safemode.extension     = 30000
15/01/15 14:32:32 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
15/01/15 14:32:32 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total heap and retry cache entry expir
y time is 600000 millis
15/01/15 14:32:32 INFO util.GSet: Computing capacity for map Namenode Retry Cache
15/01/15 14:32:32 INFO util.GSet: VM type       = 64-bit
15/01/15 14:32:32 INFO util.GSet: 0.029999999329447746% max memory = 966.7 MB
15/01/15 14:32:32 INFO util.GSet: capacity      = 2^15 = 32768 entries
15/01/15 14:32:32 INFO common.Storage: Storage directory /app/hadoop/hadoop-2.2.0/name has been successfully form
atted.
15/01/15 14:32:32 INFO namenode.FSImage: Saving image file /app/hadoop/hadoop-2.2.0/name/current/fsimage.ckpt_000
0000000000000000 using no compression
15/01/15 14:32:33 INFO namenode.FSImage: Image file /app/hadoop/hadoop-2.2.0/name/current/fsimage.ckpt_0000000000
000000000 of size 198 bytes saved in 0 seconds.
15/01/15 14:32:33 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
15/01/15 14:32:33 INFO util.ExitUtil: Exiting with status 0
15/01/15 14:32:33 INFO namenode.NameNode: SHUTDOWN_MSG:
/************************************************************
SHUTDOWN_MSG: Shutting down NameNode at hadoop1/192.168.0.61
************************************************************/
```

2.2.2 启动 HDFS

$cd /app/hadoop/hadoop-2.2.0/sbin

$./start-dfs.sh

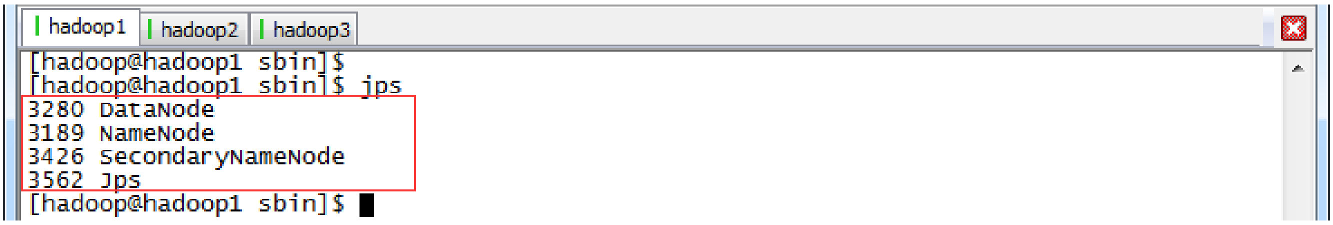


```
[hadoop@hadoop1 ~]$
[hadoop@hadoop1 ~]$ cd /app/hadoop/hadoop-2.2.0/sbin
[hadoop@hadoop1 sbin]$ ./start-dfs.sh
Starting namenodes on [hadoop1]
hadoop1: starting namenode, logging to /app/hadoop/hadoop-2.2.0/logs/hadoop-hadoop-namenode-hadoop1.out
hadoop1: starting datanode, logging to /app/hadoop/hadoop-2.2.0/logs/hadoop-hadoop-datanode-hadoop1.out
hadoop2: starting datanode, logging to /app/hadoop/hadoop-2.2.0/logs/hadoop-hadoop-datanode-hadoop2.out
hadoop3: starting datanode, logging to /app/hadoop/hadoop-2.2.0/logs/hadoop-hadoop-datanode-hadoop3.out
Starting secondary namenodes [hadoop1]
hadoop1: starting secondarynamenode, logging to /app/hadoop/hadoop-2.2.0/logs/hadoop-hadoop-secondaryname
node-hadoop1.out
[hadoop@hadoop1 sbin]$
```

如果服务器操作系统为 32 位时，出现问题 3.1 异常，可以参考解决

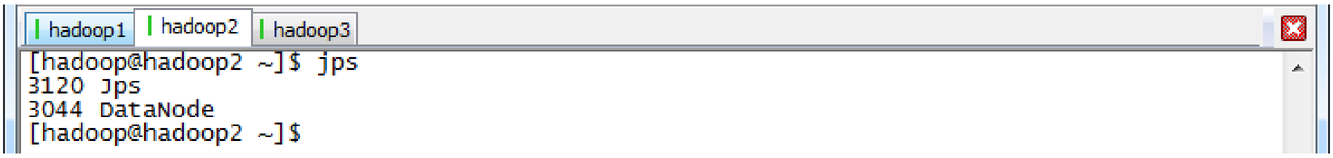
2.2.3 验证 HDFS 启动

此时在 hadoop1 上面运行的进程有：NameNode、SecondaryNameNode 和 DataNode



```
[hadoop@hadoop1 sbin]$
[hadoop@hadoop1 sbin]$ jps
3280 DataNode
3189 NameNode
3426 SecondaryNameNode
3562 Jps
[hadoop@hadoop1 sbin]$
```

hadoop2 和 hadoop3 上面运行的进程有：NameNode 和 DataNode

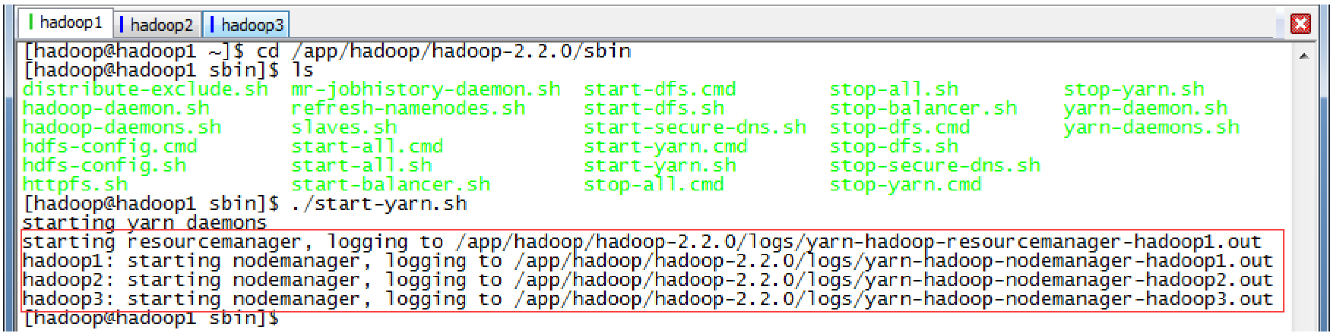


```
[hadoop@hadoop2 ~]$ jps
3120 Jps
3044 DataNode
[hadoop@hadoop2 ~]$
```

2.2.4 启动 YARN

$cd /app/hadoop/hadoop-2.2.0/sbin

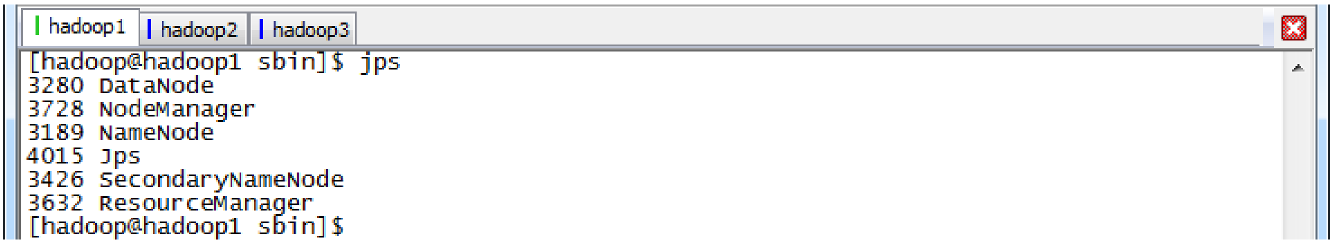
$./start-yarn.sh



```
[hadoop@hadoop1 ~]$ cd /app/hadoop/hadoop-2.2.0/sbin
[hadoop@hadoop1 sbin]$ ls
distribute-exclude.sh  mr-jobhistory-daemon.sh  start-dfs.cmd         stop-all.sh           stop-yarn.sh
hadoop-daemon.sh       refresh-namenodes.sh     start-dfs.sh          stop-balancer.sh      yarn-daemon.sh
hadoop-daemons.sh      slaves.sh                start-secure-dns.sh   stop-dfs.cmd          yarn-daemons.sh
hdfs-config.cmd        start-all.cmd            start-yarn.cmd        stop-dfs.sh
hdfs-config.sh         start-all.sh             start-yarn.sh         stop-secure-dns.sh
httpfs.sh              start-balancer.sh        stop-all.cmd          stop-yarn.cmd
[hadoop@hadoop1 sbin]$ ./start-yarn.sh
starting yarn daemons
starting resourcemanager, logging to /app/hadoop/hadoop-2.2.0/logs/yarn-hadoop-resourcemanager-hadoop1.out
hadoop1: starting nodemanager, logging to /app/hadoop/hadoop-2.2.0/logs/yarn-hadoop-nodemanager-hadoop1.out
hadoop2: starting nodemanager, logging to /app/hadoop/hadoop-2.2.0/logs/yarn-hadoop-nodemanager-hadoop2.out
hadoop3: starting nodemanager, logging to /app/hadoop/hadoop-2.2.0/logs/yarn-hadoop-nodemanager-hadoop3.out
[hadoop@hadoop1 sbin]$
```

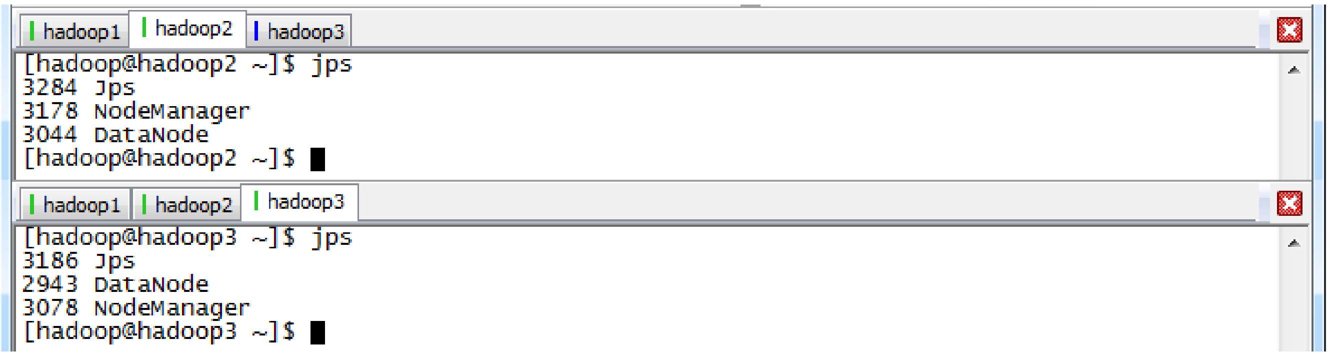
2.2.5 验证 YARN 启动

此时在 hadoop1 上运行的进程有：NameNode、SecondaryNameNode、DataNode、NodeManager 和 ResourceManager



```
[hadoop@hadoop1 sbin]$ jps
3280 DataNode
3728 NodeManager
3189 NameNode
4015 Jps
3426 SecondaryNameNode
3632 ResourceManager
[hadoop@hadoop1 sbin]$
```

hadoop2 和 hadoop3 上面运行的进程有：NameNode、DataNode 和 NodeManager



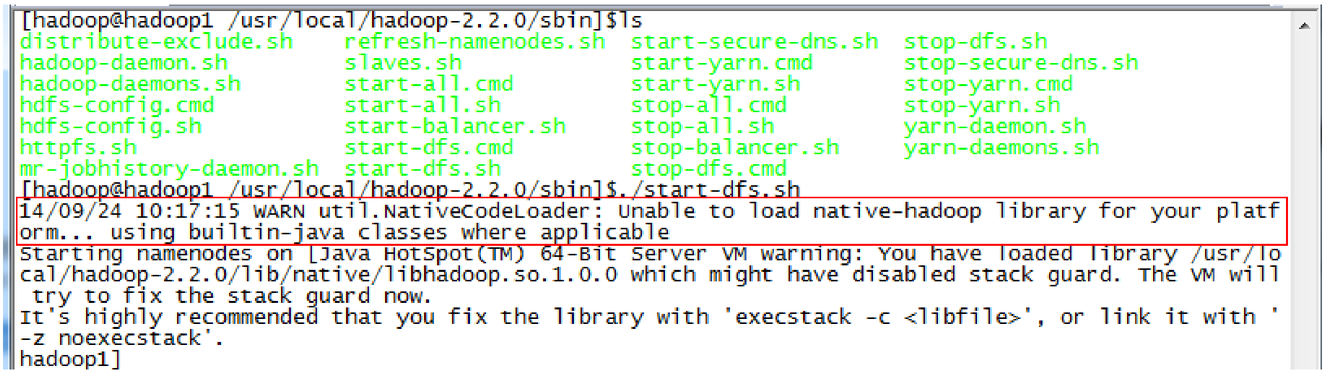
```
[hadoop@hadoop2 ~]$ jps
3284 Jps
3178 NodeManager
3044 DataNode
[hadoop@hadoop2 ~]$
```

```
[hadoop@hadoop3 ~]$ jps
3186 Jps
2943 DataNode
3078 NodeManager
[hadoop@hadoop3 ~]$
```

3 问题解决

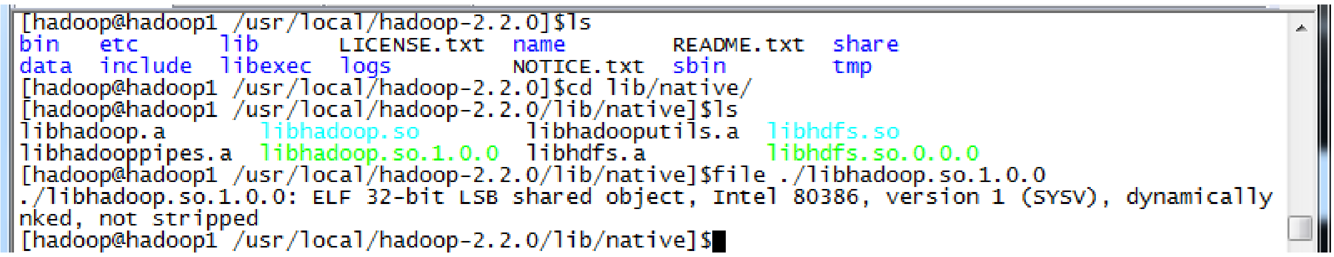
3.1 CentOS 64bit 安装 Hadoop2.2.0 中出现文件编译位数异常

在安装 hadoop2.2.0 过程中出现如下异常：Unable to load native-hadoop library for your platform... using builtin-java classes where applicable



```
[hadoop@hadoop1 /usr/local/hadoop-2.2.0/sbin]$ls
distribute-exclude.sh    refresh-namenodes.sh  start-secure-dns.sh  stop-dfs.sh
hadoop-daemon.sh         slaves.sh             start-yarn.cmd       stop-secure-dns.sh
hadoop-daemons.sh        start-all.cmd         start-yarn.sh        stop-yarn.cmd
hdfs-config.cmd          start-all.sh          stop-all.cmd         stop-yarn.sh
hdfs-config.sh           start-balancer.sh     stop-all.sh          yarn-daemon.sh
httpfs.sh                start-dfs.cmd         stop-balancer.sh     yarn-daemons.sh
mr-jobhistory-daemon.sh  start-dfs.sh          stop-dfs.cmd
[hadoop@hadoop1 /usr/local/hadoop-2.2.0/sbin]$./start-dfs.sh
14/09/24 10:17:15 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platf
orm... using builtin-java classes where applicable
Starting namenodes on [Java HotSpot(TM) 64-Bit Server VM warning: You have loaded library /usr/lo
cal/hadoop-2.2.0/lib/native/libhadoop.so.1.0.0 which might have disabled stack guard. The VM will
 try to fix the stack guard now.
It's highly recommended that you fix the library with 'execstack -c <libfile>', or link it with '
-z noexecstack'.
hadoop1]
```

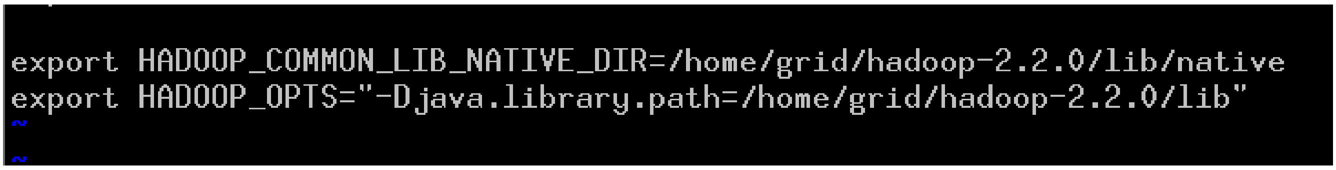
通过分析是由于 lib/native 目录中有些文件是在 32 位编译，无法适应 CentOS 64 位环境造成



```
[hadoop@hadoop1 /usr/local/hadoop-2.2.0]$ls
bin   etc      lib      LICENSE.txt  name        README.txt  share
data  include  libexec  logs         NOTICE.txt  sbin        tmp
[hadoop@hadoop1 /usr/local/hadoop-2.2.0]$cd lib/native/
[hadoop@hadoop1 /usr/local/hadoop-2.2.0/lib/native]$ls
libhadoop.a       libhadoop.so        libhadooputils.a  libhdfs.so
libhadooppipes.a  libhadoop.so.1.0.0  libhdfs.a         libhdfs.so.0.0.0
[hadoop@hadoop1 /usr/local/hadoop-2.2.0/lib/native]$file ./libhadoop.so.1.0.0
./libhadoop.so.1.0.0: ELF 32-bit LSB shared object, Intel 80386, version 1 (SYSV), dynamically
nked, not stripped
[hadoop@hadoop1 /usr/local/hadoop-2.2.0/lib/native]$
```

有两种办法解决：

- 重新编译 hadoop，然后重新部署
- 暂时办法是修改配置，忽略有问题的文件



```
export HADOOP_COMMON_LIB_NATIVE_DIR=/home/grid/hadoop-2.2.0/lib/native
export HADOOP_OPTS="-Djava.library.path=/home/grid/hadoop-2.2.0/lib"
~
~
```

3.2 编译 Hadoop2.2.0 出现代码异常

目前的 2.2.0 的 Source Code 压缩包解压出来的 code 有个 bug 需要 patch 后才能编译。否则编译 hadoop-auth 会提示下面错误：

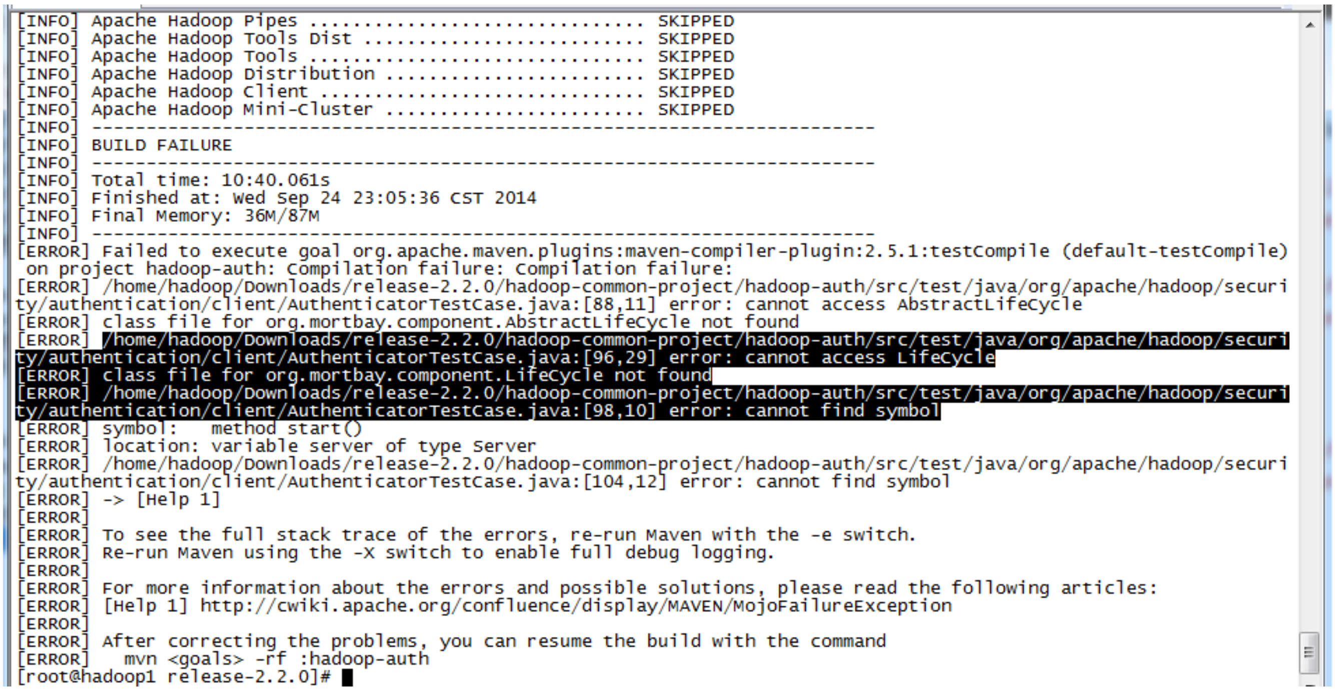
[ERROR] Failed to execute goal org.apache.maven.plugins:maven-compiler-plugin:2.5.1:testCompile (default-testCompile) on project hadoop-auth: Compilation failure: Compilation failure:

[ERROR] /home/hadoop/Downloads/release-2.2.0/hadoop-common-project/hadoop-auth/src/test/java/org/apache/hadoop/security/authentication/client/AuthenticatorTestCase.java:[88,11] error: cannot access AbstractLifeCycle

[ERROR] class file for org.mortbay.component.AbstractLifeCycle not found

[ERROR] /home/hadoop/Downloads/release-2.2.0/hadoop-common-project/hadoop-auth/src/test/java/org/apache/hadoop/security/authentication/client/AuthenticatorTestCase.java:[96,29] error: cannot access LifeCycle

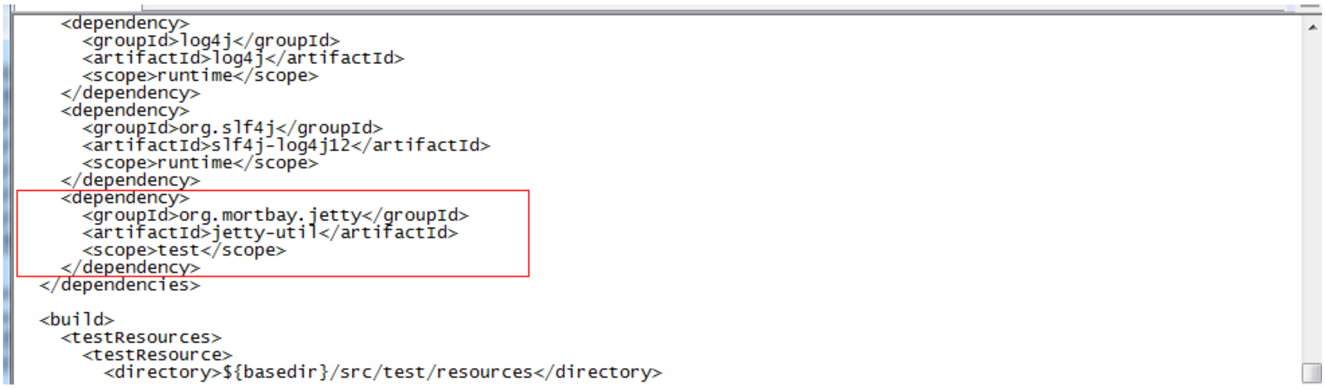
[ERROR] class file for org.mortbay.component.LifeCycle not found



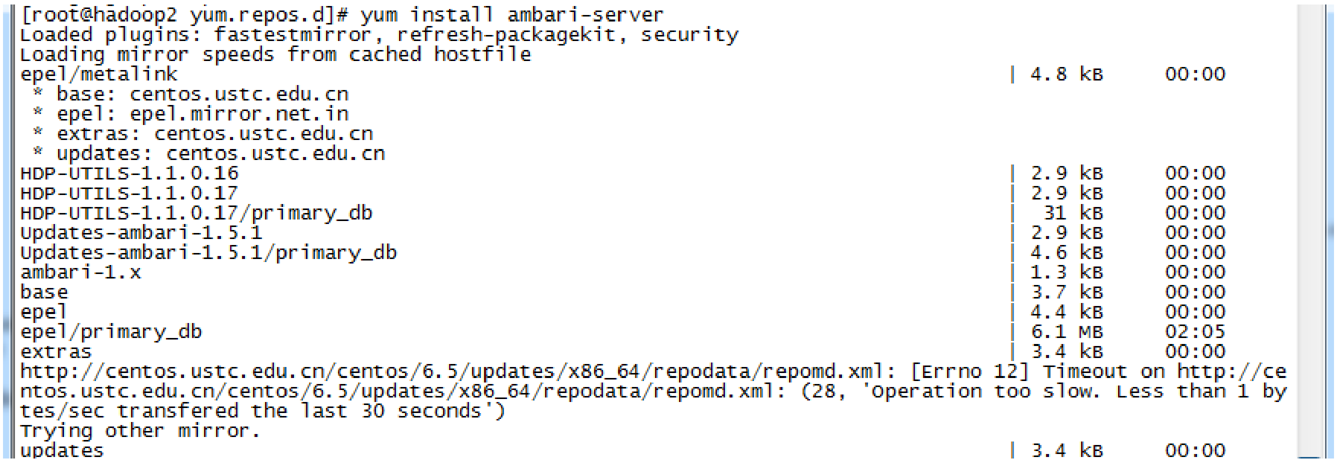
```
[INFO] Apache Hadoop Pipes ............................... SKIPPED
[INFO] Apache Hadoop Tools Dist .......................... SKIPPED
[INFO] Apache Hadoop Tools ............................... SKIPPED
[INFO] Apache Hadoop Distribution ........................ SKIPPED
[INFO] Apache Hadoop Client .............................. SKIPPED
[INFO] Apache Hadoop Mini-Cluster ........................ SKIPPED
[INFO] ------------------------------------------------------------------------
[INFO] BUILD FAILURE
[INFO] ------------------------------------------------------------------------
[INFO] Total time: 10:40.061s
[INFO] Finished at: Wed Sep 24 23:05:36 CST 2014
[INFO] Final Memory: 36M/87M
[INFO] ------------------------------------------------------------------------
[ERROR] Failed to execute goal org.apache.maven.plugins:maven-compiler-plugin:2.5.1:testCompile (default-testCompile)
 on project hadoop-auth: Compilation failure: Compilation failure:
[ERROR] /home/hadoop/Downloads/release-2.2.0/hadoop-common-project/hadoop-auth/src/test/java/org/apache/hadoop/securi
ty/authentication/client/AuthenticatorTestCase.java:[88,11] error: cannot access AbstractLifeCycle
[ERROR] class file for org.mortbay.component.AbstractLifeCycle not found
[ERROR] /home/hadoop/Downloads/release-2.2.0/hadoop-common-project/hadoop-auth/src/test/java/org/apache/hadoop/securi
ty/authentication/client/AuthenticatorTestCase.java:[96,29] error: cannot access LifeCycle
[ERROR] class file for org.mortbay.component.LifeCycle not found
[ERROR] /home/hadoop/Downloads/release-2.2.0/hadoop-common-project/hadoop-auth/src/test/java/org/apache/hadoop/securi
ty/authentication/client/AuthenticatorTestCase.java:[98,10] error: cannot find symbol
[ERROR] symbol:   method start()
[ERROR] location: variable server of type Server
[ERROR] /home/hadoop/Downloads/release-2.2.0/hadoop-common-project/hadoop-auth/src/test/java/org/apache/hadoop/securi
ty/authentication/client/AuthenticatorTestCase.java:[104,12] error: cannot find symbol
[ERROR] -> [Help 1]
[ERROR]
[ERROR] To see the full stack trace of the errors, re-run Maven with the -e switch.
[ERROR] Re-run Maven using the -X switch to enable full debug logging.
[ERROR]
[ERROR] For more information about the errors and possible solutions, please read the following articles:
[ERROR] [Help 1] http://cwiki.apache.org/confluence/display/MAVEN/MojoFailureException
[ERROR]
[ERROR] After correcting the problems, you can resume the build with the command
[ERROR]   mvn <goals> -rf :hadoop-auth
[root@hadoop1 release-2.2.0]# 
```

直接修改 hadoop-common-project/hadoop-auth/pom.xml，其实就是少了一个包，添加一个 dependency：

```
<dependency>
     <groupId>org.mortbay.jetty</groupId>
     <artifactId>jetty-util</artifactId>
     <scope>test</scope>
</dependency>
```



```
    <dependency>
      <groupId>log4j</groupId>
      <artifactId>log4j</artifactId>
      <scope>runtime</scope>
    </dependency>
    <dependency>
      <groupId>org.slf4j</groupId>
      <artifactId>slf4j-log4j12</artifactId>
      <scope>runtime</scope>
    </dependency>
    <dependency>
      <groupId>org.mortbay.jetty</groupId>
      <artifactId>jetty-util</artifactId>
      <scope>test</scope>
    </dependency>
  </dependencies>

  <build>
    <testResources>
      <testResource>
        <directory>${basedir}/src/test/resources</directory>
```



```
[root@hadoop2 yum.repos.d]# yum install ambari-server
Loaded plugins: fastestmirror, refresh-packagekit, security
Loading mirror speeds from cached hostfile
epel/metalink                                                                | 4.8 kB     00:00
 * base: centos.ustc.edu.cn
 * epel: epel.mirror.net.in
 * extras: centos.ustc.edu.cn
 * updates: centos.ustc.edu.cn
HDP-UTILS-1.1.0.16                                                           | 2.9 kB     00:00
HDP-UTILS-1.1.0.17                                                           | 2.9 kB     00:00
HDP-UTILS-1.1.0.17/primary_db                                                |  31 kB     00:00
Updates-ambari-1.5.1                                                         | 2.9 kB     00:00
Updates-ambari-1.5.1/primary_db                                              | 4.6 kB     00:00
ambari-1.x                                                                   | 1.3 kB     00:00
base                                                                         | 3.7 kB     00:00
epel                                                                         | 4.4 kB     00:00
epel/primary_db                                                              | 6.1 MB     02:05
extras                                                                       | 3.4 kB     00:00
http://centos.ustc.edu.cn/centos/6.5/updates/x86_64/repodata/repomd.xml: [Errno 12] Timeout on http://ce
ntos.ustc.edu.cn/centos/6.5/updates/x86_64/repodata/repomd.xml: (28, 'Operation too slow. Less than 1 by
tes/sec transfered the last 30 seconds')
Trying other mirror.
updates                                                                      | 3.4 kB     00:00
```

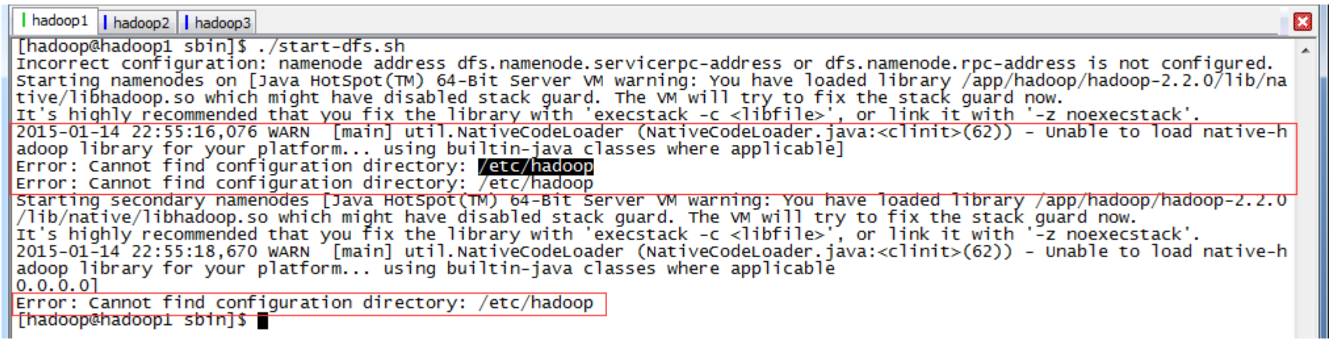
3.3 安装 Hadoop2.2.0 出现不能找到/etc/hadoop 目录异常

在安装过程中启动 HDFS 出现如下错误：

2015-01-14 22:55:16,076 WARN [main] util.NativeCodeLoader (NativeCodeLoader.java:<clinit>(62)) - Unable to load native-hadoop library for your platform... using builtin-java classes where applicable]

Error: Cannot find configuration directory: /etc/hadoop

Error: Cannot find configuration directory: /etc/hadoop

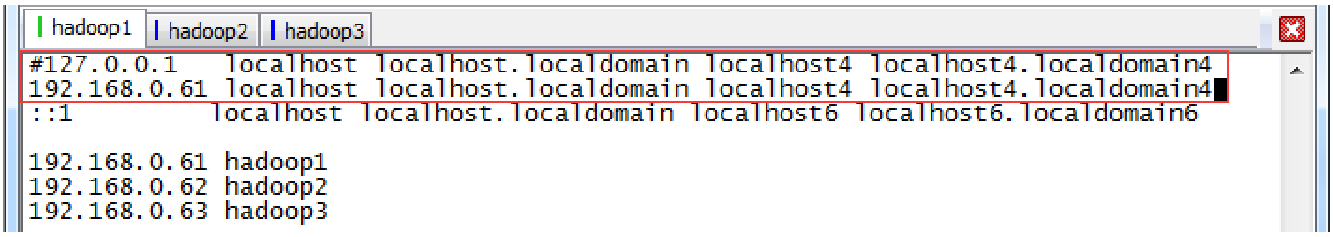


```
[hadoop@hadoop1 sbin]$ ./start-dfs.sh
Incorrect configuration: namenode address dfs.namenode.servicerpc-address or dfs.namenode.rpc-address is not configured.
Starting namenodes on [Java HotSpot(TM) 64-Bit Server VM warning: You have loaded library /app/hadoop/hadoop-2.2.0/lib/na
tive/libhadoop.so which might have disabled stack guard. The VM will try to fix the stack guard now.
It's highly recommended that you fix the library with 'execstack -c <libfile>', or link it with '-z noexecstack'.
2015-01-14 22:55:16,076 WARN  [main] util.NativeCodeLoader (NativeCodeLoader.java:<clinit>(62)) - Unable to load native-h
adoop library for your platform... using builtin-java classes where applicable]
Error: Cannot find configuration directory: /etc/hadoop
Error: Cannot find configuration directory: /etc/hadoop
Starting secondary namenodes [Java HotSpot(TM) 64-Bit Server VM warning: You have loaded library /app/hadoop/hadoop-2.2.0
/lib/native/libhadoop.so which might have disabled stack guard. The VM will try to fix the stack guard now.
It's highly recommended that you fix the library with 'execstack -c <libfile>', or link it with '-z noexecstack'.
2015-01-14 22:55:18,670 WARN  [main] util.NativeCodeLoader (NativeCodeLoader.java:<clinit>(62)) - Unable to load native-h
adoop library for your platform... using builtin-java classes where applicable
0.0.0.0]
Error: Cannot find configuration directory: /etc/hadoop
[hadoop@hadoop1 sbin]$
```

127.0.0.1   localhost

改为

192.168.0.61   localhost



```
#127.0.0.1    localhost localhost.localdomain localhost4 localhost4.localdomain4
192.168.0.61  localhost localhost.localdomain localhost4 localhost4.localdomain4
::1           localhost localhost.localdomain localhost6 localhost6.localdomain6

192.168.0.61  hadoop1
192.168.0.62  hadoop2
192.168.0.63  hadoop3
```

重启机器即可