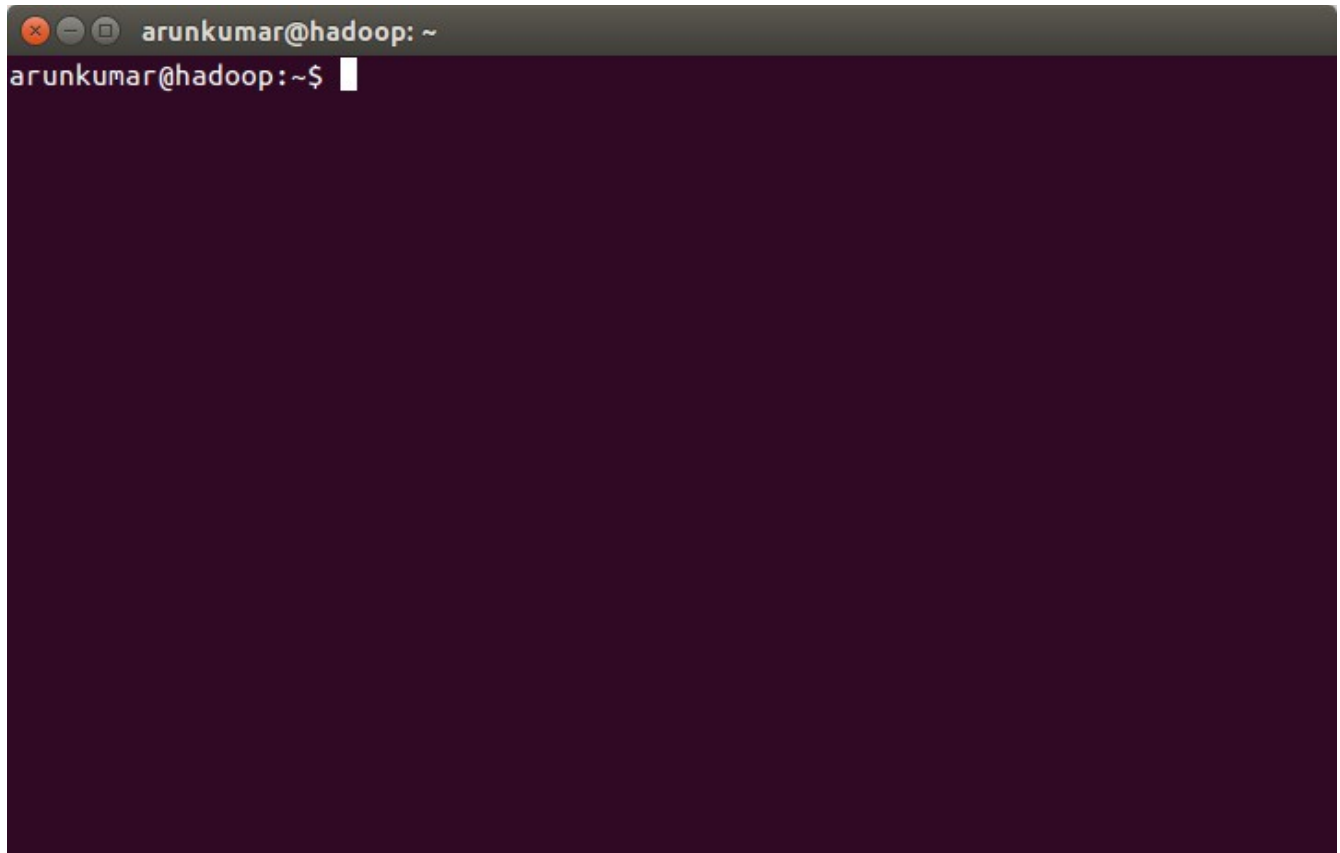


Hadoop Single node Pseudo distributed mode installation

1. Open Terminal



2. Install open SSH server

```
sudo apt-get install openssh-server
```

```
arunkumar@hadoop: ~  
arunkumar@hadoop:~$ sudo apt-get install openssh-server  
[sudo] password for arunkumar:  
Reading package lists... Done  
Building dependency tree  
Reading state information... Done  
The following extra packages will be installed:  
  libck-connector0 ncurses-term openssh-sftp-server python-requests  
  python-urllib3 ssh-import-id  
Suggested packages:  
  rssh molly-guard monkeysphere  
The following NEW packages will be installed:  
  libck-connector0 ncurses-term openssh-server openssh-sftp-server  
  python-requests python-urllib3 ssh-import-id  
0 upgraded, 7 newly installed, 0 to remove and 0 not upgraded.  
Need to get 698 kB of archives.  
After this operation, 3,832 kB of additional disk space will be used.  
Do you want to continue? [Y/n]
```

3. Yes to continue. You will end up in a similar screen

```
arunkumar@hadoop: ~  
Unpacking python-requests (2.2.1-1) ...  
Selecting previously unselected package ssh-import-id.  
Preparing to unpack .../ssh-import-id_3.21-0ubuntu1_all.deb ...  
Unpacking ssh-import-id (3.21-0ubuntu1) ...  
Processing triggers for man-db (2.6.7.1-1) ...  
Processing triggers for ureadahead (0.100.0-16) ...  
ureadahead will be reprofiled on next reboot  
Processing triggers for ufw (0.34~rc-0ubuntu2) ...  
Setting up libck-connector0:amd64 (0.4.5-3.1ubuntu2) ...  
Setting up ncurses-term (5.9+20140118-1ubuntu1) ...  
Setting up openssh-sftp-server (1:6.6p1-2ubuntu2) ...  
Setting up openssh-server (1:6.6p1-2ubuntu2) ...  
Creating SSH2 RSA key; this may take some time ...  
Creating SSH2 DSA key; this may take some time ...  
Creating SSH2 ECDSA key; this may take some time ...  
Creating SSH2 ED25519 key; this may take some time ...  
ssh start/running, process 3409  
Setting up python-urllib3 (1.7.1-1build1) ...  
Setting up python-requests (2.2.1-1) ...  
Setting up ssh-import-id (3.21-0ubuntu1) ...  
Processing triggers for libc-bin (2.19-0ubuntu6.3) ...  
Processing triggers for ureadahead (0.100.0-16) ...  
Processing triggers for ufw (0.34~rc-0ubuntu2) ...  
arunkumar@hadoop:~$
```

4. Generate SSH Key pair

commands : ssh-keygen

Enter file in which to save the key (/home/arunkumar/.ssh/id_rsa): leave as default
– do not enter anything

Enter passphrase (empty for no passphrase) : leave it as empty as it required for automatic login

Enter same passphrase again: again no input required. Just enter.

```
arunkumar@hadoop: ~  
arunkumar@hadoop:~$ ssh-keygen  
Generating public/private rsa key pair.  
Enter file in which to save the key (/home/arunkumar/.ssh/id_rsa):  
Created directory '/home/arunkumar/.ssh'.  
Enter passphrase (empty for no passphrase):  
Enter same passphrase again:  
Your identification has been saved in /home/arunkumar/.ssh/id_rsa.  
Your public key has been saved in /home/arunkumar/.ssh/id_rsa.pub.  
The key fingerprint is:  
b8:d9:08:f1:8e:c8:91:26:94:f9:01:c4:b9:4b:f5:2b arunkumar@hadoop  
The key's randomart image is:  
+--[ RSA 2048 ]-----+  
|oo.                    |  
| o+.                   |  
| +o.o                  |  
|.o...+ .              |  
|o =.. + S             |  
| = E = =              |  
|  o o = .              |  
|                        |  
+-----+  
arunkumar@hadoop:~$ ^C  
arunkumar@hadoop:~$
```

5. Add the generated key to the authorized key file to use as a password-less login.

```
cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
```

```
arunkumar@hadoop: ~
Generating public/private rsa key pair.
Enter file in which to save the key (/home/arunkumar/.ssh/id_rsa):
Created directory '/home/arunkumar/.ssh'.
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/arunkumar/.ssh/id_rsa.
Your public key has been saved in /home/arunkumar/.ssh/id_rsa.pub.
The key fingerprint is:
b8:d9:08:f1:8e:c8:91:26:94:f9:01:c4:b9:4b:f5:2b arunkumar@hadoop
The key's randomart image is:
+--[ RSA 2048 ]-----+
|oo.                    |
| o+.                   |
| +o.o                  |
|.o...+ .               |
|o =.. + S              |
| = E = =               |
|  o o = .              |
|                        |
+-----+
arunkumar@hadoop:~$ ^C
arunkumar@hadoop:~$ cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
arunkumar@hadoop:~$
```

6. try logging in to localhost

ssh localhost

type 'yes' for the question about connection

Now you can login without a password-less

7. Now install java

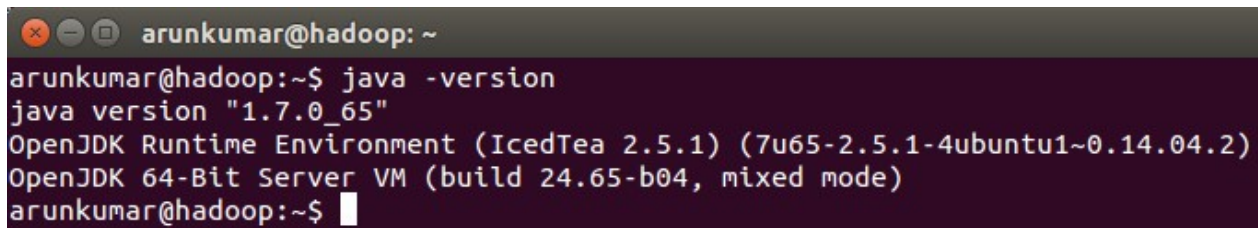
sudo apt-get install openjdk-7-jdk

```
arunkumar@hadoop: ~  
Last login: Sun Sep 14 19:32:48 2014 from localhost  
arunkumar@hadoop:~$  
arunkumar@hadoop:~$ sudo apt-get install openjdk-7-jdk  
[sudo] password for arunkumar:  
Reading package lists... Done  
Building dependency tree  
Reading state information... Done  
The following extra packages will be installed:  
  ca-certificates-java fonts-dejavu-extra icedtea-7-jre-jamvm java-common  
  libatk-wrapper-java libatk-wrapper-java-jni libbonobo2-0 libbonobo2-common  
  libgconf2-4 libgif4 libgnome2-0 libgnome2-bin libgnome2-common  
  libgnomevfs2-0 libgnomevfs2-common libice-dev libidl-common libidl0  
  libnss3-1d liborbit-2-0 liborbit2 libpthread-stubs0-dev libsm-dev libx11-dev  
  libx11-doc libxau-dev libxcb1-dev libxdmcp-dev libxt-dev openjdk-7-jre  
  openjdk-7-jre-headless tzdata-java x11proto-core-dev x11proto-input-dev  
  x11proto-kb-dev xorg-sgml-doctools xtrans-dev  
Suggested packages:  
  default-jre equivs libbonobo2-bin desktop-base libgnomevfs2-bin  
  libgnomevfs2-extra gamin fam gnome-mime-data libice-doc libsm-doc libxcb-doc  
  libxt-doc openjdk-7-demo openjdk-7-source visualvm icedtea-7-plugin  
  sun-java6-fonts fonts-ipafont-gothic fonts-ipafont-mincho ttf-wqy-microhei  
  ttf-wqy-zenhei ttf-telugu-fonts ttf-oriya-fonts ttf-kannada-fonts  
  ttf-bengali-fonts
```

Yes to any question asked. Final screen after installation.

```
arunkumar@hadoop: ~  
Adding debian:Sonera_Class_2_Root_CA.pem  
Adding debian:Entrust.net_Premium_2048_Secure_Server_CA.pem  
Adding debian:GeoTrust_Universal_CA.pem  
Adding debian:NetLock_Express_Class_C_Root.pem  
Adding debian:Swisscom_Root_EV_CA_2.pem  
Adding debian:America_Online_Root_Certification_Authority_2.pem  
Adding debian:AddTrust_Qualified_Certificates_Root.pem  
Adding debian:Entrust_Root_Certification_Authority.pem  
Adding debian:Digital_Signature_Trust_Co._Global_CA_3.pem  
Adding debian:CA_Disig.pem  
Adding debian:QuoVadis_Root_CA.pem  
Adding debian:Equifax_Secure_Global_eBusiness_CA.pem  
Adding debian:T-TeleSec_GlobalRoot_Class_3.pem  
done.  
Setting up libatk-wrapper-java-jni:amd64 (0.30.4-4) ...  
Setting up libgnome2-bin (2.32.1-4ubuntu1) ...  
Setting up libgnome2-0:amd64 (2.32.1-4ubuntu1) ...  
Processing triggers for libc-bin (2.19-0ubuntu6.3) ...  
Processing triggers for ca-certificates (20130906ubuntu2) ...  
Updating certificates in /etc/ssl/certs... 0 added, 0 removed; done.  
Running hooks in /etc/ca-certificates/update.d....  
done.  
done.  
arunkumar@hadoop:~$
```

8. check the java version

A terminal window with a dark purple background and a grey title bar. The title bar contains three window control icons (close, minimize, maximize) and the text 'arunkumar@hadoop: ~'. The terminal shows the command 'java -version' being executed, with the following output: 'java version "1.7.0_65"', 'OpenJDK Runtime Environment (IcedTea 2.5.1) (7u65-2.5.1-4ubuntu1~0.14.04.2)', and 'OpenJDK 64-Bit Server VM (build 24.65-b04, mixed mode)'. The prompt 'arunkumar@hadoop:~\$' is followed by a white cursor.

```
arunkumar@hadoop:~$ java -version
java version "1.7.0_65"
OpenJDK Runtime Environment (IcedTea 2.5.1) (7u65-2.5.1-4ubuntu1~0.14.04.2)
OpenJDK 64-Bit Server VM (build 24.65-b04, mixed mode)
arunkumar@hadoop:~$
```

10. Now its the time to download hadoop. I am going to install hadoop 1.2.0

Download a stable version from the link :

<http://archive.apache.org/dist/hadoop/core/hadoop-1.2.0/>

11. You can also use command line to download the file.

```
sudo wget http://archive.apache.org/dist/hadoop/core/hadoop-1.2.0/hadoop-1.2.0-  
bin.tar.gz
```



```
arunkumar@hadoop: ~  
arunkumar@hadoop:~$ sudo wget http://archive.apache.org/dist/hadoop/core/hadoop-  
1.2.0/hadoop-1.2.0-bin.tar.gz  
--2014-09-14 19:52:42-- http://archive.apache.org/dist/hadoop/core/hadoop-1.2.0/  
hadoop-1.2.0-bin.tar.gz  
Resolving archive.apache.org (archive.apache.org)... 192.87.106.229, 140.211.11.  
131, 2001:610:1:80bc:192:87:106:229  
Connecting to archive.apache.org (archive.apache.org)|192.87.106.229|:80... conn  
ected.  
HTTP request sent, awaiting response... 200 OK  
Length: 37650688 (36M) [application/x-gzip]  
Saving to: 'hadoop-1.2.0-bin.tar.gz'  
  
0% [          ] 236,320    70.6KB/s  eta 8m 53s
```

12. unzip the downloaded file

```
tar -zxvf hadoop-1.2.0-bin.tar.gz
```

```
arunkumar@hadoop: ~  
688]  
  
arunkumar@hadoop:~$ tar -zxvf hadoop-1.2.0-bin.tar.gz  
hadoop-1.2.0/  
hadoop-1.2.0/.eclipse.templates/  
hadoop-1.2.0/.eclipse.templates/.externalToolBuilders/  
hadoop-1.2.0/.eclipse.templates/.launches/  
hadoop-1.2.0/bin/  
hadoop-1.2.0/c++/  
hadoop-1.2.0/c++/Linux-amd64-64/  
hadoop-1.2.0/c++/Linux-amd64-64/include/  
hadoop-1.2.0/c++/Linux-amd64-64/include/hadoop/  
hadoop-1.2.0/c++/Linux-amd64-64/lib/  
hadoop-1.2.0/c++/Linux-i386-32/  
hadoop-1.2.0/c++/Linux-i386-32/include/  
hadoop-1.2.0/c++/Linux-i386-32/include/hadoop/  
hadoop-1.2.0/c++/Linux-i386-32/lib/  
hadoop-1.2.0/conf/  
hadoop-1.2.0/contrib/  
hadoop-1.2.0/contrib/datajoin/  
hadoop-1.2.0/contrib/failmon/  
hadoop-1.2.0/contrib/gridmix/  
hadoop-1.2.0/contrib/hdfsproxy/  
hadoop-1.2.0/contrib/hdfsproxy/bin/
```

13. type 'dir' and make sure the file is unzipped correctly.

```
arunkumar@hadoop: ~  
hadoop-1.2.0/sbin/hadoop-setup-applications.sh  
hadoop-1.2.0/sbin/hadoop-setup-conf.sh  
hadoop-1.2.0/sbin/hadoop-setup-hdfs.sh  
hadoop-1.2.0/sbin/hadoop-setup-single-node.sh  
hadoop-1.2.0/sbin/hadoop-validate-setup.sh  
hadoop-1.2.0/sbin/update-hadoop-env.sh  
hadoop-1.2.0/src/contrib/ec2/bin/cmd-hadoop-cluster  
hadoop-1.2.0/src/contrib/ec2/bin/create-hadoop-image  
hadoop-1.2.0/src/contrib/ec2/bin/delete-hadoop-cluster  
hadoop-1.2.0/src/contrib/ec2/bin/hadoop-ec2  
hadoop-1.2.0/src/contrib/ec2/bin/hadoop-ec2-env.sh  
hadoop-1.2.0/src/contrib/ec2/bin/hadoop-ec2-init-remote.sh  
hadoop-1.2.0/src/contrib/ec2/bin/image/create-hadoop-image-remote  
hadoop-1.2.0/src/contrib/ec2/bin/image/ec2-run-user-data  
hadoop-1.2.0/src/contrib/ec2/bin/launch-hadoop-cluster  
hadoop-1.2.0/src/contrib/ec2/bin/launch-hadoop-master  
hadoop-1.2.0/src/contrib/ec2/bin/launch-hadoop-slaves  
hadoop-1.2.0/src/contrib/ec2/bin/list-hadoop-clusters  
hadoop-1.2.0/src/contrib/ec2/bin/terminate-hadoop-cluster  
arunkumar@hadoop:~$ dir  
Desktop      Dropbox      hadoop-1.2.0-bin.tar.gz  Public  
Documents    examples.desktop  Music                Templates  
Downloads    hadoop-1.2.0     Pictures             Videos  
arunkumar@hadoop:~$
```

14. copy the file to user directory and rename to 'hadoop'

```
sudo cp -r hadoop-1.2.0 /usr/local/hadoop
```

```
arunkumar@hadoop: ~  
hadoop-1.2.0/sbin/hadoop-setup-conf.sh  
hadoop-1.2.0/sbin/hadoop-setup-hdfs.sh  
hadoop-1.2.0/sbin/hadoop-setup-single-node.sh  
hadoop-1.2.0/sbin/hadoop-validate-setup.sh  
hadoop-1.2.0/sbin/update-hadoop-env.sh  
hadoop-1.2.0/src/contrib/ec2/bin/cmd-hadoop-cluster  
hadoop-1.2.0/src/contrib/ec2/bin/create-hadoop-image  
hadoop-1.2.0/src/contrib/ec2/bin/delete-hadoop-cluster  
hadoop-1.2.0/src/contrib/ec2/bin/hadoop-ec2  
hadoop-1.2.0/src/contrib/ec2/bin/hadoop-ec2-env.sh  
hadoop-1.2.0/src/contrib/ec2/bin/hadoop-ec2-init-remote.sh  
hadoop-1.2.0/src/contrib/ec2/bin/image/create-hadoop-image-remote  
hadoop-1.2.0/src/contrib/ec2/bin/image/ec2-run-user-data  
hadoop-1.2.0/src/contrib/ec2/bin/launch-hadoop-cluster  
hadoop-1.2.0/src/contrib/ec2/bin/launch-hadoop-master  
hadoop-1.2.0/src/contrib/ec2/bin/launch-hadoop-slaves  
hadoop-1.2.0/src/contrib/ec2/bin/list-hadoop-clusters  
hadoop-1.2.0/src/contrib/ec2/bin/terminate-hadoop-cluster  
arunkumar@hadoop:~$ dir  
Desktop      Dropbox      hadoop-1.2.0-bin.tar.gz  Public  
Documents    examples.desktop  Music                Templates  
Downloads    hadoop-1.2.0     Pictures             Videos  
arunkumar@hadoop:~$ sudo cp -r hadoop-1.2.0 /usr/local/hadoop  
arunkumar@hadoop:~$
```

15. Check the content of the file

`dir /usr/local/hadoop`

```
arunkumar@hadoop: ~  
arunkumar@hadoop:~$ dir /usr/local/hadoop  
bin          hadoop-ant-1.2.0.jar      hadoop-tools-1.2.0.jar  NOTICE.txt  
build.xml    hadoop-client-1.2.0.jar   ivy                     README.txt  
c++          hadoop-core-1.2.0.jar     ivy.xml                 sbin  
CHANGES.txt hadoop-examples-1.2.0.jar lib                      share  
conf         hadoop-miniclust-1.2.0.jar libexec                 src  
contrib      hadoop-test-1.2.0.jar     LICENSE.txt             webapps  
arunkumar@hadoop:~$
```

16. Add the HOME location (local HADOOP directory) to the PATH variable in bashrc file.

```
sudo vi $HOME/.bashrc
```

Add the below lines at the end of the file

```
export HADOOP_PREFIX=/usr/local/hadoop  
export PATH=$PATH:$HADOOP_PREFIX/bin
```

```
arunkumar@hadoop: ~  
# enable programmable completion features (you don't need to enable  
# this, if it's already enabled in /etc/bash.bashrc and /etc/profile  
# sources /etc/bash.bashrc).  
if ! shopt -oq posix; then  
  if [ -f /usr/share/bash-completion/bash_completion ]; then  
    . /usr/share/bash-completion/bash_completion  
  elif [ -f /etc/bash_completion ]; then  
    . /etc/bash_completion  
  fi  
fi  
export HADOOP_PREFIX=/usr/local/hadoop  
export PATH=$PATH:$HADOOP_PREFIX/bin  
~  
~  
~  
~  
~  
~  
~  
~  
~  
~
```

17. Execute the bash once to make the changes in bashrc

exec bash

```
arunkumar@hadoop: ~  
' )"  
  
# Alias definitions.  
# You may want to put all your additions into a separate file like  
# ~/.bash_aliases, instead of adding them here directly.  
# See /usr/share/doc/bash-doc/examples in the bash-doc package.  
  
if [ -f ~/.bash_aliases ]; then  
    . ~/.bash_aliases  
fi  
  
# enable programmable completion features (you don't need to enable  
# this, if it's already enabled in /etc/bash.bashrc and /etc/profile  
# sources /etc/bash.bashrc).  
if ! shopt -oq posix; then  
    if [ -f /usr/share/bash-completion/bash_completion ]; then  
        . /usr/share/bash-completion/bash_completion  
    elif [ -f /etc/bash_completion ]; then  
        . /etc/bash_completion  
    fi  
fi  
export HADOOP_PREFIX=/usr/local/hadoop  
export PATH=$PATH:$HADOOP_PREFIX/bin  
arunkumar@hadoop:~$ exec bash
```

18. check the \$PATH for updated path

\$PATH

```
arunkumar@hadoop: ~  
# See /usr/share/doc/bash-doc/examples in the bash-doc package.  
  
if [ -f ~/.bash_aliases ]; then  
    . ~/.bash_aliases  
fi  
  
# enable programmable completion features (you don't need to enable  
# this, if it's already enabled in /etc/bash.bashrc and /etc/profile  
# sources /etc/bash.bashrc).  
if ! shopt -oq posix; then  
    if [ -f /usr/share/bash-completion/bash_completion ]; then  
        . /usr/share/bash-completion/bash_completion  
    elif [ -f /etc/bash_completion ]; then  
        . /etc/bash_completion  
    fi  
fi  
export HADOOP_PREFIX=/usr/local/hadoop  
export PATH=$PATH:$HADOOP_PREFIX/bin  
arunkumar@hadoop:~$ exec bash  
arunkumar@hadoop:~$ $PATH  
bash: /usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/bin:/usr/games:/u  
sr/local/games:/usr/local/hadoop/bin: No such file or directory  
arunkumar@hadoop:~$ ^C  
arunkumar@hadoop:~$
```

19. Now we can start configuring the hadoop. All the configuration files are located in the conf folder as shown below

```
dir /usr/local/hadoop/conf
```



```
arunkumar@hadoop: ~  
arunkumar@hadoop:~$ dir /usr/local/hadoop/conf  
capacity-scheduler.xml      hadoop-policy.xml          slaves  
configuration.xml           hdfs-site.xml             ssl-client.xml.example  
core-site.xml               log4j.properties          ssl-server.xml.example  
fair-scheduler.xml          mapred-queue-acls.xml      taskcontroller.cfg  
hadoop-env.sh               mapred-site.xml           task-log4j.properties  
hadoop-metrics2.properties  masters  
arunkumar@hadoop:~$
```

20. we can setup java and disable IPV6. This can be achieved by modifying the hadoop-env.sh file.

```
sudo vi /usr/local/hadoop/conf/hadoop-env.sh
```

uncomment the exporting of JAVA_HOME variable :

```
export JAVA_HOME=/usr/lib/jvm/java-1.7.0-openjdk-amd64
```

```
arunkumar@hadoop: ~  
# Set Hadoop-specific environment variables here.  
  
# The only required environment variable is JAVA_HOME. All others are  
# optional. When running a distributed configuration it is best to  
# set JAVA_HOME in this file, so that it is correctly defined on  
# remote nodes.  
  
# The java implementation to use. Required.  
export JAVA_HOME=/usr/lib/jvm/java-1.7.0-openjdk-amd64  
  
# Extra Java CLASSPATH elements. Optional.  
# export HADOOP_CLASSPATH=  
  
# The maximum amount of heap to use, in MB. Default is 1000.  
# export HADOOP_HEAPSIZE=2000  
  
# Extra Java runtime options. Empty by default.  
# export HADOOP_OPTS=-server  
  
# Command specific options appended to HADOOP_OPTS when specified  
export HADOOP_NAMENODE_OPTS="-Dcom.sun.management.jmxremote $HADOOP_NAMENODE_OPTS"  
@
```

21. Disable ipv6

```
export HADOOP_OPTS=-Djava.net.preferIPv4Stack=true
```

```

arunkumar@hadoop: ~
# Set Hadoop-specific environment variables here.

# The only required environment variable is JAVA_HOME. All others are
# optional. When running a distributed configuration it is best to
# set JAVA_HOME in this file, so that it is correctly defined on
# remote nodes.

# The java implementation to use. Required.
export JAVA_HOME=/usr/lib/jvm/java-1.7.0-openjdk-amd64

# Extra Java CLASSPATH elements. Optional.
# export HADOOP_CLASSPATH=

# The maximum amount of heap to use, in MB. Default is 1000.
# export HADOOP_HEAPSIZE=2000

# Extra Java runtime options. Empty by default.
export HADOOP_OPTS=-Djava.net.preferIPv4Stack=true

# Command specific options appended to HADOOP_OPTS when specified
export HADOOP_NAMENODE_OPTS="-Dcom.sun.management.jmxremote $HADOOP_NAMENODE_OPTS"
@

```

22. Configure the name node (core-site.xml)

```
sudo vi /usr/local/hadoop/conf/core-site.xml
```

Add the below property and values

```

<configuration>
<property>
<name>fs.default.name</name>
<value>hdfs://hadoop:10001</value>
</property>

<property>
<name>hadoop.tmp.dir</name>
<value>/usr/local/hadoop/tmp</value>
</property>

```

```
arunkumar@hadoop: ~  
<?xml version="1.0"?>  
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>  
  
<!-- Put site-specific property overrides in this file. -->  
  
<configuration>  
  <property>  
    <name>fs.default.name</name>  
    <value>hdfs://hadoop:10001</value>  
  </property>  
  
  <property>  
    <name>hadoop.tmp.dir</name>  
    <value>/usr/local/hadoop/tmp</value>  
  </property>  
</configuration>  
~  
~  
~  
~  
~  
~  
~
```

23. Configure the mapred-site.xml (job tracker)

```
sudo vi /usr/local/hadoop/conf/mapred-site.xml
```

```
<configuration>  
<property>  
<name>mapred.job.tracker</name>  
<value>hadoop:10002</value>  
</property>  
</configuration>
```



```
arunkumar@hadoop: ~  
arunkumar@hadoop:~$ sudo vi /usr/local/hadoop/conf/hadoop-env.sh  
arunkumar@hadoop:~$ sudo vi /usr/local/hadoop/conf/core-site.xml  
arunkumar@hadoop:~$ sudo vi /usr/local/hadoop/conf/mapred-site.xml  
arunkumar@hadoop:~$ sudo mkdir /usr/local/hadoop/tmp  
arunkumar@hadoop:~$ ^C  
arunkumar@hadoop:~$
```

25. change ownership and permission of the directory

```
sudo chown arunkumar /usr/local/hadoop/tmp  
sudo chmod 777 /usr/local/hadoop/
```

```
arunkumar@hadoop: ~  
arunkumar@hadoop:~$ sudo vi /usr/local/hadoop/conf/hadoop-env.sh  
arunkumar@hadoop:~$ sudo vi /usr/local/hadoop/conf/core-site.xml  
arunkumar@hadoop:~$ sudo vi /usr/local/hadoop/conf/mapred-site.xml  
arunkumar@hadoop:~$ sudo mkdir /usr/local/hadoop/tmp  
arunkumar@hadoop:~$ ^C  
arunkumar@hadoop:~$ sudo chown arunkumar /usr/local/hadoop/tmp  
arunkumar@hadoop:~$ sudo chmod 777 /usr/local/hadoop/tmp  
arunkumar@hadoop:~$
```

26. Format namenode

hadoop namenode -format

```

arunkumar@hadoop: ~
14/09/14 20:47:11 INFO util.GSet: recommended=2097152, actual=2097152
14/09/14 20:47:11 INFO namenode.FSNamesystem: fsOwner=arunkumar
14/09/14 20:47:11 INFO namenode.FSNamesystem: supergroup=supergroup
14/09/14 20:47:11 INFO namenode.FSNamesystem: isPermissionEnabled=true
14/09/14 20:47:11 INFO namenode.FSNamesystem: dfs.block.invalidate.limit=100
14/09/14 20:47:11 INFO namenode.FSNamesystem: isAccessTokenEnabled=false accessK
eyUpdateInterval=0 min(s), accessTokenLifetime=0 min(s)
14/09/14 20:47:11 INFO namenode.FSEditLog: dfs.namenode.edits.toleration.length
= 0
14/09/14 20:47:11 INFO namenode.NameNode: Caching file names occuring more than
10 times
14/09/14 20:47:11 INFO common.Storage: Image file of size 115 saved in 0 seconds
.
14/09/14 20:47:11 INFO namenode.FSEditLog: closing edit log: position=4, editlog
=/usr/local/hadoop/tmp/dfs/name/current/edits
14/09/14 20:47:11 INFO namenode.FSEditLog: close success: truncate to 4, editlog
=/usr/local/hadoop/tmp/dfs/name/current/edits
14/09/14 20:47:11 INFO common.Storage: Storage directory /usr/local/hadoop/tmp/d
fs/name has been successfully formatted.
14/09/14 20:47:11 INFO namenode.NameNode: SHUTDOWN_MSG:
/*****
SHUTDOWN_MSG: Shutting down NameNode at hadoop/127.0.1.1
*****/
arunkumar@hadoop:~$

```

27. start the hadoop

start-all.sh


```
arunkumar@hadoop: ~  
arunkumar@hadoop:~$ start-all.sh  
starting namenode, logging to /usr/local/hadoop/libexec/../logs/hadoop-arunkumar-namenode-hadoop.out  
localhost: starting datanode, logging to /usr/local/hadoop/libexec/../logs/hadoop-arunkumar-datanode-hadoop.out  
localhost: starting secondarynamenode, logging to /usr/local/hadoop/libexec/../logs/hadoop-arunkumar-secondarynamenode-hadoop.out  
starting jobtracker, logging to /usr/local/hadoop/libexec/../logs/hadoop-arunkumar-jobtracker-hadoop.out  
localhost: starting tasktracker, logging to /usr/local/hadoop/libexec/../logs/hadoop-arunkumar-tasktracker-hadoop.out  
arunkumar@hadoop:~$ jps  
12896 SecondaryNameNode  
13138 TaskTracker  
13263 Jps  
12981 JobTracker  
12753 DataNode  
12616 NameNode  
arunkumar@hadoop:~$ clear
```

28. Check whether all the daemons are running correctly!

jps

```
arunkumar@hadoop: ~  
arunkumar@hadoop:~$ jps  
12896 SecondaryNameNode  
13138 TaskTracker  
13295 Jps  
12981 JobTracker  
12753 DataNode  
12616 NameNode  
arunkumar@hadoop:~$
```

29. Stopping all the daemons

stop-all.sh

```
arunkumar@hadoop: ~  
arunkumar@hadoop:~$ jps  
12896 SecondaryNameNode  
13138 TaskTracker  
13295 Jps  
12981 JobTracker  
12753 DataNode  
12616 NameNode  
arunkumar@hadoop:~$ stop-all.sh  
stopping jobtracker  
localhost: stopping tasktracker  
stopping namenode  
localhost: stopping datanode  
localhost: stopping secondarynamenode  
arunkumar@hadoop:~$
```

30. Now let see the web UI. Start all the daemons - start-all.sh

Namenode : <http://hadoop:50070/dfshealth.jsp>

Hadoop NameNode h... x

hadoop:50070/dfshealth.jsp

Apps Log On How to Get Free M Amazon Student W List of United Stat MyChart - Home New Tab DMV behind-the-v Other bookmarks

Started: Sun Sep 14 20:55:49 PDT 2014
Version: 1.2.0, r1479473
Compiled: Mon May 6 06:59:37 UTC 2013 by hortonfo
Upgrades: There are no upgrades in progress.

[Browse the filesystem](#)
[Namenode Logs](#)

Cluster Summary

8 files and directories, 2 blocks = 10 total. Heap Size is 57 MB / 889 MB (6%)

| | | |
|---------------------------------------|---|-----------|
| Configured Capacity | : | 7.97 GB |
| DFS Used | : | 40 KB |
| Non DFS Used | : | 7.4 GB |
| DFS Remaining | : | 586.78 MB |
| DFS Used% | : | 0 % |
| DFS Remaining% | : | 7.19 % |
| Live Nodes | : | 1 |
| Dead Nodes | : | 0 |
| Decommissioning Nodes | : | 0 |
| Number of Under-Replicated Blocks | : | 1 |

NameNode Storage:

| Storage Directory | Type | State |
|--------------------------------|-----------------|--------|
| /usr/local/hadoop/tmp/dfs/name | IMAGE_AND_EDITS | Active |

hadoop-2.5.1-....tar.gz Show all downloads...

31. UI for jobtracker – map reduce

<http://hadoop:50030/jobtracker.jsp>

Hadoop NameNode h...
hadoop Hadoop Map/...

hadoop:50030/jobtracker.jsp

Apps
Log On
How to Get Free M...
Amazon Student
W List of United Sta...
MyChart - Home
New Tab
DMV behind-the-v...
Other bookmarks

Quick Links

hadoop Hadoop Map/Reduce Administration

State: RUNNING
Started: Sun Sep 14 20:55:56 PDT 2014
Version: 1.2.0, r1479473
Compiled: Mon May 6 06:59:37 UTC 2013 by hortonfo
Identifier: 201409142055
SafeMode: OFF

Cluster Summary (Heap Size is 56.5 MB/889 MB)

| Running Map Tasks | Running Reduce Tasks | Total Submissions | Nodes | Occupied Map Slots | Occupied Reduce Slots | Reserved Map Slots | Reserved Reduce Slots | Map Task Capacity | Reduce Task Capacity | Avg. Tasks/Node | Blacklisted Nodes | Graylisted Nodes | Excluded Nodes |
|-------------------|----------------------|-------------------|-------------------|--------------------|-----------------------|--------------------|-----------------------|-------------------|----------------------|-----------------|-------------------|-------------------|-------------------|
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 2 | 4.00 | 0 | 0 | 0 |

Scheduling Information

| Queue Name | State | Scheduling Information |
|-------------------------|---------|------------------------|
| default | running | N/A |

Filter (Jobid, Priority, User, Name)
Example: 'user:smith 3200' will filter by 'smith' only in the user field and '3200' in all fields

Running Jobs

hadoop-2.5.1-....tar.gz

Show all downloads...

32. UI for tasktracker

<http://hadoop:50060/tasktracker.jsp>

tracker_hadoop:localhost/127.0.0.1:35782 Task Tracker Status



Version: 1.2.0, r1479473

Compiled: Mon May 6 06:59:37 UTC 2013 by hortonfo

Running tasks

| | | | |
|---------------|--------|----------|--------|
| Task Attempts | Status | Progress | Errors |
|---------------|--------|----------|--------|

Non-Running Tasks

| | |
|---------------|--------|
| Task Attempts | Status |
|---------------|--------|

Tasks from Running Jobs

| | | | |
|---------------|--------|----------|--------|
| Task Attempts | Status | Progress | Errors |
|---------------|--------|----------|--------|

Local Logs

[Log](#) directory

This is [Apache Hadoop](#) release 1.2.0

hadoop-2.5.1-....tar.gz

Show all downloads... x