

## Contents

Environment .....	2
Install Hosts.....	2
Install JDK.....	5
Set non - password login.....	8
Hadoop Settings.....	11
Set Hadoop configuration files .....	11
Test Hadoop.....	20

## Environment

1. Under Cloud OS
2. 3 VM Ubuntu 16.04.1 GUI
3. 1 master and 2 slaves

## Install Hosts

Create 3 virtual machines. As we need 1 master and 2 slaves, please name the computers as “master-id”, “slave1-id” and “slave2-id”, such as “master-123456”, “slave1-123456” and “slave2-123456”. The image is “Ubuntu 16.04-GUI”.

H3C CloudOS

Dashboard Cloud Services Cloud Management Cloud O&M

Cloud Services | Compute & Storage | VMs

**Basic Information**

VM Name \* master-123456

VM Alias Please enter VM alias

VMs \* 1 You can create up to 99 VMs in one operation.

Tenancy \* Permanent Custom

Description

Authentication Policy Generate Manual Key Authentication Image Default Password

Specify VM Version version-8

Scripts

[More Settings](#)

Create Cancel

Owner 123

Resource Zone uic

Physical Host

Image Ubuntu16.04-GUI

Specifications 2 Cores , 4 GB , Comm...

Storage System Disk ( 40 GB )

Data Disk

Private Network yewu(172.31.12.0/22)

Security Group default

VM Name master-123456

VM Alias

VMs 1

Tenancy Permanent

Authentication Generate Password

Policy

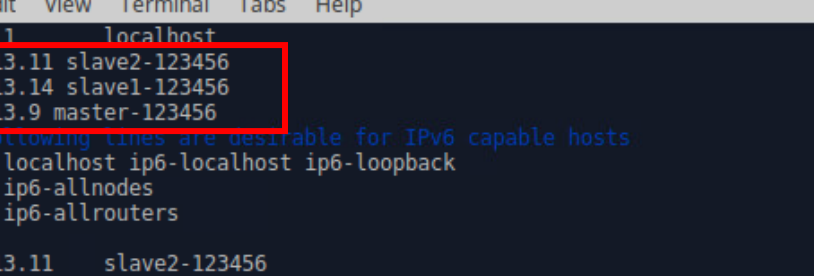
Quantity 0.00% Used/All 0/4pcs

CPUs 0.00% Used/All 0/8Cores

Memory 0.00% Used/All 0/24GB



**Remember** this operation for **vim** or **vi** editing. “w” means “write”, “q” means “quit”, “!” means “obligatory”. If you just want to read and exit, just input “:q” then press “Enter”.

[illegible]

The screenshot shows a terminal window titled "Terminal - uic@slave2-123456: ~". The terminal displays the content of the `/etc/hosts` file. The first four lines are highlighted with a red box:

```
127.0.0.1    localhost
172.31.13.11 slave2-123456
172.31.13.14 slave1-123456
172.31.13.9  master-123456
```

Below these lines, there is a comment in blue text: "# The following lines are desirable for IPv6 capable hosts". This is followed by several IPv6 entries:

```
:::1        localhost ip6-localhost ip6-loopback
ff02::1     ip6-allnodes
ff02::2     ip6-allrouters
```

At the bottom of the terminal, the command `172.31.13.11 slave2-123456` is entered, and the prompt `~` is visible. The status bar at the bottom right shows "1,1" and "All".

## Install JDK

### 1. Install JDK

\$sudo apt install openjdk-8-jdk

```
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  ca-certificates-java fonts-dejavu-extra java-common libgif7 libice-dev
  libpthread-stubs0-dev libsm-dev libx11-dev libx11-doc libxau-dev libxcb1-dev
  libxdmcp-dev libxt-dev openjdk-8-jdk-headless openjdk-8-jre
  openjdk-8-jre-headless x11proto-core-dev x11proto-input-dev x11proto-kb-dev
  xorg-sgml-doctools xtrans-dev
Suggested packages:
  default-jre libice-doc libsm-doc libxcb-doc libxt-doc openjdk-8-demo
  openjdk-8-source visualvm icedtea-8-plugin fonts-ipafont-gothic
  fonts-ipafont-mincho fonts-wqy-microhei fonts-wqy-zenhei fonts-indic
The following NEW packages will be installed:
  ca-certificates-java fonts-dejavu-extra java-common libgif7 libice-dev
  libpthread-stubs0-dev libsm-dev libx11-dev libx11-doc libxau-dev libxcb1-dev
  libxdmcp-dev libxt-dev openjdk-8-jdk openjdk-8-jdk-headless openjdk-8-jre
  openjdk-8-jre-headless x11proto-core-dev x11proto-input-dev x11proto-kb-dev
  xorg-sgml-doctools xtrans-dev
0 upgraded, 22 newly installed, 0 to remove and 0 not upgraded.
Need to get 42.2 MB of archives.
After this operation, 167 MB of additional disk space will be used.
Do you want to continue? [Y/n] y
```

Input "Y" and press "Enter".

```
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/wsgen to provide
 /usr/bin/wsgen (wsgen) in auto mode
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/jcmd to provide
 /usr/bin/jcmd (jcmd) in auto mode
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/jarsigner to pr
 ovide /usr/bin/jarsigner (jarsigner) in auto mode
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/jmap to provide
 /usr/bin/jmap (jmap) in auto mode
Setting up openjdk-8-jdk:amd64 (8u222-b10-1ubuntu1~16.04.1) ...
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/appletviewer to
 provide /usr/bin/appletviewer (appletviewer) in auto mode
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/jconsole to pro
 vide /usr/bin/jconsole (jconsole) in auto mode
Processing triggers for libc-bin (2.23-0ubuntu1) ...
Processing triggers for ca-certificates (20170717~16.04.2) ...
Updating certificates in /etc/ssl/certs...
0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...

done.
done.
uic@master-123456:~$
```

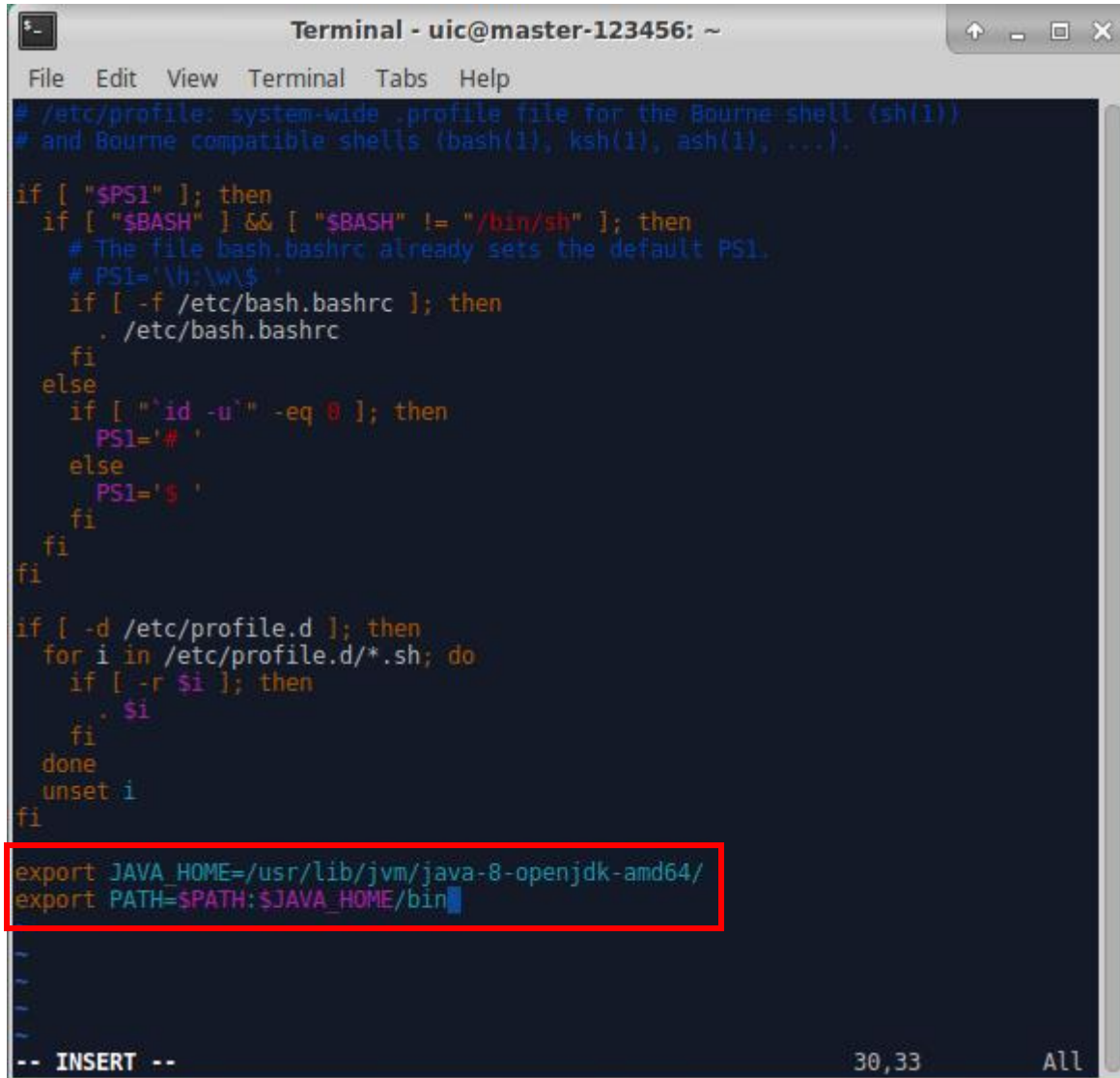
### 2. Modify the system path file

\$ sudo vim /etc/profile

Set the environment variables (operation mentioned above -- Insert, Esc, :wq!, Enter)

```
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64/
```

```
export PATH=$PATH:$JAVA_HOME/bin
```



A terminal window titled "Terminal - uic@master-123456: ~" with a menu bar (File, Edit, View, Terminal, Tabs, Help). The terminal shows the contents of the `/etc/profile` file, which is a system-wide profile file for the Bourne shell. The file contains several conditional blocks for setting the `PS1` prompt and sourcing files from `/etc/profile.d`. At the bottom of the file, the environment variables `JAVA_HOME` and `PATH` are being set. These two lines are highlighted with a red rectangular box. The terminal status bar at the bottom shows "-- INSERT --", "30,33", and "All".

```
# /etc/profile: system-wide .profile file for the Bourne shell (sh(1))
# and Bourne compatible shells (bash(1), ksh(1), ash(1), ...).

if [ "$PS1" ]; then
  if [ "$BASH" ] && [ "$BASH" != "/bin/sh" ]; then
    # The file bash.bashrc already sets the default PS1.
    # PS1='\h:\w\$ '
    if [ -f /etc/bash.bashrc ]; then
      . /etc/bash.bashrc
    fi
  else
    if [ "`id -u`" -eq 0 ]; then
      PS1='# '
    else
      PS1='$ '
    fi
  fi
fi

if [ -d /etc/profile.d ]; then
  for i in /etc/profile.d/*.sh; do
    if [ -r $i ]; then
      . $i
    fi
  done
  unset i
fi

export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64/
export PATH=$PATH:$JAVA_HOME/bin
```

3. Put the environment variable into effect immediately

```
$ source /etc/profile
```

4. Verify

```
$ java -version
```

```
Terminal - uic@master-123456: ~
File Edit View Terminal Tabs Help
/usr/bin/xjc (xjc) in auto mode
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/rmic to provide
/usr/bin/rmic (rmic) in auto mode
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/jstatd to provide
/usr/bin/jstatd (jstatd) in auto mode
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/jdb to provide
/usr/bin/jdb (jdb) in auto mode
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/serialver to provide
/usr/bin/serialver (serialver) in auto mode
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/wsgen to provide
/usr/bin/wsgen (wsgen) in auto mode
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/jcmd to provide
/usr/bin/jcmd (jcmd) in auto mode
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/jarsigner to provide
/usr/bin/jarsigner (jarsigner) in auto mode
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/jmap to provide
/usr/bin/jmap (jmap) in auto mode
Setting up openjdk-8-jdk:amd64 (8u222-b10-1ubuntu1~16.04.1) ...
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/appletviewer to provide
/usr/bin/appletviewer (appletviewer) in auto mode
update-alternatives: using /usr/lib/jvm/java-8-openjdk-amd64/bin/jconsole to provide
/usr/bin/jconsole (jconsole) in auto mode
Processing triggers for libc-bin (2.23-0ubuntu11) ...
Processing triggers for ca-certificates (20170717~16.04.2) ...
Updating certificates in /etc/ssl/certs...
0 added, 0 removed; done.
Running hooks in /etc/ca-certificates/update.d...

done.
done.
uic@master-123456:~$ sudo vim /etc/profile
uic@master-123456:~$ source /etc/profile
uic@master-123456:~$ java -version
openjdk version "1.8.0_222"
OpenJDK Runtime Environment (build 1.8.0_222-8u222-b10-1ubuntu1~16.04.1-b10)
OpenJDK 64-Bit Server VM (build 25.222-b10, mixed mode)
uic@master-123456:~$
```

It is okay.

Then install JDK on the two slave hosts.

## Set non - password login

# Why and how to set ssh non-password login:

<http://www.linuxidc.com/Linux/2015-03/114709.htm>

1. Generate the public key of SSH.

```
$ ssh-keygen -t rsa -P '' -f ~/.ssh/id_rsa
```

```
uic@master-123456:~$ ssh-keygen -t rsa -P '' -f ~/.ssh/id_rsa
Generating public/private rsa key pair.
Created directory '/home/uic/.ssh'.
Your identification has been saved in /home/uic/.ssh/id_rsa.
Your public key has been saved in /home/uic/.ssh/id_rsa.pub.
The key fingerprint is:
SHA256:LtYk99YWcu6LIkkK6eTw0LhVwBPoAJ0cPrWwIRmhRfU uic@master-123456
The key's randomart image is:
+---[RSA 2048]---+
|*OB+o          |
|==+* o         |
|+ B . E        |
|. +           |
|  . . S . o    |
|  . . * . = .  |
|0+.0  o o o +  |
|0+. . . . . +  |
|+=. . . . . o.  |
+---[SHA256]-----+
uic@master-123456:~$
```

2. Copy and send files

```
$ ssh-copy-id slave1-123456
```

```
uic@master-123456:~$ ssh-copy-id slave1-123456
The authenticity of host 'slave1-123456 (172.31.13.14)' can't be established.
ECDSA key fingerprint is SHA256:dIBK3tkau8bIQmfEdY8Fme541jFxDwWezlywuIRftco.
Are you sure you want to continue connecting (yes/no)? yes
/usr/bin/ssh-copy-id: INFO: attempting to log in with the new key(s), to filter
out any that are already installed
/usr/bin/ssh-copy-id: INFO: 1 key(s) remain to be installed -- if you are prompt
ed now it is to install the new keys
uic@slave1-123456's password:

Number of key(s) added: 1

Now try logging into the machine, with:  "ssh 'slave1-123456'"
and check to make sure that only the key(s) you wanted were added.

uic@master-123456:~$
```



\$ ssh-copy-id slave2-123456

```
uic@master-123456:~$ ssh-copy-id slave2-123456
The authenticity of host 'slave2-123456 (172.31.13.11)' can't be established.
ECDSA key fingerprint is SHA256:dIBK3tkau8bIQmfEdY8Fme54ljFxDwWezlywuIRftco.
Are you sure you want to continue connecting (yes/no)? yes
/usr/bin/ssh-copy-id: INFO: attempting to log in with the new key(s), to filter
out any that are already installed
/usr/bin/ssh-copy-id: INFO: 1 key(s) remain to be installed -- if you are prompt
ed now it is to install the new keys
uic@slave2-123456's password:

Number of key(s) added: 1

Now try logging into the machine, with:  "ssh 'slave2-123456'"
and check to make sure that only the key(s) you wanted were added.

uic@master-123456:~$
```

Do not forget connect the master self

\$ ssh-copy-id localhost

```
The authenticity of host 'localhost (::1)' can't be established.
ECDSA key fingerprint is SHA256:dIBK3tkau8bIQmfEdY8Fme54ljFxDwWezlywuIRftco.
Are you sure you want to continue connecting (yes/no)? yes
/usr/bin/ssh-copy-id: INFO: attempting to log in with the new key(s), to filter
out any that are already installed
/usr/bin/ssh-copy-id: INFO: 1 key(s) remain to be installed -- if you are prompt
ed now it is to install the new keys
uic@localhost's password:

Number of key(s) added: 1

Now try logging into the machine, with:  "ssh 'localhost'"
and check to make sure that only the key(s) you wanted were added.
```

Then do the same operations **on the two slave hosts**.

On slave1-123456:

\$ ssh-keygen -t rsa -P "" -f ~/.ssh/id\_rsa

\$ ssh-copy-id master-123456

\$ ssh-copy-id slave2-123456

On slave2-123456:

\$ ssh-keygen -t rsa -P "" -f ~/.ssh/id\_rsa

\$ ssh-copy-id master-123456

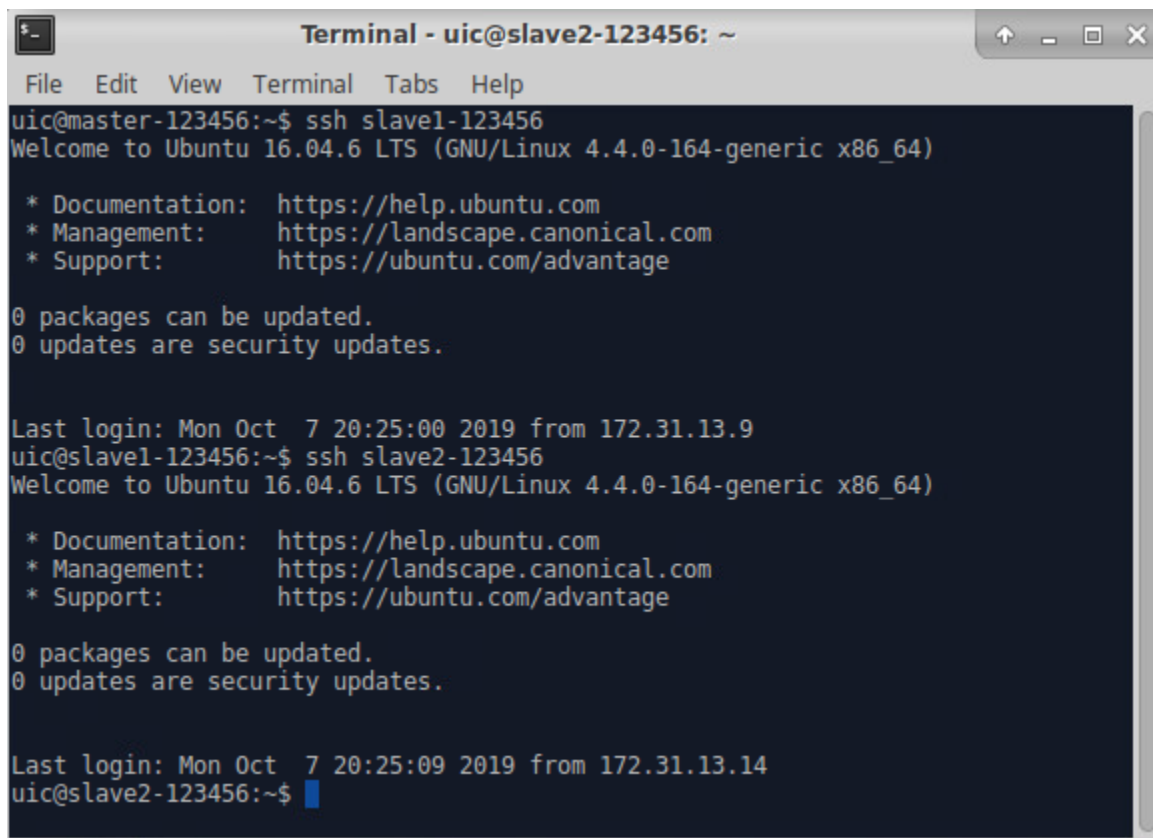
\$ ssh-copy-id slave1-123456

3. Test the ssh connection from master to slaves.

On master:

```
$ ssh slave1-123456
```

```
$ ssh slave2-123456
```

A terminal window titled "Terminal - uic@slave2-123456: ~" with a menu bar (File, Edit, View, Terminal, Tabs, Help) and window controls. The terminal shows an SSH session from master-123456 to slave2-123456. The output includes the Ubuntu 16.04.6 LTS welcome message, links for documentation, management, and support, and package update information. The session ends with the last login timestamp and the user prompt.

```
uic@master-123456:~$ ssh slave1-123456
Welcome to Ubuntu 16.04.6 LTS (GNU/Linux 4.4.0-164-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:       https://ubuntu.com/advantage

0 packages can be updated.
0 updates are security updates.

Last login: Mon Oct  7 20:25:00 2019 from 172.31.13.9
uic@slave1-123456:~$ ssh slave2-123456
Welcome to Ubuntu 16.04.6 LTS (GNU/Linux 4.4.0-164-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:    https://landscape.canonical.com
 * Support:       https://ubuntu.com/advantage

0 packages can be updated.
0 updates are security updates.

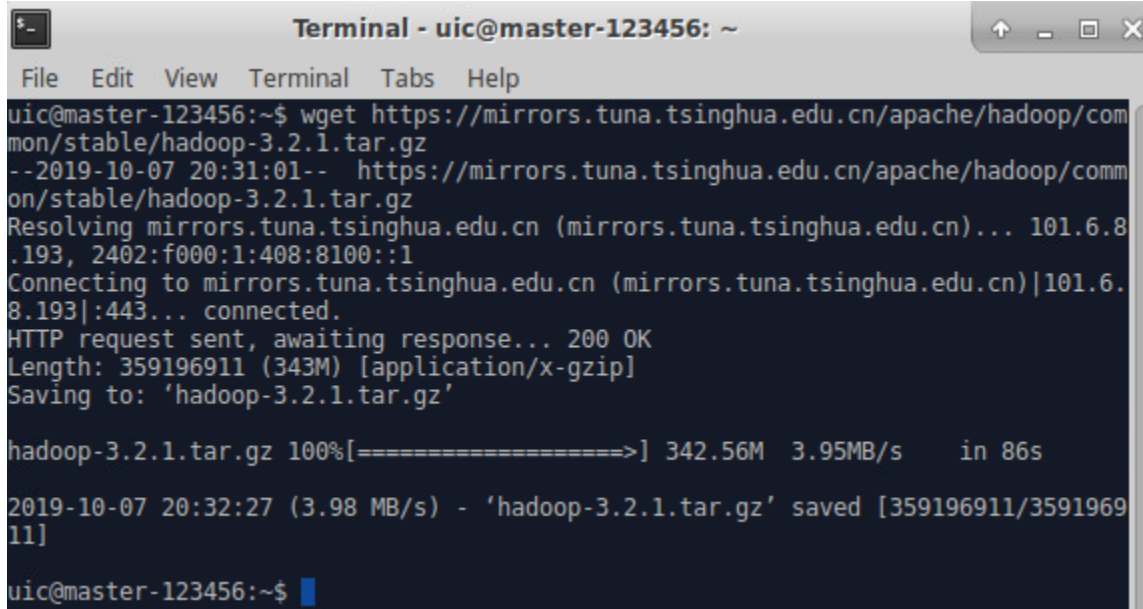
Last login: Mon Oct  7 20:25:09 2019 from 172.31.13.14
uic@slave2-123456:~$
```

It is okay.

## Hadoop Settings

### 1. Download

\$ wget https://mirrors.tuna.tsinghua.edu.cn/apache/hadoop/common/stable/hadoop-3.2.1.tar.gz

A terminal window titled "Terminal - uic@master-123456: ~" showing the execution of the wget command to download the Hadoop 3.2.1 tar.gz file from a mirror. The output shows the file size (343M), the download speed (3.95MB/s), and the time taken (86s).

```
uic@master-123456:~$ wget https://mirrors.tuna.tsinghua.edu.cn/apache/hadoop/common/stable/hadoop-3.2.1.tar.gz
--2019-10-07 20:31:01-- https://mirrors.tuna.tsinghua.edu.cn/apache/hadoop/common/stable/hadoop-3.2.1.tar.gz
Resolving mirrors.tuna.tsinghua.edu.cn (mirrors.tuna.tsinghua.edu.cn)... 101.6.8.193, 2402:f000:1:408:8100::1
Connecting to mirrors.tuna.tsinghua.edu.cn (mirrors.tuna.tsinghua.edu.cn)|101.6.8.193|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 359196911 (343M) [application/x-gzip]
Saving to: 'hadoop-3.2.1.tar.gz'

hadoop-3.2.1.tar.gz 100%[=====>] 342.56M  3.95MB/s   in 86s

2019-10-07 20:32:27 (3.98 MB/s) - 'hadoop-3.2.1.tar.gz' saved [359196911/359196911]

uic@master-123456:~$
```

### 2. Unzip the file

We put it in the default directory (/home/uic/hadoop-3.2.1)

\$ tar -zxvf hadoop-3.2.1.tar.gz

## Set Hadoop configuration files

Set the environment variables

\$ cd ~/hadoop-3.2.1/etc/hadoop

\$ sudo vim /etc/profile

export HADOOP\_HOME=/home/uic/hadoop-3.2.1

export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64/

export PATH=\$PATH:\$JAVA\_HOME/bin:\$HADOOP\_HOME/bin:\$HADOOP\_HOME/sbin

```
Terminal - uic@master-123456: ~/hadoop-3.2.1/etc/hadoop
File Edit View Terminal Tabs Help

# /etc/profile: system-wide .profile file for the Bourne shell (sh(1))
# and Bourne compatible shells (bash(1), ksh(1), ash(1), ...).

if [ "$PS1" ]; then
  if [ "$BASH" ] && [ "$BASH" != "/bin/sh" ]; then
    # The file bash.bashrc already sets the default PS1.
    # PS1='\h:\w\$ '
    if [ -f /etc/bash.bashrc ]; then
      . /etc/bash.bashrc
    fi
  else
    if [ "`id -u`" -eq 0 ]; then
      PS1='# '
    else
      PS1='$ '
    fi
  fi
fi

if [ -d /etc/profile.d ]; then
  for i in /etc/profile.d/*.sh; do
    if [ -r $i ]; then
      . $i
    fi
  done
  unset i
fi

export HADOOP_HOME=/home/uic/hadoop-3.2.1
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64/
export PATH=$PATH:$JAVA_HOME/bin:$HADOOP_HOME/bin:$HADOOP_HOME/sbin

-- INSERT --
```

1. Configure [hadoop-env.sh](#) and [yarn-env.sh](#)

```
$ cd ~/hadoop-3.2.1/etc/hadoop
```

```
$ sudo vim hadoop-env.sh
```

```
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
```

```
Terminal - uic@master-123456: ~/hadoop-3.2.1/etc/hadoop
File Edit View Terminal Tabs Help

# Licensed to the Apache Software Foundation (ASF) under one
# or more contributor license agreements. See the NOTICE file
# distributed with this work for additional information
# regarding copyright ownership. The ASF licenses this file
# to you under the Apache License, Version 2.0 (the
# "License"); you may not use this file except in compliance
# with the License. You may obtain a copy of the License at
#
# http://www.apache.org/licenses/LICENSE-2.0
#
# Unless required by applicable law or agreed to in writing, software
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the License.

# Set Hadoop-specific environment variables here.

##
## THIS FILE ACTS AS THE MASTER FILE FOR ALL HADOOP PROJECTS.
## SETTINGS HERE WILL BE READ BY ALL HADOOP COMMANDS. THEREFORE,
## ONE CAN USE THIS FILE TO SET YARN, HDFS, AND MAPREDUCE
## CONFIGURATION OPTIONS INSTEAD OF xxx-env.sh.
##
## Precedence rules:
##
## {yarn-env.sh|hdfs-env.sh} > hadoop-env.sh > hard-coded defaults
##
## {YARN_xyz|HDFS_xyz} > HADOOP_xyz > hard-coded defaults
##
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
# Many of the options here are built from the perspective that users
# may want to provide OVERWRITING values on the command line.
# For example:
#
"hadoop-env.sh" 417L, 16285C 5,1 Top
```

\$ sudo vim yarn-env.sh

export JAVA\_HOME=/usr/lib/jvm/java-8-openjdk-amd64

```
# Supplemental options for privileged registry DNS
# By default, Hadoop uses jsvc which needs to know to launch a
# server jvm.
# export YARN_REGISTRYDNS_SECURE_EXTRA_OPTS="-jvm server"
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
###
# YARN Services parameters
###
# Directory containing service examples
# export YARN_SERVICE_EXAMPLES_DIR = $HADOOP_YARN_HOME/$share/hadoop/yarn/yarn-service-examples
```

Note: we need create the following folders for the following configuration:

~/hadoop-3.2.1/tmp

~/hadoop-3.2.1/hdfs/name

~/hadoop-3.2.1/hdfs/data

\$ cd ~/hadoop-3.2.1

\$ mkdir tmp

\$ mkdir -p hdfs

2. core-site.xml

\$ cd ~/hadoop-3.2.1/etc/hadoop

\$ sudo vim core-site.xml

<configuration>

<property>

<name>fs.defaultFS</name>

<value>hdfs://master-123456:9000</value>

</property>

<property>

<name>hadoop.tmp.dir</name>

<value>/home/uic/hadoop-3.2.1/tmp</value>

<description>Abase for other temporary directories.</description>

</property>

</configuration>

```
Terminal - uic@master-123456: ~/hadoop-3.2.1/etc/hadoop
File Edit View Terminal Tabs Help
<!--
Licensed under the Apache License, Version 2.0 (the "License");
you may not use this file except in compliance with the License.
You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->
<!-- Put site-specific property overrides in this file. -->
<configuration>
  <property>
    <name>fs.defaultFS</name>
    <value>hdfs://master-123456:9000</value>
  </property>
  <property>
    <name>hadoop.tmp.dir</name>
    <value>/home/uic/hadoop-3.2.1/tmp</value>
    <description>Abase for other temporary directories.</description>
  </property>
</configuration>
-
```

### 3. mapred-site.xml

```
$ sudo vim mapred-site.xml
```

```
<configuration>

  <property>

    <name>mapreduce.framework.name</name>

    <value>yarn</value>

  </property>

  <property>

    <name>mapred.job.tracker</name>

    <value>master-123456:49001</value>

  </property>

  <property>
```

```
<name>mapred.job.dir</name>
<value>/home/uic/hadoop-3.2.1/var</value>
</property>
<property>
  <name>yarn.app.mapreduce.am.env</name>
  <value>HADOOP_MAPRED_HOME=/home/uic/hadoop-3.2.1</value>
</property>
<property>
  <name>mapreduce.map.env</name>
  <value>HADOOP_MAPRED_HOME=/home/uic/hadoop-3.2.1</value>
</property>
<property>
  <name>mapreduce.reduce.env</name>
  <value>HADOOP_MAPRED_HOME=/home/uic/hadoop-3.2.1</value>
</property>
</configuration>
```



```

# Add site specific property overrides in this file.
<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
  <property>
    <name>mapred.job.tracker</name>
    <value>master-123456:49001</value>
  </property>
  <property>
    <name>mapred.job.dir</name>
    <value>/home/uic/hadoop-3.2.1/var</value>
  </property>

  <property>
    <name>yarn.app.mapreduce.am.env</name>
    <value>HADOOP_MAPRED_HOME=/home/uic/hadoop-3.2.1</value>
  </property>

  <property>
    <name>mapreduce.map.env</name>
    <value>HADOOP_MAPRED_HOME=/home/uic/hadoop-3.2.1</value>
  </property>

  <property>
    <name>mapreduce.reduce.env</name>
    <value>HADOOP_MAPRED_HOME=/home/uic/hadoop-3.2.1</value>
  </property>
</configuration>

```

#### 4. hdfs-site.xml

\$ sudo vim hdfs-site.xml

```

<configuration>

  <property>

    <name>dfs.replication</name>

    <value>2</value>

  </property>

  <property>

    <name>dfs.namenode.name.dir</name>

    <value>file:/home/uic/hadoop-3.2.1/hdfs/name</value>

  </property>

  <property>

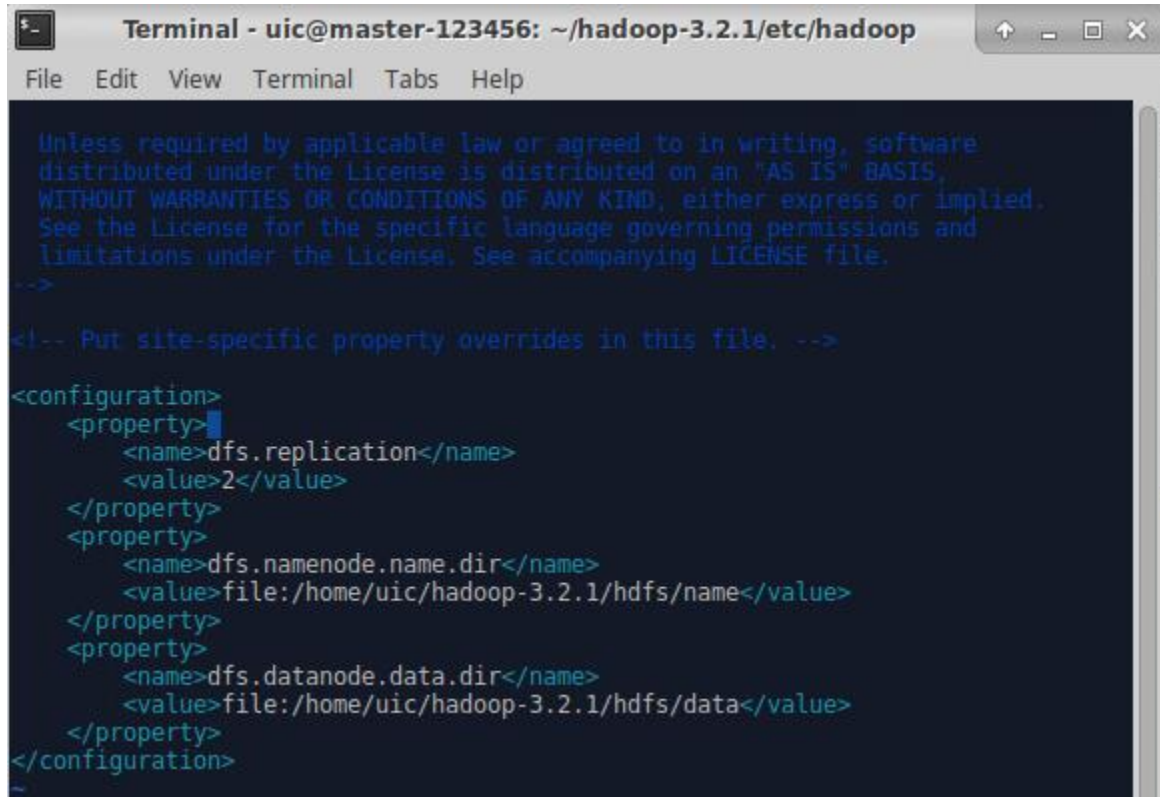
    <name>dfs.datanode.data.dir</name>

```

```
<value>file:/home/uic/hadoop-3.2.1/hdfs/data</value>
```

```
</property>
```

```
</configuration>
```



```
Terminal - uic@master-123456: ~/hadoop-3.2.1/etc/hadoop
File Edit View Terminal Tabs Help

Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->

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

<configuration>
  <property>
    <name>dfs.replication</name>
    <value>2</value>
  </property>
  <property>
    <name>dfs.namenode.name.dir</name>
    <value>file:/home/uic/hadoop-3.2.1/hdfs/name</value>
  </property>
  <property>
    <name>dfs.datanode.data.dir</name>
    <value>file:/home/uic/hadoop-3.2.1/hdfs/data</value>
  </property>
</configuration>
```

## 5. yarn-site.xml

```
$ sudo vim yarn-site.xml
```

```
<configuration>
```

```
<property>
```

```
<name>yarn.resourcemanager.hostname</name>
```

```
<value>master-123456</value>
```

```
</property>
```

```
<property>
```

```
<name>yarn.nodemanager.aux-services</name>
```

```
<value>mapreduce_shuffle</value>
```

```
</property>
```

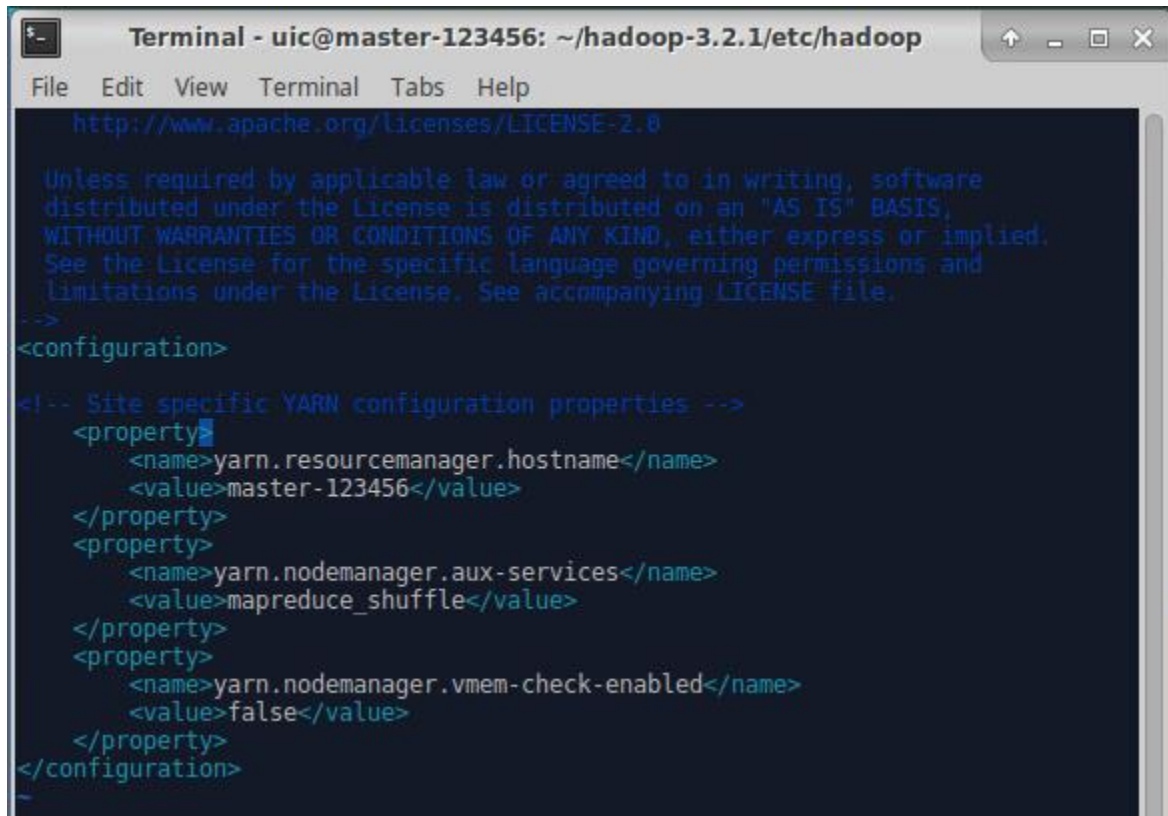
```
<property>

  <name>yarn.nodemanager.vmem-check-enabled</name>

  <value>>false</value>

</property>

</configuration>
```

A terminal window titled "Terminal - uic@master-123456: ~/hadoop-3.2.1/etc/hadoop" with a menu bar (File, Edit, View, Terminal, Tabs, Help). The terminal content shows the Apache License 2.0 text, followed by XML configuration tags for YARN properties. The cursor is positioned at the end of the configuration block.

```
http://www.apache.org/licenses/LICENSE-2.0

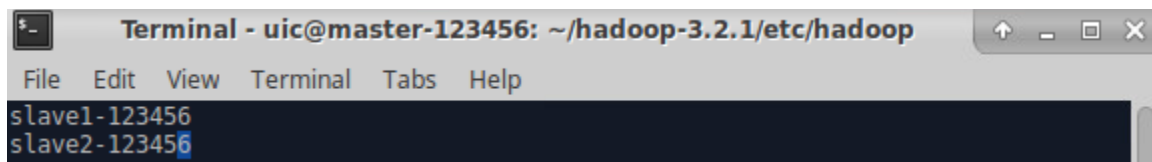
Unless required by applicable law or agreed to in writing, software
distributed under the License is distributed on an "AS IS" BASIS,
WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
See the License for the specific language governing permissions and
limitations under the License. See accompanying LICENSE file.
-->
<configuration>
<!-- Site specific YARN configuration properties -->
  <property>
    <name>yarn.resourcemanager.hostname</name>
    <value>master-123456</value>
  </property>
  <property>
    <name>yarn.nodemanager.aux-services</name>
    <value>mapreduce_shuffle</value>
  </property>
  <property>
    <name>yarn.nodemanager.vmem-check-enabled</name>
    <value>>false</value>
  </property>
</configuration>
~
```

6. Set slave nodesw

```
$sudo vim workers
```

```
slave1-123456
```

```
slave2-123456
```

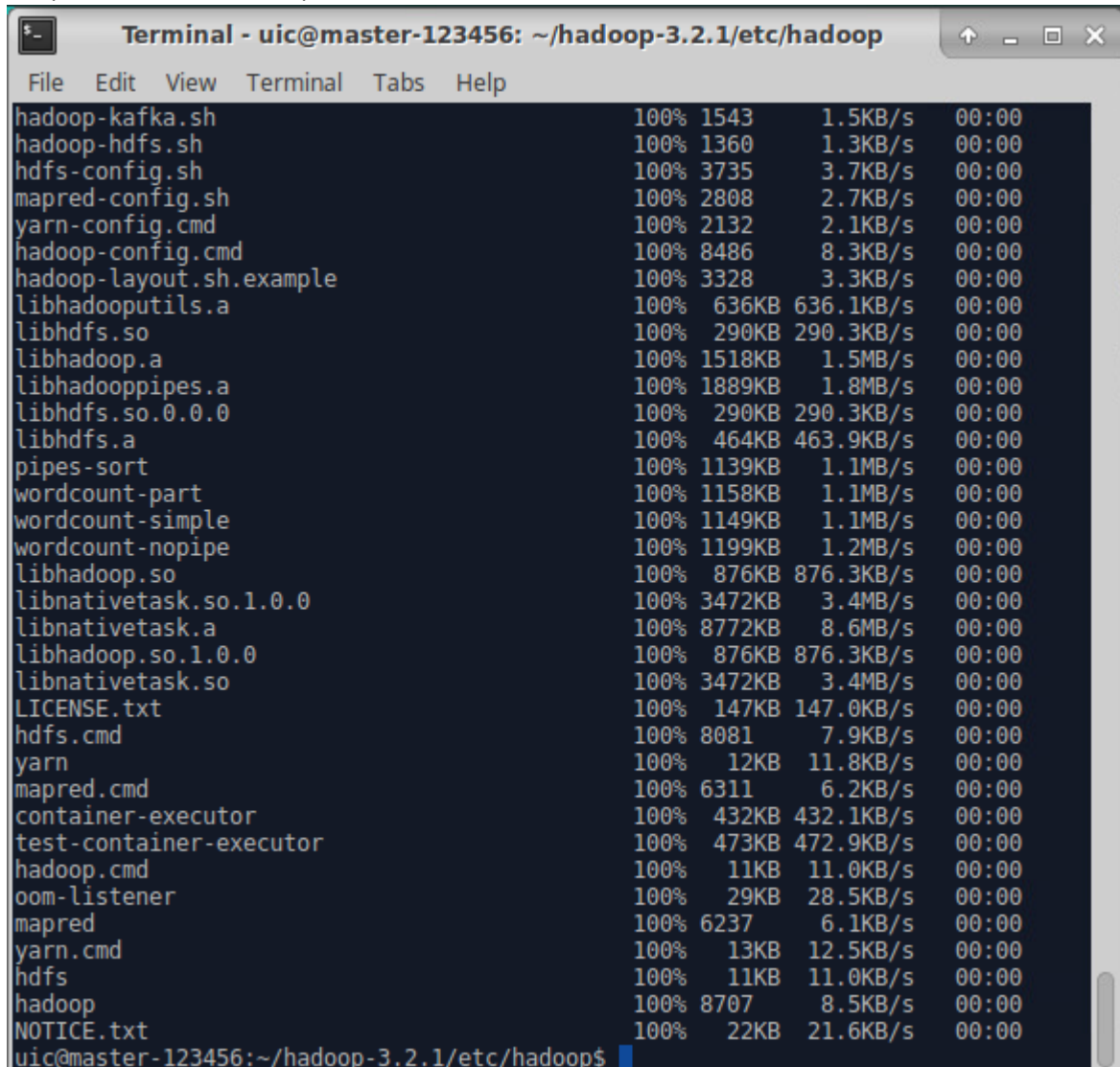
A terminal window titled "Terminal - uic@master-123456: ~/hadoop-3.2.1/etc/hadoop" with a menu bar (File, Edit, View, Terminal, Tabs, Help). The terminal content shows two lines of text representing slave node configurations.

```
slave1-123456
slave2-123456
```

## 7. Copy files to slaves

```
$ scp -r /home/uic/hadoop-3.2.1/ uic@slave1-123456:~
```

```
$ scp -r /home/uic/hadoop-3.2.1/ uic@slave2-123456:~
```



A terminal window titled "Terminal - uic@master-123456: ~/hadoop-3.2.1/etc/hadoop" displays the output of the scp command. It shows a list of files being transferred to slave nodes, with columns for progress percentage, file size, transfer speed, and time remaining. The files listed include configuration scripts, binaries, and command-line tools.

File	Progress	Size	Speed	Time
hadoop-kafka.sh	100%	1543	1.5KB/s	00:00
hadoop-hdfs.sh	100%	1360	1.3KB/s	00:00
hdfs-config.sh	100%	3735	3.7KB/s	00:00
mapred-config.sh	100%	2808	2.7KB/s	00:00
yarn-config.cmd	100%	2132	2.1KB/s	00:00
hadoop-config.cmd	100%	8486	8.3KB/s	00:00
hadoop-layout.sh.example	100%	3328	3.3KB/s	00:00
libhadooputils.a	100%	636KB	636.1KB/s	00:00
libhdfs.so	100%	290KB	290.3KB/s	00:00
libhadoop.a	100%	1518KB	1.5MB/s	00:00
libhadooppipes.a	100%	1889KB	1.8MB/s	00:00
libhdfs.so.0.0.0	100%	290KB	290.3KB/s	00:00
libhdfs.a	100%	464KB	463.9KB/s	00:00
pipes-sort	100%	1139KB	1.1MB/s	00:00
wordcount-part	100%	1158KB	1.1MB/s	00:00
wordcount-simple	100%	1149KB	1.1MB/s	00:00
wordcount-nopipe	100%	1199KB	1.2MB/s	00:00
libhadoop.so	100%	876KB	876.3KB/s	00:00
libnativetask.so.1.0.0	100%	3472KB	3.4MB/s	00:00
libnativetask.a	100%	8772KB	8.6MB/s	00:00
libhadoop.so.1.0.0	100%	876KB	876.3KB/s	00:00
libnativetask.so	100%	3472KB	3.4MB/s	00:00
LICENSE.txt	100%	147KB	147.0KB/s	00:00
hdfs.cmd	100%	8081	7.9KB/s	00:00
yarn	100%	12KB	11.8KB/s	00:00
mapred.cmd	100%	6311	6.2KB/s	00:00
container-executor	100%	432KB	432.1KB/s	00:00
test-container-executor	100%	473KB	472.9KB/s	00:00
hadoop.cmd	100%	11KB	11.0KB/s	00:00
oom-listener	100%	29KB	28.5KB/s	00:00
mapred	100%	6237	6.1KB/s	00:00
yarn.cmd	100%	13KB	12.5KB/s	00:00
hdfs	100%	11KB	11.0KB/s	00:00
hadoop	100%	8707	8.5KB/s	00:00
NOTICE.txt	100%	22KB	21.6KB/s	00:00

uic@master-123456:~/hadoop-3.2.1/etc/hadoop\$

## Test Hadoop

### 1. Format distributed file system

```
$ cd ~
```

```
$ hdfs namenode -format
```

```
Terminal - uic@master-123456: ~
File Edit View Terminal Tabs Help
uic@master-123456:~$ hdfs namenode -format
2019-10-08 09:09:43,027 INFO namenode.NameNode: STARTUP_MSG:
/*****
STARTUP_MSG: Starting NameNode
STARTUP_MSG: host = master-123456/172.31.13.9
STARTUP_MSG: args = [-format]
STARTUP_MSG: version = 3.2.1
STARTUP_MSG: classpath = /home/uic/hadoop-3.2.1/etc/hadoop:/home/uic/hadoop-3.2.1/share/hadoop/common/lib/commons-beanutils-1.9.3.jar:/home/uic/hadoop-3.2.1/share/hadoop/common/lib/paranamer-2.3.jar:/home/uic/hadoop-3.2.1/share/hadoop/common/lib/animal-sniffer-annotations-1.17.jar:/home/uic/hadoop-3.2.1/share/hadoop/common/lib/jersey-server-1.19.jar:/home/uic/hadoop-3.2.1/share/hadoop/common/lib/kerb-client-1.0.1.jar:/home/uic/hadoop-3.2.1/share/hadoop/common/lib/htrace-core4-4.1.0-incubating.jar:/home/uic/hadoop-3.2.1/share/hadoop/common/lib/jul-to-slf4j-1.7.25.jar:/home/uic/hadoop-3.2.1/share/hadoop/common/lib/jetty-http-9.3.24.jar
```

2. Start

\$ start-all.sh

```
uic@master-123456:~$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as uic in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [master-123456]
Starting datanodes
Starting secondary namenodes [master-123456]
Starting resourcemanager
Starting nodemanagers
```

3. Verify if it is successful

On master host, there should be:

NameNode, SecondaryNameNode, Jps and ResourceManager.

\$ jps

```
uic@master-123456:~$ jps
2560 ResourceManager
2085 NameNode
2970 Jps
2335 SecondaryNameNode
```

On slave hosts, there should be: Jps, DataNode and NodeManager.

\$ jps

```
uic@slave1-123456:~$ jps
2385 Jps
2137 NodeManager
1996 DataNode
```

It is successful.

#### 4. Stop

\$ stop-all.sh

```
uic@master-123456:~$ stop-all.sh
WARNING: Stopping all Apache Hadoop daemons as uic in 10 seconds.
WARNING: Use CTRL-C to abort.
Stopping namenodes on [master-123456]
Stopping datanodes
Stopping secondary namenodes [master-123456]
Stopping nodemanagers
Stopping resourcemanager
uic@master-123456:~$
```

#### Note:

If perform “hdfs namenode –format” more than one time, you may encounter the problem that DataNode cannot start.

```
uic@slave1-123456:~$ jps
16529 Jps
16362 NodeManager
```

Try the following command: \$ rm -rf /home/uic/hadoop-3.2.1/hdfs/\* and then perform the format operation again.

#### 5. test hdfs commands

\$ hdfs dfs -mkdir -p /user/uic/hadoop/  
[Create the directories under HDFS](#)

\$ hdfs dfs -put ~/Desktop/test1.txt /user/uic/hadoop/test2.txt  
[Upload a file to a specified directory and rename it.](#)

\$ hdfs dfs -cat /user/uic/hadoop/test2.txt  
[View the contents of the in directory on HDFS.](#)

\$ hdfs dfs -ls /user/uic/hadoop/  
[List the files under HDFS.](#)

\$ hdfs dfs -get /user/uic/hadoop/test2.txt ~/Desktop  
[Get the file from HDFS](#)  
[Like put, you can manipulate files and directories.](#)

```
$ hdfs dfs -rm /user/uic/hadoop/test2.txt
```

Delete the specified file from HDFS.

```
$ hdfs dfsadmin -report
```

View basic statistics for HDFS.

## 6. Web interface

Once the Hadoop cluster is up and running check the web-ui of the components as described below:

To check the namenode information: <http://master-123456:9870>

To check the resource manager information: <http://master-123456:8088>