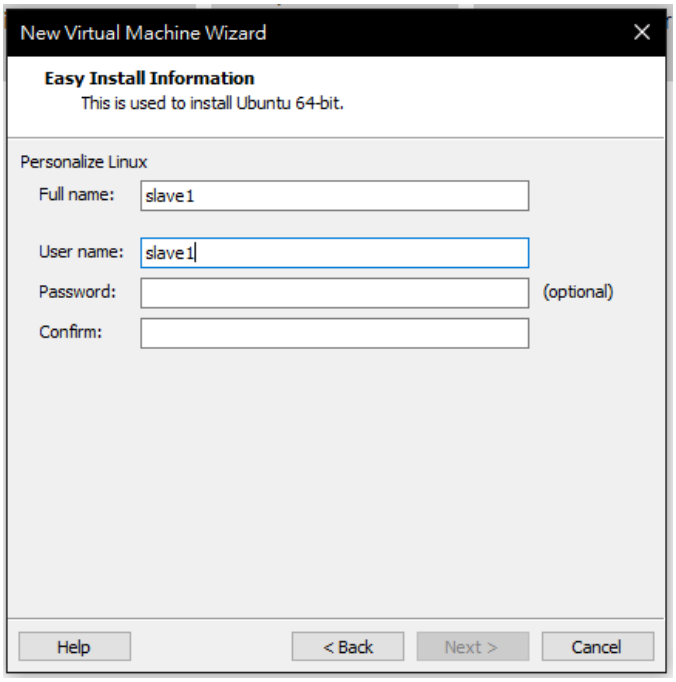
##
#Hadoop 多節點安裝
##

※ 如果 gedit 無法開啟文件，請重開 terminal 再試一次。

新增 slave 主機：(設置如下)

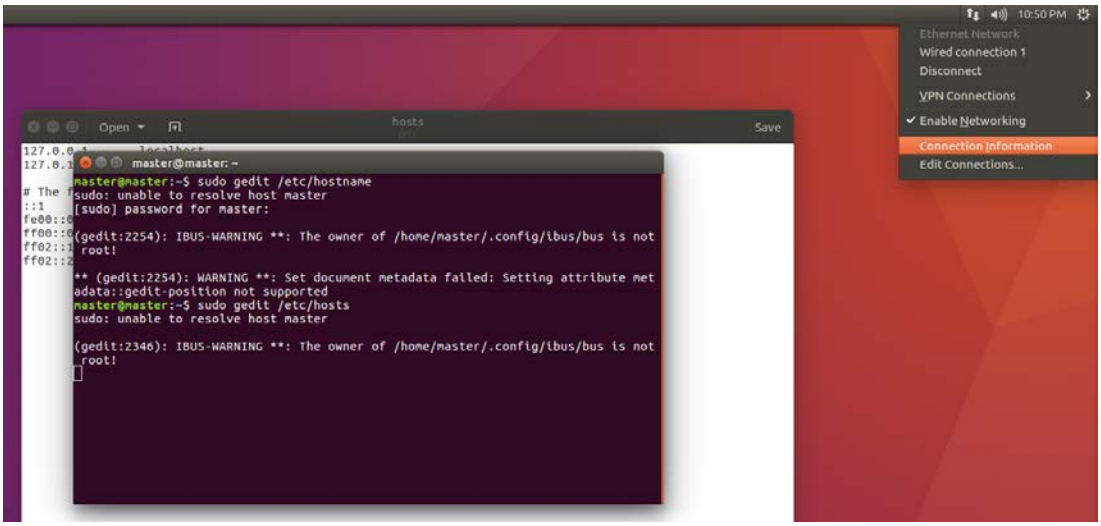


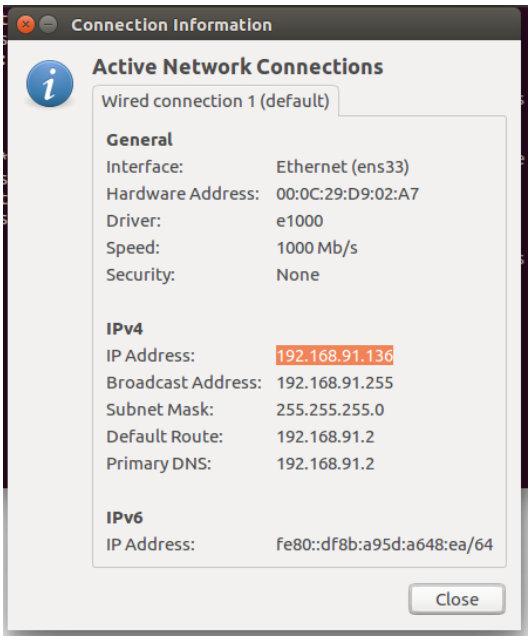
1. 修改所有主機名稱：master 節點設置為 master，slave 節點設置為 slave1。

sudo gedit /etc/hostname

設置完須重開虛擬機

2. 查詢每台主機的 IP 位址：





3. 編輯所有主機之/etc/hosts：

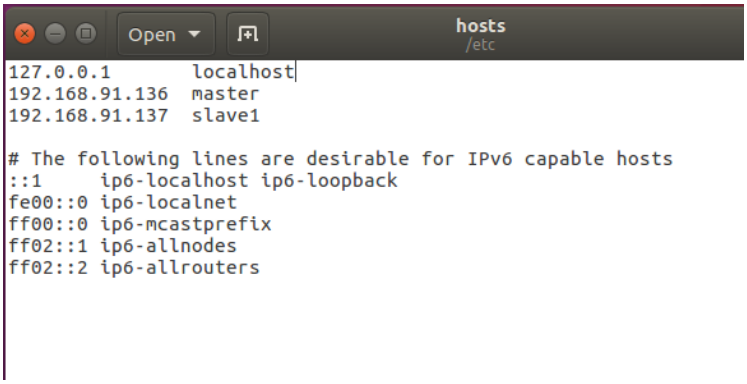
sudo gedit /etc/hosts

將 IP 位址、hostname 對應，寫入/etc/hosts

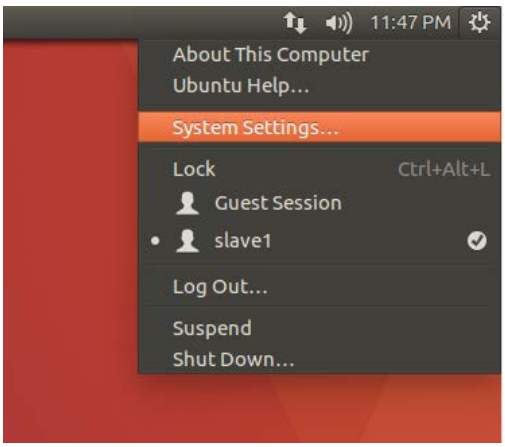
192.168.91.136　　master

192.168.91.137　　slave1

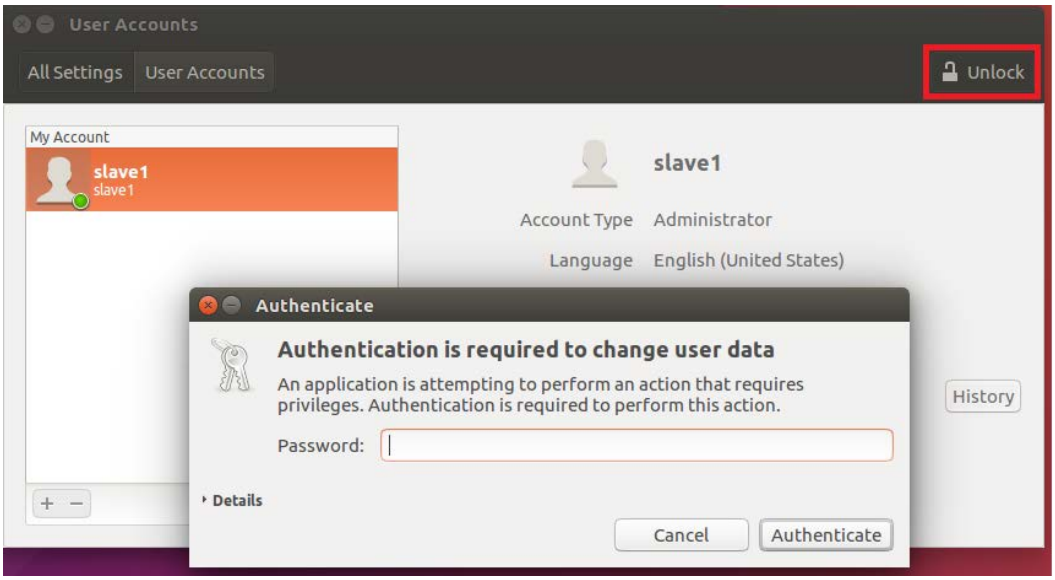
(IP 位址)　　(hostname)



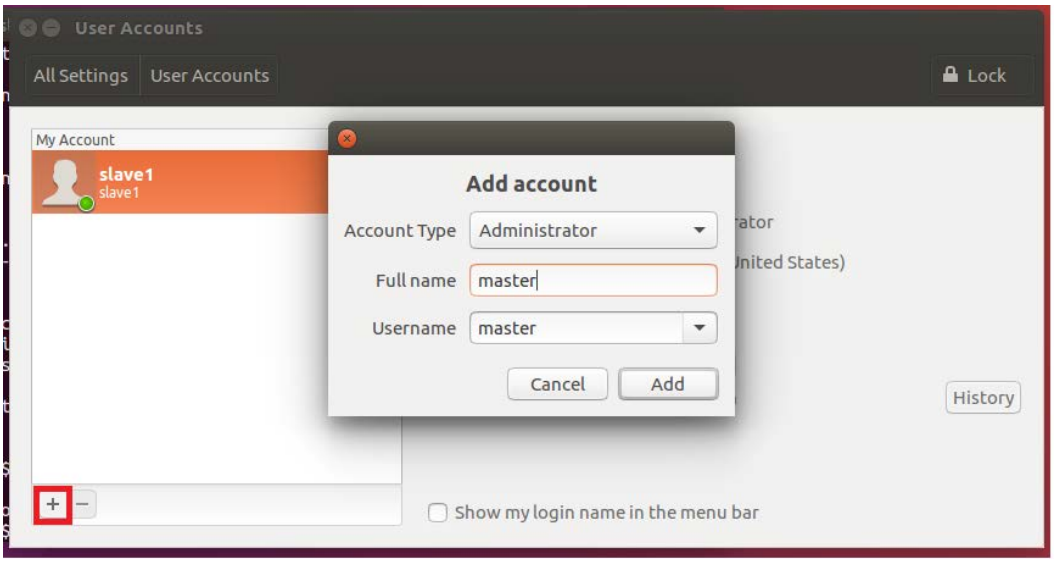
4. 在 slave 建立新使用者


11:47 PM
About This Computer
Ubuntu Help...
System Settings...
Lock
Ctrl+Alt+L
Guest Session
slave1
Log Out...
Suspend
Shut Down...

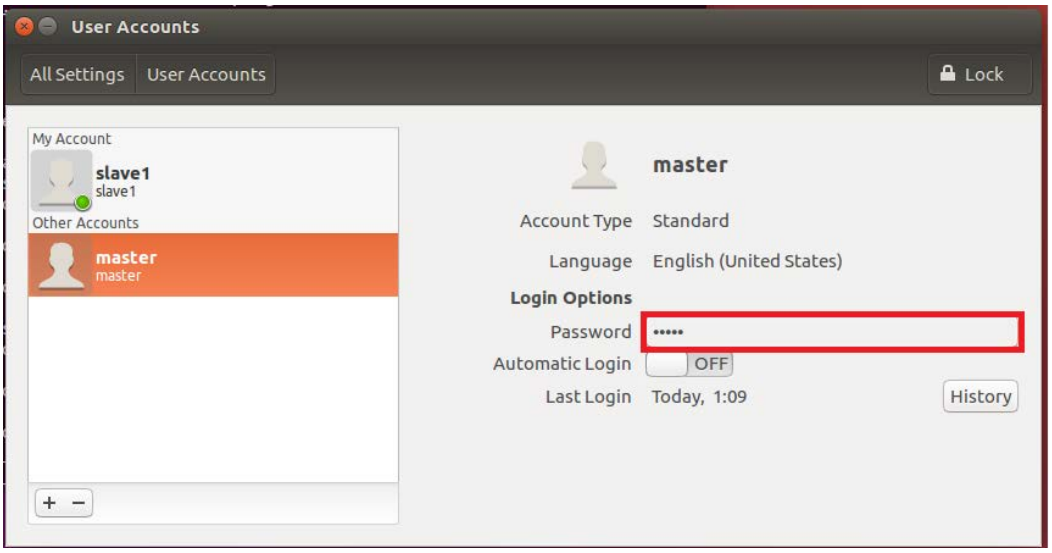
System Settings
All Settings
Personal
Appearance
Brightness & Lock
Language Support
Online Accounts
Security & Privacy
Text Entry
Hardware
Bluetooth
Color
Displays
Keyboard
Mouse & Touchpad
Network
Power
Printers
Sound
Wacom Tablet
System
Backups
Details
Software & Updates
Time & Date
Universal Access
User Accounts


User Accounts
All Settings
User Accounts
Unlock
My Account
slave1
slave1
slave1
Account Type
Administrator
Language
English (United States)
History
Authenticate
Authentication is required to change user data
An application is attempting to perform an action that requires privileges. Authentication is required to perform this action.
Password:
Details
Cancel
Authenticate

Name 須設為 master



設置密碼：



5. 在 slave 安裝 ssh：

sudo apt-get install openssh-server

sudo /etc/init.d/ssh start

sudo su master(更換使用者)

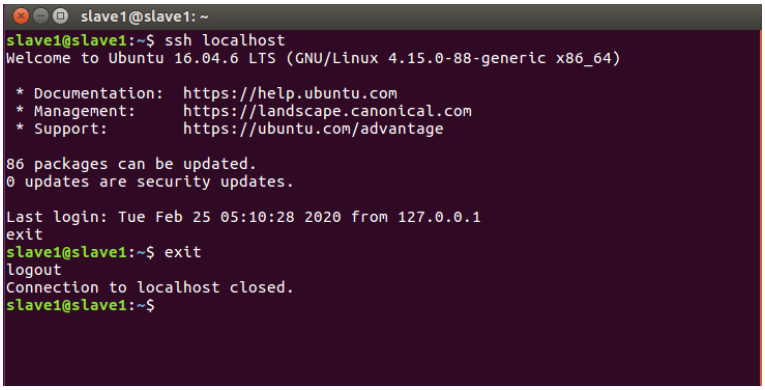
cd ~

ssh-keygen -t rsa -P ""

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

ssh localhost

exit



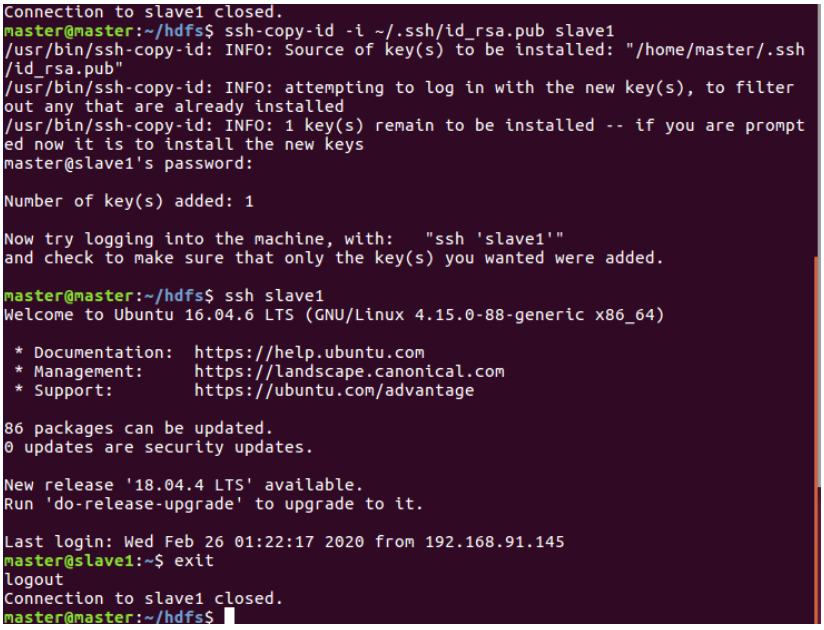
6. 在 master 輸入指令：

ssh-copy-id -i ~/.ssh/id_rsa.pub slave1

7. 在 master 用 ssh 登入 slave：

ssh slave1

exit



8. 在 master 下指令修改如下四個檔案，core-site.xml, mapred-site.xml, hdfs-site.xml, yarn-site.xml：

sudo gedit /opt/hadoop/etc/hadoop/core-site.xml

#core-site.xml，指定 namenode

```
<configuration>
    <property>
        <name>fs.defaultFS</name>
        <value>hdfs://master:9000</value>
    </property>
    <property>
        <name>hadoop.tmp.dir</name>
        <value>file:/home/master/tmp</value>
        <description>Abase for other temporary directories.</description>
    </property>
</configuration>
```

****

sudo gedit /opt/hadoop/etc/hadoop/mapred-site.xml

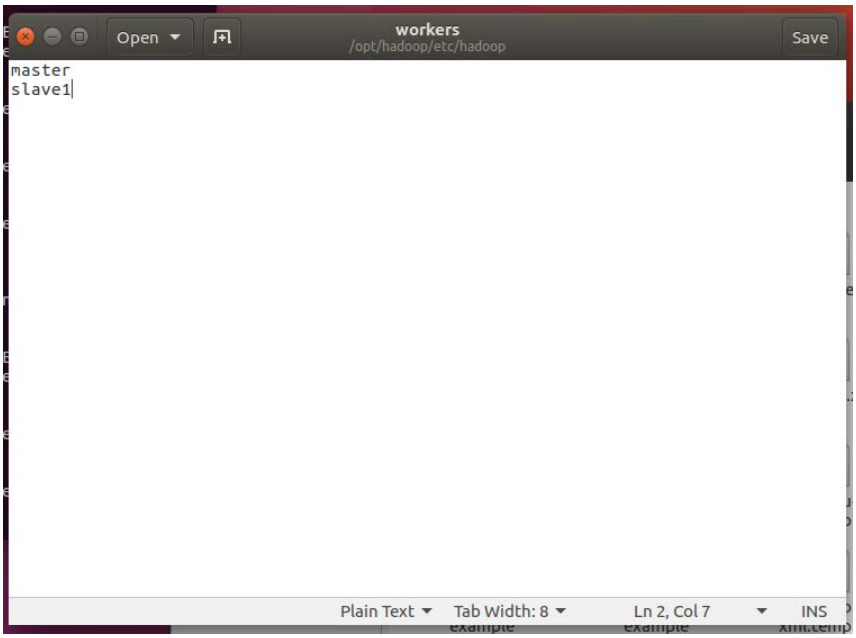
```
#mapred-site.xml 指定 jobtracker
<configuration>
   <property>
      <name>mapreduce.framework.name</name>
      <value>yarn</value>
   </property>
   <property>
      <name>mapreduce.jobhistory.address</name>
      <value>master:10020</value>
   </property>
   <property>
      <name>mapreduce.jobhistory.webapp.address</name>
      <value>master:19888</value>
   </property>
   <property>
      <name>yarn.app.mapreduce.am.env</name>
      <value>HADOOP_MAPRED_HOME=/opt/hadoop</value>
   </property>
   <property>
      <name>mapreduce.map.env</name>
      <value>HADOOP_MAPRED_HOME=/opt/hadoop</value>
   </property>
   <property>
      <name>mapreduce.reduce.env</name>
      <value>HADOOP_MAPRED_HOME=/opt/hadoop</value>
   </property>
</configuration>
**********************************************
sudo gedit /opt/hadoop/etc/hadoop/yarn-site.xml
#yarn-site.xml
<configuration>
   <property>
      <name>yarn.resourcemanager.hostname</name>
      <value>master</value>
   </property>
   <property>
      <name>yarn.nodemanager.resource.memory-mb</name>
      <value>10240</value>
   </property>
   <property>
      <name>yarn.nodemanager.aux-services</name>
      <value>mapreduce_shuffle</value>
```

```
    </property>
</configuration>
**********************************************
sudo gedit /opt/hadoop/etc/hadoop/hdfs-site.xml
#hdfs-site.xml
<configuration>
    <property>
        <name>dfs.namenode.secondary.http-address</name>
        <value>master:50090</value>
    </property>
    <property>
        <name>dfs.replication</name>
        <!--注意此處，因為有兩個 slave 節點，所以由 1 改為 2 -->
        <value>2</value>
    </property>
    <property>
        <name>dfs.namenode.name.dir</name>
        <value>file:/home/master/hdfs/name</value>
    </property>
    <property>
        <name>dfs.datanode.data.dir</name>
        <value>file:/home/master/hdfs/data</value>
    </property>
</configuration>
**********************************************
```

9. 在 master 下指令增加節點：

sudo gedit /opt/hadoop/etc/hadoop/workers

寫入 master、slave 主機名稱



10. 在 master 下指令，將 master 的環境複製到 slave 中

cd /opt

壓縮 hadoop：

sudo tar -zcf ./hadoop.tar.gz ./hadoop

傳送給 slave 端：(@前應放 slave 的使用者名稱；@後應放 slave 的主機名稱)

scp ./hadoop.tar.gz slave1@slave1:~/Downloads

scp -r jdk slave1@slave1:~/Downloads

11. 在 slave 下指令配置 slave 的環境

su slave1

sudo gedit /etc/sudoers

在 root ALL=(ALL:ALL) ALL

下增加 master ALL=(ALL:ALL) ALL

```
# Cmnd alias specification

# User privilege specification
root    ALL=(ALL:ALL) ALL
master  ALL=(ALL:ALL) ALL

# Members of the admin group may gain root privileges
```

sudo su master (使用者變換)

sudo gedit ~/.bashrc (配置環境變數)

將以下程式碼加在檔案最下面並且儲存

export CLASS_PATH="/opt/jdk/lib"

export JAVA_HOME="/opt/jdk"

export HADOOP_HOME="/opt/hadoop"

export MAVEN_HOME="/opt/maven"

export MAHOUT_HOME="/opt/mahout"

export PATH=$PATH:$JAVA_HOME/bin:$HADOOP_HOME/bin:$HADOOP_HOME/sbin:$MAVEN_HOME/bin:$MAHOUT_HOME/bin

export HADOOP_CONF_DIR="/opt/hadoop/conf"

export HADOOP_MAPRED_HOME="/opt/hadoop"

export HADOOP_COMMON_HOME="/opt/hadoop"

export HADOOP_HDFS_HOME="/opt/hadoop"

export HADOOP_LOCAL="/opt/hadoop"

export YARN_HOME="/opt/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 CLASS_PATH="/opt/jdk/lib"
export JAVA_HOME="/opt/jdk"
export HADOOP_HOME="/opt/hadoop"
export MAVEN_HOME="/opt/maven"
export MAHOUT_HOME="/opt/mahout"
export PATH=$PATH:$JAVA_HOME/bin:$HADOOP_HOME/bin:$HADOOP_HOME/sbin:$MAVEN_HOME/bin:$MAHOUT_HOME/
bin
export HADOOP_CONF_DIR="/opt/hadoop/conf"
export HADOOP_MAPRED_HOME="/opt/hadoop"
export HADOOP_COMMON_HOME="/opt/hadoop"
export HADOOP_HDFS_HOME="/opt/hadoop"
export HADOOP_LOCAL="/opt/hadoop"
export YARN_HOME="/opt/hadoop"
```

sh ▾ Tab Width: 8 ▾ Ln 1, Col 1 ▾ INS

source ~/.bashrc

sudo su slave1 (使用者變換)

在 slave 下指令配置 java：

cd ~

cd Downloads

sudo mv jdk /opt/jdk (移到/opt 目錄下)

在 slave 下指令配置 hadoop：

cd ~

cd Downloads

tar zxvf hadoop.tar.gz (解壓縮)

sudo mv hadoop /opt/hadoop (移到/opt 目錄下)

sudo chmod -R 777 /opt/hadoop/logs (更改權限)

sudo su master (使用者變換)

source /opt/hadoop/etc/hadoop/hadoop-env.sh (讓配置生效)

java -version (測試安裝正常)

hadoop version (測試安裝正常)

```
master@slave1:/home/slave1$ source /opt/hadoop/etc/hadoop/hadoop-env.sh
master@slave1:/home/slave1$ java -version
java version "1.8.0_241"
Java(TM) SE Runtime Environment (build 1.8.0_241-b07)
Java HotSpot(TM) 64-Bit Server VM (build 25.241-b07, mixed mode)
master@slave1:/home/slave1$ hadoop version
Hadoop 3.2.1
Source code repository https://gitbox.apache.org/repos/asf/hadoop.git -r b3cbbb4
67e22ea829b3808f4b7b01d07e0bf3842
Compiled by rohithsharmaks on 2019-09-10T15:56Z
Compiled with protoc 2.5.0
From source with checksum 776eaf9eee9c0ffc370bcbc1888737
This command was run using /opt/hadoop/share/hadoop/common/hadoop-common-3.2.1.j
ar
master@slave1:/home/slave1$
```

12. 在 master 下指令啟用 hadoop (要先刪除 home/hdfs 資料夾內的 name、data)

source /opt/hadoop/etc/hadoop/hadoop-env.sh

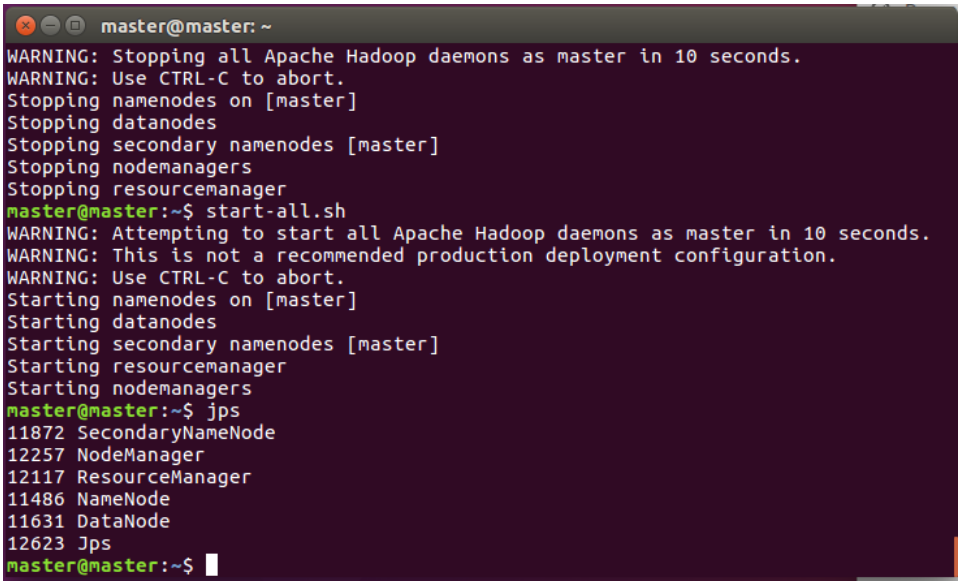
hdfs namenode -format

source /opt/hadoop/etc/hadoop/hadoop-env.sh

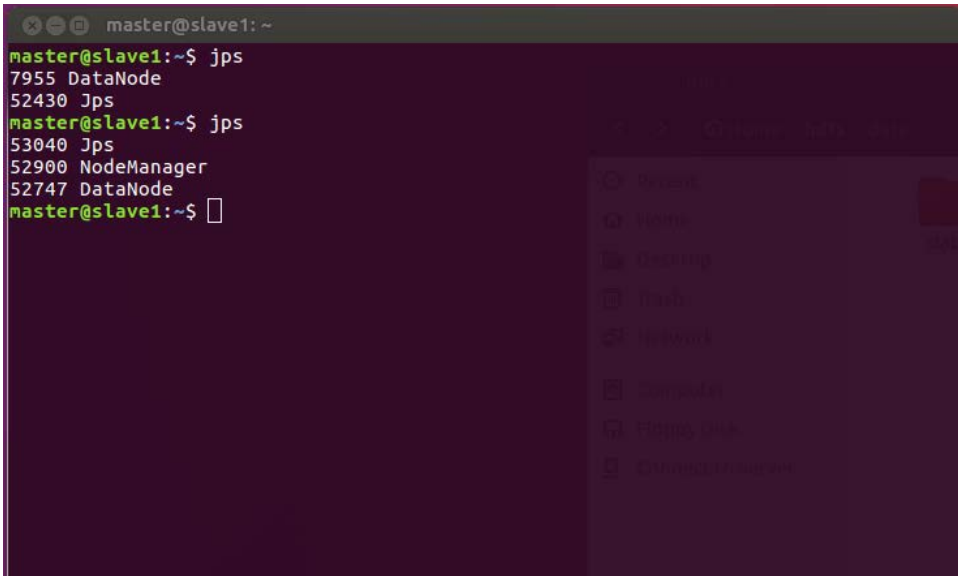
start-all.sh

完成畫面：

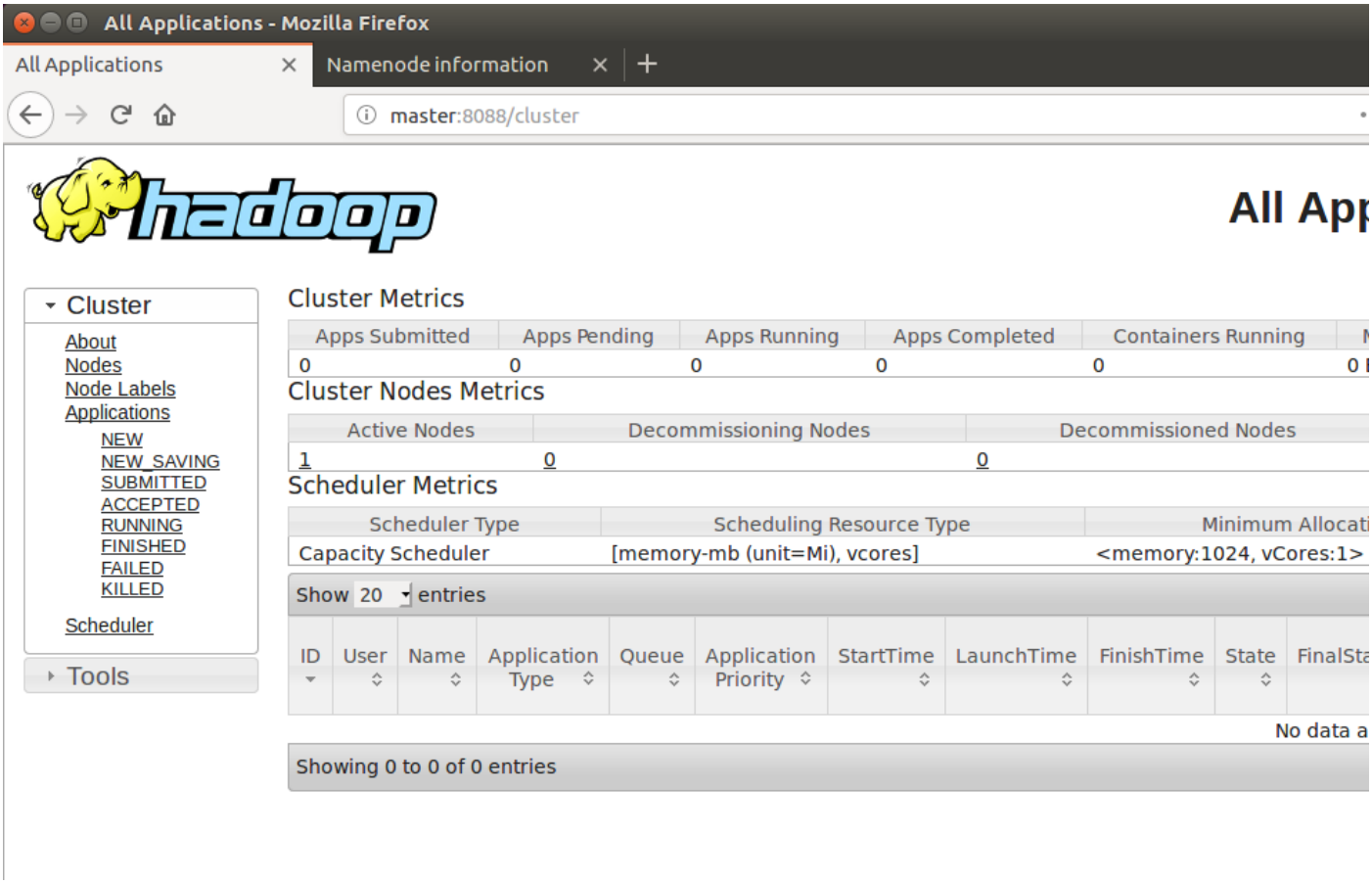
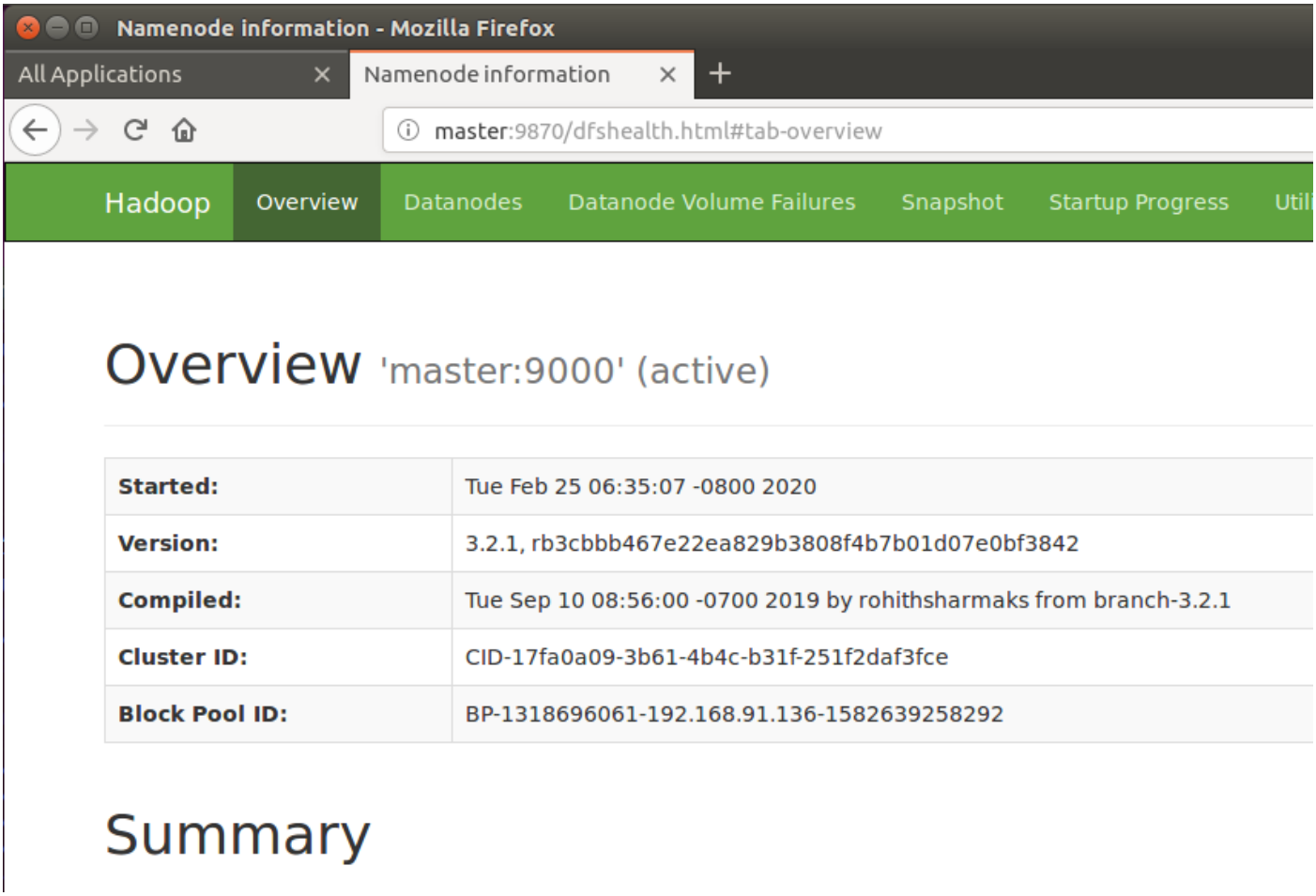
master 主機：



slave1 主機：



試試以下 2 個網址有無正常運作，http://master:8088



http://master:9870