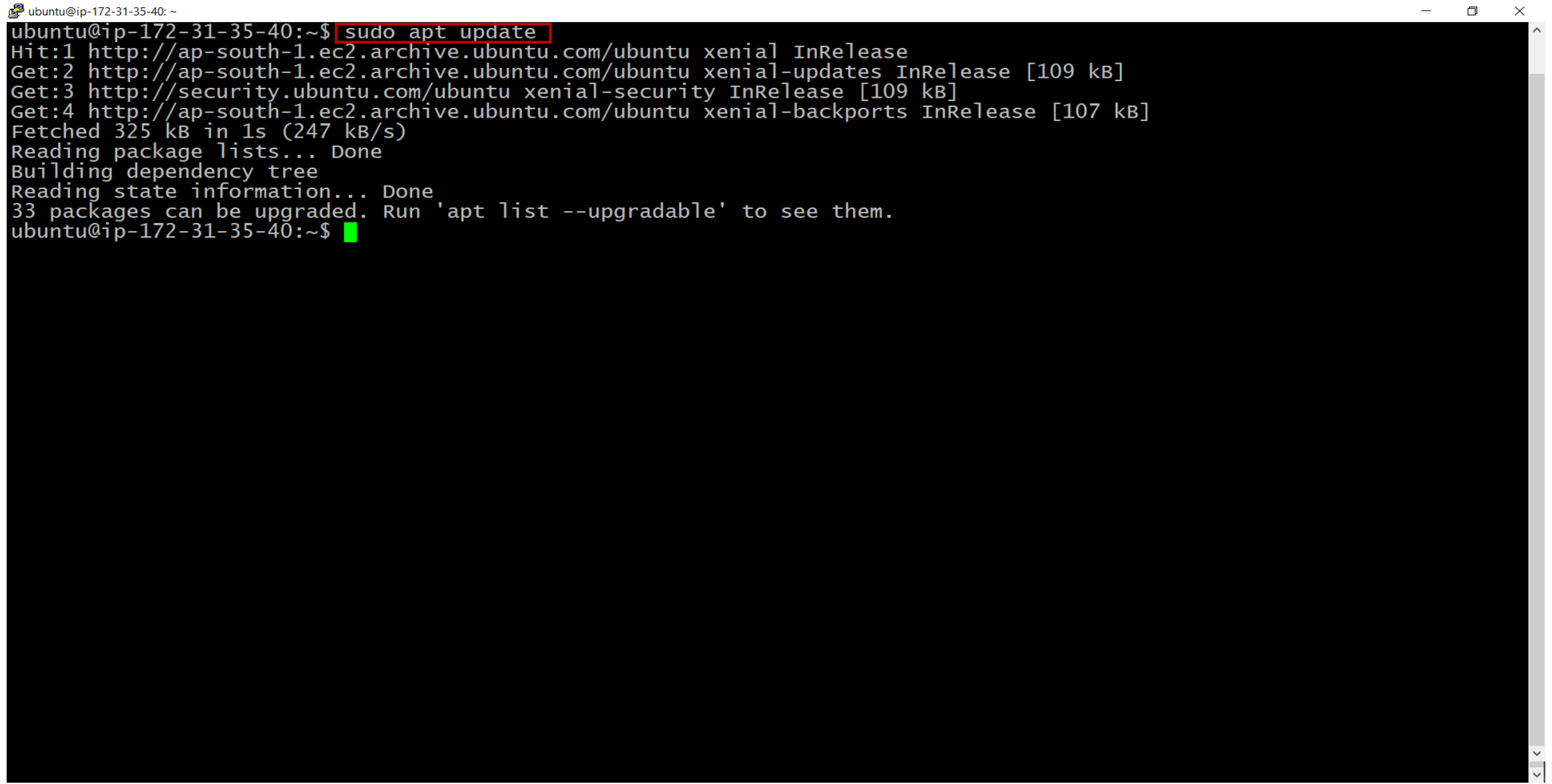




# Single Node Cluster Installation

# 01) sudo apt-get update



```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ sudo apt update  
Hit:1 http://ap-south-1.ec2.archive.ubuntu.com/ubuntu xenial InRelease  
Get:2 http://ap-south-1.ec2.archive.ubuntu.com/ubuntu xenial-updates InRelease [109 kB]  
Get:3 http://security.ubuntu.com/ubuntu xenial-security InRelease [109 kB]  
Get:4 http://ap-south-1.ec2.archive.ubuntu.com/ubuntu xenial-backports InRelease [107 kB]  
Fetched 325 kB in 1s (247 kB/s)  
Reading package lists... Done  
Building dependency tree  
Reading state information... Done  
33 packages can be upgraded. Run 'apt list --upgradable' to see them.  
ubuntu@ip-172-31-35-40:~$
```

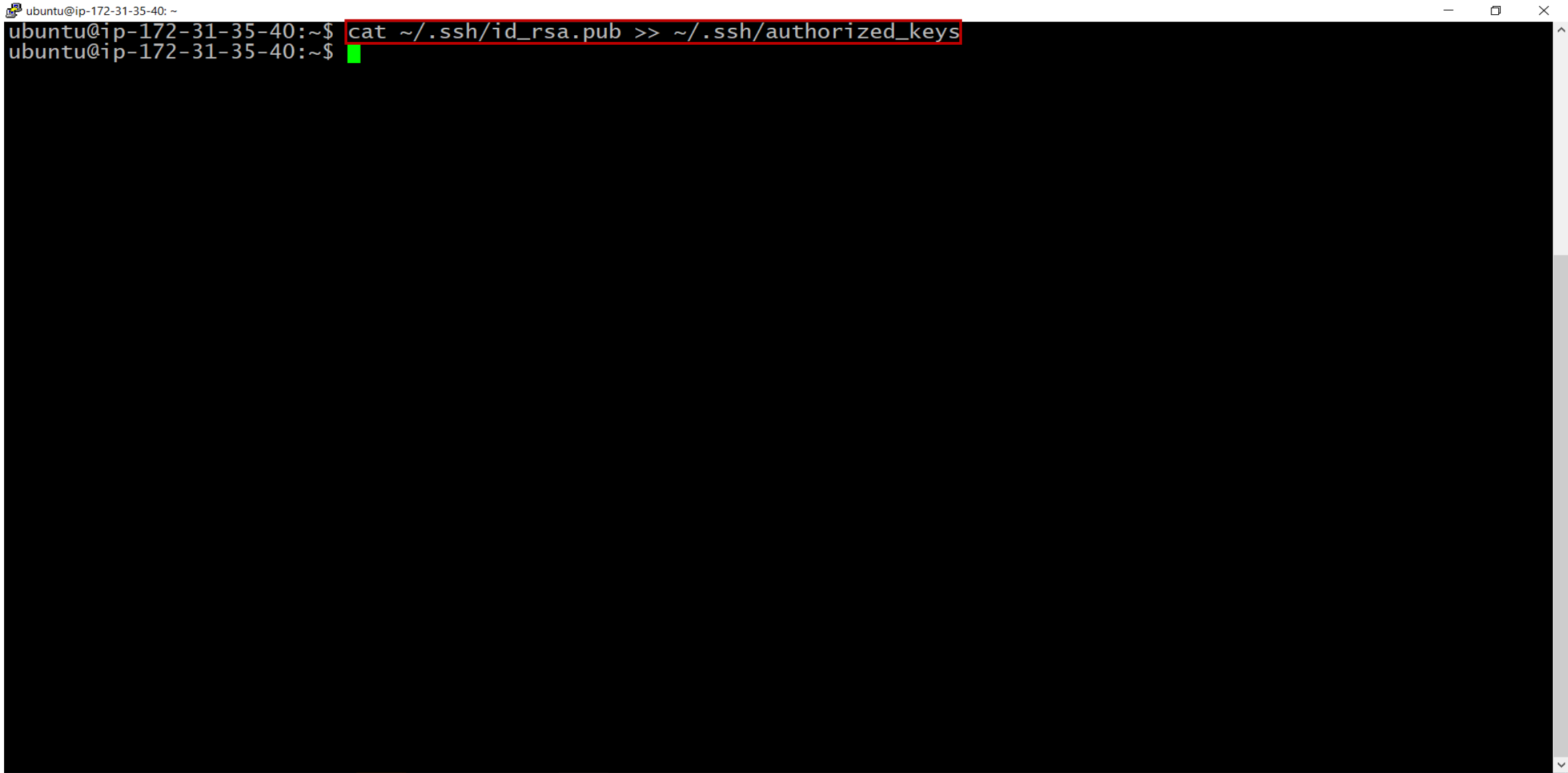
press enter

## 02) ssh-keygen

```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ ssh-keygen  
Generating public/private rsa key pair.  
Enter file in which to save the key (/home/ubuntu/.ssh/id_rsa):  
Enter passphrase (empty for no passphrase):  
Enter same passphrase again:  
Your identification has been saved in /home/ubuntu/.ssh/id_rsa.  
Your public key has been saved in /home/ubuntu/.ssh/id_rsa.pub.  
The key fingerprint is:  
SHA256:Emgqfy4zfyT2GyON50Rhb79s/CxMxOHUseIJs12iSnc ubuntu@ip-172-31-35-40  
The key's randomart image is:  
+---[RSA 2048]---+  
|                 .  
|      .          o .  
|    o = o+o .  
|  o . O =+o  
|. . . * E.O  
| o .O*.+ .  
|..=+* +.  
| +o =oo .=o  
| =o.o .ooo  
+---[SHA256]-----+  
ubuntu@ip-172-31-35-40:~$
```

press enter 6 times

03) cat ~/.ssh/id\_rsa.pub >> ~/.ssh/authorized\_keys

A terminal window with a black background and white text. The title bar at the top shows 'ubuntu@ip-172-31-35-40: ~' and standard window controls. The command 'cat ~/.ssh/id\_rsa.pub >> ~/.ssh/authorized\_keys' is entered and highlighted with a red box. The prompt 'ubuntu@ip-172-31-35-40:~\$' is visible on the first line, and the same prompt followed by a green cursor is on the second line.

```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys  
ubuntu@ip-172-31-35-40:~$
```

press enter

## 04) ssh localhost

```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ ssh localhost  
Welcome to Ubuntu 16.04.7 LTS (GNU/Linux 4.4.0-1128-aws x86_64)  
  
* Documentation:  https://help.ubuntu.com  
* Management:    https://landscape.canonical.com  
* Support:        https://ubuntu.com/advantage  
  
18 packages can be updated.  
2 of these updates are security updates.  
To see these additional updates run: apt list --upgradable  
  
New release '18.04.6 LTS' available.  
Run 'do-release-upgrade' to upgrade to it.  
  
*** System restart required ***  
Last login: Thu Oct 14 08:39:06 2021 from 152.57.204.27  
ubuntu@ip-172-31-35-40:~$
```

press enter

## 05) sudo apt-get install openjdk-8-jdk -y

```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ sudo apt-get install openjdk-8-jdk -y  
Reading package lists... Done  
Building dependency tree  
Reading state information... Done  
The following additional packages will be installed:  
  ca-certificates-java fontconfig fontconfig-config fonts-dejavu-core fonts-dejavu-extra hicolor-icon-theme  
  java-common libasound2 libasound2-data libasyns0 libatk1.0-0 libatk1.0-data libavahi-client3  
  libavahi-common-data libavahi-common3 libcairo2 libcups2 libdatrie1 libdrm-amdgpu1 libdrm-intel1 libdrm-nouveau2  
  libdrm-radeon1 libflac8 libfontconfig1 libgdk-pixbuf2.0-0 libgdk-pixbuf2.0-common libgif7 libgl1-mesa-dri  
  libgl1-mesa-glx libglapi-mesa libgraphite2-3 libgtk2.0-0 libgtk2.0-bin libgtk2.0-common libharfbuzz0b libice-dev  
  libice6 libjbig0 libjpeg-turbo8 libjpeg8 liblcms2-2 libllvm6.0 libnspr4 libnss3 libnss3-nssdb libogg0  
  libpango-1.0-0 libpangocairo-1.0-0 libpangoft2-1.0-0 libpciaccess0 libpcsclite1 libpixman-1-0  
  libpthread-stubs0-dev libpulse0 libsensors4 libsm-dev libsm6 libsndfile1 libthai-data libthai0 libtiff5  
  libtxc-dxtn-s2tc0 libvorbis0a libvorbisenc2 libx11-dev libx11-doc libx11-xcb1 libxau-dev libxcb-dri2-0  
  libxcb-dri3-0 libxcb-glx0 libxcb-present0 libxcb-render0 libxcb-shm0 libxcb-sync1 libxcb1-dev libxcomposite1  
  libxcursor1 libxdamage1 libxdmcp-dev libxf86vm3 libxi6 libxinerama1 libxrandr2 libxrender1 libxshmfence1 libxt-dev  
  libxt6 libxtst6 libxxf86vm1 openjdk-8-jdk-headless openjdk-8-jre openjdk-8-jre-headless x11-common  
  x11proto-core-dev x11proto-input-dev x11proto-kb-dev xorg-sgml-doctools xtrans-dev  
Suggested packages:  
  default-jre libasound2-plugins alsa-utils cups-common librsvg2-common gvfs libice-doc liblcms2-utils pscd  
  pulseaudio lm-sensors libsm-doc libxcb-doc libxt-doc openjdk-8-demo openjdk-8-source visualvm icedtea-8-plugin  
  libnss-mdns fonts-ipafont-gothic fonts-ipafont-mincho fonts-wqy-microhei fonts-wqy-zenhei fonts-indic  
The following NEW packages will be installed:  
  ca-certificates-java fontconfig fontconfig-config fonts-dejavu-core fonts-dejavu-extra hicolor-icon-theme  
  java-common libasound2 libasound2-data libasyns0 libatk1.0-0 libatk1.0-data libavahi-client3  
  libavahi-common-data libavahi-common3 libcairo2 libcups2 libdatrie1 libdrm-amdgpu1 libdrm-intel1 libdrm-nouveau2  
  libdrm-radeon1 libflac8 libfontconfig1 libgdk-pixbuf2.0-0 libgdk-pixbuf2.0-common libgif7 libgl1-mesa-dri  
  libgl1-mesa-glx libglapi-mesa libgraphite2-3 libgtk2.0-0 libgtk2.0-bin libgtk2.0-common libharfbuzz0b libice-dev  
  libice6 libjbig0 libjpeg-turbo8 libjpeg8 liblcms2-2 libllvm6.0 libnspr4 libnss3 libnss3-nssdb libogg0  
  libpango-1.0-0 libpangocairo-1.0-0 libpangoft2-1.0-0 libpciaccess0 libpcsclite1 libpixman-1-0  
  libpthread-stubs0-dev libpulse0 libsensors4 libsm-dev libsm6 libsndfile1 libthai-data libthai0 libtiff5  
  libtxc-dxtn-s2tc0 libvorbis0a libvorbisenc2 libx11-dev libx11-doc libx11-xcb1 libxau-dev libxcb-dri2-0  
  libxcb-dri3-0 libxcb-glx0 libxcb-present0 libxcb-render0 libxcb-shm0 libxcb-sync1 libxcb1-dev libxcomposite1  
  libxcursor1 libxdamage1 libxdmcp-dev libxf86vm3 libxi6 libxinerama1 libxrandr2 libxrender1 libxshmfence1 libxt-dev  
  libxt6 libxtst6 libxxf86vm1 openjdk-8-jdk openjdk-8-jdk-headless openjdk-8-jre openjdk-8-jre-headless x11-common  
  x11proto-core-dev x11proto-input-dev x11proto-kb-dev xorg-sgml-doctools xtrans-dev  
0 upgraded, 100 newly installed, 0 to remove and 16 not upgraded.
```

press enter

06) `wget https://archive.apache.org/dist/hadoop/common/hadoop-1.2.1/hadoop-1.2.1.tar.gz`

```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ wget https://archive.apache.org/dist/hadoop/common/hadoop-1.2.1/hadoop-1.2.1.tar.gz  
--2021-10-14 06:46:07-- https://archive.apache.org/dist/hadoop/common/hadoop-1.2.1/hadoop-1.2.1.tar.gz  
Resolving archive.apache.org (archive.apache.org)... 138.201.131.134, 2a01:4f8:172:2ec5::2  
Connecting to archive.apache.org (archive.apache.org)|138.201.131.134|:443... connected.  
HTTP request sent, awaiting response... 200 OK  
Length: 63851630 (61M) [application/x-gzip]  
Saving to: 'hadoop-1.2.1.tar.gz'  
  
hadoop-1.2.1.tar.gz      100%[=====>] 60.89M  12.2MB/s   in 5.7s  
2021-10-14 06:46:13 (10.7 MB/s) - 'hadoop-1.2.1.tar.gz' saved [63851630/63851630]  
ubuntu@ip-172-31-35-40:~$ █
```

press enter

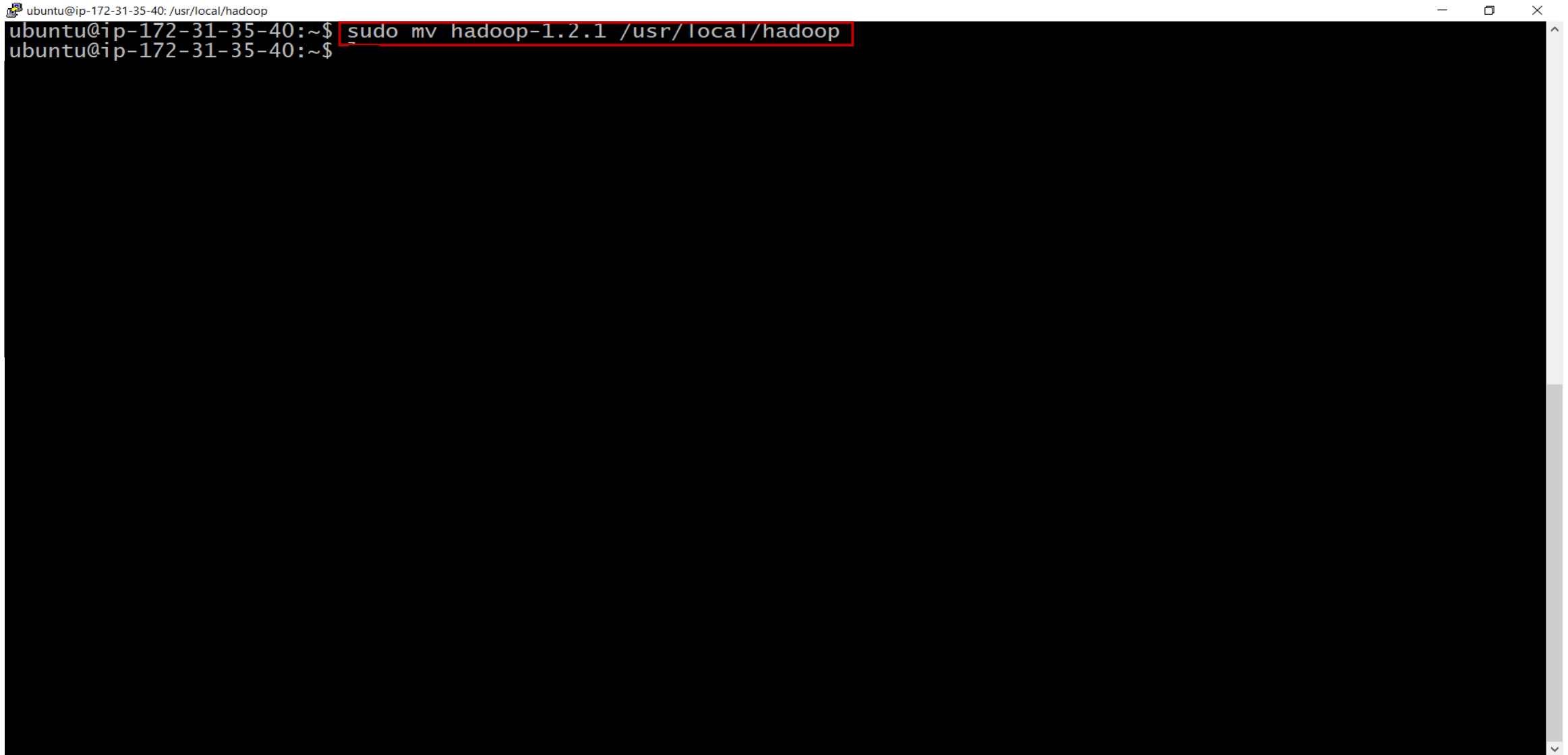
## 07) tar -xzvf hadoop-1.2.1.tar.gz

```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ tar -xzvf hadoop-1.2.1.tar.gz  
hadoop-1.2.1/  
hadoop-1.2.1/.eclipse.templates/  
hadoop-1.2.1/.eclipse.templates/.externalToolBuilders/  
hadoop-1.2.1/.eclipse.templates/.launches/  
hadoop-1.2.1/bin/  
hadoop-1.2.1/c++/  
hadoop-1.2.1/c++/Linux-amd64-64/  
hadoop-1.2.1/c++/Linux-amd64-64/include/  
hadoop-1.2.1/c++/Linux-amd64-64/include/hadoop/  
hadoop-1.2.1/c++/Linux-amd64-64/lib/
```

press enter



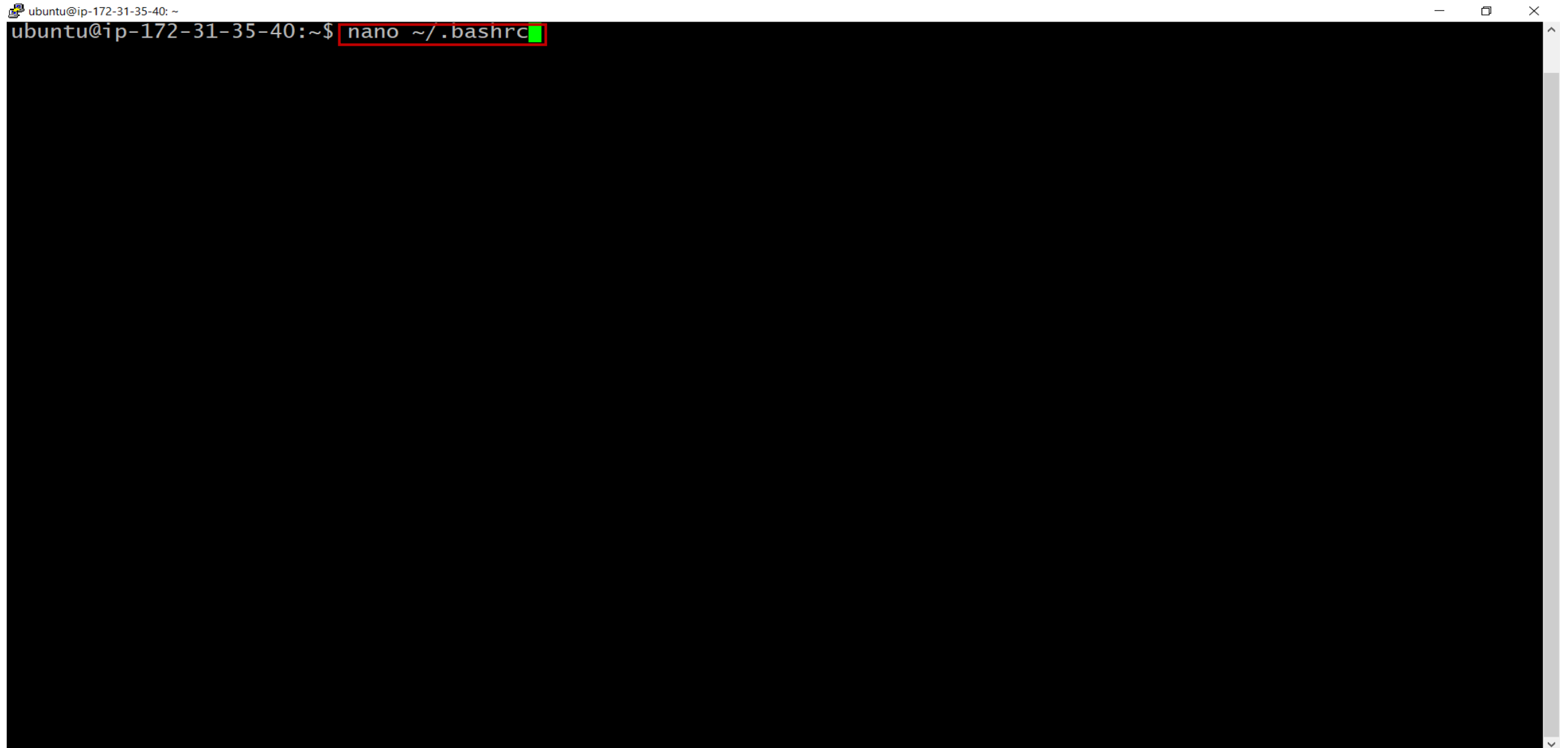
## 08) sudo mv hadoop-1.2.1 /usr/local/hadoop

A terminal window with a black background and white text. The title bar at the top reads 'ubuntu@ip-172-31-35-40: /usr/local/hadoop'. The terminal shows two lines of text: 'ubuntu@ip-172-31-35-40:~\$' followed by 'sudo mv hadoop-1.2.1 /usr/local/hadoop' on the same line. The command is highlighted with a red rectangular box. Below this, a new prompt 'ubuntu@ip-172-31-35-40:~\$' is visible. The rest of the terminal area is empty.

```
ubuntu@ip-172-31-35-40: /usr/local/hadoop
ubuntu@ip-172-31-35-40:~$ sudo mv hadoop-1.2.1 /usr/local/hadoop
ubuntu@ip-172-31-35-40:~$
```

press enter

## 09) nano ~/.bashrc



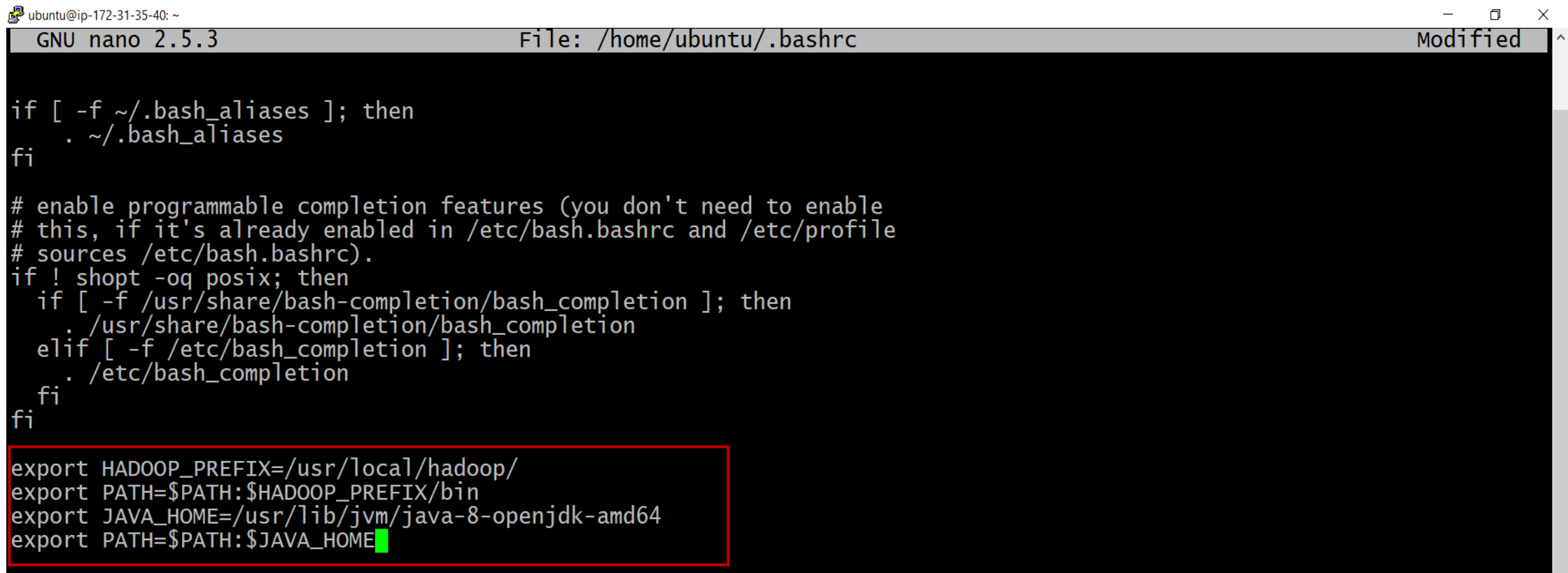
A terminal window with a black background and white text. The title bar at the top shows 'ubuntu@ip-172-31-35-40: ~'. The command prompt is 'ubuntu@ip-172-31-35-40:~\$'. The command 'nano ~/.bashrc' has been entered and is highlighted with a red rectangular box. A green cursor is positioned at the end of the command. The terminal window has standard window controls (minimize, maximize, close) in the top right corner.

```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ nano ~/.bashrc
```

press enter

10) Scroll down at the bottom area & paste the commands

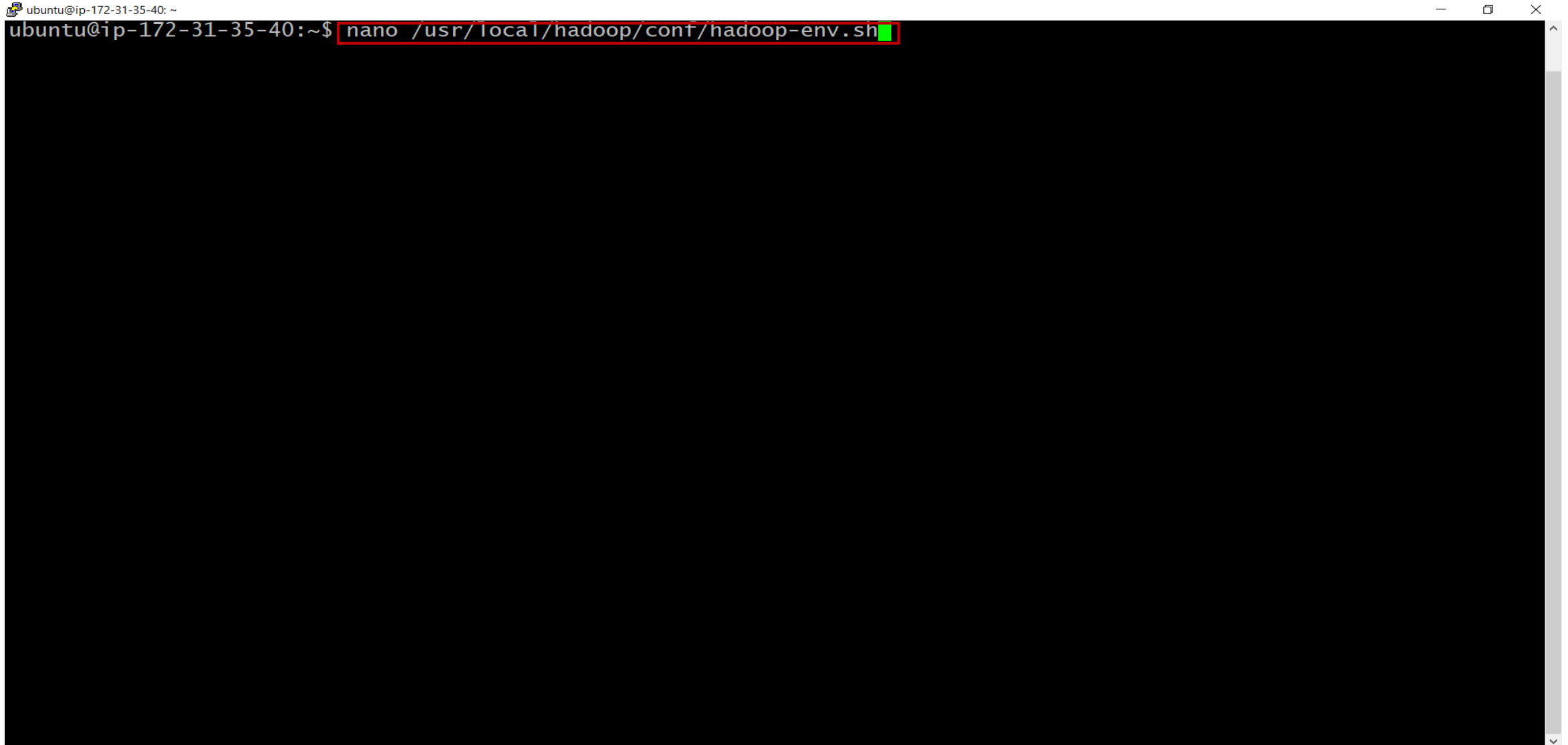
```
export HADOOP_PREFIX=/usr/local/hadoop/  
export PATH=$PATH:$HADOOP_PREFIX/bin  
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64  
export PATH=$PATH:$JAVA_HOME
```



```
ubuntu@ip-172-31-35-40: ~  
GNU nano 2.5.3 File: /home/ubuntu/.bashrc Modified  
  
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  
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64  
export PATH=$PATH:$JAVA_HOME
```

ctrl + O (to save) > Enter > ctrl + X (to exit)

# 1 1) nano /usr/local/hadoop/conf/hadoop-env.sh

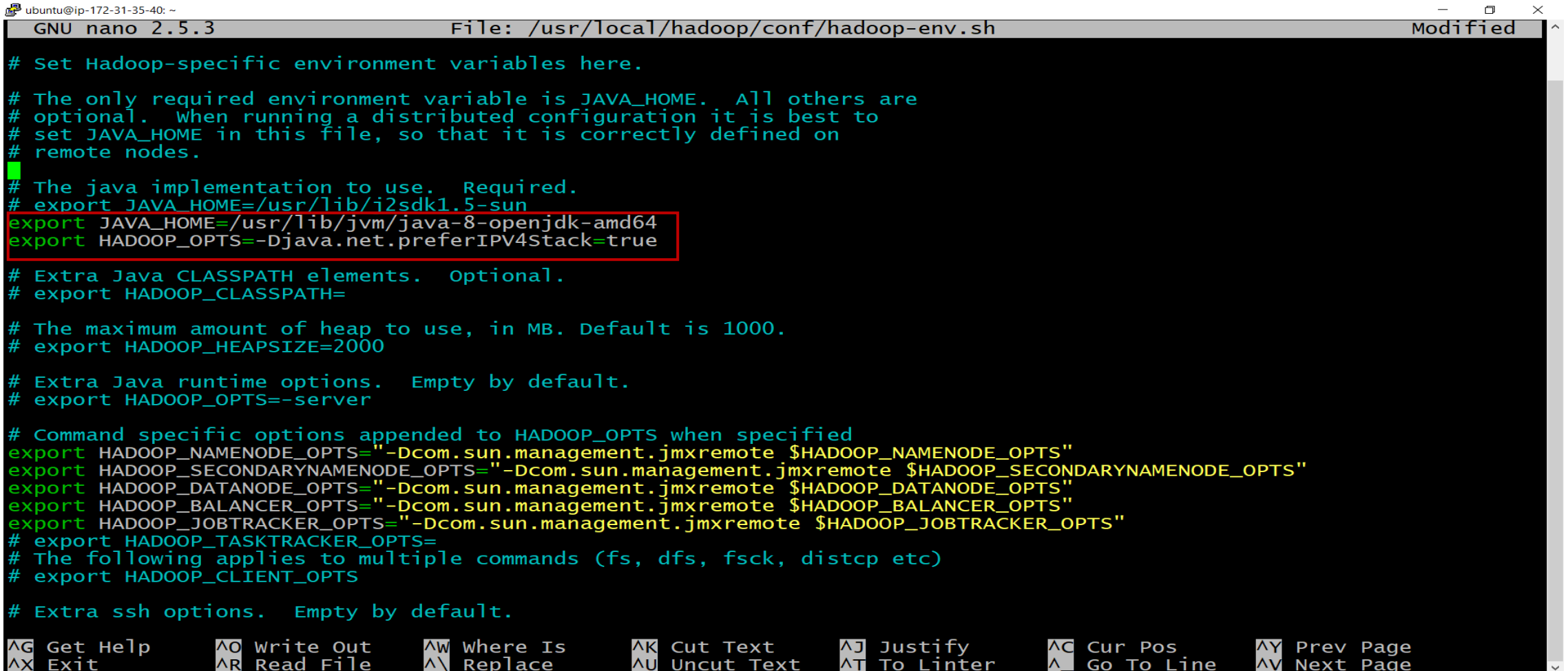


A terminal window with a black background and white text. The title bar at the top shows 'ubuntu@ip-172-31-35-40: ~' and standard window controls. The command prompt 'ubuntu@ip-172-31-35-40:~\$' is followed by the command 'nano /usr/local/hadoop/conf/hadoop-env.sh', which is highlighted with a red rectangular box. A green cursor is positioned at the end of the command. The rest of the terminal area is empty.

press enter

12) Paste the commands in the mentioned area as per the below image

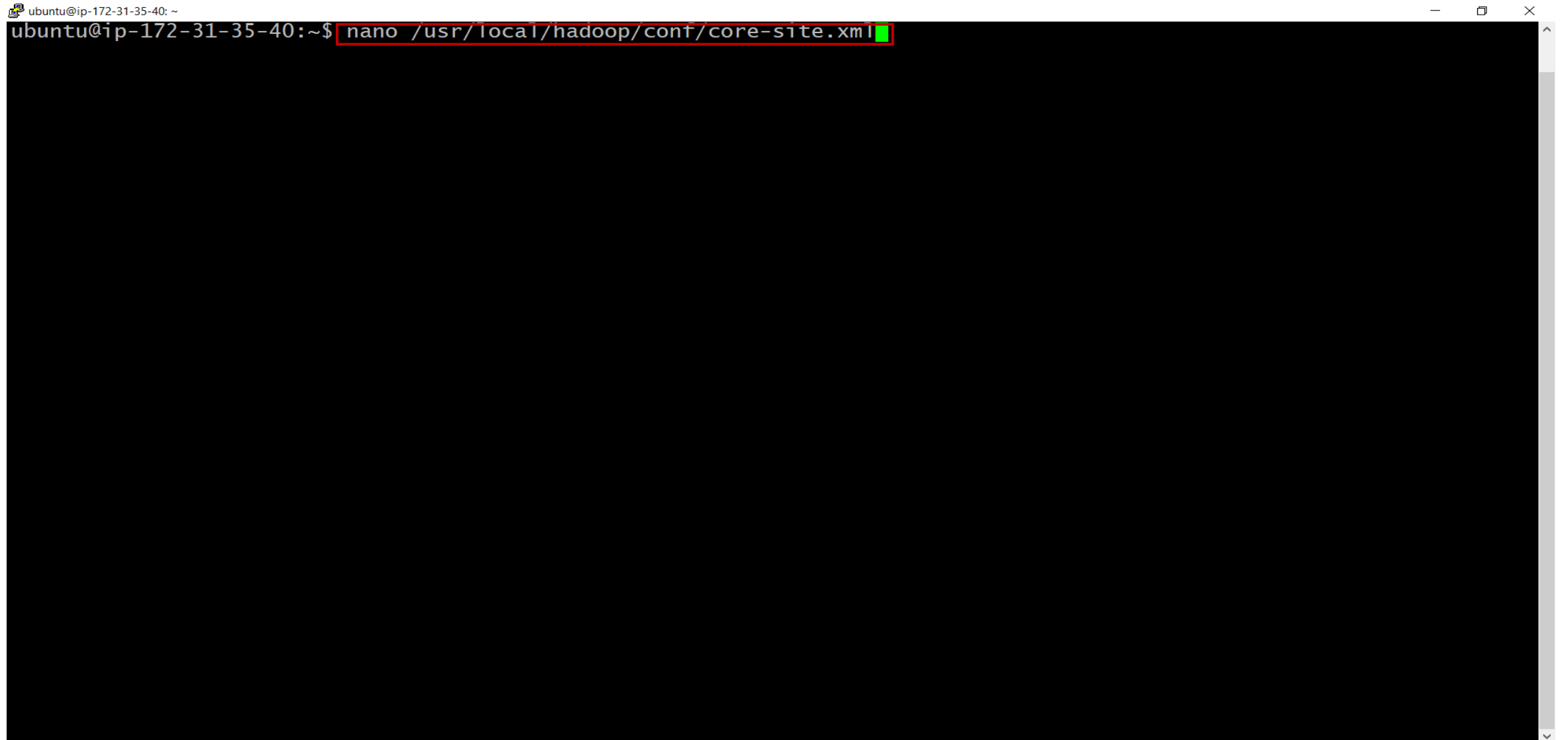
```
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
export HADOOP_OPTS=-Djava.net.preferIPv4Stack=true
```



```
ubuntu@ip-172-31-35-40: ~
GNU nano 2.5.3 File: /usr/local/hadoop/conf/hadoop-env.sh Modified
# 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/j2sdk1.5-sun
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
export HADOOP_OPTS=-Djava.net.preferIPv4Stack=true
# 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"
export HADOOP_SECONDARYNAMENODE_OPTS="-Dcom.sun.management.jmxremote $HADOOP_SECONDARYNAMENODE_OPTS"
export HADOOP_DATANODE_OPTS="-Dcom.sun.management.jmxremote $HADOOP_DATANODE_OPTS"
export HADOOP_BALANCER_OPTS="-Dcom.sun.management.jmxremote $HADOOP_BALANCER_OPTS"
export HADOOP_JOBTRACKER_OPTS="-Dcom.sun.management.jmxremote $HADOOP_JOBTRACKER_OPTS"
# export HADOOP_TASKTRACKER_OPTS=
# The following applies to multiple commands (fs, dfs, fsck, distcp etc)
# export HADOOP_CLIENT_OPTS
# Extra ssh options. Empty by default.
^G Get Help ^O Write Out ^W Where Is ^K Cut Text ^J Justify ^C Cur Pos ^Y Prev Page
^X Exit ^R Read File ^\ Replace ^U Uncut Text ^T To Linter ^_ Go To Line ^V Next Page
```

ctrl + O (to save) > Enter > ctrl + X (to exit)

# 13) nano /usr/local/hadoop/conf/core-site.xml

A terminal window with a black background and white text. The title bar at the top shows 'ubuntu@ip-172-31-35-40: ~' and standard window controls. The command prompt 'ubuntu@ip-172-31-35-40:~\$' is followed by the command 'nano /usr/local/hadoop/conf/core-site.xml' which is highlighted with a red rectangular box. A green cursor is positioned at the end of the command. The rest of the terminal area is empty.

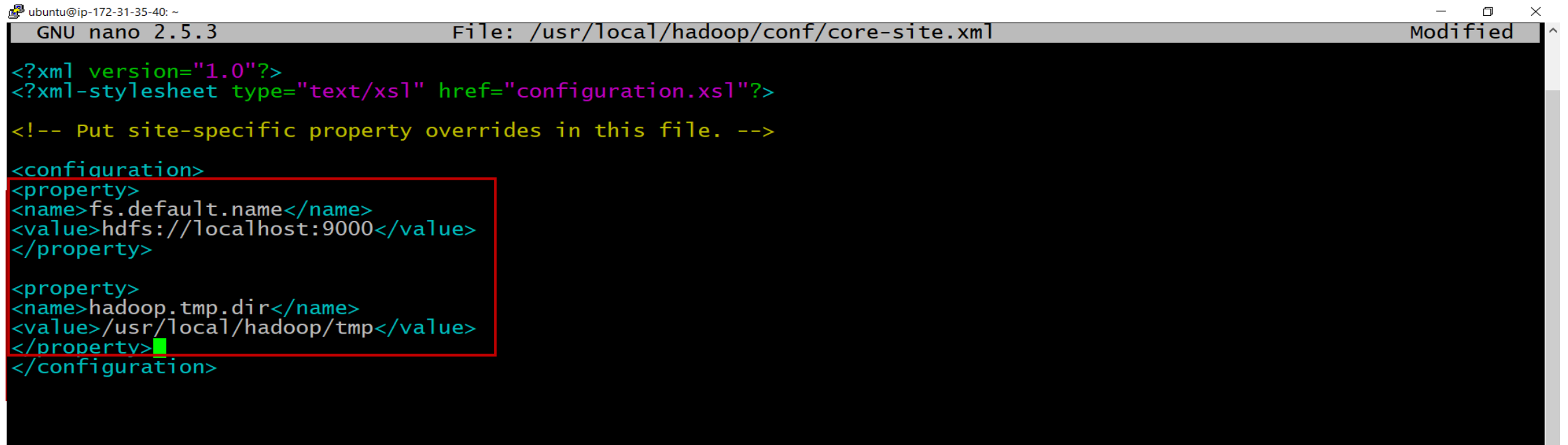
```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ nano /usr/local/hadoop/conf/core-site.xml
```

press enter

- 14) Paste the following commands in-between the <configuration> & </configuration>  
(If you create your own DNS, you need to rename the localhost to your hostname e.g is)

```
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:9000</value>
</property>

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



```
ubuntu@ip-172-31-35-40: ~
GNU nano 2.5.3 File: /usr/local/hadoop/conf/core-site.xml Modified
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

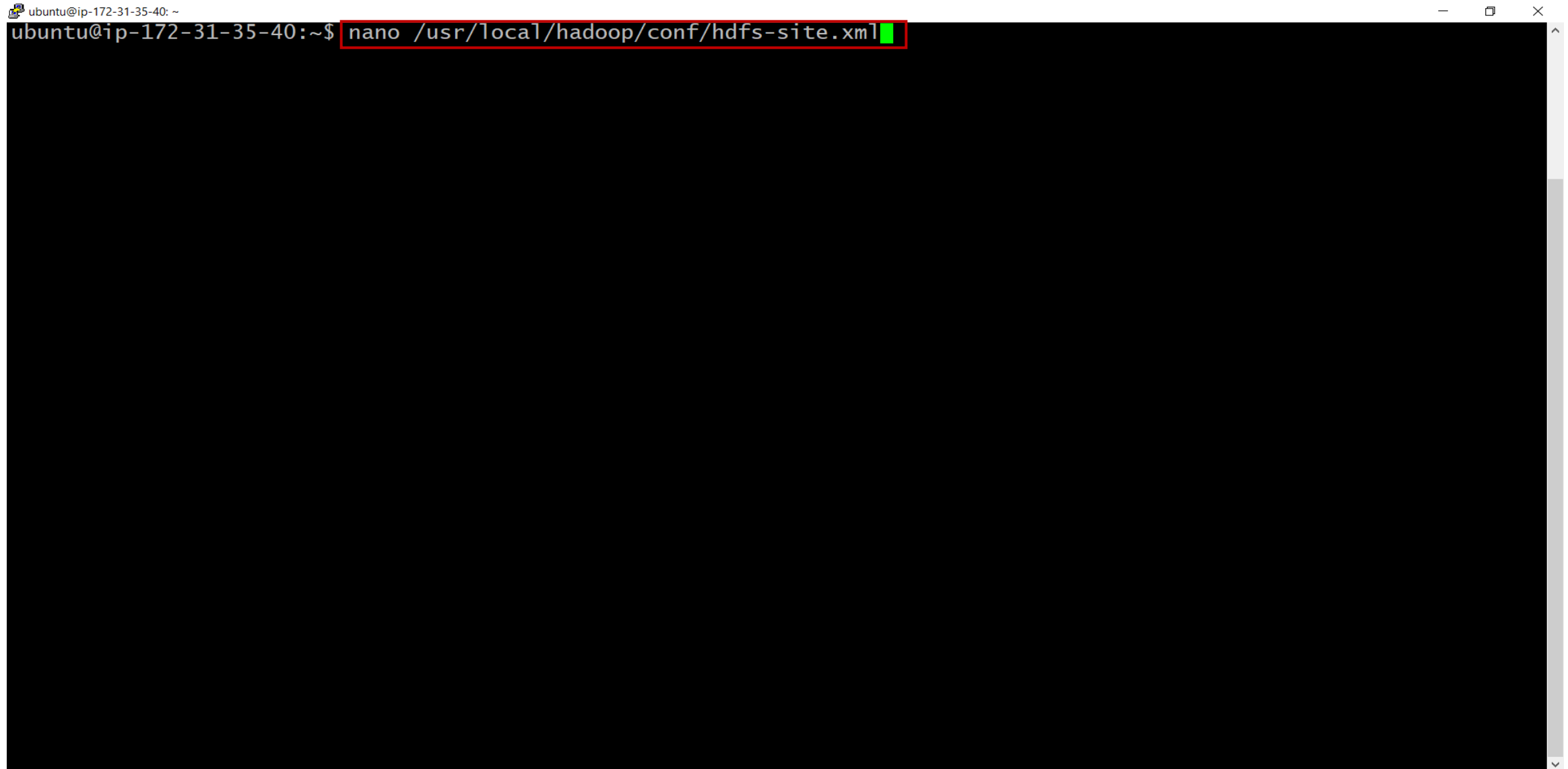
<!-- Put site-specific property overrides in this file. -->

<configuration>
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:9000</value>
</property>

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

ctrl + O (to save) > Enter > ctrl + X (to exit)

## 15) nano /usr/local/hadoop/conf/hdfs-site.xml

A terminal window with a black background and white text. The title bar at the top shows a small icon, the text 'ubuntu@ip-172-31-35-40: ~', and window control buttons (minimize, maximize, close). The terminal content shows the prompt 'ubuntu@ip-172-31-35-40:~\$' followed by the command 'nano /usr/local/hadoop/conf/hdfs-site.xml'. The command is highlighted with a red rectangular box, and a green cursor is positioned at the end of the command. The rest of the terminal area is empty.

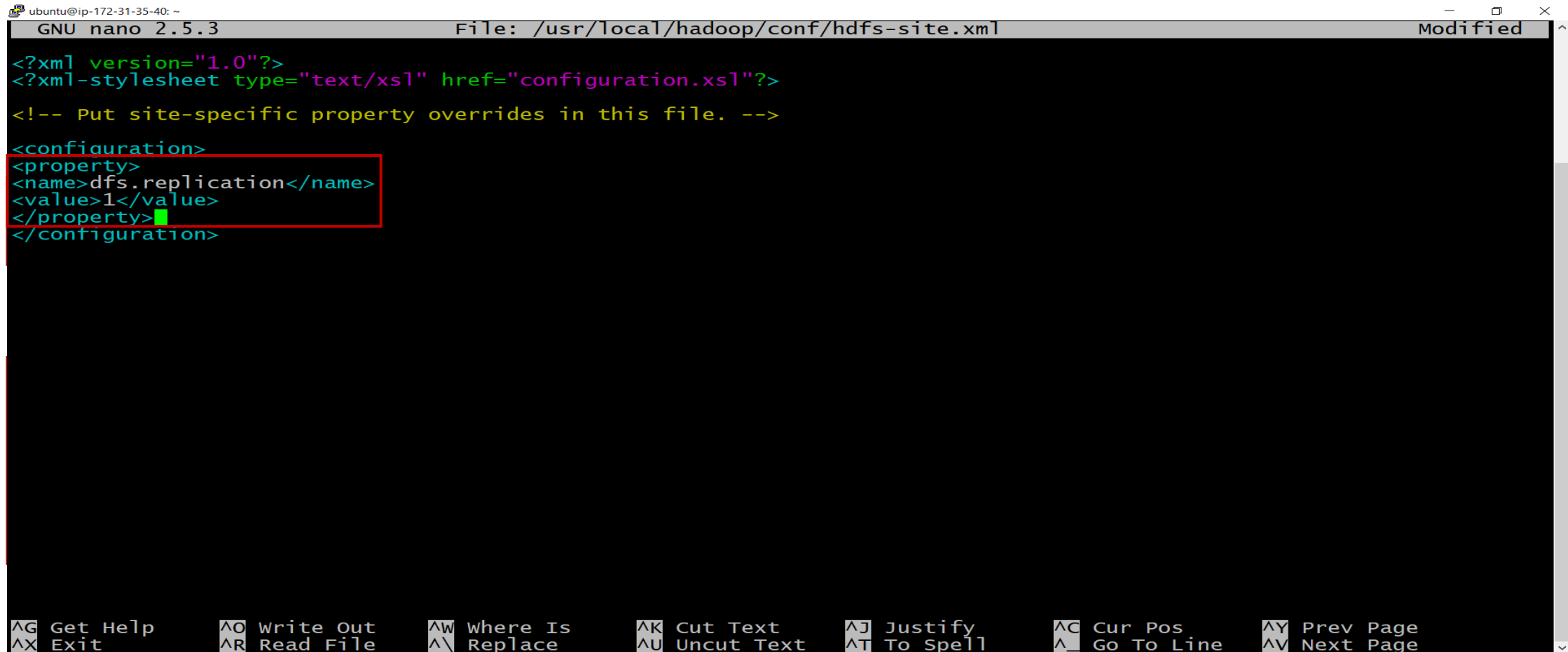
```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ nano /usr/local/hadoop/conf/hdfs-site.xml
```

press enter



16) Paste the following commands in-between the <configuration> & </configuration>

```
<property>
  <name>dfs.replication</name>
  <value>1</value>
</property>
```

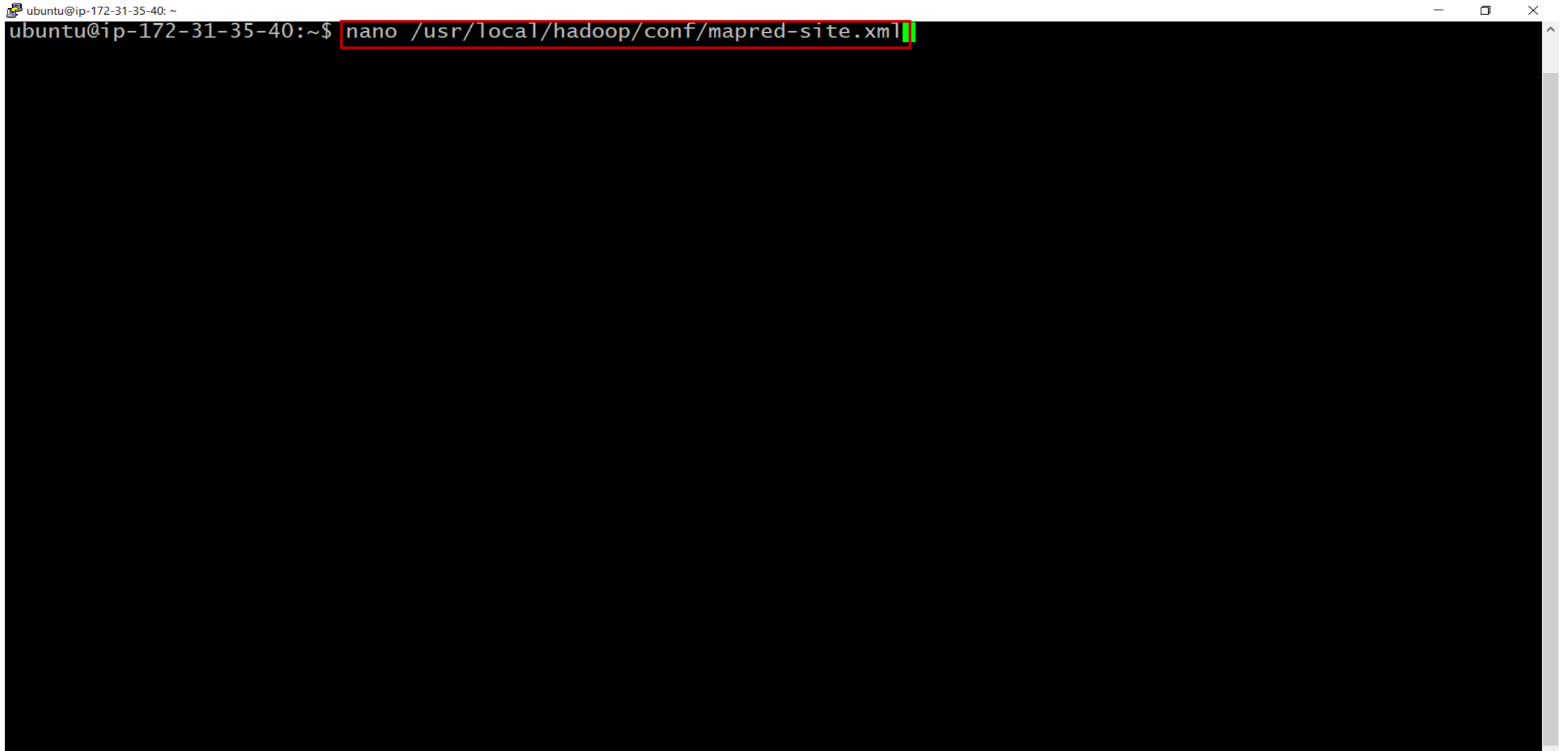


```
ubuntu@ip-172-31-35-40: ~
GNU nano 2.5.3 File: /usr/local/hadoop/conf/hdfs-site.xml Modified
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!-- Put site-specific property overrides in this file. -->
<configuration>
<property>
<name>dfs.replication</name>
<value>1</value>
</property>
</configuration>
```

^G Get Help ^O Write Out ^W Where Is ^K Cut Text ^J Justify ^C Cur Pos ^Y Prev Page  
^X Exit ^R Read File ^\ Replace ^U Uncut Text ^T To Spell ^\_ Go To Line ^V Next Page

ctrl + O (to save) > Enter > ctrl + X (to exit)

# 17) nano /usr/local/hadoop/conf/mapred-site.xml

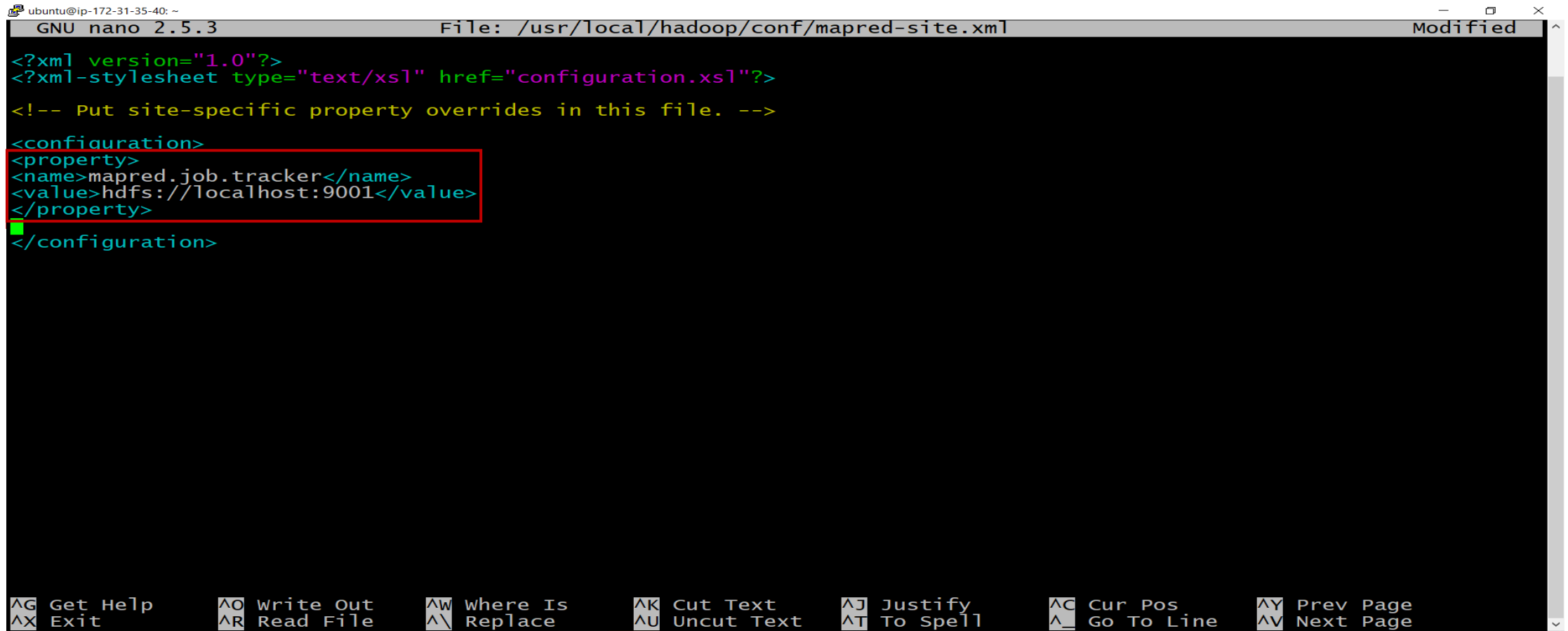
A terminal window with a black background and white text. The title bar at the top shows 'ubuntu@ip-172-31-35-40: ~' and standard window controls. The command prompt 'ubuntu@ip-172-31-35-40:~\$' is followed by the command 'nano /usr/local/hadoop/conf/mapred-site.xml' which is highlighted with a red rectangular box. A green cursor is positioned at the end of the command. The rest of the terminal area is empty.

```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ nano /usr/local/hadoop/conf/mapred-site.xml
```

press enter

- 18) Paste the following commands in-between the `<configuration>` & `</configuration>`  
(If you create your own DNS, you need to rename the localhost to your hostname e.g is)

```
<property>
<name>mapred.job.tracker</name>
<value>hdfs://localhost:9001</value>
</property>
```



```
ubuntu@ip-172-31-35-40: ~
GNU nano 2.5.3 File: /usr/local/hadoop/conf/mapred-site.xml Modified
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!-- Put site-specific property overrides in this file. -->
<configuration>
<property>
<name>mapred.job.tracker</name>
<value>hdfs://localhost:9001</value>
</property>
</configuration>
```

^G Get Help ^O Write Out ^W Where Is ^K Cut Text ^J Justify ^C Cur Pos ^Y Prev Page  
^X Exit ^R Read File ^\ Replace ^U Uncut Text ^T To Spell ^\_ Go To Line ^V Next Page

ctrl + O (to save) > Enter > ctrl + X (to exit)

## 19) mkdir /usr/local/hadoop/tmp

ubuntu@ip-172-31-35-40: ~

```
ubuntu@ip-172-31-35-40:~$ mkdir /usr/local/hadoop/tmp  
ubuntu@ip-172-31-35-40:~$
```

press enter

## 20) exec bash

ubuntu@ip-172-31-35-40: ~

```
ubuntu@ip-172-31-35-40:~$ exec bash  
ubuntu@ip-172-31-35-40:~$
```

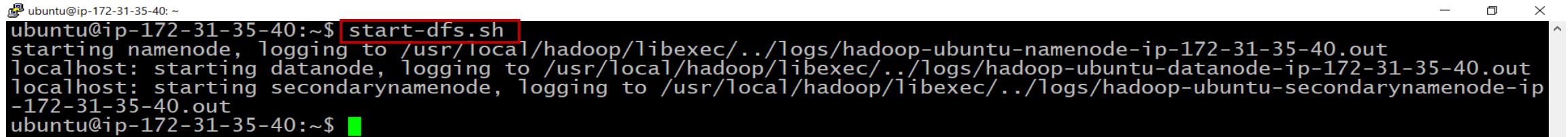
press enter

## 21) hadoop namenode -format

```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ hadoop namenode -format  
21/10/14 07:35:43 INFO namenode.NameNode: STARTUP_MSG:  
/*****  
STARTUP_MSG: Starting NameNode  
STARTUP_MSG: host = ip-172-31-35-40/172.31.35.40  
STARTUP_MSG: args = [-format]  
STARTUP_MSG: version = 1.2.1  
STARTUP_MSG: build = https://svn.apache.org/repos/asf/hadoop/common/branches/branch-1.2 -r 1503152; compiled by 'mat  
tf' on Mon Jul 22 15:23:09 PDT 2013  
STARTUP_MSG: java = 1.8.0_292  
*****/  
21/10/14 07:35:43 INFO util.GSet: Computing capacity for map BlocksMap  
21/10/14 07:35:43 INFO util.GSet: VM type = 64-bit  
21/10/14 07:35:43 INFO util.GSet: 2.0% max memory = 1013645312  
21/10/14 07:35:43 INFO util.GSet: capacity = 2^21 = 2097152 entries  
21/10/14 07:35:43 INFO util.GSet: recommended=2097152, actual=2097152  
21/10/14 07:35:44 INFO namenode.FSNamesystem: fsOwner=ubuntu  
21/10/14 07:35:44 INFO namenode.FSNamesystem: supergroup=supergroup  
21/10/14 07:35:44 INFO namenode.FSNamesystem: isPermissionEnabled=true  
21/10/14 07:35:44 INFO namenode.FSNamesystem: dfs.block.invalidate.limit=100  
21/10/14 07:35:44 INFO namenode.FSNamesystem: isAccessTokenEnabled=false accessKeyUpdateInterval=0 min(s), accessToken  
Lifetime=0 min(s)  
21/10/14 07:35:44 INFO namenode.FSEditLog: dfs.namenode.edits.toleration.length = 0  
21/10/14 07:35:44 INFO namenode.NameNode: Caching file names occurring more than 10 times  
21/10/14 07:35:44 INFO common.Storage: Image file /usr/local/hadoop/tmp/dfs/name/current/fsimage of size 112 bytes sav  
ed in 0 seconds.  
21/10/14 07:35:44 INFO namenode.FSEditLog: closing edit log: position=4, editlog=/usr/local/hadoop/tmp/dfs/name/curren  
t/edits  
21/10/14 07:35:44 INFO namenode.FSEditLog: close success: truncate to 4, editlog=/usr/local/hadoop/tmp/dfs/name/curren  
t/edits  
21/10/14 07:35:44 INFO common.Storage: Storage directory /usr/local/hadoop/tmp/dfs/name has been successfully formatte  
d.  
21/10/14 07:35:44 INFO namenode.NameNode: SHUTDOWN_MSG:  
/*****  
SHUTDOWN_MSG: Shutting down NameNode at ip-172-31-35-40/172.31.35.40  
*****/  
ubuntu@ip-172-31-35-40:~$
```

press enter

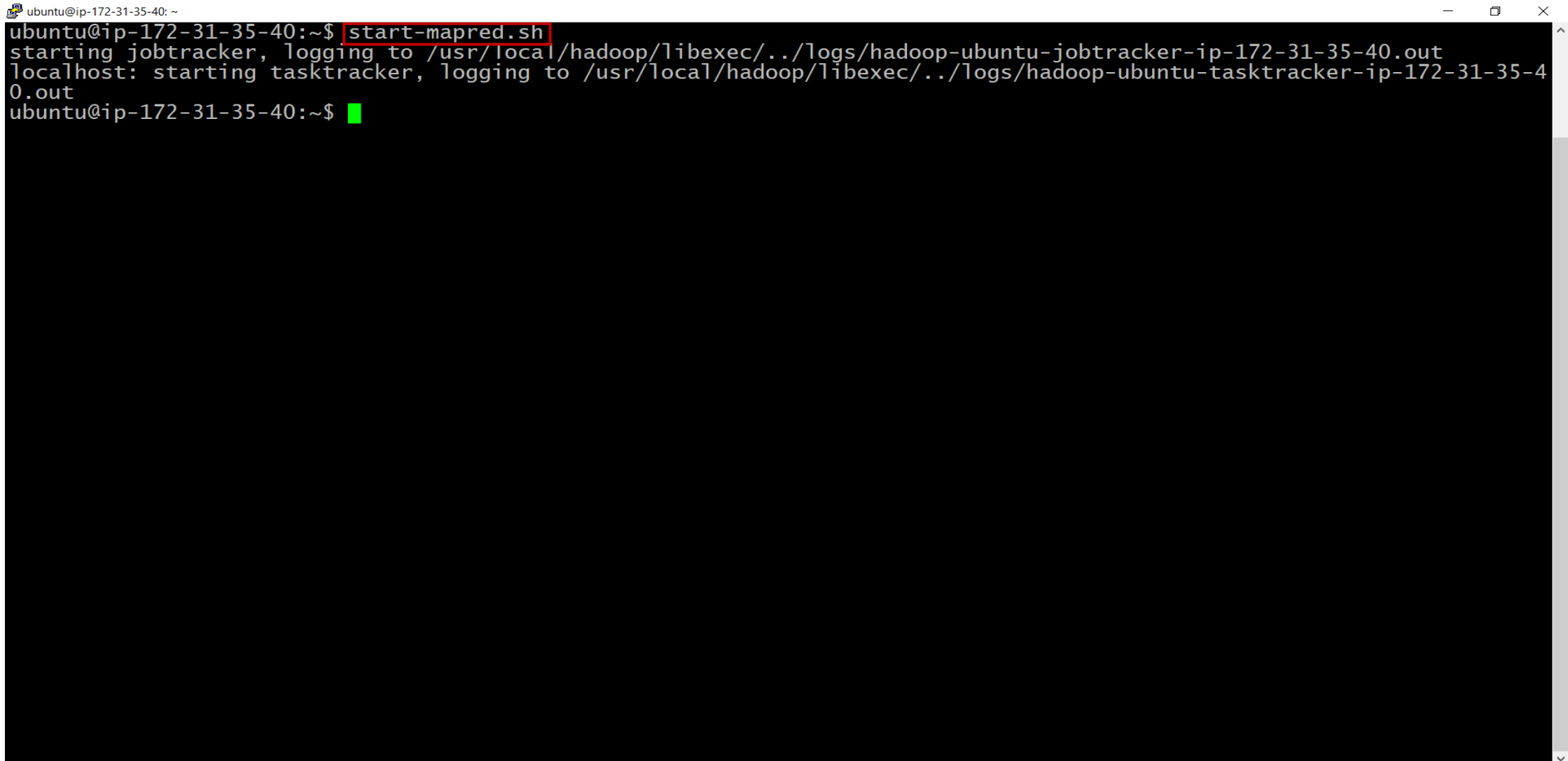
## 22) start-dfs.sh

A terminal window with a black background and white text. The title bar at the top shows 'ubuntu@ip-172-31-35-40: ~' and standard window controls. The command 'start-dfs.sh' is entered and highlighted with a red box. The output shows the starting of namenode, datanode, and secondarynamenode with their respective log paths. A green cursor is visible at the end of the last line.

```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ start-dfs.sh  
starting namenode, logging to /usr/local/hadoop/libexec/./logs/hadoop-ubuntu-namenode-ip-172-31-35-40.out  
localhost: starting datanode, logging to /usr/local/hadoop/libexec/./logs/hadoop-ubuntu-datanode-ip-172-31-35-40.out  
localhost: starting secondarynamenode, logging to /usr/local/hadoop/libexec/./logs/hadoop-ubuntu-secondarynamenode-ip-172-31-35-40.out  
ubuntu@ip-172-31-35-40:~$
```

press enter

## 23) start-mapred.sh

A terminal window with a black background and white text. The window title is 'ubuntu@ip-172-31-35-40: ~'. The prompt is 'ubuntu@ip-172-31-35-40:~\$'. The command 'start-mapred.sh' is entered and highlighted with a red box. The output shows the starting of the jobtracker and tasktracker, with their respective log paths. The prompt returns to 'ubuntu@ip-172-31-35-40:~\$' with a green cursor.

```
ubuntu@ip-172-31-35-40:~$ start-mapred.sh
starting jobtracker, logging to /usr/local/hadoop/libexec/../logs/hadoop-ubuntu-jobtracker-ip-172-31-35-40.out
localhost: starting tasktracker, logging to /usr/local/hadoop/libexec/../logs/hadoop-ubuntu-tasktracker-ip-172-31-35-40.out
ubuntu@ip-172-31-35-40:~$
```

press enter



## 24) jps

```
ubuntu@ip-172-31-35-40: ~  
ubuntu@ip-172-31-35-40:~$ jps  
26144 TaskTracker  
25729 DataNode  
25892 SecondaryNameNode  
26005 JobTracker  
26405 Jps  
25583 NameNode  
ubuntu@ip-172-31-35-40:~$
```

press enter

## 25) Open the security groups

The screenshot displays the AWS Management Console interface for the 'ap-south-1' region. The left-hand navigation pane shows the 'Instances' section expanded, with 'Instances' selected. The main content area shows a list of instances, with 'kida\_02' (ID: i-087e2cd356526e2d1) selected. Below the list, the 'Instance: i-087e2cd356526e2d1 (kida\_02)' details page is open. The 'Security' tab is selected, and the 'Security groups' section is expanded, showing a single security group: 'sg-01e71f1ed24be1fe6 (launch-wizard-17)'. A red box highlights this security group name, and a red circle with the number '2' is placed next to it. Another red circle with the number '1' is placed over the 'Security' tab label. The top of the console shows the 'Instances (1/1)' header and a search bar. The bottom of the console shows a 'Filter rules' search bar.

ap-south-1.console.aws.amazon.com/ec2/v2/home?region=ap-south-1#Instances:

Services

Search for services, features, marketplace products, and docs [Alt+S]

Anuj Gogate Mumbai Support

New EC2 Experience Tell us what you think

EC2 Dashboard

EC2 Global View

Events

Tags

Limits

Instances

Instances New

Instance Types

Launch Templates

Spot Requests

Savings Plans

Reserved Instances New

Dedicated Hosts

Capacity Reservations

Images

AMIs

Elastic Block Store

Volumes

Snapshots

Instances (1/1) Info

Filter instances

<input checked="" type="checkbox"/>	Name	Instance ID	Instance state	Instance type	Status check	Alarm status	Availability Zone	Public IPv4 DNS
<input checked="" type="checkbox"/>	kida_02	i-087e2cd356526e2d1	Running	t2.micro	2/2 checks passed		ap-south-1a	ec2-13-233-148-

Instance: i-087e2cd356526e2d1 (kida\_02)

Details Security Networking Storage Status checks Monitoring Tags

Security details

IAM Role	Owner ID	Launch time
-	509149891835	Wed Oct 13 2021 17:45:17 GMT+0530 (India Standard Time)

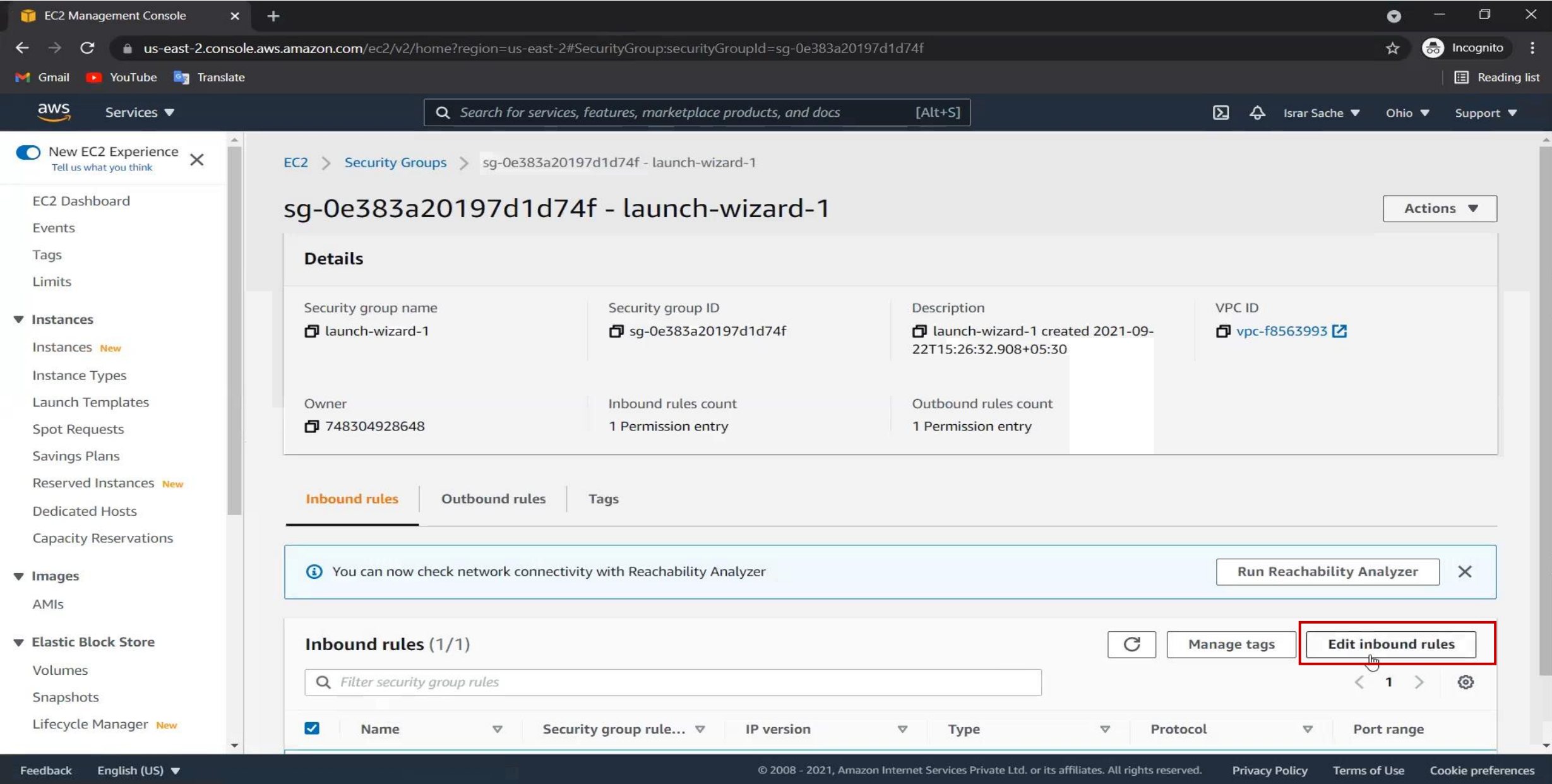
Security groups

sg-01e71f1ed24be1fe6 (launch-wizard-17)

Inbound rules

Filter rules

# 26) Click on edit inbound rules



## 27) Click on Add rule for 4 times

The screenshot shows the AWS Management Console interface for editing inbound rules on a security group. The breadcrumb trail indicates the path: EC2 > Security Groups > sg-0e383a20197d1d74f - launch-wizard-1 > Edit inbound rules. The main heading is 'Edit inbound rules' with an 'Info' link. Below the heading is a descriptive sentence: 'Inbound rules control the incoming traffic that's allowed to reach the instance.'

The 'Inbound rules' section contains a table with the following columns: Security group rule ID, Type, Protocol, Port range, Source, and Description - optional. A single rule is listed with ID 'sgr-0d187120c1797e931', Type 'SSH', Protocol 'TCP', Port range '22', and Source 'Custom'. Below the table, the 'Add rule' button is highlighted with a red rectangular box, and a mouse cursor is pointing at it. To the right of the table, there is a search bar, a text input field, and a 'Delete' button. Below the search bar, a dropdown menu shows '0.0.0.0/0' with a close button.

At the bottom right of the console, there are three buttons: 'Cancel', 'Preview changes', and 'Save rules'.

The footer of the console includes 'Feedback', 'English (US)', and copyright information: '© 2008 - 2021, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved.' It also links to 'Privacy Policy', 'Terms of Use', and 'Cookie preferences'.

## 28) Configure the port range to 50070, 50030, 50075, 50010 & change there source to My Ip and click on save rules

The screenshot shows the AWS EC2 Management Console interface for configuring inbound security group rules. The browser address bar indicates the URL: `us-east-2.console.aws.amazon.com/ec2/v2/home?region=us-east-2#ModifyInboundSecurityGroupRules:securityGroupId=sg-0e383a20197d1d74f`. The page title is "Inbound rules" with an "Info" link. The table lists the following rules:

Security group rule ID	Type	Protocol	Port range	Source	Description - optional	Actions
sgr-0d187120c1797e931	SSH	TCP	22	Custom	0.0.0.0/0	Delete
-	Custom TCP	TCP	50070	My IP	103.47.153.162/32	Delete
-	Custom TCP	TCP	50030	My IP	103.47.153.162/32	Delete
-	Custom TCP	TCP	50075	My IP	103.47.153.162/32	Delete
-	Custom TCP	TCP	50010	My IP	103.47.153.162/32	Delete

At the bottom of the console, there are three buttons: "Cancel", "Preview changes", and "Save rules". The "Save rules" button is highlighted with a red box.

# 29) Now you have successfully modified on the security groups

EC2 Management Console

us-east-2.console.aws.amazon.com/ec2/v2/home?region=us-east-2#SecurityGroup:group-id=sg-0e383a20197d1d74f

Gmail YouTube Translate

aws Services

Search for services, features, marketplace products, and docs [Alt+S]

Israr Sache Ohio Support

New EC2 Experience

EC2 Dashboard

Events

Tags

Limits

Instances

Instance Types

Launch Templates

Spot Requests

Savings Plans

Reserved Instances

Dedicated Hosts

Capacity Reservations

Images

AMIs

Elastic Block Store

Volumes

Snapshots

Lifecycle Manager

Inbound security group rules successfully modified on security group (sg-0e383a20197d1d74f | launch-wizard-1)

Details

EC2 > Security Groups > sg-0e383a20197d1d74f - launch-wizard-1

sg-0e383a20197d1d74f - launch-wizard-1

Actions

Details

Security group name	Security group ID	Description	VPC ID
launch-wizard-1	sg-0e383a20197d1d74f	launch-wizard-1 created 2021-09-22T15:26:32.908+05:30	vpc-f8563993
Owner	Inbound rules count	Outbound rules count	
748304928648	5 Permission entries	1 Permission entry	

Inbound rules

Outbound rules

Tags

You can now check network connectivity with Reachability Analyzer

Run Reachability Analyzer

Inbound rules (5)

Manage tags

Edit inbound rules

us-east-2.console.aws.amazon.com/ec2/v2/home?region=us-east-2#LaunchTemplates:

© 2008 - 2021, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved.

Privacy Policy

Terms of Use

Cookie preferences

30) To review, open Incognito mode and launch yourpublicIPv4address:50070

← → ↻ ⚠ Not secure | 13.233.148.222:50070/dfshealth.jsp ☆ ⚙ 🌐 ⋮

# NameNode 'localhost:9000'

Started: Thu Oct 14 07:42:02 UTC 2021

Version: 1.2.1, r1503152

Compiled: Mon Jul 22 15:23:09 PDT 2013 by mattf

Upgrades: There are no upgrades in progress.

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

## Cluster Summary

8 files and directories, 1 blocks = 9 total. Heap Size is 23.57 MB / 966.69 MB (2%)

Configured Capacity

:

7.69 GB

DFS Used

:

40 KB

Non DFS Used

:

1.88 GB

DFS Remaining

:

5.81 GB

DFS Used%

:

0 %

DFS Remaining%

:

75.56 %

[Live Nodes](#)

:

1

[Dead Nodes](#)

:

0

[Decommissioning Nodes](#)

:

0

Number of Under-Replicated Blocks

:

0

## NameNode Storage:

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

THU OCT 14 10:04:10 UTC 2021

31) Also, open your public IPv4 address:50030

← → ↺ ⚠ Not secure | 13.233.148.222:50030/jobtracker.jsp ☆ ⚙ 👤 ⋮

Quick Links

# localhost Hadoop Map/Reduce Administration

State: RUNNING  
Started: Thu Oct 14 07:44:00 UTC 2021  
Version: 1.2.1, r1503152  
Compiled: Mon Jul 22 15:23:09 PDT 2013 by mattf  
Identifier: 202110140744  
SafeMode: OFF

---

## Cluster Summary (Heap Size is 15.5 MB/966.69 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	<a href="#">1</a>	0	0	0	0	2	2	4.00	<a href="#">0</a>	<a href="#">0</a>	<a href="#">0</a>

---

## Scheduling Information

Queue Name	State	Scheduling Information
<a href="#">default</a>	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

none

---

## Retired Jobs