

TRƯỜNG ĐẠI HỌC XÂY DỰNG HÀ NỘI
KHOA CÔNG NGHỆ THÔNG TIN



BIG DATA

Cài đặt Hadoop, MongoDB, Redis

Lớp 64CS3	
Họ và Tên	Ngô Đức Thịnh
MSSV	189464
Giảng viên hướng dẫn	ThS. Nguyễn Đình Quý

Hà Nội, 05/2022

Mục lục

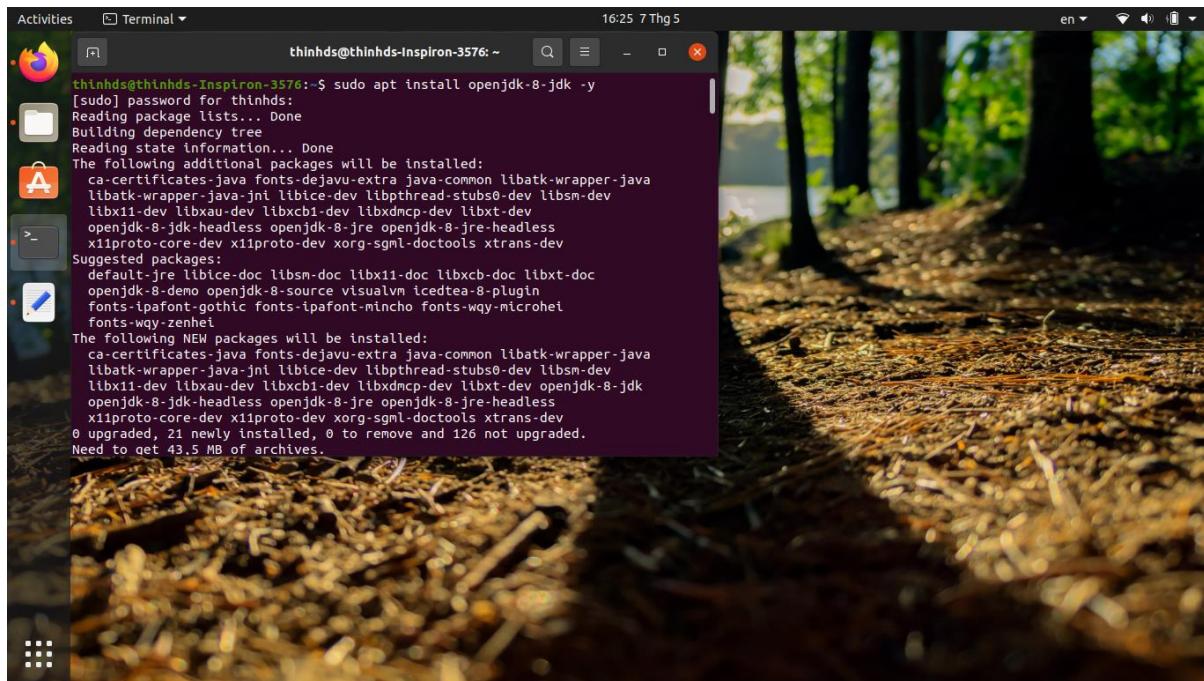
Chương 1: Cài đặt Hadoop, MongoDB, Redis	3
1.1. Cài đặt Hadoop	3
1.2. Cài đặt MongoDB	13
1.3. Cài đặt Redis	18
Chương 2: Thực nghiệm các thao tác cơ bản trong MongoDB, Redis	21
2.1. Thực nghiệm MongoDB	21
2.2. Thực nghiệm Redis	21
Chương 3: So sánh Mongo và Redis	24
3.1. Cấu trúc database	24
3.2. Ứng dụng	24
3.3. Sử dụng thế nào	24
3.4. Nhược điểm	25

Chương 1: Cài đặt Hadoop, MongoDB, Redis

1.1. Cài đặt Hadoop

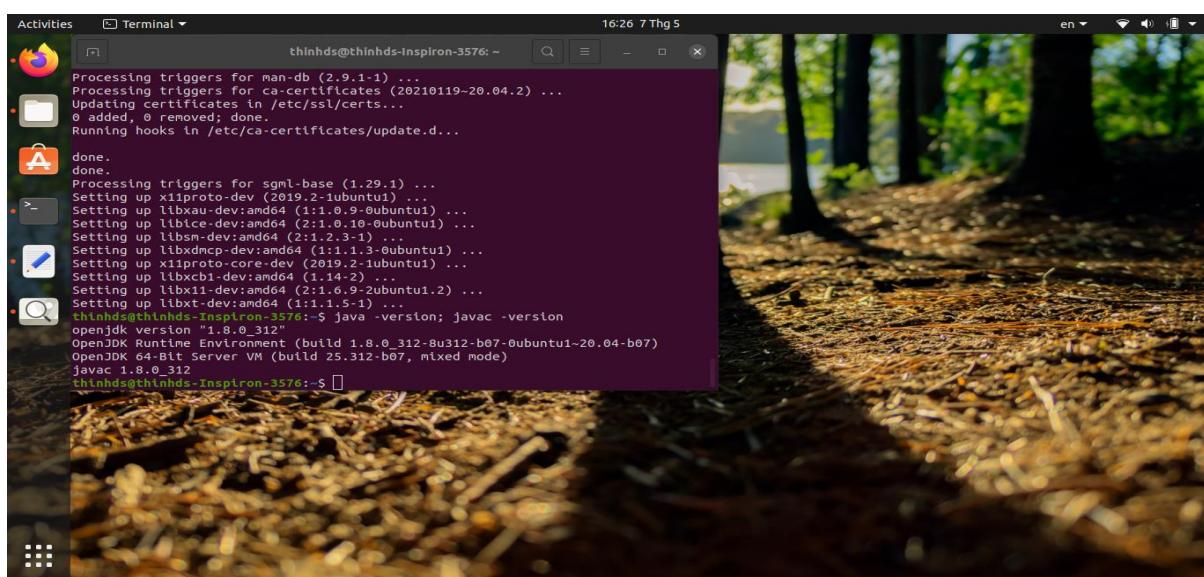
Bước 1: Cài đặt java

```
sudo apt install openjdk-8-jdk -y
```



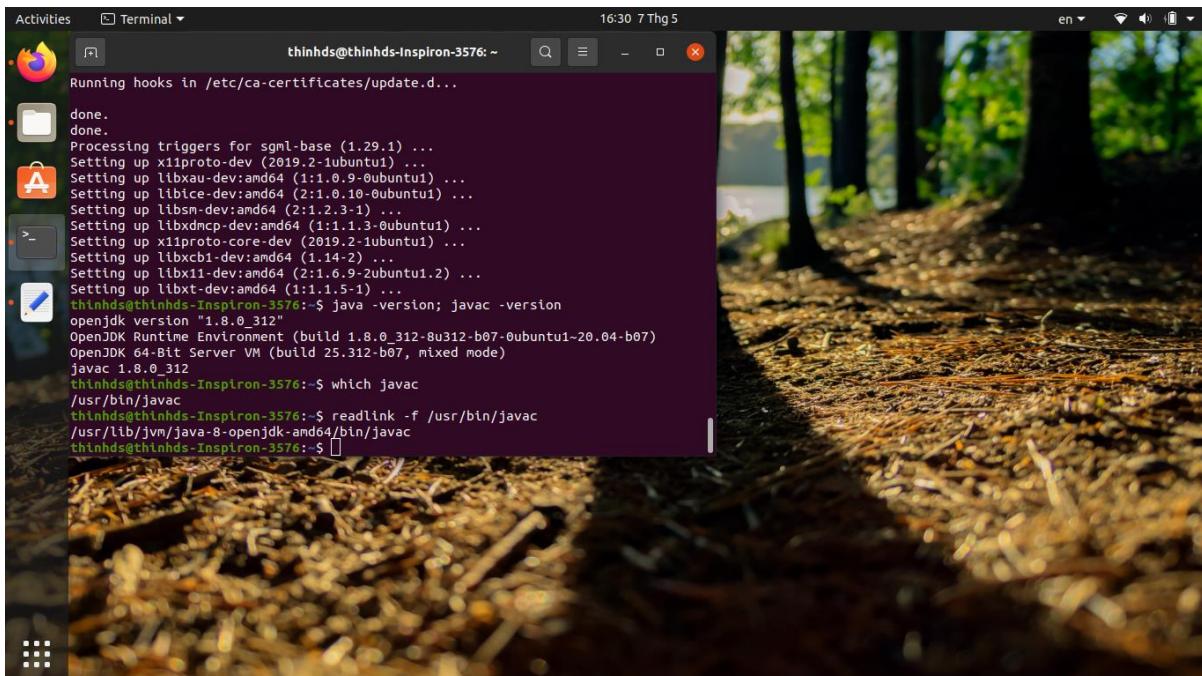
Bước 2: Kiểm tra cài đặt java

```
java -version; javac -version
```



Bước 3: Kiểm tra đường dẫn

```
which javac  
readlink -f /usr/bin/javac
```

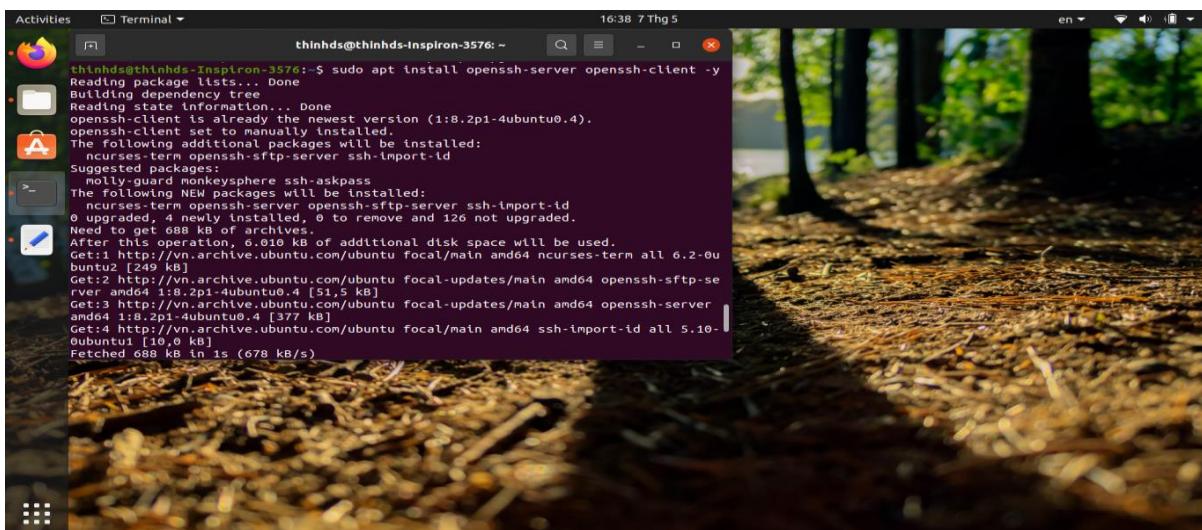


Bước 4: Tải Hadoop

Các bạn vào link: <https://hadoop.apache.org/releases.html> để tải phiên bản hadoop về máy. Ta tải bản hadoop 3.3.2 binary về máy và giải nén vào thư mục Home.

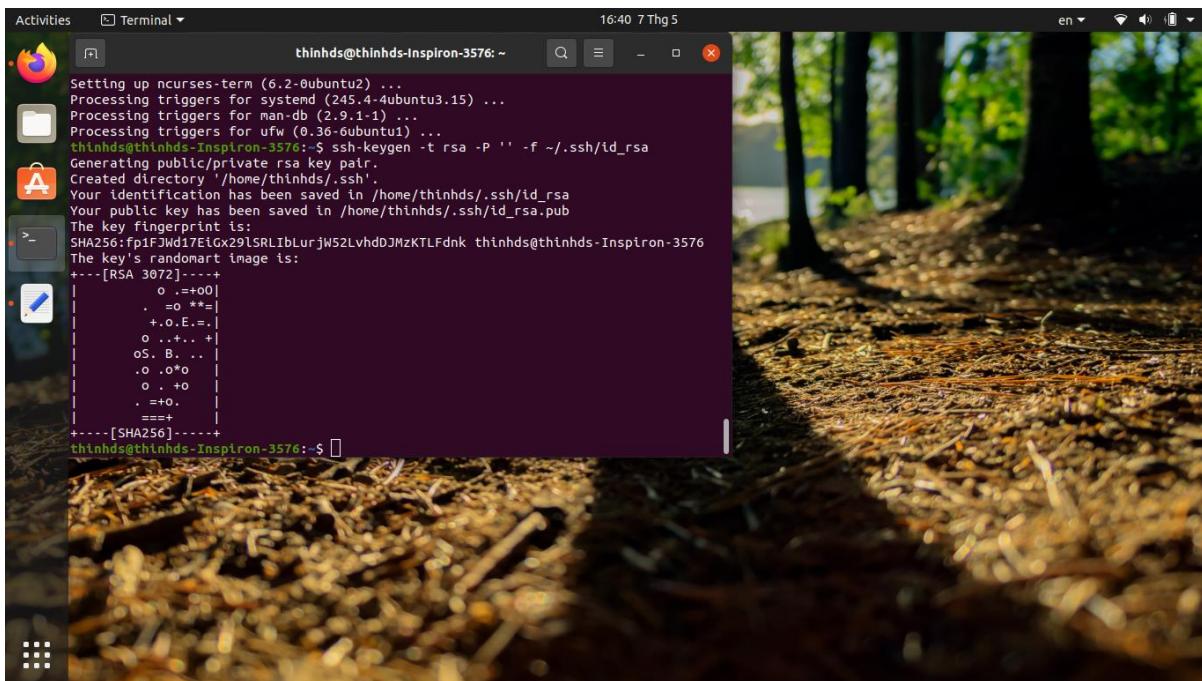
Bước 5: Cài đặt openSSH

```
sudo apt install openssh-server openssh-client -y
```



Bước 6: Tạo và cài đặt SSH Certificates

```
ssh-keygen -t rsa -P "" -f ~/.ssh/id_rsa
```

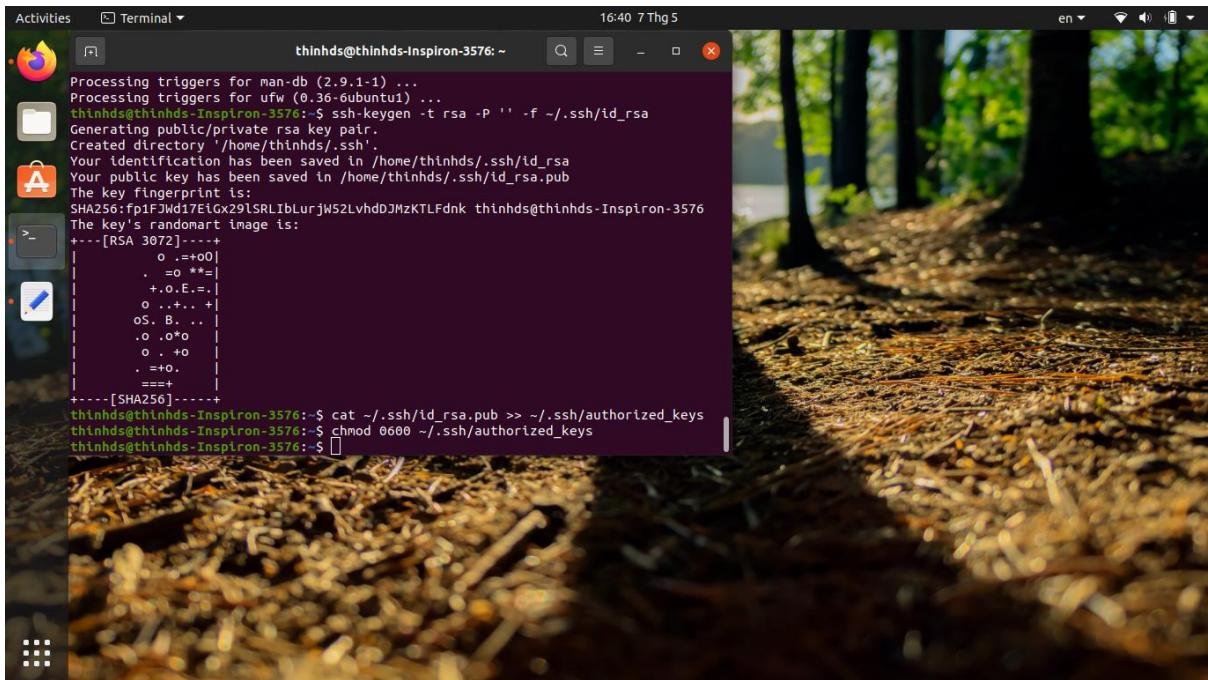


Lưu lại thông tin:

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

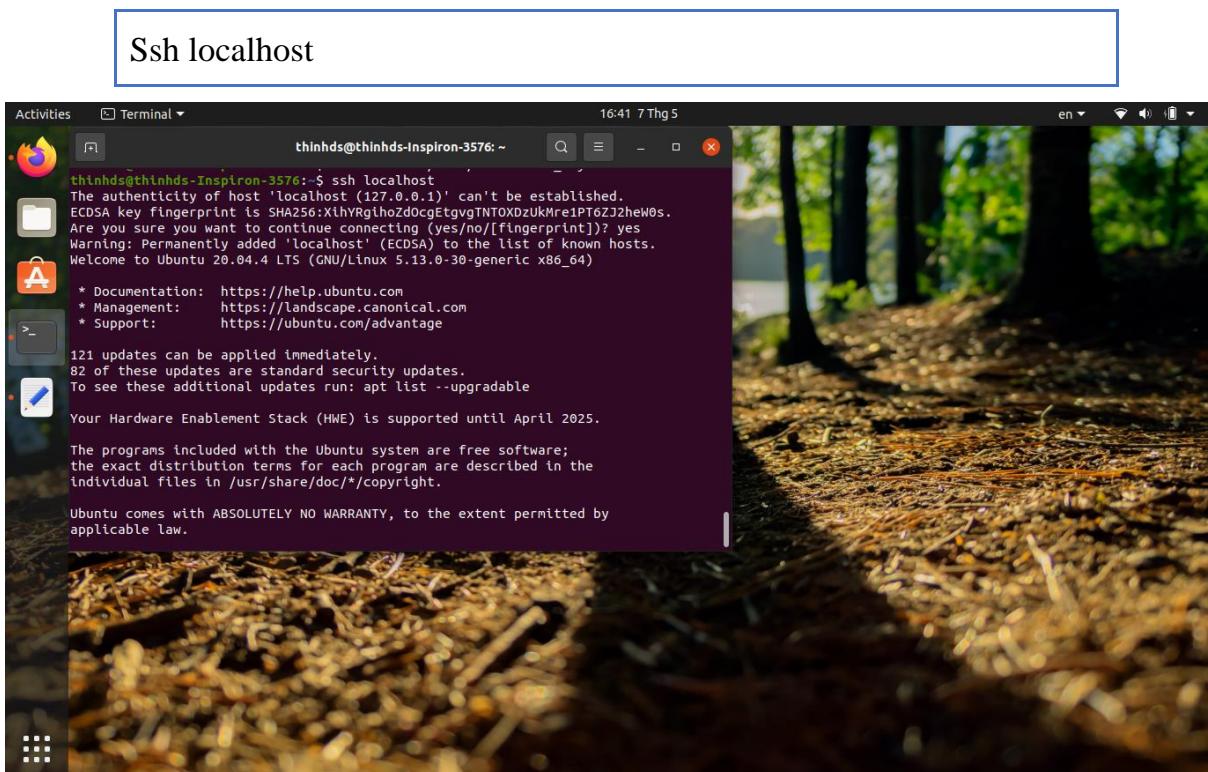
Cấp quyền cho user:

```
chmod 0600 ~/.ssh/authorized_keys
```



```
Activities Terminal 16:40 7 Thg 5
thinhds@thinhds-Inspiron-3576: ~
Processing triggers for man-db (2.9.1-1) ...
Processing triggers for ufw (0.36-0ubuntu1) ...
thinhds@thinhds-Inspiron-3576: ~$ ssh-keygen -t rsa -P '' -f ~/.ssh/id_rsa
Generating public/private rsa key pair.
Created directory '/home/thinhds/.ssh'.
Your identification has been saved in /home/thinhds/.ssh/id_rsa
Your public key has been saved in /home/thinhds/.ssh/id_rsa.pub
The key fingerprint is:
SHA256:fp1FjWd17EiGx29LSRLibLurjW52LvhDJDmzKTLFdNk thinhds@thinhds-Inspiron-3576
The key's randomart image is:
+---[RSA 3072]----+
|          o =+o |
|         .o **= |
|        +o.E.=. |
|       o ..+.. + |
|      oS. B... |
|     .o .o*o |
|    o . +o |
|   . =+o. |
|  ==+=+ |
+---[SHA256]-----+
thinhds@thinhds-Inspiron-3576: ~$ cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
thinhds@thinhds-Inspiron-3576: ~$ chmod 0600 ~/.ssh/authorized_keys
thinhds@thinhds-Inspiron-3576: ~$
```

Kiểm tra ssh:



Ssh localhost

```
Activities Terminal 16:41 7 Thg 5
thinhds@thinhds-Inspiron-3576: ~
thinhds@thinhds-Inspiron-3576: ~$ ssh localhost
The authenticity of host 'localhost (127.0.0.1)' can't be established.
ECDSA key fingerprint is SHA256:XihYRgihoZdOcgEtvgvINTOXDzukhre1PT6ZJ2heW0s.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'localhost' (ECDSA) to the list of known hosts.
Welcome to Ubuntu 20.04.4 LTS (GNU/Linux 5.13.0-38-generic x86_64)

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

121 updates can be applied immediately.
82 of these updates are standard security updates.
To see these additional updates run: apt list --upgradable

Your Hardware Enablement Stack (HWE) is supported until April 2025.

The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.
```

Địa chỉ localhost ‘127.0.0.1’, kiểm tra localhost của máy nếu khác thì sửa lại.

Bước 7: Cấu hình lại các file trong hadoop

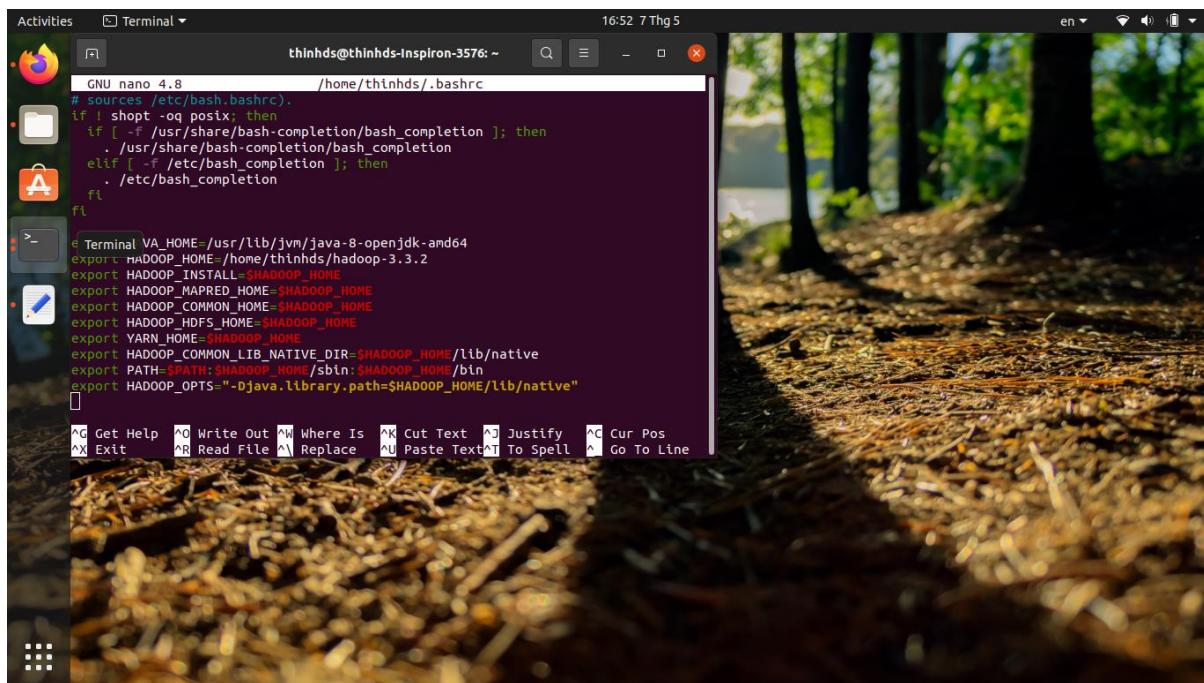
Cấu hình file bashrc:

```
nano ~/.bashrc
```

Copy vào cuối file để setup các đường dẫn env

```
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
export HADOOP_HOME=/home/thinhds/hadoop-3.3.2
export HADOOP_INSTALL=$HADOOP_HOME
export HADOOP_MAPRED_HOME=$HADOOP_HOME
export HADOOP_COMMON_HOME=$HADOOP_HOME
export HADOOP_HDFS_HOME=$HADOOP_HOME
export YARN_HOME=$HADOOP_HOME
export
HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export PATH=$PATH:$HADOOP_HOME/sbin:$HADOOP_HOME/bin
export HADOOP_OPTS="-
Djava.library.path=$HADOOP_HOME/lib/native"
```

Lưu ý: tên user và phiên bản hadoop sẽ khác ở mỗi máy. Và lưu lại file khi đã hoàn thành thêm đường dẫn.

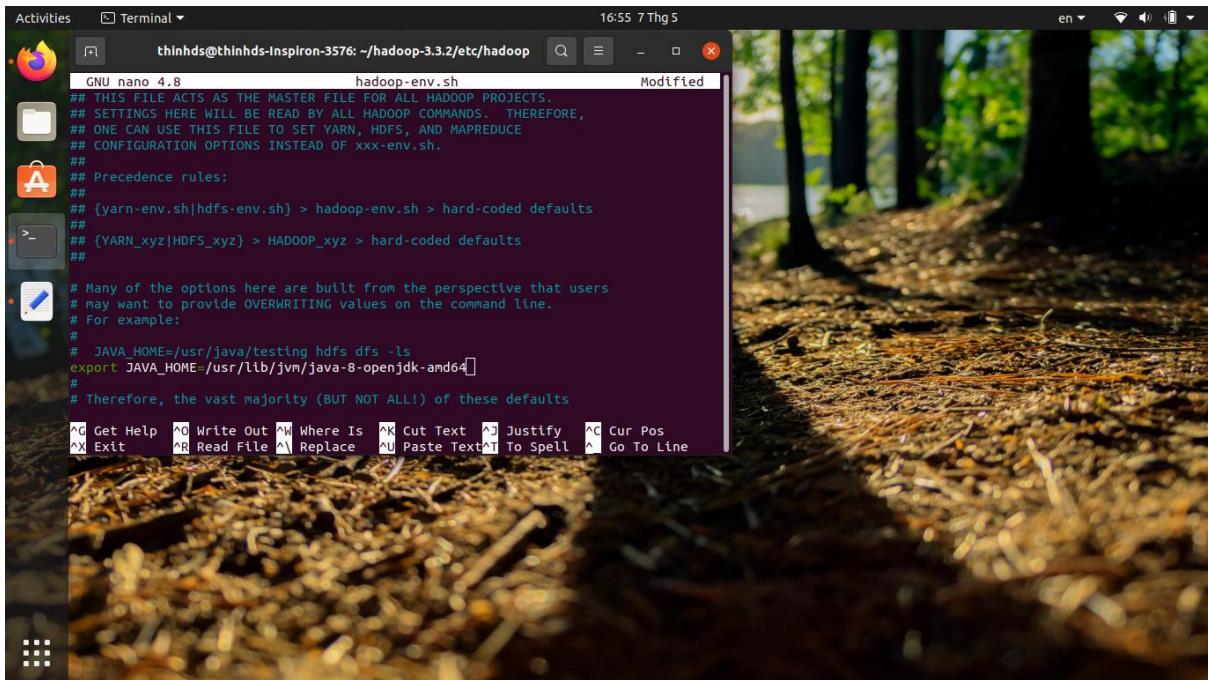


Cấu hình file hadoop-env.sh:

```
nano hadoop-env.sh
```

Thêm địa chỉ JAVA_HOME cho hadoop:

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



Cấu hình file core-site.xml:

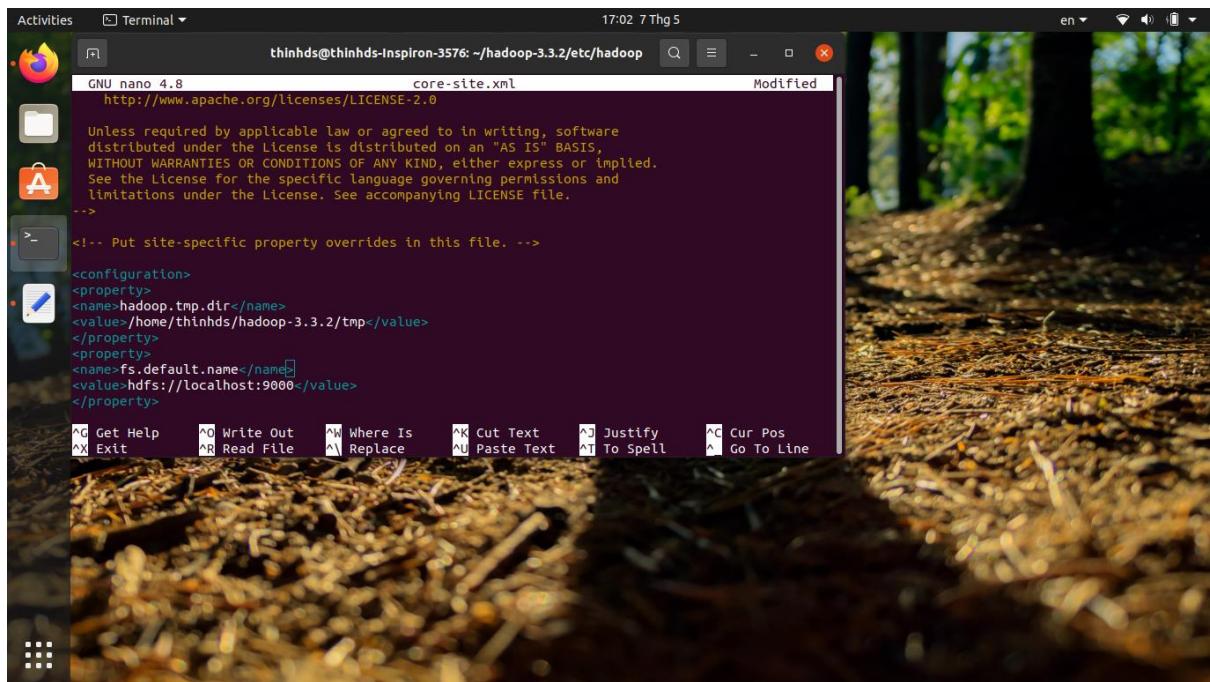
Tạo folder trung chuyển tmp rồi cấp quyền tất cả user truy cập vào nó.

```
mkdir -p /home/thinhds/hadoop-3.3.2/tmp
```

Vào file core-site.xml và copy lệnh vào thẻ configuration.

```
nano core-site.xml
```

```
<property>
<name>hadoop.tmp.dir</name>
<value>/home/thinhds/hadoop-3.3.2/tmp</value>
</property>
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:9000</value>
</property>
```



A screenshot of a Linux desktop environment showing a terminal window titled "core-site.xml". The terminal displays the contents of the file, which includes the Apache License header and configuration for HDFS and YARN. The terminal window has a dark theme with white text. The desktop background is a photograph of a forest floor with sunlight filtering through the trees.

```
GNU nano 4.8 core-site.xml Modified
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.

-->

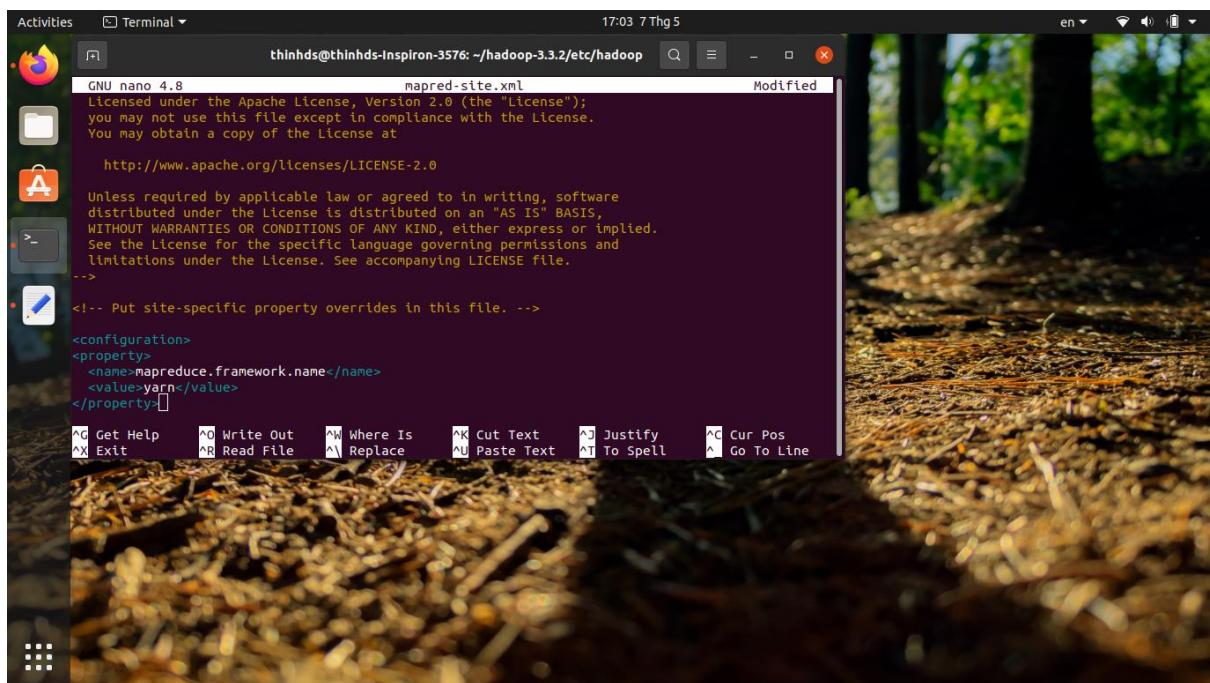
<!-- Put site-specific property overrides in this file. -->

<configuration>
<property>
<name>hadoop.tmp.dir</name>
<value>/home/thinhds/hadoop-3.3.2/tmp</value>
</property>
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:9000</value>
</property>
```

Cấu hình file mapred-site.xml:

nano mapred-site.xml

```
<property>
<name>mapreduce.framework.name</name>
<value>yarn</value>
</property>
```



A screenshot of a Linux desktop environment showing a terminal window titled "mapred-site.xml". The terminal displays the contents of the file, which includes the Apache License header and configuration for the MapReduce framework. The terminal window has a dark theme with white text. The desktop background is a photograph of a forest floor with sunlight filtering through the trees.

```
GNU nano 4.8 mapred-site.xml Modified
http://www.apache.org/licenses/LICENSE-2.0

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>mapreduce.framework.name</name>
<value>yarn</value>
</property>
```

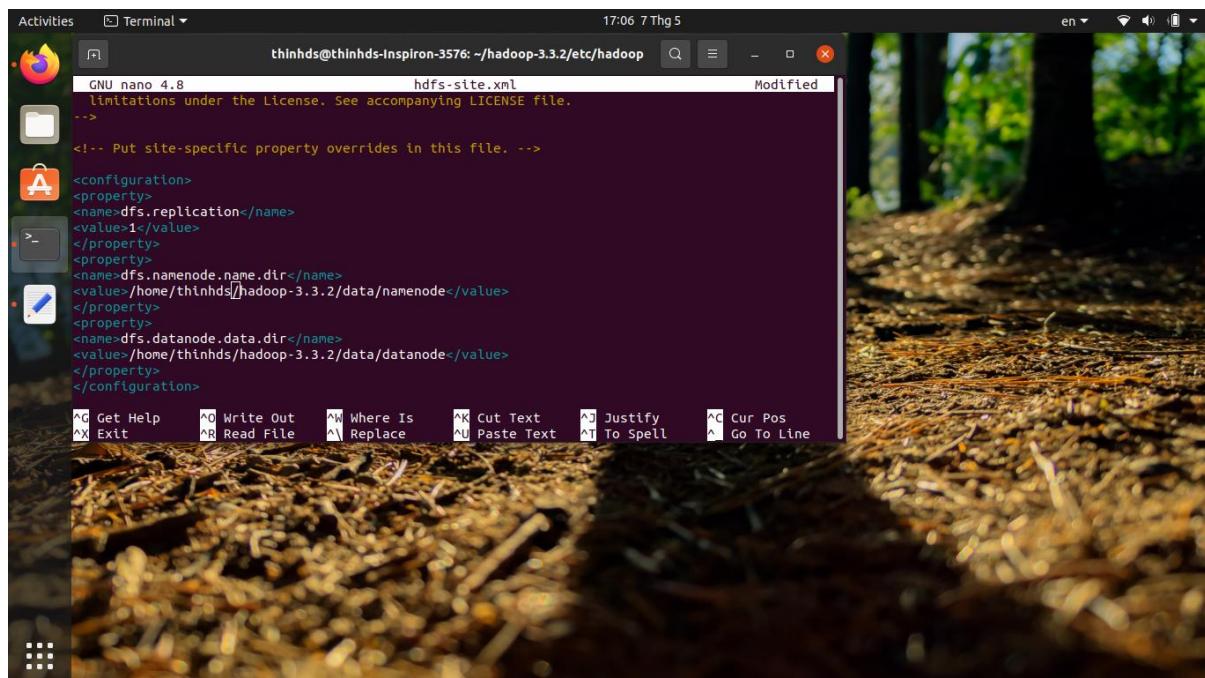
Cấu hình file hdfs-site.xml:

Trước tiên, ta đi tạo 2 folder namenode và datanode

```
nano hdfs-site.xml
```

```
mkdir -p /home/thinhds/hadoop-3.3.2/data/namenode  
mkdir -p /home/thinhds/hadoop-3.3.2/data/datanode
```

```
<property>  
<name>dfs.replication</name>  
<value>1</value>  
</property>  
<property>  
<name>dfs.namenode.name.dir</name>  
<value>/home/thinhds/hadoop-3.3.2/data/namenode</value>  
</property>  
<property>  
<name>dfs.datanode.data.dir</name>  
<value>/home/thinhds/hadoop-3.3.2/data/datanode</value>  
</property>
```



Cấu hình file yarn-site.xml:

```
nano yarn-site.xml
```

```
<property>
  <name>yarn.nodemanager.aux-services</name>
  <value>mapreduce_shuffle</value>
</property>
<property>
  <name>yarn.nodemanager.aux-services.mapreduce.shuffle.class</name>
  <value>org.apache.hadoop.mapred.ShuffleHandler</value>
</property>
<property>
  <name>yarn.resourcemanager.hostname</name>
  <value>127.0.0.1</value>
</property>
<property>
  <name>yarn.acl.enable</name>
  <value>0</value>
</property>
<property>
  <name>yarn.nodemanager.env-whitelist</name>
  <value>JAVA_HOME,HADOOP_COMMON_HOME,HADOOP_HDFS_HOME,HADOOP_CONF_DIR,CLASSPATH_PERPEND_DISTCACHE,HADOOP_YARN_HOME,HADOOP_MAPRED_HOME</value>
</property>
```

Bước 8: Format file hệ thống Hadoop mới

hdfs namenode -format

```
Activities Terminal 17:11 7 Thg 5
thinhsds@thinhsds-Inspiron-3576: ~/hadoop-3.3.2/etc/hadoop
thinhsds@thinhsds-Inspiron-3576:~/hadoop-3.3.2/etc/hadoop$ hdfs namenode -format
WARNING: /home/thinhsds/hadoop-3.3.2/logs does not exist. Creating.
2022-05-07 17:10:49,557 INFO namenode.NameNode: STARTUP_MSG:
/*****
STARTUP_MSG: Starting NameNode
STARTUP_MSG: host = thinhsds-Inspiron-3576/127.0.1.1
STARTUP_MSG: args = [-format]
STARTUP_MSG: version = 3.3.2
STARTUP_MSG: classpath = /home/thinhsds/hadoop-3.3.2/etc/hadoop:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/curator-client-4.2.0.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/nimbus-jose-jwt-9.8.1.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/jetty-io-9.4.43.v20210629.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/kerby-config-1.0.1.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/jersey-json-1.19.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/asm-5.0.4.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/jersey-servlet-1.19.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/curator-recipes-4.2.0.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/token-provider-1.0.1.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/kerby-pkix-1.0.1.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/gson-2.8.9.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/commons-beanutils-1.9.4.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/jax-b-apt-2.2.11.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/avro-1.7.7.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/jackson-xc-1.9.13.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/protobuf-java-2.5.0.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/failover-eaccess-1.0.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/jackson-core-asl-1.9.13.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/jetty-xml-9.4.43.v20210629.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/jersey-server-1.19.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/jsc-0.1.55.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/guava-27.0.jre.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/jackson-jaxrs-1.9.13.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/common/lib/httpcore-4.4.13.jar:/home/thinhsds/hadoop-3.3.2/share/hadoop/c
```

Bước 9: Chạy namenode,datanode, secondary namenode

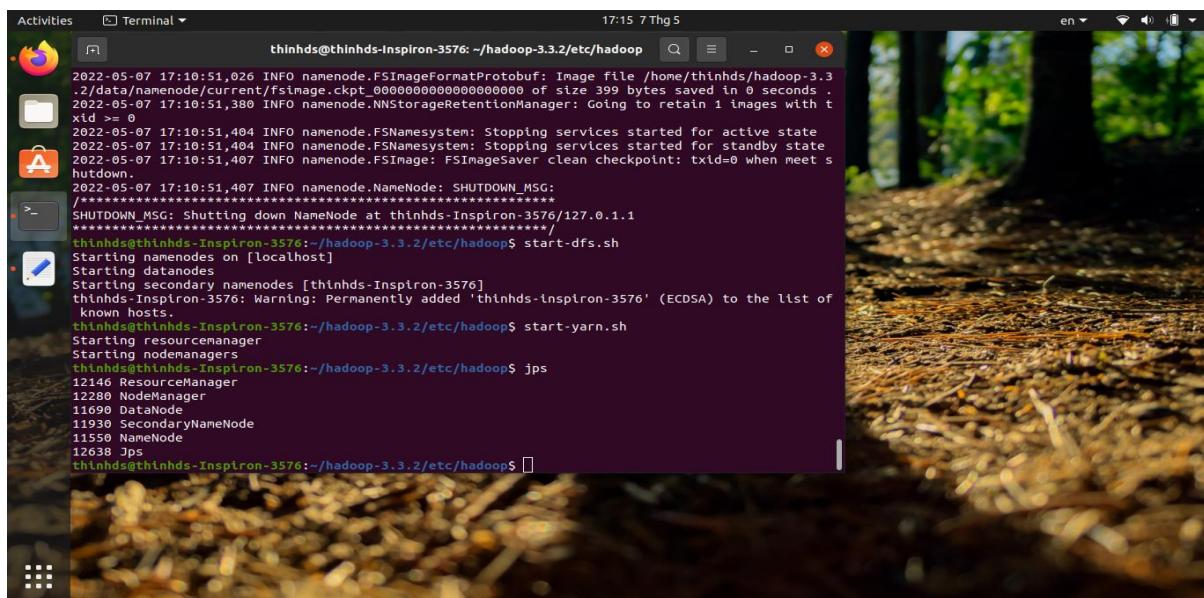
```
start-dfs.sh
```

Bước 10: chạy yarn : resourcemanager, nodemanagers

```
start-yarn.sh
```

Bước 11: Kiểm tra liên kết

```
jps
```



The screenshot shows a terminal window titled "Terminal" with the command "start-dfs.sh" being run. The output of the command shows the startup of NameNodes, DataNodes, and SecondaryNameNodes. Following this, the command "start-yarn.sh" is run, starting the ResourceManager and NodeManagers. Finally, the "jps" command is run to list active Java processes, which includes the ResourceManager (12146), NodeManager (12280), DataNode (11690), SecondaryNameNode (11930), and NameNode (11550). The terminal window is set against a background image of a forest floor.

```
Activities Terminal 17:15 7 Thg 5
thinhds@thinhds-Inspiron-3576: ~/hadoop-3.3.2/etc/hadoop
2022-05-07 17:10:51,026 INFO namenode.FSImageFormatProtobuf: Image file /home/thinhds/hadoop-3.3.2/data/namenode/current/fsimage.ckpt_0000000000000000 of size 399 bytes saved in 0 seconds
2022-05-07 17:10:51,380 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with t
xid >= 0
2022-05-07 17:10:51,404 INFO namenode.FSNamesystem: Stopping services started for active state
2022-05-07 17:10:51,407 INFO namenode.FSNamesystem: Stopping services started for standby state
2022-05-07 17:10:51,407 INFO namenode.FSImageSaver clean checkpoint: txid=0 when meet s
utdown.
2022-05-07 17:10:51,407 INFO namenode.NameNode: SHUTDOWN_MSG:
*****SHUTDOWN_MSG: Shutting down NameNode at thinhds-Inspiron-3576/127.0.1.1
*****
thinhds@thinhds-Inspiron-3576:~/hadoop-3.3.2/etc/hadoop$ start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [thinhds-Inspiron-3576]
thinhds-Inspiron-3576: Warning: Permanently added 'thinhds-inspiron-3576' (ECDSA) to the list of known hosts.
thinhds@thinhds-Inspiron-3576:~/hadoop-3.3.2/etc/hadoop$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
thinhds@thinhds-Inspiron-3576:~/hadoop-3.3.2/etc/hadoop$ jps
12146 ResourceManager
12280 NodeManager
11690 DataNode
11930 SecondaryNameNode
11550 NameNode
12638 Jps
thinhds@thinhds-Inspiron-3576:~/hadoop-3.3.2/etc/hadoop$
```

Lưu ý: nếu không chạy đủ 6 chương trình thì có thể bị lỗi trong lúc cài đặt như tạo file, cấp quyền cho file, hoặc sai đường dẫn đến các env trong file bashrc.

Bước 12: Kiểm tra HDFS Web UI

Bước 13: Dừng hadoop

stop-all.sh

```
thinhds@thinhds-Inspiron-3576:~/hadoop-3.3.2/etc/hadoop$ stop-all.sh
WARNING: Stopping all Apache Hadoop daemons as thinhds in 10 seconds.
WARNING: Use CTRL-C to abort.
Stopping namenodes on [localhost]
Stopping datanodes
Stopping secondary namenodes [thinhds-Inspiron-3576]
Stopping nodemanagers
Stopping resourcemanager
```

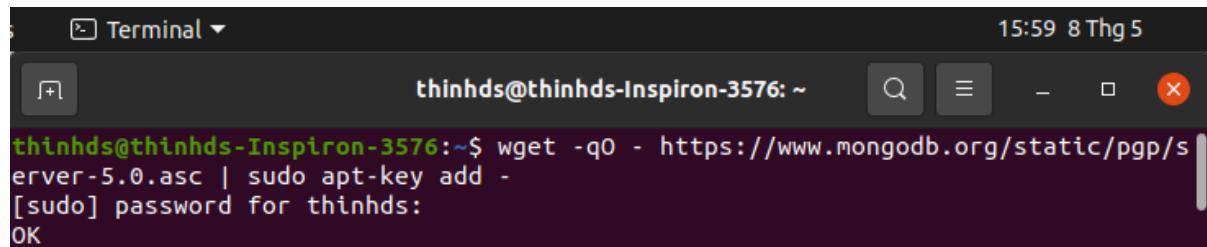
1.2. Cài đặt MongoDB

Ta sẽ có hai cách cài đặt mongo server: cài trên terminal và trực tiếp trên website tải file.

Trên Terminal:

Bước 1: Ta sẽ truy cập link và lấy khóa công khai GPG.

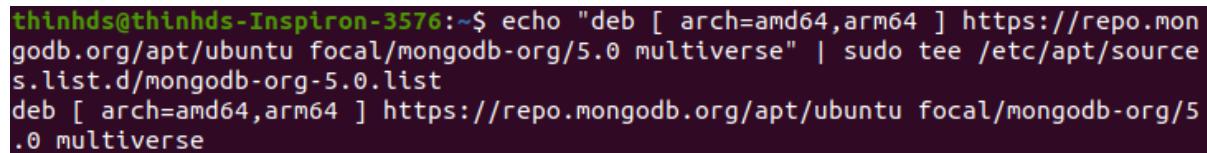
```
wget -qO - https://www.mongodb.org/static/pgp/server-5.0.asc | sudo apt-key add -
```



```
thinhds@thinhds-Inspiron-3576:~$ wget -qO - https://www.mongodb.org/static/pgp/server-5.0.asc | sudo apt-key add -  
[sudo] password for thinhds:  
OK
```

Bước 2: Tạo /etc/apt/sources.list.d/mongodb-org-5.0.list tệp cho Ubuntu 20.04 (Focal):

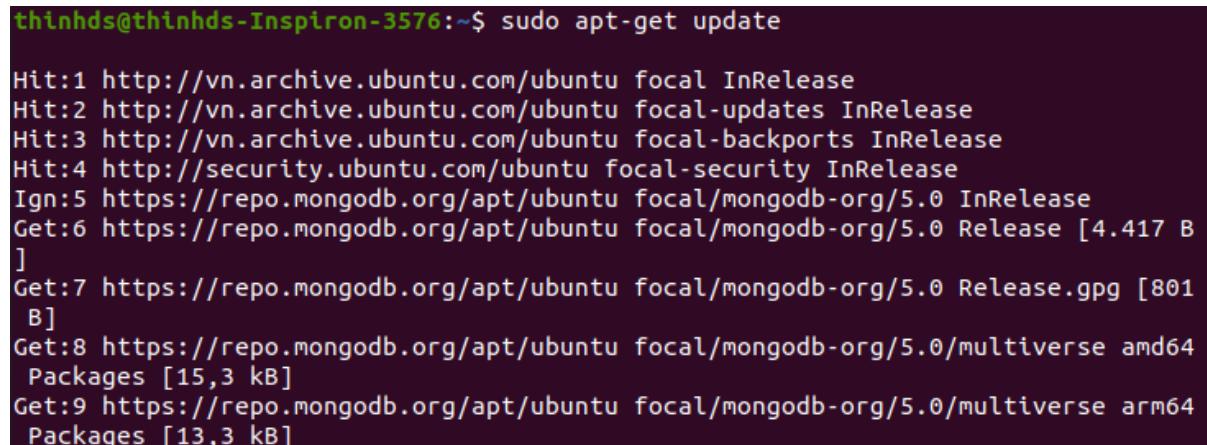
```
echo "deb [ arch=amd64,arm64 ] https://repo.mongodb.org/apt/ubuntu focal/mongodb-org/5.0 multiverse" | sudo tee /etc/apt/sources.list.d/mongodb-org-5.0.list
```



```
thinhds@thinhds-Inspiron-3576:~$ echo "deb [ arch=amd64,arm64 ] https://repo.mongodb.org/apt/ubuntu focal/mongodb-org/5.0 multiverse" | sudo tee /etc/apt/sources.list.d/mongodb-org-5.0.list  
deb [ arch=amd64,arm64 ] https://repo.mongodb.org/apt/ubuntu focal/mongodb-org/5.0 multiverse
```

Bước 3: Tải lại gói cơ sở dữ liệu cục bộ

```
sudo apt-get update
```



```
thinhds@thinhds-Inspiron-3576:~$ sudo apt-get update  
Hit:1 http://vn.archive.ubuntu.com/ubuntu focal InRelease  
Hit:2 http://vn.archive.ubuntu.com/ubuntu focal-updates InRelease  
Hit:3 http://vn.archive.ubuntu.com/ubuntu focal-backports InRelease  
Hit:4 http://security.ubuntu.com/ubuntu focal-security InRelease  
Ign:5 https://repo.mongodb.org/apt/ubuntu focal/mongodb-org/5.0 InRelease  
Get:6 https://repo.mongodb.org/apt/ubuntu focal/mongodb-org/5.0 Release [4.417 B]  
Get:7 https://repo.mongodb.org/apt/ubuntu focal/mongodb-org/5.0 Release.gpg [801 B]  
Get:8 https://repo.mongodb.org/apt/ubuntu focal/mongodb-org/5.0/multiverse amd64 Packages [15,3 kB]  
Get:9 https://repo.mongodb.org/apt/ubuntu focal/mongodb-org/5.0/multiverse arm64 Packages [13,3 kB]
```

Bước 4: Tải mongo server

```
sudo apt-get install -y mongodb-org
```

```
thinhdः@thinhdः-Inspiron-3576:~$ sudo apt-get install -y mongodb-org
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following packages were automatically installed and are no longer required:
 cabextract fonts-wine glib-networking:i386 gstreamer1.0-plugins-base:i386
 gstreamer1.0-plugins-good:i386 gstreamer1.0-x:i386 i965-va-driver:i386
 icoutils intel-media-va-driver:i386 jq libaa1:i386 libaom0:i386
 libapparmor1:i386 libasn1-8-heimdal:i386 libasound2:i386
 libasound2-plugins:i386 libasyncns0:i386 libatomic1:i386
 libavahi-client3:i386 libavahi-common-data:i386 libavahi-common3:i386
 libavc1394-0:i386 libavcodec58:i386 libavutil56:i386 libbrotli1:i386
 libbsds0:i386 libcaca0:i386 libcairo-gobject2:i386 libcairo2:i386
 libcap2:i386 libcapi20-3:i386 libcdparanoia0:i386 libcodec2-0.9:i386
 libcurl2:i386 libcurl3-gnutls:i386 libdatrie1:i386 libdbus-1-3:i386
 libdrm-amdgpu1:i386 libdrm-intel1:i386 libdrm-nouveau2:i386
```

Bước 5: Khởi động MongoDB

```
sudo systemctl start mongod
```

```
thinhds@thinhds-Inspiron-3576:~$ sudo systemctl start mongod
```

Bước 6: Kiểm tra hoạt động của MongoDB

```
sudo systemctl status mongod
```

```
thinhds@thinhds-Inspiron-3576:~$ sudo systemctl status mongod
● mongod.service - MongoDB Database Server
    Loaded: loaded (/lib/systemd/system/mongod.service; disabled; vendor prese>
    Active: active (running) since Sun 2022-05-08 15:47:41 +07; 9s ago
      Docs: https://docs.mongodb.org/manual
  Main PID: 87468 (mongod)
    Memory: 112.8M
   CGroup: /system.slice/mongod.service
           └─87468 /usr/bin/mongod --config /etc/mongod.conf

May 08 15:47:41 thinhds-Inspiron-3576 systemd[1]: Started MongoDB Database Se>
```

Trên Website:

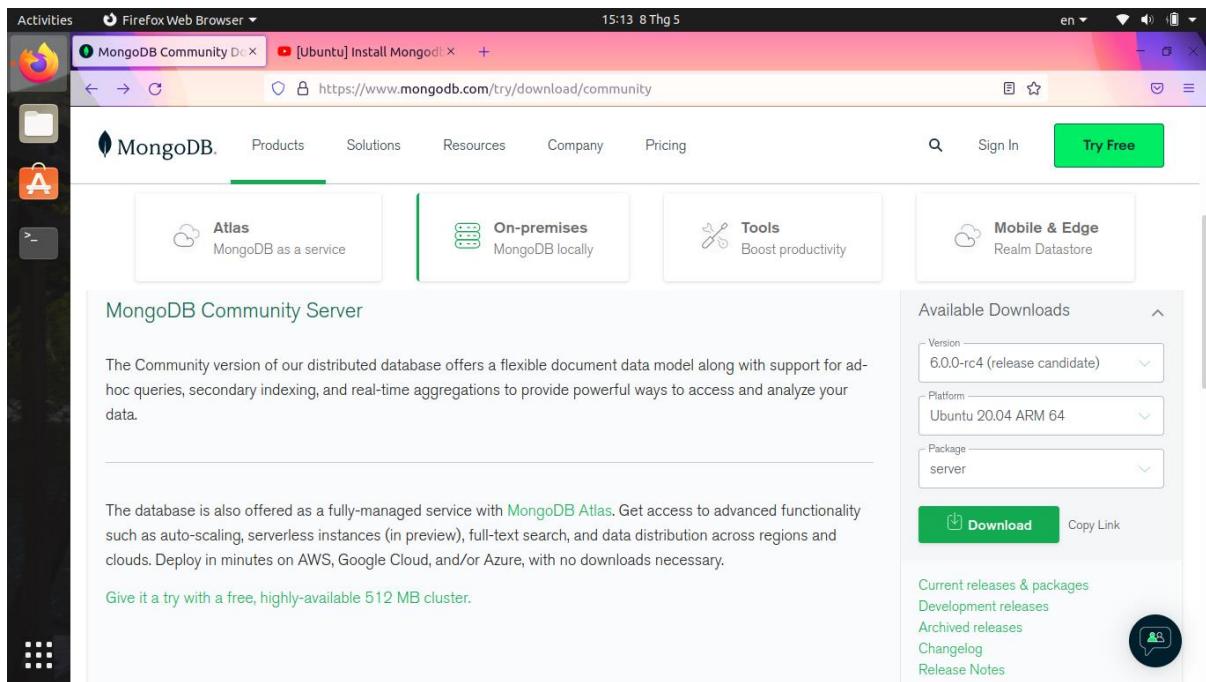
Bước 1: Vào trang <https://www.mongodb.com/try>

The screenshot shows the MongoDB website's "Try MongoDB Atlas Prod" page. At the top, there are four deployment options: "Atlas" (MongoDB as a service), "On-premises" (MongoDB locally), "Tools" (Boost productivity), and "Mobile & Edge" (Realm Datastore). Below these, a section titled "Choose which type of deployment is best for you" is displayed. A call-to-action button "Try MongoDB Atlas Products" is present, along with fields for signing up with Google or entering company and work email information.

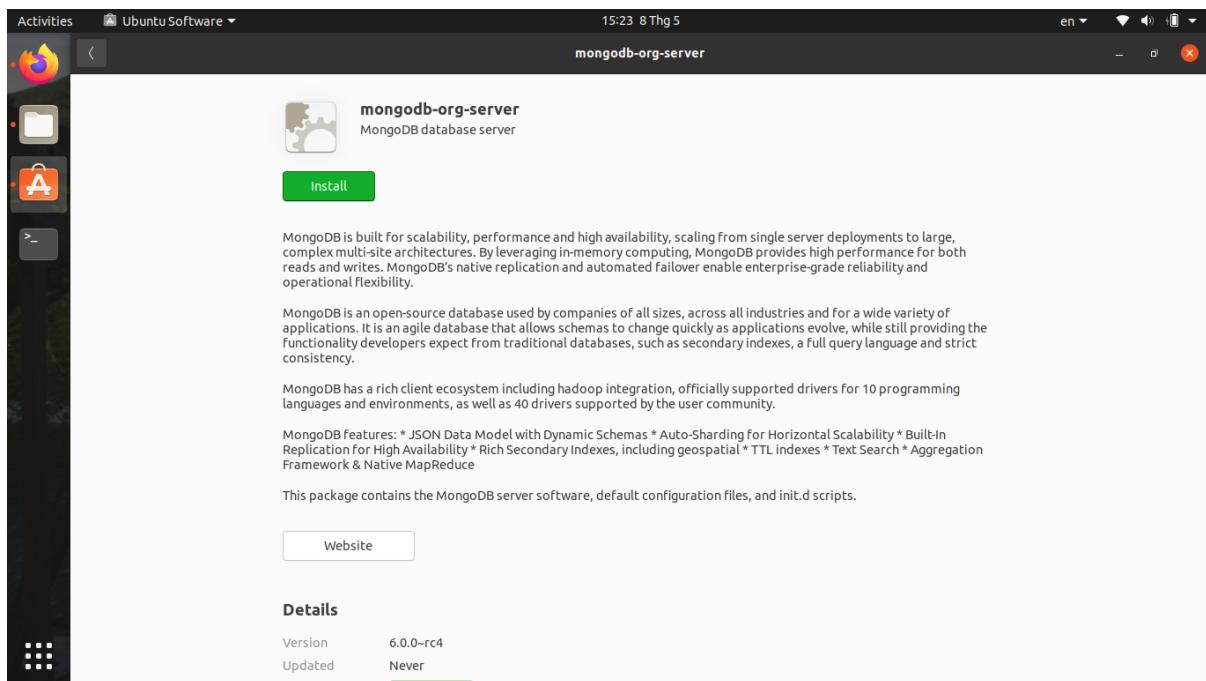
Bước 2: Chọn On-premises và hiện lên trang, ta đi chọn tiếp MongoDB Community Server

The screenshot shows the "Download MongoDB Locally" section of the MongoDB website. It lists several download options under "MongoDB Enterprise Server", "MongoDB Community Server", "MongoDB Ops Manager", and "MongoDB Enterprise Kubernetes Operator". The "On-premises" option from the previous screen is highlighted.

Bước 3: Chọn phiên bản thích hợp với máy tính và tải xuống ở phần Available Download



Bước 4: Mở file và cài đặt.



Nếu như muốn cài thêm giao diện mongoDB thì thực hiện các bước như sau:

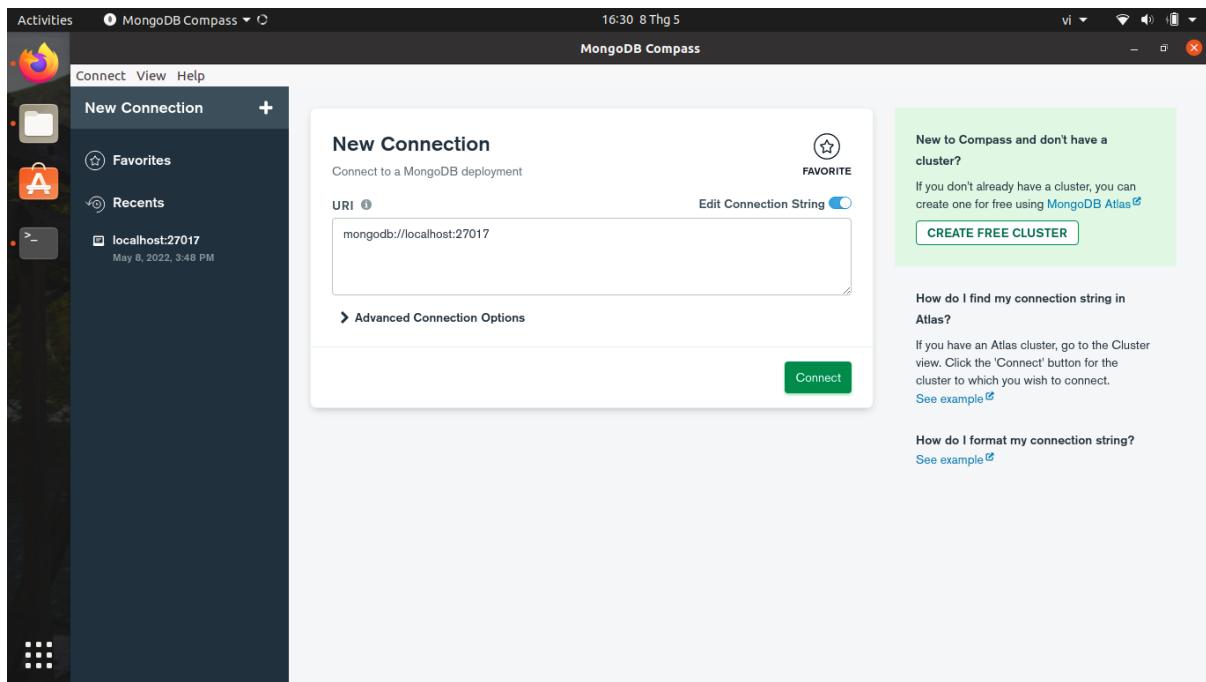
Bước 1: Vào trang <https://www.mongodb.com/try> giống như trên và chọn tool. Tiếp đến là chọn phiên bản thích hợp tải xuống.

The screenshot shows a Firefox browser window with the URL <https://www.mongodb.com/try/download/compass>. The page displays the MongoDB homepage with sections for Atlas, On-premises, Tools, and Mobile & Edge. The 'Tools' section is highlighted. Below it, the 'MongoDB Compass' section describes the tool for exploring and manipulating databases. It mentions three versions: a full version with all features, a read-only version without write or delete capabilities, and an isolated edition whose sole network connection is to the MongoDB instance. A 'Documentation' link is provided. To the right, there's a 'Available Downloads' sidebar with dropdown menus for Version (1.31.2 Stable), Platform (Ubuntu 64-bit (14.04+)), and Package (deb). A large green 'Download' button is visible, along with links for 'Documentation' and 'Archived releases'. A small icon of a person with a speech bubble is in the bottom right corner.

Bước 2: Mở file và đi cài đặt

The screenshot shows the Ubuntu Software Center interface. A search bar at the top has 'mongodb-compass' typed into it. Below the search bar, a card for the 'mongodb-compass' package is displayed. It features a thumbnail of the MongoDB logo, the package name, a brief description 'The MongoDB GUI', and a large green 'Install' button. Below the card, a 'Website' button is shown. Under the heading 'Details', a table provides information about the package: Version 1.31.2, Updated Never, License Proprietary, Source [mongodb-compass_1.31.2_amd64.deb](#), Installed Size 71,8 MB, and Download Size 0 bytes.

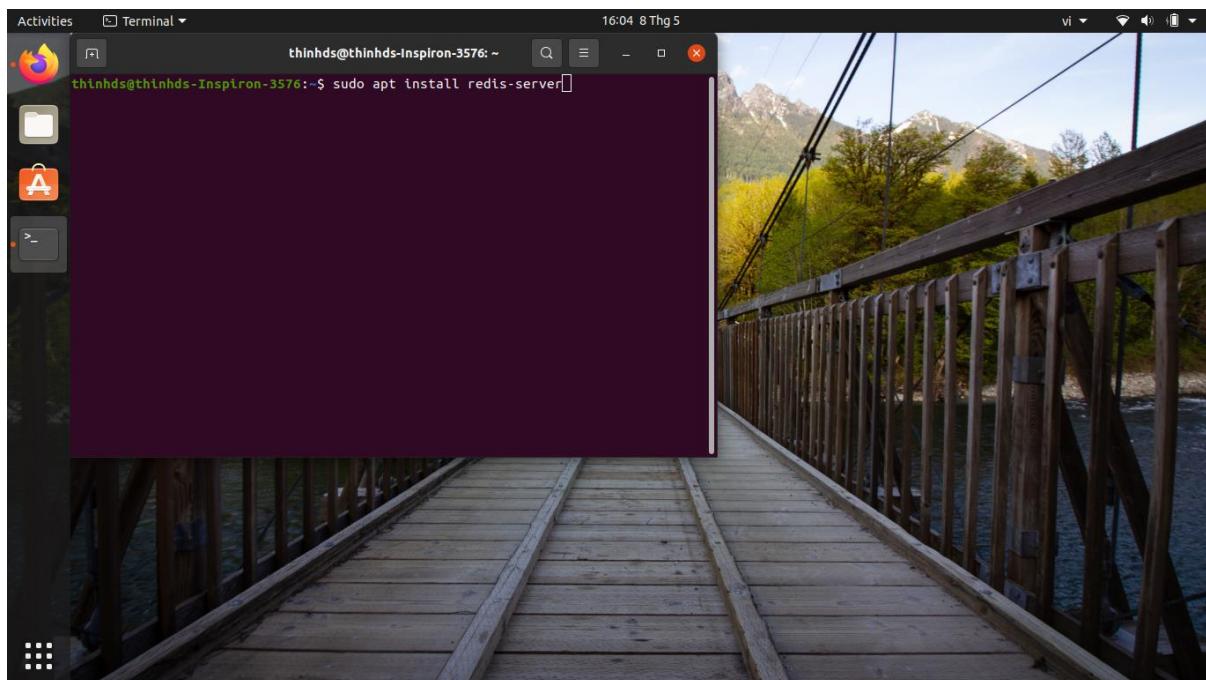
Bước 3: Mở MongoDB lên và connect và xong hãy đi tạo database hay thêm data tùy mục đích.



1.3. Cài đặt Redis

Bước 1: Cài đặt Redis

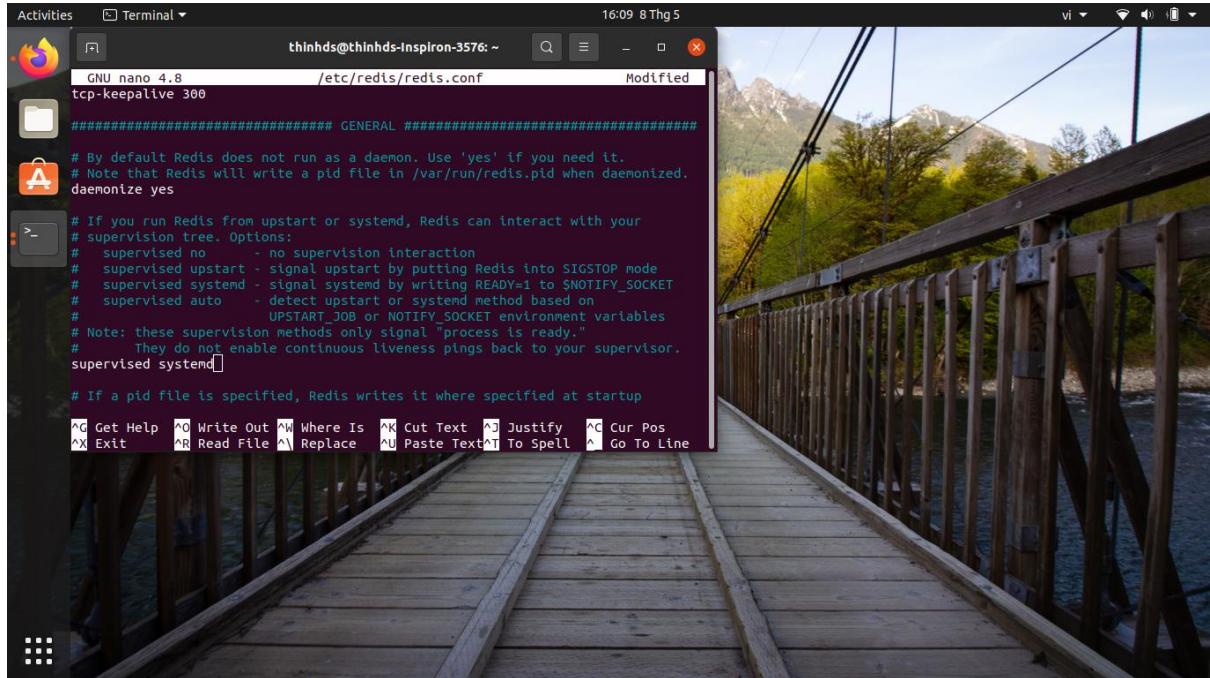
```
sudo apt install redis-server
```



Bước 2: Chính sửa một số trong file redis.conf

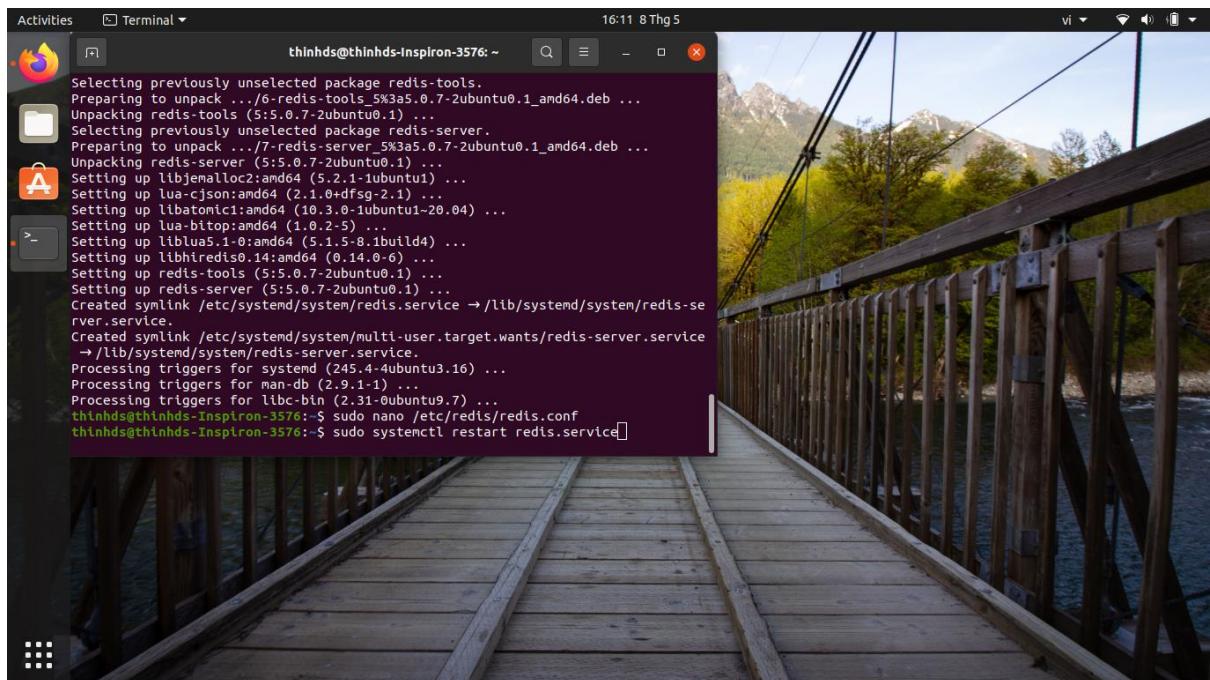
```
sudo nano /etc/redis/redis.conf
```

Tìm supervised và nó có mặc định là ‘no’ ta thay nó thành hệ thống khởi tạo ‘systemd’.

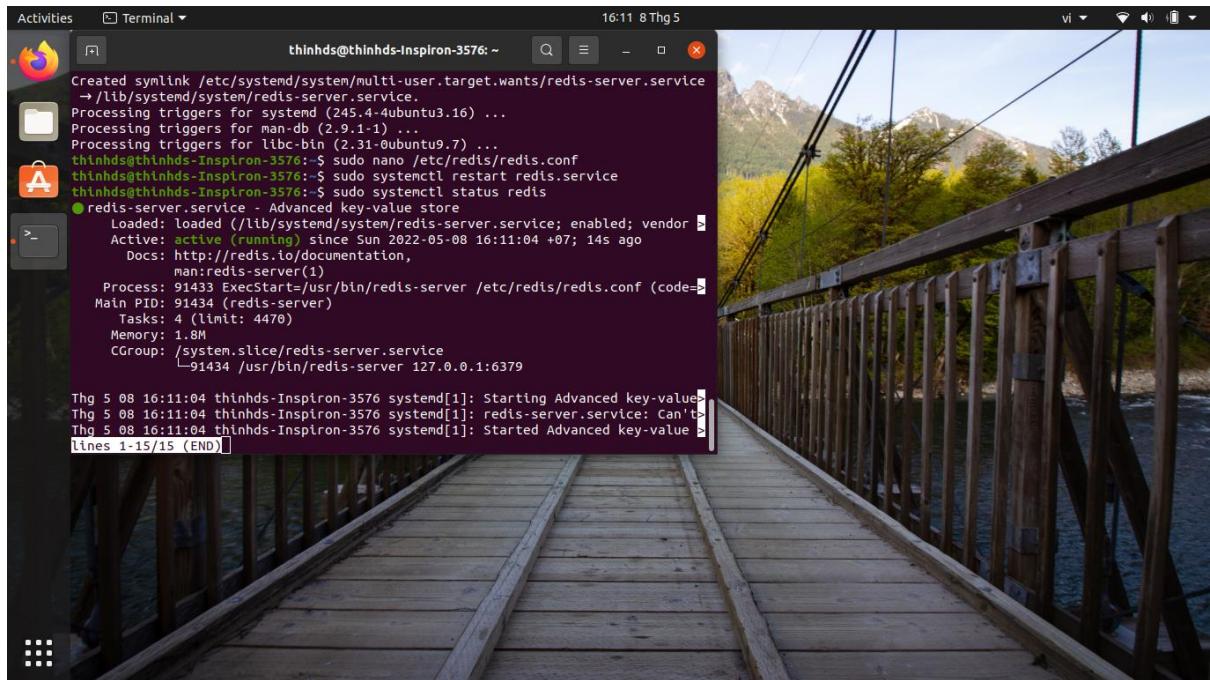


Bước 3: Ta di khởi tạo lại Redis

```
sudo systemctl restart redis.service
```



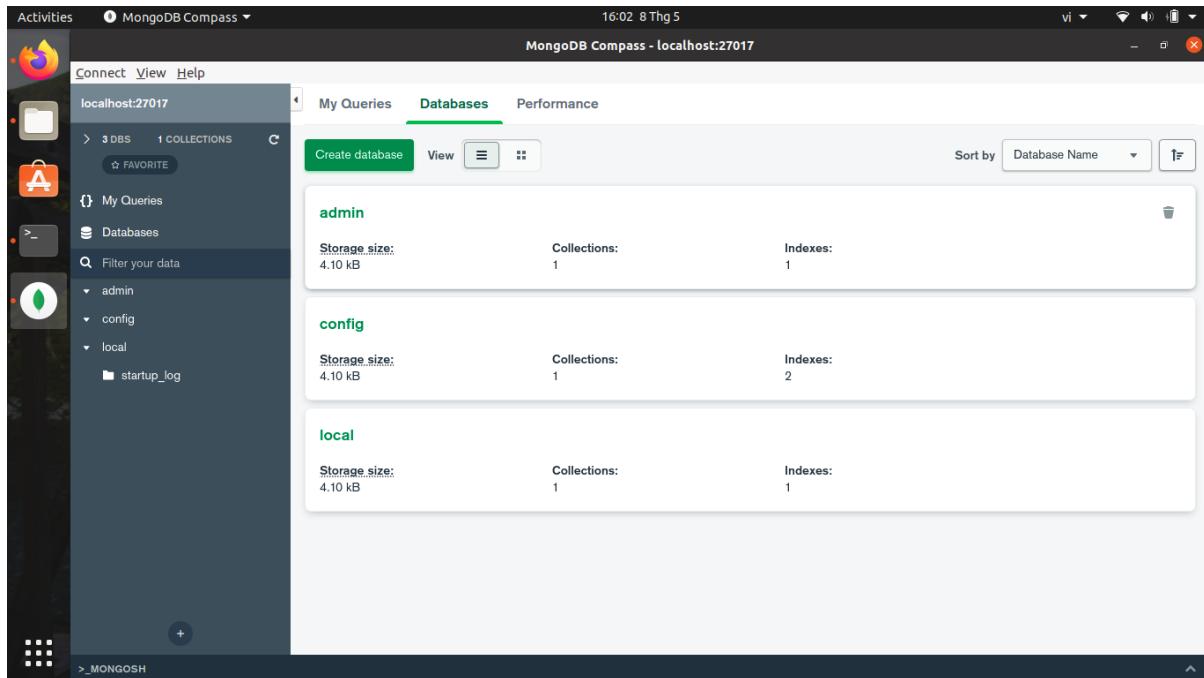
Bước 4: Kiểm tra Redis hoạt động không.



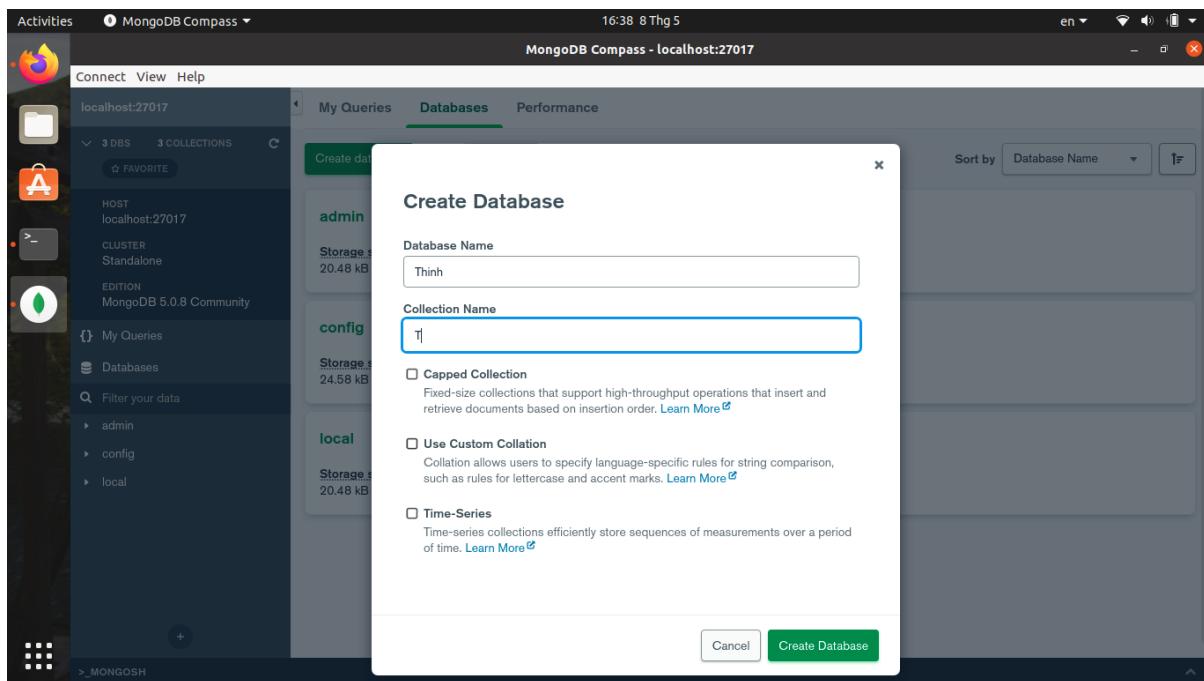
Vậy là hiển thị ‘active (running)’, Redis đang hoạt động. Và cài xong.

Chương 2: Thực nghiệm các thao tác cơ bản trong MongoDB, Redis

2.1. Thực nghiệm MongoDB



Vào MongoDB và thực hiện tạo database và các thao tác khác.



2.2. Thực nghiệm Redis

Vào Terminal thực hiện câu lệnh redis-cli và ping

```
redis-cli
```

The screenshot shows a Linux desktop interface with a terminal window and a web browser window.

Terminal Window:

```
thinhds@thinhds-Inspiron-3576:~$ redis-cli
127.0.0.1:6379> ping
PONG
127.0.0.1:6379> 
```

Web Browser (Mozilla Firefox):

The browser is displaying a page about Redis. The visible text includes:

- A "ping" command example.
- A note about Redis being alive.
- A "get" command example.
- A note about retrieving stored values.
- A "set" command example.
- A note about setting keys.

RELATED sections mention "Initial Server Setup with Ubuntu 12.04" and "How To Install Ruby on Rails on Ubuntu 12.04 LTS (Precise Pangolin) with RVM".

Thực hiện các thao tác tạo cơ sở dữ liệu trên terminal. Và sau đây là một số lệnh cơ bản kiểu String trong Redis.

The screenshot shows a Linux desktop interface with a terminal window.

Terminal Window:

```
thinhds@thinhds-Inspiron-3576:~$ redis-cli
127.0.0.1:6379> ping
PONG
127.0.0.1:6379> set name "thinh"
OK
127.0.0.1:6379> set age 21
OK
127.0.0.1:6379> get name
"thinh"
127.0.0.1:6379> get age
"21"
127.0.0.1:6379> del name
(integer) 1
127.0.0.1:6379> get name
(nil)
127.0.0.1:6379> del age
(integer) 1
127.0.0.1:6379> get age
(nil)
127.0.0.1:6379> 
```

set	Thiết lập giá trị cho key
get	Lấy giá trị value của key
del	Xóa key và giá trị của nó

Còn một số thao tác thực hiện tạo cơ sở dữ liệu thì có thể tham khảo thêm trên link:

<https://redis.io/commands/>

The screenshot shows the Redis Commands page within a Firefox browser window. The URL in the address bar is <https://redis.io/commands/>. The page title is "Commands". At the top, there are navigation links: GET STARTED, DOCS, COMMANDS, RESOURCES, and COMMUNITY, along with a "Download" button and a "Try Redis Cloud" button. Below the header, there is a search bar with the placeholder "Search" and a filter dropdown labeled "(filter by group)". The main content area displays a grid of Redis commands categorized under "ACL". The commands listed are:

ACL	ACL CAT [category...]	ACL DELUSER user...	ACL DRYRUN usern...
A container for Access List Control commands	List the ACL categories or the commands inside a category	Remove the specified ACL users and the associated rules	Returns whether the user can execute the given command without executing the command.
ACL GENPASS [bits]	ACL GETUSER user...	ACL HELP	ACL LIST
Generate a a pseudorandom secure password to	Get the rules for a specific ACL user	Show helpful text about the different subcommands	List the current ACL rules in ACL config file format

Chương 3: So sánh Mongo và Redis

3.1. Cấu trúc database

Redis	MongoDB
<ul style="list-style-type: none"> - Key-value store. - Redis là một mã nguồn mở, BSD licensed, lưu trữ dữ liệu dưới dạng key-value trong bộ nhớ. - Nó thường được gọi là máy chủ cấu trúc dữ liệu vì các key có thể chứa chuỗi, băm, danh sách, set và sorted-set. - Cấu trúc này khác với các cơ sở dữ liệu quan hệ có các bảng, hàng và cột; có thể gây khó khăn cho một số bộ đệm SQL truyền thông tới master 	<ul style="list-style-type: none"> - Document store. - MongoDB lưu trữ dữ liệu trong các document tương tự như JSON và có thể thay đổi về cấu trúc, cung cấp một schema linh hoạt, năng động. - Trong Mongod collection của document không có cấu trúc được xác định trước và các column có thể thay đổi cho các document khác nhau. - MongoDB cũng được thiết kế để có tính sẵn sàng và khả năng mở rộng cao, với tính năng sao chép và tự động tích hợp sẵn.

3.2. Ứng dụng

Redis	MongoDB
<ul style="list-style-type: none"> - Cơ sở dữ liệu NoSQL. - Dễ cài đặt, sử dụng, deploy và duy trì. - Lưu trữ dữ liệu trong bộ nhớ nên cho hiệu năng cao và tốc độ nhanh. - Mã nguồn mở, ổn định, chi phí hiệu quả. - Tính mở rộng và sẵn sàng cao, sao chép bất đồng bộ, ... - Tích hợp dễ dàng với Sidekiq cho Rails background job. - Cấu trúc dữ liệu đa dạng, ... 	<ul style="list-style-type: none"> - Một trong những document store phổ biến nhất. - NoSQL. - Dễ dàng setup, sử dụng cũng như duy trì. - Mã nguồn mở, miễn phí, đáng tin cậy có tốc độ và hiệu năng nhanh. - Khả năng linh hoạt, mở rộng, sẵn sàng cao, có khả năng auto-sharding. - Hỗ trợ đầy đủ index, tốc độ cập nhật nhanh chóng.

3.3. Sử dụng thế nào

Redis	MongoDB
<ul style="list-style-type: none"> - Dữ liệu key-value. - Caching. - Lo lắng và có thời gian cân nhắc thiết kế cơ sở dữ liệu. - Thực sự cần hiệu năng cao. - Không quan tâm nhiều tới scaling, kích thước dữ liệu ổn định. 	<ul style="list-style-type: none"> - Dữ liệu phức tạp, lớn, đang phát triển. - Linh hoạt, dễ để học. - Prototyping, Startups, Hackathons - Không quan tâm lắm về cơ sở dữ liệu. - Cần thay đổi schema nhanh chóng. - Các hệ thống realtime (thời gian thực) yêu cầu phản hồi nhanh, bigdata với

	<p>yêu cầu truy vấn nhanh, có tần suất write/insert lớn.</p> <ul style="list-style-type: none"> - Sử dụng làm search engine.
--	---

3.4. Nhược điểm

Redis	MongoDB
<ul style="list-style-type: none"> - Vì Redis sử dụng RAM làm bộ nhớ cho mình nên khi lượng file cache lớn thì sẽ dẫn đến trường hợp thiếu RAM cho Server. - Không thể truy vấn trực tiếp các object. 	<ul style="list-style-type: none"> - Không hỗ trợ JOIN. - Dữ liệu dễ bị sai do không có ràng buộc. - Sử dụng nhiều bộ nhớ và giới hạn kích thước bản ghi (< 16Mb).