**Cài đặt Hadoop 3.3.4 Single Node**

Chuẩn bị: Các bạn cần chuẩn bị file iso cài đặt ubuntu server 22.04

[Link iso ubuntu-server-22.04](https://drive.google.com/file/d/1t2YcUkOeeFLzPoOQq7CF5Y-DKMXexlwp/view?usp=sharing)

Chú ý khi cài đặt ubuntu-server-22.04 cần cài đặt gói openssh-server:

$ sudo apt-get install openssh-server

$ sudo systemctl restart ssh

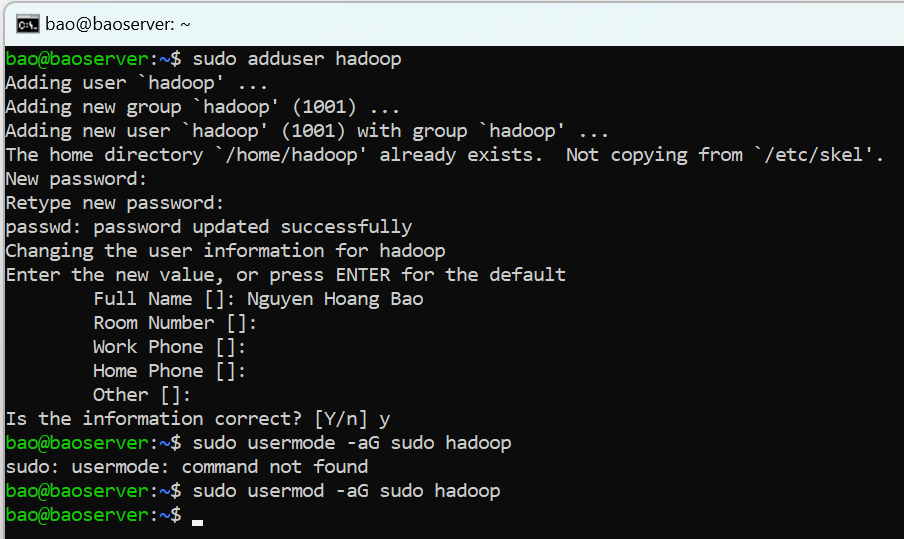
Video hương dẫn cài ubuntu server 22.04

Link: <https://www.youtube.com/watch?v=xfT2gqwHyO0>

**Bước 1: Tạo tài khoản hadoop và cấp quyền sudo cho tài khoản hadoop**

$ sudo adduser hadoop

$ sudo usermod -aG sudo hadoop

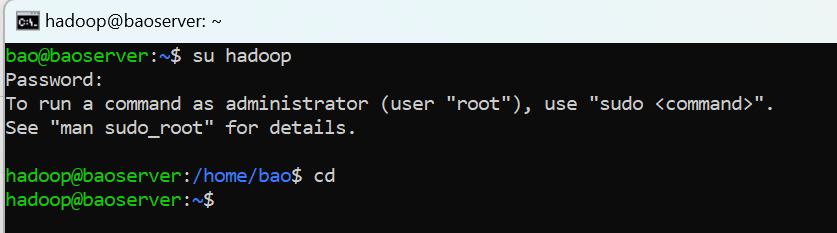


$ sudo reboot

**Bước 2: đăng nhập vào tài khoản hadoop**

$ su hadoop

$ cd

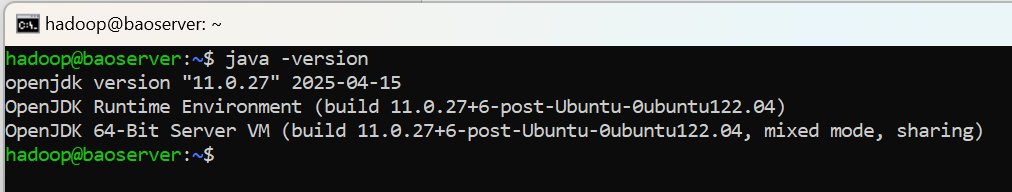


**Bước 3: Cài đặt java**

$ sudo apt update

$ sudo apt install openjdk-11-jdk -y

$ java -version

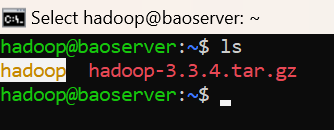


**Bước 4: Tải và giải nén gói hadoop-3.3.4**

$ wget <https://archive.apache.org/dist/hadoop/common/hadoop-3.3.4/hadoop-3.3.4.tar.gz>

$ tar -xzvf hadoop-3.3.4.tar.gz

$ mv hadoop-3.3.4 /home/hadoop/hadoop



Bước 5: Thiết lập môi trường

$ vim ~/.bashrc

Dán đoạn code này ở cuối file:

# Java home

export JAVA\_HOME=/usr/lib/jvm/java-11-openjdk-amd64

# Hadoop variables

export HADOOP\_HOME=/home/hadoop/hadoop

export HADOOP\_INSTALL=$HADOOP\_HOME

export HADOOP\_MAPRED\_HOME=$HADOOP\_HOME

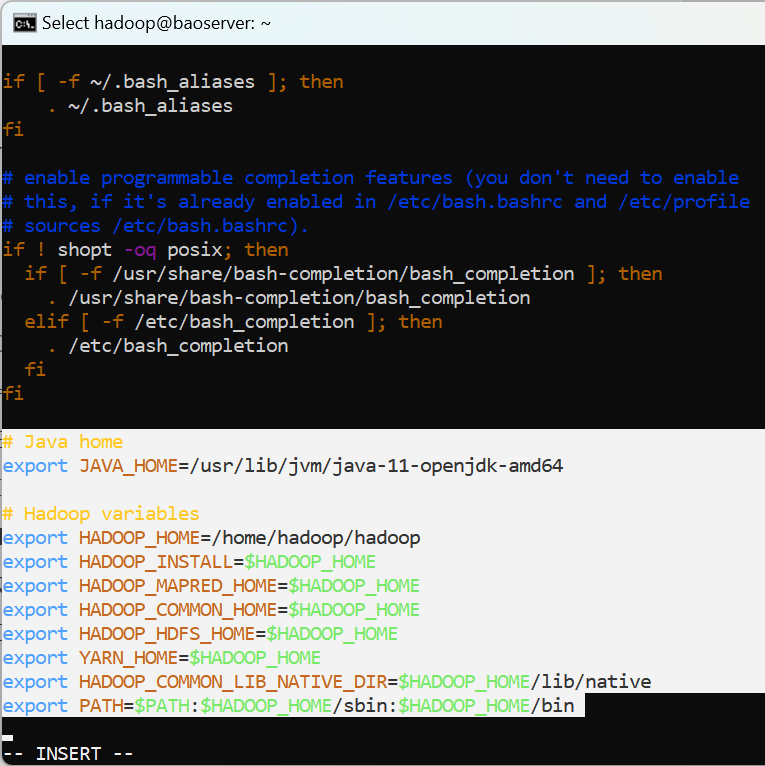
export HADOOP\_COMMON\_HOME=$HADOOP\_HOME

export HADOOP\_HDFS\_HOME=$HADOOP\_HOME

export YARN\_HOME=$HADOOP\_HOME

export HADOOP\_COMMON\_LIB\_NATIVE\_DIR=$HADOOP\_HOME/lib/native

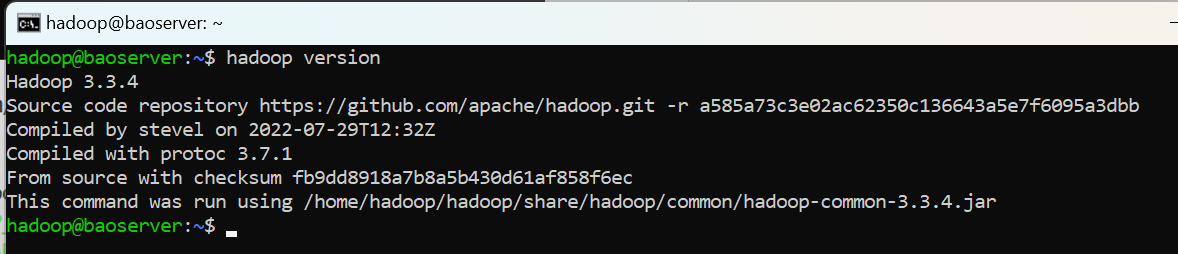
export PATH=$PATH:$HADOOP\_HOME/sbin:$HADOOP\_HOME/bin



$ source ~/.bashrc

**Bước 6: Kiểm tra hadoop version**

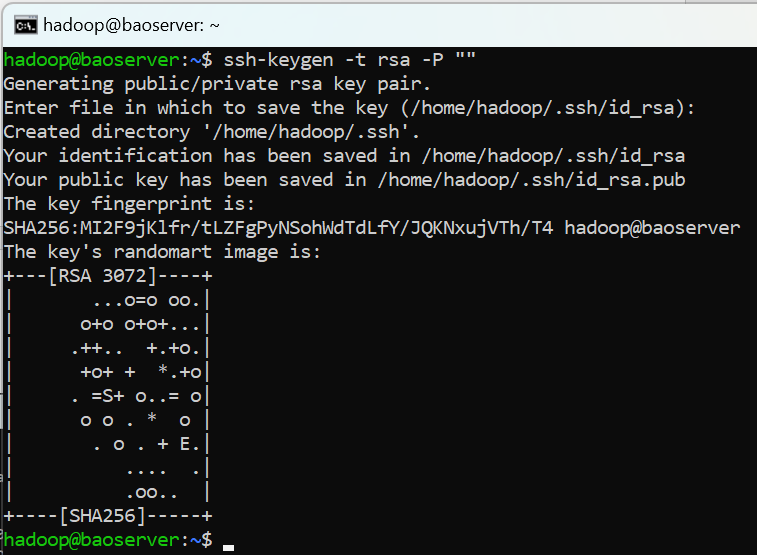
$ hadoop version



**Bước 6: Cấu hình ssh passwordless**

$ ssh-keygen -t rsa -P ""

Gặp câu hỏi “Enter file in which to save the key (/home/hadoop/.ssh/id\_rsa):” thì nhấn Enter



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

$ chmod 600 ~/.ssh/authorized\_keys

**Bước 7: Tạo thư mục lưu trữ HDFS**

$ mkdir -p ~/hadoopdata/hdfs/namenode

$ mkdir -p ~/hadoopdata/hdfs/datanode

**Bước 8: Cấu hình các file trong ~/hadoop/etc/hadoop**

Cấu hình core-site.xml:

$ vim ~/hadoop/etc/hadoop/core-site.xml

*Code cấu hình:*

<configuration>

<property>

<name>fs.defaultFS</name>

<value>hdfs://localhost:9000</value>

</property>

</configuration>

Cấu hình hdfs-site.xml:

$ vim ~/hadoop/etc/hadoop/hdfs-site.xml

*Code cấu hình:*

<configuration>

<property>

<name>dfs.replication</name>

<value>1</value>

</property>

<property>

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

<value>file:///home/hadoop/hadoopdata/hdfs/namenode</value>

</property>

<property>

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

<value>file:///home/hadoop/hadoopdata/hdfs/datanode</value>

</property>

</configuration>

Cấu hình file mapred-site.xml:

$ vim ~/hadoop/etc/hadoop/mapred-site.xml

*Code cấu hình:*

<configuration>

<property>

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

<value>yarn</value>

</property>

<property>

<name>yarn.app.mapreduce.am.env</name>

<value>HADOOP\_MAPRED\_HOME=/home/hadoop/hadoop</value>

</property>

<property>

<name>mapreduce.map.env</name>

<value>HADOOP\_MAPRED\_HOME=/home/hadoop/hadoop</value>

</property>

<property>

<name>mapreduce.reduce.env</name>

<value>HADOOP\_MAPRED\_HOME=/home/hadoop/hadoop</value>

</property>

</configuration>

Cấu hình yarn-site.xml:

$ vim ~/hadoop/etc/hadoop/yarn-site.xml

*Code cấu hình:*

<configuration>

<property>

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

<value>mapreduce\_shuffle</value>

</property>

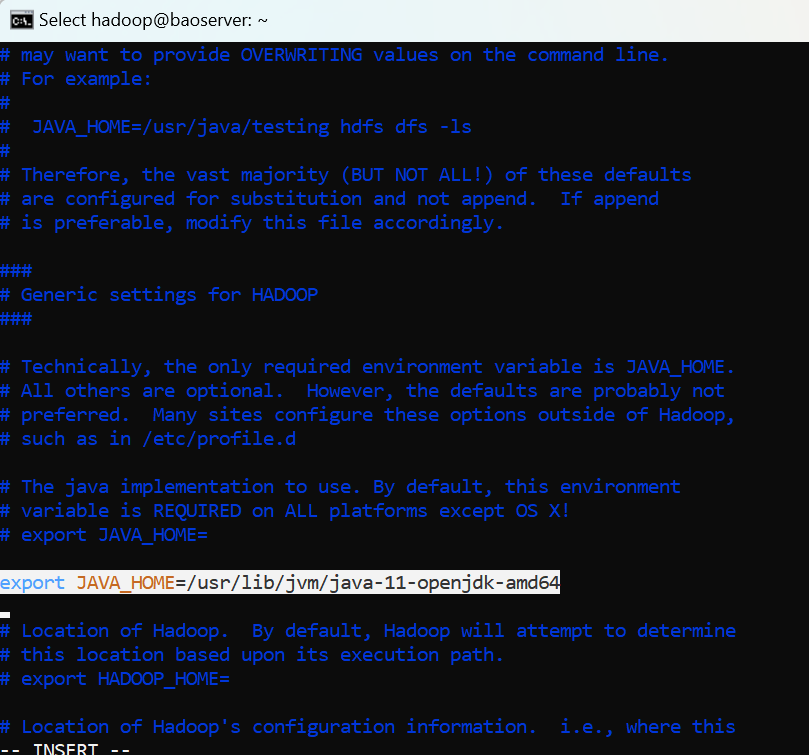
</configuration>

**Bước 9: Cấu hình java path trong hadoop**

$ vim ~/hadoop/etc/hadoop/hadoop-env.sh

*Thêm đoạn code sau vào:*

export JAVA\_HOME=/usr/lib/jvm/java-11-openjdk-amd64



**Bước 10: Format HDFS**

$ hdfs namenode -format

**Bươc 11: Khởi động hadoop**

$ start-dfs.sh

$ start-yarn.sh

**Bước 12: Kiểm tra dịch vụ**

$ jps

