

- ssh keygen

```
To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

ubuntu@ip-172-31-39-208:~$ 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:
SHA256:3Dj+fgvEChQ5F4GFeokOcxB2ZoPc5K/Zz4DOWgU2TYA ubuntu@ip-172-31-39-208
The key's randomart image is:
+---[RSA 3072]-----+
|  .+*B*=+.          |
|  .*=Eo.o           |
```

- ```
+---[SHA256]-----+
ubuntu@ip-172-31-39-208:~$ cd .ssh
ubuntu@ip-172-31-39-208:~/.ssh$ ls
authorized_keys id_rsa id_rsa.pub
ubuntu@ip-172-31-39-208:~/.ssh$ cat id_rsa.pub >> authorized_keys
ubuntu@ip-172-31-39-208:~/.ssh$ cd
ubuntu@ip-172-31-39-208:~$ ssh localhost
The authenticity of host 'localhost (127.0.0.1)' can't be established.
ED25519 key fingerprint is SHA256:a/nbdis1fNgtJiNlpZEe7Hk9wS1/7OdS6Zk6OiAfJME.
This key's fingerprint is not known by any other names.
Building dependency tree... Done
Reading state information... Done
129 packages can be upgraded. Run 'apt list --upgradable' to see them.
ubuntu@ip-172-31-39-208:~$ sudo apt install openjdk-17-jdk -y
```
- ```
Running kernel seems to be up-to-date.

No services need to be restarted.

No containers need to be restarted.

No user sessions are running outdated binaries.

No VM guests are running outdated hypervisor (qemu) binaries on this host.
ubuntu@ip-172-31-39-208:~$ nano .bashrc
ubuntu@ip-172-31-39-208:~$ wget https://archive.apache.org/dist/hadoop/core/hadoop-3.3.1/hadoop-3.3.1.tar.gz
--2023-10-09 10:16:50-- https://archive.apache.org/dist/hadoop/core/hadoop-3.3.1/hadoop-3.3.1.tar.gz
Resolving archive.apache.org (archive.apache.org)... 65.108.204.189, 2a01:4f9:1a:a084::2
Connecting to archive.apache.org (archive.apache.org)|65.108.204.189|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 605187279 (577M) [application/x-gzip]
Saving to: 'hadoop-3.3.1.tar.gz'

hadoop-3.3.1.tar.gz      0%[
ubuntu@ip-172-31-39-208:~$ wget https://archive.apache.org/dist/hadoop/core/hadoop-3.3.1/hadoop-3.3.1.tar.gz
--2023-10-09 10:16:50-- https://archive.apache.org/dist/hadoop/core/hadoop-3.3.1/hadoop-3.3.1.tar.gz
Resolving archive.apache.org (archive.apache.org)... 65.108.204.189, 2a01:4f9:1a:a084::2
Connecting to archive.apache.org (archive.apache.org)|65.108.204.189|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 605187279 (577M) [application/x-gzip]
Saving to: 'hadoop-3.3.1.tar.gz'

hadoop-3.3.1.tar.gz      100%[=====] 577.15M  1.38MB/s  in 15m 15s

2023-10-09 10:32:05 (646 KB/s) - 'hadoop-3.3.1.tar.gz' saved [605187279/605187279]
```
- ```
2023-10-09 10:32:05 (646 KB/s) - 'hadoop-3.3.1.tar.gz' saved [605187279/605187279]

ubuntu@ip-172-31-39-208:~$ ls
hadoop-3.3.1.tar.gz
ubuntu@ip-172-31-39-208:~$ tar -xvzf hadoop-3.3.1.tar.gz
```
- ```
hadoop-3.3.1/include/ConfigurationUtil.h
hadoop-3.3.1/include/StringUtils.h
hadoop-3.3.1/include/hdfs.h
hadoop-3.3.1/include/Pipes.h
ubuntu@ip-172-31-39-208:~$ sudo mv hadoop-3.3.1 /usr/local/hadoop
ubuntu@ip-172-31-39-208:~$ ls
hadoop-3.3.1
ubuntu@ip-172-31-39-208:~$
```

```

hadoop-3.3.1.tar.gz
ubuntu@ip-172-31-39-208:~$ cd /usr/local/hadoop/
ubuntu@ip-172-31-39-208:~$ ls
LICENSE-binary LICENSE.txt NOTICE-binary NOTICE.txt README.txt
ubuntu@ip-172-31-39-208:~$ cd /usr/local/hadoop/etc/
ubuntu@ip-172-31-39-208:~$ ls
hadoop
ubuntu@ip-172-31-39-208:~$ cd /usr/local/hadoop/etc/
ubuntu@ip-172-31-39-208:~$ cd /usr/local/hadoop/etc/hadoop/
hadoop
ubuntu@ip-172-31-39-208:~$ cd /usr/local/hadoop/etc/hadoop/
ubuntu@ip-172-31-39-208:~$ ls
capacity-scheduler.xml      hadoop-policy.xml      kms-acls.xml            mapred-queues.xml.template  yarn-env.cmd
configuration.xml          hadoop-user-functions.sh.example  kms-env.sh              mapred-site.xml            yarn-env.sh
container-executor.cfg      hdfs-rbf-site.xml       kms-log4j.properties   mapred-site.xml            yarn-site.xml
core-site.xml              hdfs-site.xml           kms-site.xml            ssl-client.xml.example     yarnservice-log4j.properties
hadoop-env.cmd             https-env.sh            log4j.properties       ssl-server.xml.example     user_ec_policies.xml.template
hadoop-env.sh              https-log4j.properties   mapred-env.cmd          user_ec_policies.xml.template  workers
hadoop-metrics2.properties https-site.xml           mapred-env.sh
ubuntu@ip-172-31-39-208:~$ nano /usr/local/hadoop/etc/hadoop/hadoop-env.sh
ubuntu@ip-172-31-39-208:~$ nano /usr/local/hadoop/etc/hadoop/core-site.xml
ubuntu@ip-172-31-39-208:~$ nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml
ubuntu@ip-172-31-39-208:~$ nano /usr/local/hadoop/etc/hadoop/mapred-site.xml
ubuntu@ip-172-31-39-208:~$ nano /usr/local/hadoop/etc/hadoop/workers

```

```

ubuntu@ip-172-31-39-208:~$ cd /usr/local/hadoop/etc/hadoop/
ubuntu@ip-172-31-39-208:~$ hadoop namenode -format
hadoop: command not found
ubuntu@ip-172-31-39-208:~$ exec bash
ubuntu@ip-172-31-39-208:~$ hadoop namenode -format
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
WARNING: Use of this script to execute namenode is deprecated.
WARNING: Attempting to execute replacement "hdfs namenode" instead.

WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
WARNING: /usr/local/hadoop/logs does not exist. Creating.
2023-10-09 10:52:32,181 INFO namenode.NameNode: STARTUP MSG:
/*****
STARTUP_MSG: Starting NameNode
STARTUP_MSG: host = ip-172-31-39-208/172.31.39.208
STARTUP_MSG: args = [-format]
STARTUP_MSG: version = 3.3.1
STARTUP_MSG: classpath = /usr/local/hadoop/etc/hadoop:/usr/local/hadoop/share/hadoop/common/lib/curator-recipes-4.2.0.
/kerb-identity-1.0.1.jar:/usr/local/hadoop/share/hadoop/common/lib/kerby-asn1-1.0.1.jar:/usr/local/hadoop/share/hadoop/c
adpool/share/hadoop/common/lib/kerb-crypto-1.0.1.jar:/usr/local/hadoop/share/hadoop/common/lib/jersey-core-1.19.jar:/usr/

```

```

*****/
ubuntu@ip-172-31-39-208:~$ start-dfs.sh
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
Starting namenodes on [localhost]
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
Starting datanodes
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
Starting secondary namenodes [ip-172-31-39-208]
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
ip-172-31-39-208: Warning: Permanently added 'ip-172-31-39-208' (ED25519) to the list of known hosts.
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
ubuntu@ip-172-31-39-208:~$

```

```

--2023-10-09 11:12:07-- https://raw.githubusercontent.com/ErikSchierboom/sentencegenerator/master/sample.txt
Resolving raw.githubusercontent.com (raw.githubusercontent.com)... 185.199.111.133, 185.199.108.133, 18
Connecting to raw.githubusercontent.com (raw.githubusercontent.com)|185.199.111.133|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 4332499 (4.1M) [text/plain]
Saving to: 'the-king-james-bible.txt'

the-king-james-bible.txt      100%[=====]

2023-10-09 11:12:07 (50.2 MB/s) - 'the-king-james-bible.txt' saved [4332499/4332499]

ubuntu@ip-172-31-39-208:~$ cat sample.txt
ubuntu@ip-172-31-39-208:~$ ls
hadoop-3.3.1.tar.gz  sample.txt  the-king-james-bible.txt
ubuntu@ip-172-31-39-208:~$

```

```

hadoop-3.3.1.tar.gz  sample.txt  the-king-james-bible.txt
ubuntu@ip-172-31-39-208:~$ jps
5922 Jps
5427 DataNode
5290 NameNode
5663 SecondaryNameNode
ubuntu@ip-172-31-39-208:~$

```

```

ubuntu@ip-172-31-39-208:~$ hdfs dfs -mkdir -p /user/ubuntu
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
ubuntu@ip-172-31-39-208:~$ hdfs dfs -mkdir input
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
ubuntu@ip-172-31-39-208:~$ hdfs dfs -put sample.txt input
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
ubuntu@ip-172-31-39-208:~$ cd /usr/local/hadoop/share/hadoop/mapreduce/
ubuntu@ip-172-31-39-208:~$ cd /usr/local/hadoop/share/hadoop/mapreduce/hadoop.jar hadoop-mapreduce-examples-3.3.1.jar wordcount input output
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
2023-10-09 11:42:13,670 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2023-10-09 11:42:13,682 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2023-10-09 11:42:13,682 INFO impl.MetricsSystemImpl: JobTracker metrics system started

```

```
Bytes Written:0
ubuntu@ip-172-31-39-208:/usr/local/hadoop/share/hadoop/mapreduce$ hdfs dfs -ls output
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
Found 2 items
-rw-r--r-- 1 ubuntu supergroup          0 2023-10-09 11:42 output/_SUCCESS
-rw-r--r-- 1 ubuntu supergroup          0 2023-10-09 11:42 output/part-r-00000
ubuntu@ip-172-31-39-208:/usr/local/hadoop/share/hadoop/mapreduce$ hdfs dfs -tail output/part-r-00000 | tail
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
ubuntu@ip-172-31-39-208:/usr/local/hadoop/share/hadoop/mapreduce$ hdfs dfs -get output output
WARNING: HADOOP_PREFIX has been replaced by HADOOP_HOME. Using value of HADOOP_PREFIX.
ubuntu@ip-172-31-39-208:/usr/local/hadoop/share/hadoop/mapreduce$ ls
hadoop-mapreduce-client-app-3.3.1.jar      hadoop-mapreduce-client-hs-plugins-3.3.1.jar  hadoop-mapreduce-client-shuffle-3.3.1.jar  lib-examples
hadoop-mapreduce-client-common-3.3.1.jar   hadoop-mapreduce-client-jobclient-3.3.1-tests.jar  hadoop-mapreduce-client-uploader-3.3.1.jar  output
hadoop-mapreduce-client-core-3.3.1.jar     hadoop-mapreduce-client-jobclient-3.3.1.jar      hadoop-mapreduce-examples-3.3.1.jar        sources
hadoop-mapreduce-client-hs-3.3.1.jar       hadoop-mapreduce-client-native-task-3.3.1.jar      jdiff
ubuntu@ip-172-31-39-208:/usr/local/hadoop/share/hadoop/mapreduce$
```