

```
pes2ug20cs389@vishwa-VirtualBox:~$ ssh localhost
Welcome to Ubuntu 22.04.1 LTS (GNU/Linux 5.15.0-46-generic x86_64)

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

0 updates can be applied immediately.

Last login: Mon Aug 22 23:25:50 2022 from 127.0.0.1
pes2ug20cs389@vishwa-VirtualBox:~$
```

```
GNU nano 6.2 /home/pes2ug20cs389/.bashrc
fi

#Hadoop Path Configs
export HADOOP_HOME=/home/$USER/hadoop-3.3.3
export HADOOP_INSTALL=$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
```

```
GNU nano 6.2 /home/pes2ug20cs389/hadoop-3.3.3/etc/hadoop/hadoop-env.sh
###

# Technically, the only required environment variable is JAVA_HOME.
# All others are optional. However, the defaults are probably not
# preferred. Many sites configure these options outside of Hadoop,
# such as in /etc/profile.d

# The java implementation to use. By default, this environment
# variable is REQUIRED on ALL platforms except OS X!
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
```

```
GNU nano 6.2 /home/pes2ug20cs389/hadoop-3.3.3/etc/hadoop/core-site.xml
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
  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>hadoop.tmp.dir</name>
  <value>/home/pes2ug20cs389/tmpdata</value>
</property>
<property>
  <name>fs.default.name</name>
  <value>hdfs://127.0.0.1:9000</value>
</property>
</configuration>
```

GNU nano 6.2 /home/pes2ug20cs389/hadoop-3.3.3/etc/hadoop/hdfs-site.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
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>dfs.name.dir</name>
  <value>/home/pes2ug20cs389/dfsdata/namenode</value>
</property>
<property>
  <name>dfs.data.dir</name>
  <value>/home/pes2ug20cs389/dfsdata/datanode</value>
</property>
<property>
  <name>dfs.replication</name>
  <value>1</value>
</property>
```

Wrote 32 lines

GNU nano 6.2 /home/pes2ug20cs389/hadoop-3.3.3/etc/hadoop/mapred-site.xml \*

```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
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>
</configuration>
```

<b>^G</b> Help	<b>^O</b> Write Out	<b>^W</b> Where Is	<b>^K</b> Cut	<b>^T</b> Execute	<b>^C</b> Location
<b>^X</b> Exit	<b>^R</b> Read File	<b>^_</b> Replace	<b>^U</b> Paste	<b>^J</b> Justify	<b>^/</b> Go To Line



```
GNU nano 6.2 /home/pes2ug20cs389/hadoop-3.3.3/etc/hadoop/yarn-site.xml
<configuration>

<!-- Site specific YARN configuration properties -->
<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,CLA
SSPATH,PERPEND_DISTCACHE,HADOOP_YARN_HOME,HADOOP_MAPRED_HOME</value>
</property>
</configuration>

```

<b>^G</b> Help	<b>^O</b> Write Out	<b>^W</b> Where Is	<b>^K</b> Cut	<b>^T</b> Execute	<b>^C</b> Location
<b>^X</b> Exit	<b>^R</b> Read File	<b>^_</b> Replace	<b>^U</b> Paste	<b>^J</b> Justify	<b>^_</b> Go To Line

```
/namenode has been successfully formatted.
2022-08-24 11:19:35,009 INFO namenode.FSImageFormatProtobuf: Saving image file /home/pes2u
g20cs389/dfsdata/namenode/current/fsimage.ckpt_00000000000000000000 using no compression
2022-08-24 11:19:35,201 INFO namenode.FSImageFormatProtobuf: Image file /home/pes2ug20cs38
9/dfsdata/namenode/current/fsimage.ckpt_00000000000000000000 of size 408 bytes saved in 0 s
econds
2022-08-24 11:19:35,215 INFO namenode.NNStorageRetentionManager: Going to retain 1 images
with txid >= 0
2022-08-24 11:19:35,265 INFO namenode.FSNamesystem: Stopping services started for active s
tate
2022-08-24 11:19:35,265 INFO namenode.FSNamesystem: Stopping services started for standby
state
2022-08-24 11:19:35,278 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when
meet shutdown.
2022-08-24 11:19:35,278 INFO namenode.NameNode: SHUTDOWN_MSG:
/*****
SHUTDOWN_MSG: Shutting down NameNode at pes2ug20cs389-VirtualBox/127.0.1.1
*****/
pes2ug20cs389@pes2ug20cs389-VirtualBox:~$
```

```
pes2ug20cs389@pes2ug20cs389-VirtualBox:~/hadoop-3.3.3/sbin$ ./start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as pes2ug20cs389 in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [pes2ug20cs389-VirtualBox]
pes2ug20cs389-VirtualBox: Warning: Permanently added 'pes2ug20cs389-virtualbox' (ED25519)
to the list of known hosts.
Starting resourcemanager
Starting nodemanagers
pes2ug20cs389@pes2ug20cs389-VirtualBox:~/hadoop-3.3.3/sbin$ jps
3409 NameNode
3527 DataNode
3912 ResourceManager
4026 NodeManager
3706 SecondaryNameNode
4350 Jps
```

```
pes2ug20cs389@pes2ug20cs389-VirtualBox:~$ hdfs dfs -cat /example/output/part-r-00000
This      1
a         1
be        1
file      1
for       1
is        1
mapreduce      1
sample     1
text       1
that       1
used       1
will       1
wordcount      1
pes2ug20cs389@pes2ug20cs389-VirtualBox:~$
```

```
pes2ug20cs389@pes2ug20cs389-VirtualBox:~/UE20CS322-H1$ hdfs dfs -cat /handson/output-tech/
part-00000
54
pes2ug20cs389@pes2ug20cs389-VirtualBox:~/UE20CS322-H1$
```