Assignment 1

```
stud@stud-OptiPlex-3060:~$ sudo apt-get update
[sudo] password for stud:
Get:1 https://download.docker.com/linux/ubuntu bionic InRelease [64.4 kB]
Hit:2 http://in.archive.ubuntu.com/ubuntu jammy InRelease
Get:3 https://download.docker.com/linux/ubuntu bionic/stable amd64 Packages [36.6 kB]
Get:4 http://security.ubuntu.com/ubuntu jammy-security InRelease [110 kB]
Get:5 http://in.archive.ubuntu.com/ubuntu jammy-updates InRelease [119 kB]
Get:6 http://in.archive.ubuntu.com/ubuntu jammy-backports InRelease [108 kB]
Get:7 http://security.ubuntu.com/ubuntu jammy-security/main amd64 Packages [799 kB]
Get:8 http://security.ubuntu.com/ubuntu jammy-security/main i386 Packages [293 kB]
Get:9 http://security.ubuntu.com/ubuntu jammy-security/main Translation-en [156 kB]
Get:10 http://security.ubuntu.com/ubuntu jammy-security/main amd64 DEP-11 Metadata [41.
```

```
Need to get 262 MB of archives.
After this operation, 413 MB of additional disk space will be used.
Do you want to continue? [Y/n] t
Abort.
stud@stud-OptiPlex-3060:~$ sudo apt install default-jdk
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following packages were automatically installed and are no longer required:
  bridge-utils ubuntu-fan
Use 'sudo apt autoremove' to remove them.
The following additional packages will be installed:
  ca-certificates-java default-jdk-headless default-jre default-jre-headless
  fonts-dejavu-extra java-common libatk-wrapper-java libatk-wrapper-java-jni
  libice-dev libpthread-stubs0-dev libsm-dev libx11-dev libxau-dev libxcb1-dev
  libxdmcp-dev libxt-dev openjdk-11-jdk openjdk-11-jdk-headless openjdk-11-jre
  openjdk-11-jre-headless x11proto-dev xorg-sgml-doctools xtrans-dev
Suggested packages:
```

```
Setting up default-jdk (2:1.11-72build2) ...
Setting up libxt-dev:amd64 (1:1.2.1-1) ...
stud@stud-OptiPlex-3060:~$ java -version
openjdk version "11.0.18" 2023-01-17
OpenJDK Runtime Environment (build 11.0.18+10-post-Ubuntu-0ubuntu122.04)
OpenJDK 64-Bit Server VM (build 11.0.18+10-post-Ubuntu-0ubuntu122.04, mixed mode, sharing)
stud@stud-OptiPlex-3060:~$ sudo addgroup hadoop
Adding group `hadoop' (GID 1001) ...
Done.
stud@stud-OptiPlex-3060:~$ sudo adduser --ingroup hadoop hdp123
Adding user `hdp123' ...
Adding new user `hdp123' (1001) with group `hadoop' ...
Creating home directory `/home/hdp123' ...
Copying files from `/etc/skel' ...
New password:
BAD PASSWORD: The password is shorter than 8 characters
Retype new password:
passwd: password updated successfully
Changing the user information for hdp123
Enter the new value, or press ENTER for the default
        Full Name []:
        Room Number []:
        Work Phone []:
        Home Phone []:
        Other []:
Is the information correct? [Y/n] y
stud@stud-OptiPlex-3060:~$ sudo adduser hdp123 sudo
Adding user `hdp123' to group `sudo' ...
Adding user hdp123 to group sudo
Done.
stud@stud-OptiPlex-3060:~$ sudo apt-get - install openssh-server
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following packages were automatically installed and are no longer required:
```

```
Created symlink /etc/systemd/system/sshd.service → /lib/systemd/system/ssh.service.
Created symlink /etc/systemd/system/multi-user.target.wants/ssh.service → /lib/systemd/sy
rescue-ssh.target is a disabled or a static unit, not starting it.
ssh.socket is a disabled or a static unit, not starting it.
Processing triggers for man-db (2.10.2-1) ...
Processing triggers for ufw (0.36.1-4build1) ...
stud@stud-OptiPlex-3060:~$ su - hdp123
Password:
To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

hdp123@stud-OptiPlex-3060:~$ ssh-keygen -t rsa -P ""
Generating public/private rsa key pair.
Enter file in which to save the key (/home/hdp123/.ssh/id_rsa):
Created directory '/home/hdp123/.ssh'.
Your identification has been saved in /home/hdp123/.ssh/id_rsa
Your public key has been saved in /home/hdp123/.ssh/id_rsa.pub
The key fingerprint is:
SHA256:hg9x4XoKOnE5oUxjw/KllWtosZGygnIwa3eKfYln/ZQ hdp123@stud-OptiPlex-3060
The key's randomart image is:
+---[RSA 3072]----+
|         .       |
| . . . . .       |
|= X = . o        |
|.% % + =         |
|*.@.B.+ S         |
|+o+=+oo*  .       |
|  .o+ =...E      |
|    .+   o        |
|         .        |
+----[SHA256]-----+
hdp123@stud-OptiPlex-3060:~$ cat $HOME/.ssh/id_rsa.pub >> $HOME/.ssh/authorized_keys
hdp123@stud-OptiPlex-3060:~$ ssh localhost
The authenticity of host 'localhost (127.0.0.1)' can't be established.
ED25519 key fingerprint is SHA256:BYXNTVQhkEa3fJOrcpX+ValHor6ERnfexBLt5lGtHzc.
```

```
84 updates could not be installed automatically. For more details,
see /var/log/unattended-upgrades/unattended-upgrades.log

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.

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

hdp123@stud-OptiPlex-3060:~$ exit
logout
Connection to localhost closed.
hdp123@stud-OptiPlex-3060:~$ pwd
/home/hdp123
hdp123@stud-OptiPlex-3060:~$ cd /home
hdp123@stud-OptiPlex-3060:/home$ sudo su
[sudo] password for hdp123:
root@stud-OptiPlex-3060:/home#
root@stud-OptiPlex-3060:/home# ls
hdp123  stud
root@stud-OptiPlex-3060:/home# su hdp123
hdp123@stud-OptiPlex-3060:/home$ ls
hdp123  stud
hdp123@stud-OptiPlex-3060:/home$ cd
hdp123@stud-OptiPlex-3060:~$ ls
snap
hdp123@stud-OptiPlex-3060:~$
```

```
gzip: stdin: invalid compressed data--format violated
tar: Unexpected EOF in archive
tar: Unexpected EOF in archive
tar: Error is not recoverable: exiting now
stud@stud-OptiPlex-3060:~/Desktop$ ls
CNS   hadoop-2.10.1   hadoop-2.10.1.tar.gz   s190088608   s190088611
stud@stud-OptiPlex-3060:~/Desktop$ sudo mv hadoop-2.10.1 /usr/local/hadoop
[sudo] password for stud:
stud@stud-OptiPlex-3060:~/Desktop$ sudo chown -R hdp123 /usr/local
stud@stud-OptiPlex-3060:~/Desktop$
```

```
hdp123@stud-OptiPlex-3060: /usr/local/hadoop/etc/hadoop      ×          stud@stud-OptiPlex-3060: ~/Desktop        ×   ⌄

hdp123@stud-OptiPlex-3060:/home$ ls
hdp123   stud
hdp123@stud-OptiPlex-3060:/home$ cd
hdp123@stud-OptiPlex-3060:~$ ls
snap
hdp123@stud-OptiPlex-3060:~$ sudo nano ~/.bashrc
[sudo] password for hdp123:
hdp123@stud-OptiPlex-3060:~$ source ~/.bashrc
hdp123@stud-OptiPlex-3060:~$ sudo nano /usr/local/hadoop/etc/hadoop/hadoop-env.sh
hdp123@stud-OptiPlex-3060:~$ cd
hdp123@stud-OptiPlex-3060:~$ su hdp123
Password:
hdp123@stud-OptiPlex-3060:~$ cd /usr/local/hadoop/etc/hadoop
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ nano core-site.xml
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ nano hdfs-site.xml
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ nano yarn-site.xml
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ sudo cp /usr/local/hadoop/etc/hadoop/mapred-site.xml.template /usr/local/hadoop/etc/ha
doop/mapred-site.xml
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ sudo nano mapred-site.xml
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ sudo mkdir -p /usr/local/hadoop_tmp
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ sudo mkdir -p /usr/local/hadoop_tmp/hdfs/namenode
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ sudo mkdir -p /usr/local/hadoop_tmp/hdfs/datanode
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ sudo chown -R hdp123 /usr/local/hadoop_tmp
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ hdfs namenode -format
23/04/28 12:48:52 INFO namenode.NameNode: STARTUP_MSG:
/************************************************************
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = stud-OptiPlex-3060/127.0.1.1
STARTUP_MSG:   args = [-format]
STARTUP_MSG:   version = 2.10.1
```

```
  GNU nano 6.2                          /usr/local/hadoop/etc/hadoop/hadoop-env.sh *
# HDFS Mover specific parameters
###
# Specify the JVM options to be used when starting the HDFS Mover.
# These options will be appended to the options specified as HADOOP_OPTS
# and therefore may override any similar flags set in HADOOP_OPTS
#
# export HADOOP_MOVER_OPTS=""

###
# Router-based HDFS Federation specific parameters
# Specify the JVM options to be used when starting the RBF Routers.
# These options will be appended to the options specified as HADOOP_OPTS
# and therefore may override any similar flags set in HADOOP_OPTS
#
# export HADOOP_DFSROUTER_OPTS=""
###


###
# Advanced Users Only!
###

# The directory where pid files are stored. /tmp by default.
# NOTE: this should be set to a directory that can only be written to by
#       the user that will run the hadoop daemons.  Otherwise there is the
#       potential for a symlink attack.
export HADOOP_PID_DIR=${HADOOP_PID_DIR}
export HADOOP_SECURE_DN_PID_DIR=${HADOOP_PID_DIR}

# A string representing this instance of hadoop. $USER by default.
export HADOOP_IDENT_STRING=$USER


export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64
```

```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>
</configuration>

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

```xml
  limitations under the License. See accompanying LICENSE file.
-->

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

<configuration>

</configuration>

<property>
  <name>dfs.replication</name>
  <value>1</value>
</property>
<property>
  <name>dfs.namenode.name.dir</name>
  <value>file:/usr/local/hadoop_tmp/hdfs/namenode</value>
</property>
<property>
 <name>dfs.datanode.data.dir</name>
 <value>file:/usr/local/hadoop_tmp/hdfs/datanode</value>
</property>
```

```xml
  See the License for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->
<configuration>

<!-- Site specific YARN configuration properties -->

</configuration>

<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>
```

```
  GNU nano 6.2                                              mapred-site.xml *


  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>


</configuration>



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

```
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ sudo nano mapred-site.xml
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ sudo mkdir -p /usr/local/hadoop_tmp
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ sudo mkdir -p /usr/local/hadoop_tmp/hdfs/namenode
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ sudo mkdir -p /usr/local/hadoop_tmp/hdfs/datanode
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ sudo chown -R hdp123 /usr/local/hadoop_tmp
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ hdfs namenode -format
23/04/28 12:48:52 INFO namenode.NameNode: STARTUP_MSG:
/************************************************************
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = stud-OptiPlex-3060/127.0.1.1
STARTUP_MSG:   args = [-format]
STARTUP_MSG:   version = 2.10.1
```

```
         at org.apache.hadoop.conf.Configuration.loadResources(Configuration.java:2755)
         at org.apache.hadoop.conf.Configuration.getProps(Configuration.java:2635)
         at org.apache.hadoop.conf.Configuration.get(Configuration.java:1097)
         at org.apache.hadoop.conf.Configuration.getTimeDuration(Configuration.java:1704)
         at org.apache.hadoop.conf.Configuration.getTimeDuration(Configuration.java:1685)
         at org.apache.hadoop.util.ShutdownHookManager.getShutdownTimeout(ShutdownHookManager.java:183)
         at org.apache.hadoop.util.ShutdownHookManager.shutdownExecutor(ShutdownHookManager.java:145)
         at org.apache.hadoop.util.ShutdownHookManager.access$300(ShutdownHookManager.java:65)
         at org.apache.hadoop.util.ShutdownHookManager$1.run(ShutdownHookManager.java:102)
Caused by: com.ctc.wstx.exc.WstxParsingException: Illegal to have multiple roots (start tag in epilog?).
 at [row,col,system-id]: [22,2,"file:/usr/local/hadoop/etc/hadoop/core-site.xml"]
         at com.ctc.wstx.sr.StreamScanner.constructWfcException(StreamScanner.java:621)
         at com.ctc.wstx.sr.StreamScanner.throwParseError(StreamScanner.java:491)
         at com.ctc.wstx.sr.StreamScanner.throwParseError(StreamScanner.java:475)
         at com.ctc.wstx.sr.BasicStreamReader.handleExtraRoot(BasicStreamReader.java:2242)
         at com.ctc.wstx.sr.BasicStreamReader.nextFromProlog(BasicStreamReader.java:2156)
         at com.ctc.wstx.sr.BasicStreamReader.closeContentTree(BasicStreamReader.java:2991)
         at com.ctc.wstx.sr.BasicStreamReader.nextFromTree(BasicStreamReader.java:2734)
         at com.ctc.wstx.sr.BasicStreamReader.next(BasicStreamReader.java:1123)
         at org.apache.hadoop.conf.Configuration.loadResource(Configuration.java:2826)
         ... 9 more
hdp123@stud-OptiPlex-3060:/usr/local/hadoop/etc/hadoop$ start-dfs.sh
23/04/28 12:49:47 ERROR conf.Configuration: error parsing conf core-site.xml
com.ctc.wstx.exc.WstxParsingException: Illegal to have multiple roots (start tag in epilog?).
 at [row,col,system-id]: [22,2,"file:/usr/local/hadoop/etc/hadoop/core-site.xml"]
         at com.ctc.wstx.sr.StreamScanner.constructWfcException(StreamScanner.java:621)
         at com.ctc.wstx.sr.StreamScanner.throwParseError(StreamScanner.java:491)
         at com.ctc.wstx.sr.StreamScanner.throwParseError(StreamScanner.java:475)
         at com.ctc.wstx.sr.BasicStreamReader.handleExtraRoot(BasicStreamReader.java:2242)
         at com.ctc.wstx.sr.BasicStreamReader.nextFromProlog(BasicStreamReader.java:2156)
```

```
cable
hdp123@stud-OptiPlex-3060:~$ start-yarn.sh
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-hdp123-resourcemanager-stud-OptiPlex-3060.out
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/
lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
Enter passphrase for key '/home/hdp123/.ssh/id_rsa':
hdp123@localhost's password:
localhost: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-hdp123-nodemanager-stud-OptiPlex-3060.out
localhost: WARNING: An illegal reflective access operation has occurred
localhost: WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/had
oop/common/lib/hadoop-auth-2.10.1.jar) to method sun.security.krb5.Config.getInstance()
localhost: WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
localhost: WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
localhost: WARNING: All illegal access operations will be denied in a future release
hdp123@stud-OptiPlex-3060:~$ jps
9968 NodeManager
9267 DataNode
9076 NameNode
9476 SecondaryNameNode
10101 Jps
9645 ResourceManager
```