ASSIGNMENT_2

Common confifuration:

1. Node1 and Node2 hadoop-env.sh

2. Node1 and node2 core-site.xml

```
GNU nano 6.2                                                        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>fs.defaultFS</name>
<value>hdfs://manager:9000</value>
</property>
</configuration>
```

```
                                       [ Read 24 lines ]
^G Help        ^O Write Out   ^W Where Is    ^K Cut        ^T Execute     ^C Location    M-U Undo       M-A Set Mark   M-] To Bracket
^X Exit        ^R Read File   ^\ Replace     ^U Paste      ^J Justify     ^/ Go To Line  M-E Redo       M-6 Copy       ^Q Where Was
```

3. Manager **yarn-site.xml**

```
GNU nano 6.2                                    yarn-site.xml
<?xml version="1.0"?>
<!--
  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.
-->
<configuration>

<property>
<name>yarn.resourcemanager.hostname</name>
<value>manager</value>
</property>
</configuration>
```
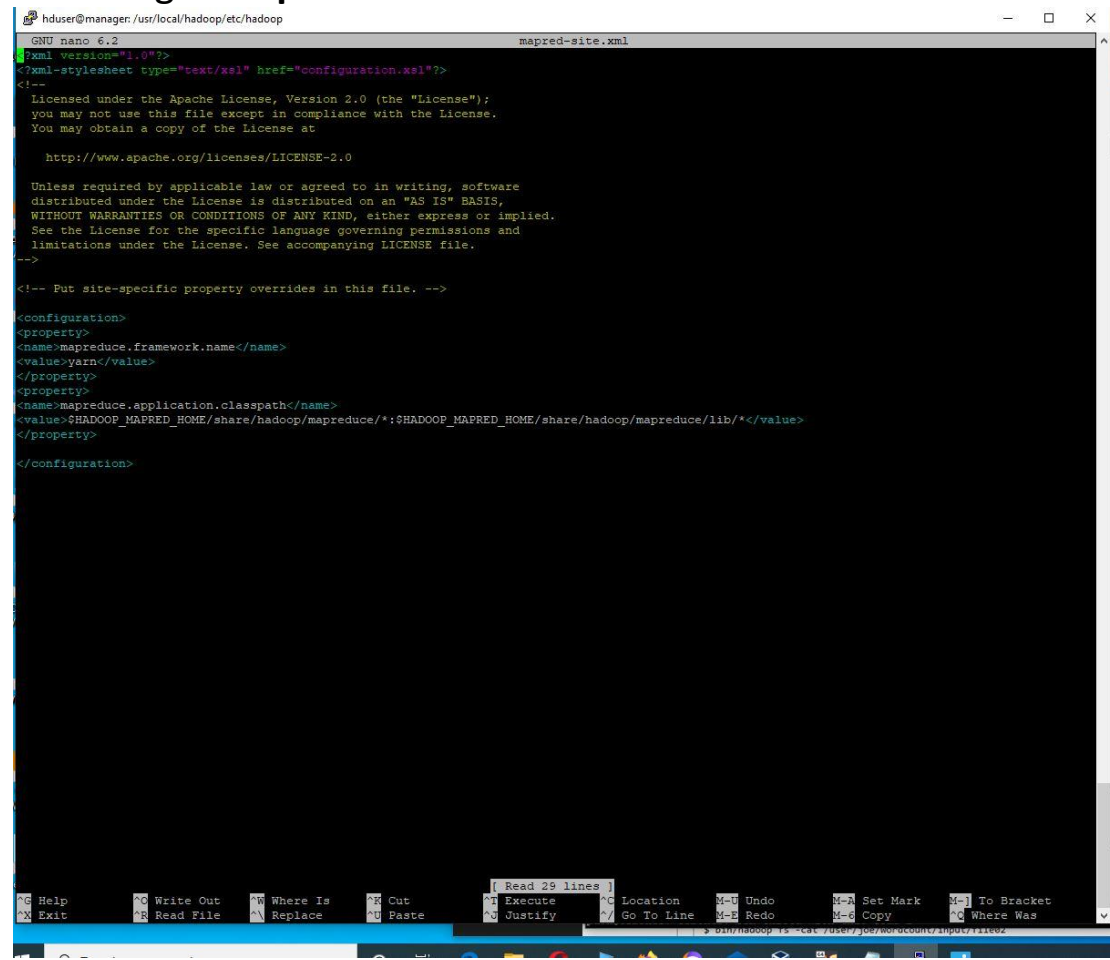
```
                                         [ Read 21 lines ]
^G Help       ^O Write Out   ^W Where Is    ^K Cut        ^T Execute    ^C Location    M-U Undo    M-A Set Mark   M-] To Bracket
^X Exit       ^R Read File   ^\ Replace     ^U Paste      ^J Justify    ^/ Go To Line  M-E Redo    M-6 Copy       ^Q Where Was
```

4. Manager **mapred-site.xml**

```
GNU nano 6.2                                              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>
<property>
<name>mapreduce.application.classpath</name>
<value>$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/*:$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/lib/*</value>
</property>

</configuration>



                                          [ Read 29 lines ]
^G Help       ^O Write Out   ^W Where Is   ^K Cut       ^T Execute    ^C Location   M-U Undo   M-A Set Mark   M-] To Bracket
^X Exit       ^R Read File   ^\ Replace    ^U Paste     ^J Justify    ^/ Go To Line M-E Redo   M-6 Copy       ^Q Where Was
```
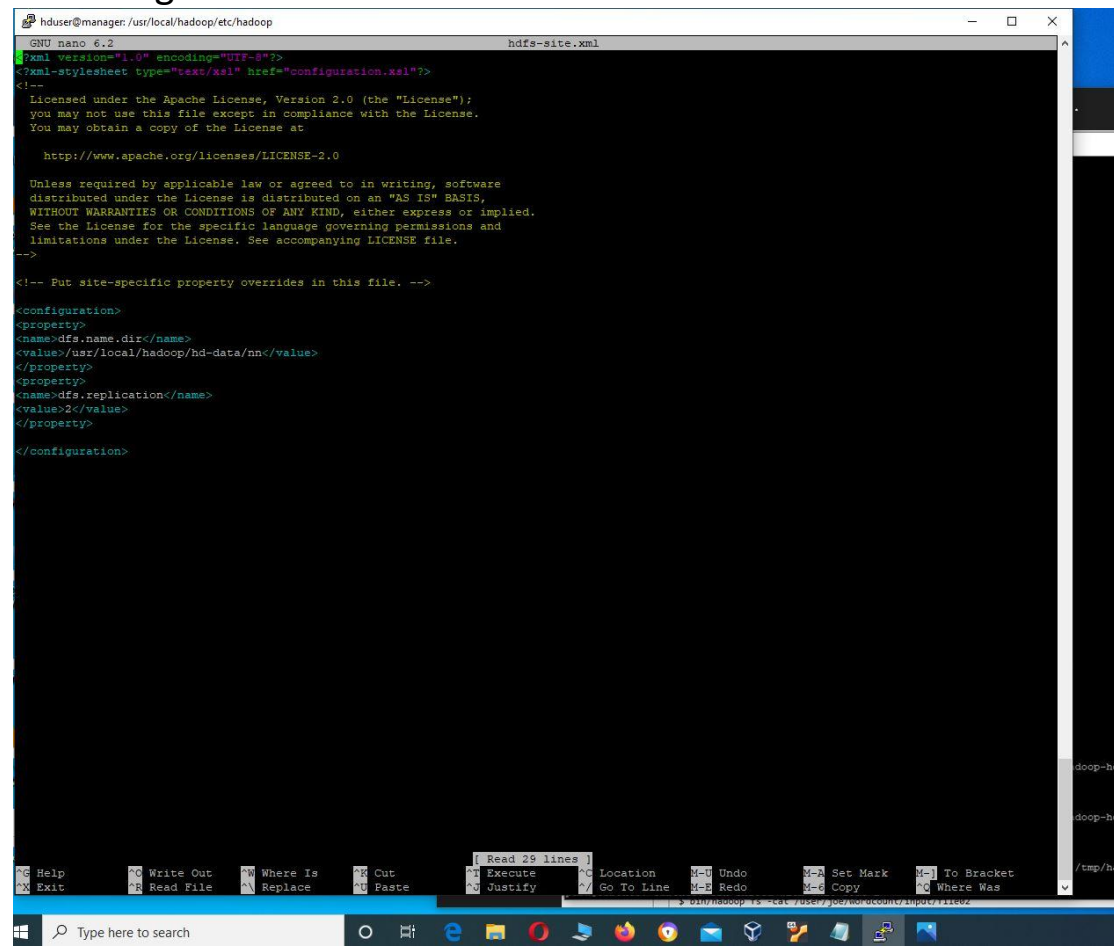
5. Manager **hdfs-site.xml**


```
GNU nano 6.2                                          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>/usr/local/hadoop/hd-data/nn</value>
</property>
<property>
<name>dfs.replication</name>
<value>2</value>
</property>


</configuration>
```

6. Manager hadoop-env.sh


```
GNU nano 6.2                                            hadoop-env.sh
# 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 HDFS_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 HDFS_DFSROUTER_OPTS=""

###
# HDFS StorageContainerManager specific parameters
###
# Specify the JVM options to be used when starting the HDFS Storage Container Manager.
# These options will be appended to the options specified as HADOOP_OPTS
# and therefore may override any similar flags set in HADOOP_OPTS
#
# export HDFS_STORAGECONTAINERMANAGER_OPTS=""

###
# Advanced Users Only!
###

#
# When building Hadoop, one can add the class paths to the commands
# via this special env var:
# export HADOOP_ENABLE_BUILD_PATHS="true"

#
# To prevent accidents, shell commands be (superficially) locked
# to only allow certain users to execute certain subcommands.
# It uses the format of (command)_(subcommand)_USER.
#
# For example, to limit who can execute the namenode command,
# export HDFS_NAMENODE_USER=hdfs
#JAVA
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64/
export JRE_HOME=/usr/lib/jvm/java-8-openjdk-amd64/jre
#Hadoop Environment Variables
export HADOOP_HOME=/usr/local/hadoop
export HADOOP_CONF_DIR=/usr/local/hadoop/etc/hadoop
export HADOOP_LOG_DIR=$HADOOP_HOME/logs
export HDFS_NAMENODE_USER=hduser
export HDFS_DATANODE_USER=hduser
export HDFS_SECONDARYNAMENODE_USER=hduser
export YARN_RESOURCEMANAGER_USER=hduser
export YARN_NODEMANAGER_USER=hduser
export YARN_NODEMANAGER_USER=hduser
# Add Hadoop bin/ directory to PATH
export PATH=$PATH:$HADOOP_HOME/sbin:$HADOOP_HOME/bin
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_HOME/lib/native
export HADOOP_OPTS="-Djava.library.path=$HADOOP_HOME/lib/native"

^G Help        ^O Write Out    ^W Where Is    ^K Cut        ^T Execute    ^C Location    M-U Undo    M-A Set Mark    M-] To Bracket
^X Exit        ^R Read File    ^\ Replace     ^U Paste      ^J Justify    ^/ Go To Line  M-E Redo    M-6 Copy        ^Q Where Was
```
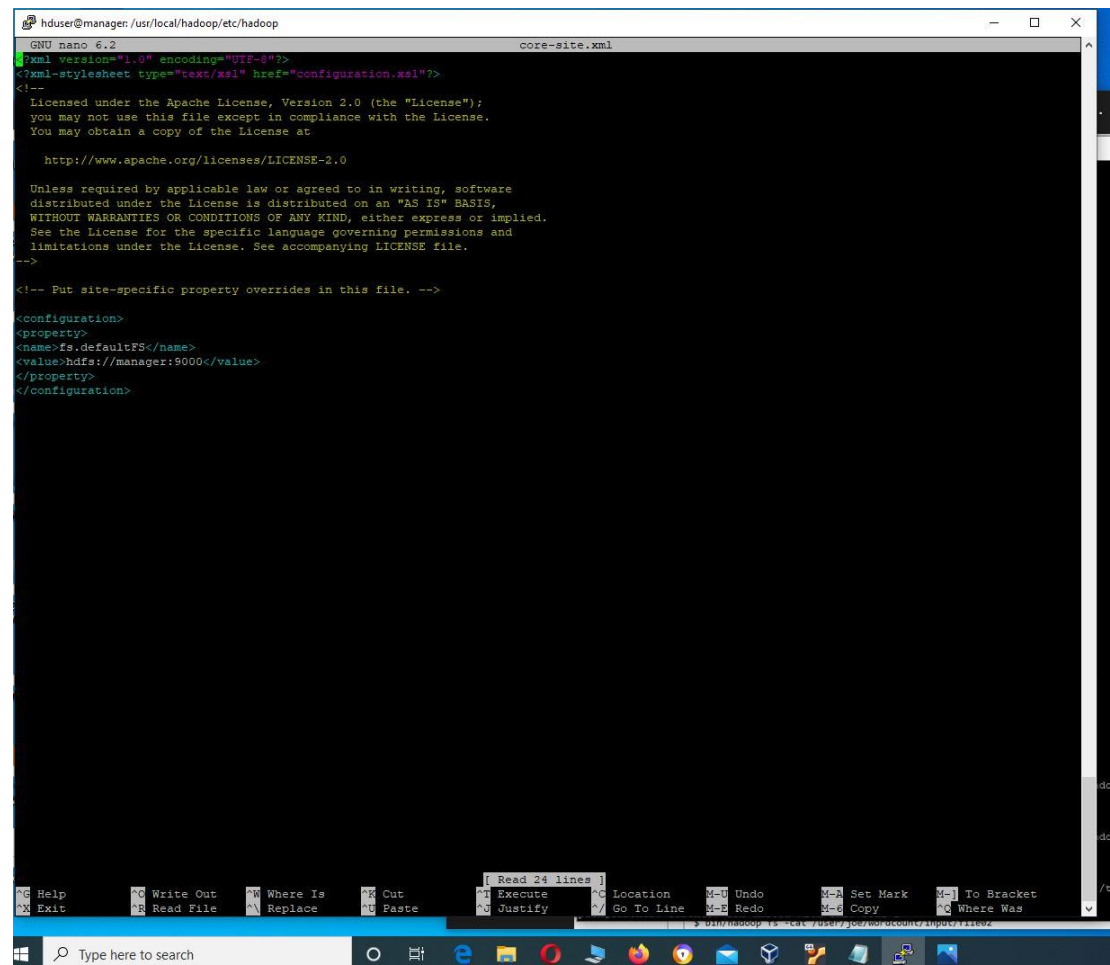
7. Manager **core-site.xml**

8. Node1 and Node2 **yarn-site.xml**

```
  GNU nano 6.2                                                yarn-site.xml
<?xml version="1.0"?>
<!--
  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.
-->
<configuration>
<property>
<name>yarn.resourcemanager.hostname</name>
<value>master</value>
</property>
<property>
<name>yarn.nodemanager.aux-services</name>
<value>mapreduce_shuffle</value>
</property>
<property>
<name>yarn.nodemanager.local-dirs</name>
<value>/usr/local/hadoop/hd-data/yarn/data</value>
</property>
<property>
<name>yarn.nodemanager.logs-dirs</name>
<value>/usr/local/hadoop/hd-data/yarn/logs</value>
</property>
<property>
<name>yarn.nodemanager.disk-health-checker.max-disk-utilization-perdisk-percentage</name>
<value>99.9</value>
</property>
<property>
<name>yarn.nodemanager.vmem-check-enabled</name>
<value>false</value>
</property>
<property>
<name>yarn.nodemanager.env-whitelist</name>
<value>JAVA_HOME,HADOOP_COMMON_HOME,HADOOP_HDFS_HOME,HADOOP_CONF_DIR,CLASSPATH_PREPEND_DISTCACHE,HADOOP_YARN_HOME,HADOOP_MAPRED_HOME</value>
</property>


</configuration>

                                          [ Read 45 lines ]
^G Help        ^O Write Out   ^W Where Is    ^K Cut         ^T Execute     ^C Location    M-U Undo       M-A Set Mark   M-] To Bracket
^X Exit        ^R Read File    ^\ Replace    ^U Paste       ^J Justify     ^/ Go To Line  M-E Redo       M-6 Copy       ^Q Where Was
```
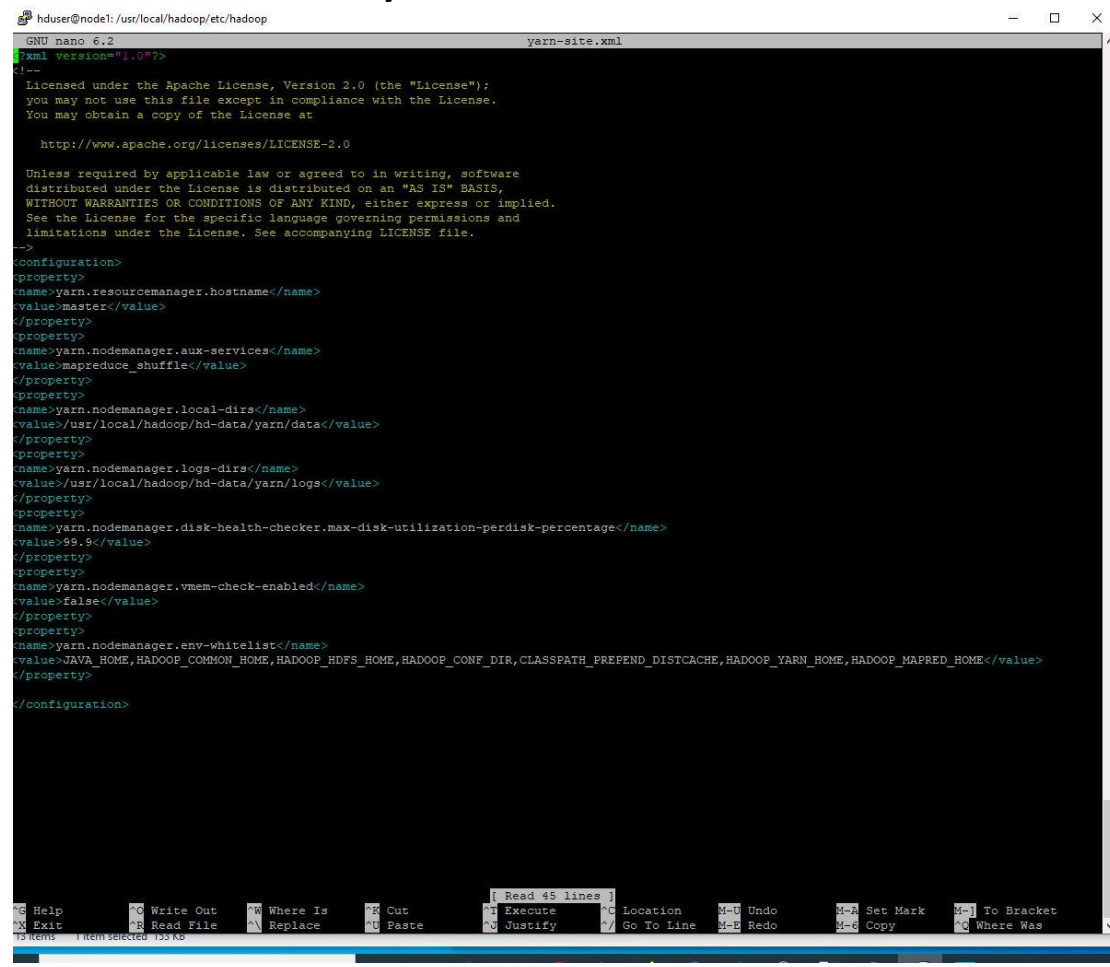
9. Node1 and Node2 **mapred-site.xml**


```
GNU nano 6.2                                        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>
<property>
<name>mapreduce.application.classpath</name>
<value>$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/*:$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/lib/*</value>
</property>

</configuration>
```

Different configuration

1. namenode **workers**

2. Node1 and Node2 **workers**

GNU nano 6.2   workers

localhost

```
^G Help        ^O Write Out    ^W Where Is     ^K Cut        ^T Execute     ^C Location    M-U Undo    M-A Set Mark   M-] To Bracket
^X Exit        ^R Read File    ^\ Replace      ^U Paste      ^J Justify     ^/ Go To Line  M-E Redo    M-6 Copy       ^Q Where Was
```

Jps

Stop

```
hduser@manager: ~                                                    —  □  ×

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

254 updates can be applied immediately.
1 of these updates is a standard security update.
To see these additional updates run: apt list --upgradable

New release '24.04.1 LTS' available.
Run 'do-release-upgrade' to upgrade to it.

Last login: Mon Oct 21 09:16:44 2024 from 192.168.56.104
hduser@node1:~$ nano yarn-site.xml
hduser@node1:~$ ls
Desktop   Documents   Downloads   hadoop-3.2.4   hadoop.tar.gz   Music   Pictures   Public   snap   Templates   Videos
hduser@node1:~$ cd /usr/local/hadoop/etc/hadoop
hduser@node1:/usr/local/hadoop/etc/hadoop$ ls
capacity-scheduler.xml        hadoop-policy.xml               kms-acls.xml            mapred-queues.xml.template   yarn-env.cmd
configuration.xsl             hadoop-user-functions.sh.example  kms-env.sh            mapred-site.xml              yarn-env.sh
container-executor.cfg        hdfs-site.xml                   kms-log4j.properties    shellprofile.d               yarnservice-log4j.properties
core-site.xml                 httpfs-env.sh                   kms-site.xml            ssl-client.xml.example       yarn-site.xml
hadoop-env.cmd                httpfs-log4j.properties         log4j.properties        ssl-server.xml.example
hadoop-env.sh                 httpfs-signature.secret         mapred-env.cmd          user_ec_policies.xml.template
hadoop-metrics2.properties    httpfs-site.xml                 mapred-env.sh           workers
hduser@node1:/usr/local/hadoop/etc/hadoop$ nano yarn-site.xml
hduser@node1:/usr/local/hadoop/etc/hadoop$ jps
6544 DataNode
7098 Jps
hduser@node1:/usr/local/hadoop/etc/hadoop$ ssh manager
hduser@manager's password:
Welcome to Ubuntu 22.04 LTS (GNU/Linux 6.8.0-47-generic x86_64)

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

253 updates can be applied immediately.
1 of these updates is a standard security update.
To see these additional updates run: apt list --upgradable

New release '24.04.1 LTS' available.
Run 'do-release-upgrade' to upgrade to it.

Last login: Mon Oct 21 09:26:10 2024 from 192.168.56.105
hduser@manager:~$ cd /usr/local/hadoop/etc/hadoop
hduser@manager:/usr/local/hadoop/etc/hadoop$ jps
7848 ResourceManager
7115 NameNode
8780 Jps
7583 SecondaryNameNode
hduser@manager:/usr/local/hadoop/etc/hadoop$ cd
hduser@manager:~$ stop-all.sh
WARNING: Stopping all Apache Hadoop daemons as hduser in 10 seconds.
WARNING: Use CTRL-C to abort.
Stopping namenodes on [manager]
Stopping datanodes
Stopping secondary namenodes [manager]
Stopping nodemanagers
Stopping resourcemanager
hduser@manager:~$
5 items
```