Hadoop Single Node Commands

sudo apt update
sudo apt install ssh
sudo apt install pdsh
nano .bashrc
export PDSH_RCMD_TYPE=ssh
ssh-keygen -t rsa -P ""
cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
ssh localhost
sudo apt install openjdk-8-jdk
java -version
sudo wget -P ~
https://archive.apache.org/dist/hadoop/common/hadoop-3.2.1/hadoop-3.2.1.tar.gz
tar xzf hadoop-3.2.1.tar.gz
mv hadoop-3.2.1 hadoop
nano ~/hadoop/etc/hadoop/hadoop-env.sh

Goes in Hadoop-env.sh

      export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64/

sudo mv hadoop /usr/local/hadoop
sudo nano /etc/environment

PATH="/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/bin:/usr/games:/usr/local/games:/usr/
local/hadoop/bin:/usr/local/hadoop/sbin"JAVA_HOME="/usr/lib/jvm/java-8-openjdk-amd64/jre"

sudo adduser hadoopuser
sudo usermod -aG hadoopuser hadoopuser
sudo chown hadoopuser:root -R /usr/local/hadoop/
sudo chmod g+rwx -R /usr/local/hadoop/
sudo adduser hadoopuser sudo

sudo nano /etc/hosts
sudo nano /etc/hostname
sudo reboot

```
su - hadoopuser
ssh-keygen -t rsa
```

```
ssh-copy-id hadoopuser@hadoop-master
```

```
sudo nano /usr/local/hadoop/etc/hadoop/core-site.xml
```

```xml
<configuration>
   <property>
      <name>fs.defaultFS</name>
      <value>hdfs://localhost:9000</value>
   </property>
   <property>
      <name>hadoop.tmp.dir</name>
      <value>/usr/local/hadoop</value>
   </property>
</configuration>
```

```
sudo nano /usr/local/hadoop/etc/hadoop/hdfs-site.xml
```

```xml
<configuration>
   <property>
      <name>dfs.replication</name>
      <value>1</value>
   </property>
</configuration>
```

```
sudo nano /usr/local/hadoop/etc/hadoop/mapred-site.xml

<configuration>
   <property>
      <name>mapreduce.framework.name</name>
      <value>yarn</value>
   </property>
   <property>
      <name>yarn.app.mapreduce.am.env</name>
      <value>HADOOP_MAPRED_HOME=/usr/local/hadoop</value>
   </property>
   <property>
      <name>mapreduce.map.env</name>
      <value>HADOOP_MAPRED_HOME=/usr/local/hadoop</value>
   </property>
   <property>
      <name>mapreduce.reduce.env</name>
      <value>HADOOP_MAPRED_HOME=/usr/local/hadoop</value>
   </property>
</configuration>
```

```
source /etc/environment
hdfs namenode -format
```

```
start-dfs.sh
```

```
nano .bashrc
```

```
export HADOOP_HOME="/usr/local/hadoop"
export HADOOP_COMMON_HOME=$HADOOP_HOME
export HADOOP_CONF_DIR=$HADOOP_HOME/etc/hadoop
```

```
export HADOOP_HDFS_HOME=$HADOOP_HOME
export HADOOP_MAPRED_HOME=$HADOOP_HOME
export HADOOP_YARN_HOME=$HADOOP_HOME
source ~/.bashrc

sudo nano /usr/local/hadoop/etc/hadoop/yarn-site.xml

<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>
</configuration>


start-yarn.sh

sudo apt remove pdsh
```