Hadoop-2 Single Node Cluster Creation Installation

1.Download Hadoop and Java

Java Download command = sudo apt install default-jdk

Download Hadoop software from google

tar -zxvf hadoop-2.9.1.tar.gz (Extract the tar file)

/////sudo apt-get install vim (Install USER Friendly Editer)

vi .bashrc (Set the java Path in your Home Path)

Hadoop environment variables

export JAVA_HOME=/usr/lib/jvm/java-1.11.0-openjdk-amd64

export HADOOP_HOME=/home/kumaran/hadoop-3.2.4

export HADOOP_INSTALL=$HADOOP_HOME

export HADOOP_MAPRED_HOME=$HADOOP_HOME

export HADOOP_COMMON_HOME=$HADOOP_HOME

export HADOOP_HDFS_HOME=$HADOOP_HOME

export HADOOP_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"

export JAVA_HOME=/home/username/jdk1.8.0_45

export PATH=HOME/bin:JAVA_HOME/bin:PATH

source .bashrc (Execute the bashrc file)

echo JAVA_HOME (Check the java path)

==========

2. Modify Hadoop Configuration Files

NAMENODE ----> core-site.xml

RESOURCE MANGER ----> mapperd-site.xml

SECONDARYNAMENODE ---->

DATANODE ----> slaves

NODEMANGER ----> slaves & yarn-site.xml

vi etc/hadoop/core-site.xml

<property>

<name>fs.default.name</name>

<value>hdfs://localhost:9000</value>

</property>

vi etc/hadoop/yarn-site.xml

<property>

<name>yarn.nodemanager.aux-services</name>

<value>mapreduce_shuffle</value>

</property>

<property>

<name>yarn.nodemanager.auxservices.mapreduce.shuffle.class</name>

<value>org.apache.hadoop.mapred.ShuffleHandler</value>

</property>

vi etc/hadoop/hdfs-site.xml

<property>

<name>dfs.replication</name>

<value>1</value>

</property>

<property>

<name>dfs.permission</name>

<value>false</value>

</property>

vi etc/hadoop/mapred-site.xml

<property>

<name>mapreduce.framework.name</name>

<value>yarn</value>

</property>

vi etc/hadoop/hadoop-env.sh

export JAVA_HOME=/usr/lib/jvm/java-1.11.0-openjdk-amd64

vi etc/hadoop/mapred-env.sh

export JAVA_HOME=/usr/lib/jvm/java-1.11.0-openjdk-amd64

vi etc/hadoop/yarn-env.sh

export JAVA_HOME=/usr/lib/jvm/java-1.11.0-openjdk-amd64

vi etc/hadoop/slaves

localhost

3.Install the ssh key

(Generates, Manages and Converts Authentication keys)

sudo apt-get install openssh-server

ssh-keygen -t rsa

(Setup passwordless ssh to localhost and to slaves)

cd .ssh

ls

cat id_rsa.pub >> authorized_keys (copy the .pub)

(Copy the id_rsa.pub from NameNode to authorized_keys in all machines)

ssh localhost

(Asking No Password)

==========

4. Format NameNode

cd hadoop-2.9.1

bin/hadoop namenode -format (Your Hadoop File System Ready)

==========

5. Start All Hadoop Related Services

sbin/start-all.sh

(Starting Daemon's For DFS & YARN)

NameNode

DataNode

SecondaryNameNode

ResourceManager

NodeManager

(check the Browser Web GUI)

NameNode - http://localhost:50070/

Resource Manager - http://localhost:8088/

==========

6.Stop All Hadoop and Yarn Related Services

sbin/stop-all.sh

7.To check the running port

sudo apt instal net-tools

sudo netstat -tulpn