Ex No 1

Downloading and installing Hadoop on Mac, Understanding different Hadoop modes, Startup scripts, Configuration files.

AIM:

To Download and install Hadoop, Understanding different Hadoop modes, Startup scripts, Configuration files.

PROCEDURE:

Step 1: Download/Update Homebrew

To set up Hadoop on a Mac OS we need to have Homebrew installed on our system. Homebrew is a package manager specifically designed for the Mac operating system making it simple to install Unix applications. Homebrew you can easily install it by following the instructions provided at this link:

https://brew.sh/

Alternatively you can quickly install Homebrew by typing the following command into your terminal:

```
/bin/bash -c "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"
```

Update the repository index of the Homebrew package installer. This can be done through the brew update command.

```
brew update
```

Add below lines to ***.bash_profile*** or ***.zprofile***

```
eval "$(/opt/homebrew/bin/brew shellenv)"
```

Step 2: Install Java 1.8 version

Confirm the correct version of java (version 8) installed on the machine. If any other version than 1.8., **make sure to install 1.8 as Hadoop 3.3.6/3.4.0 have dependency and only supported in Java 1.8 versions**

Brew command to cast Java to version 8 / install java 8 using brew
`brew install - cask homebrew/cask-versions/adoptopenjdk8`

Add below lines to ***.bash_profile*** or ***.zprofile***
`export JAVA_HOME=$(/usr/libexec/java_home)`

`java -version`

`openjdk version "1.8.0_292"`

`OpenJDK Runtime Environment (AdoptOpenJDK)(build 1.8.0_292-b10)`

`OpenJDK 64-Bit Server VM (AdoptOpenJDK)(build 25.292-b10, mixed mode)`

Step 3: Install Hadoop 3.3.6/3.4.0 using homebrew

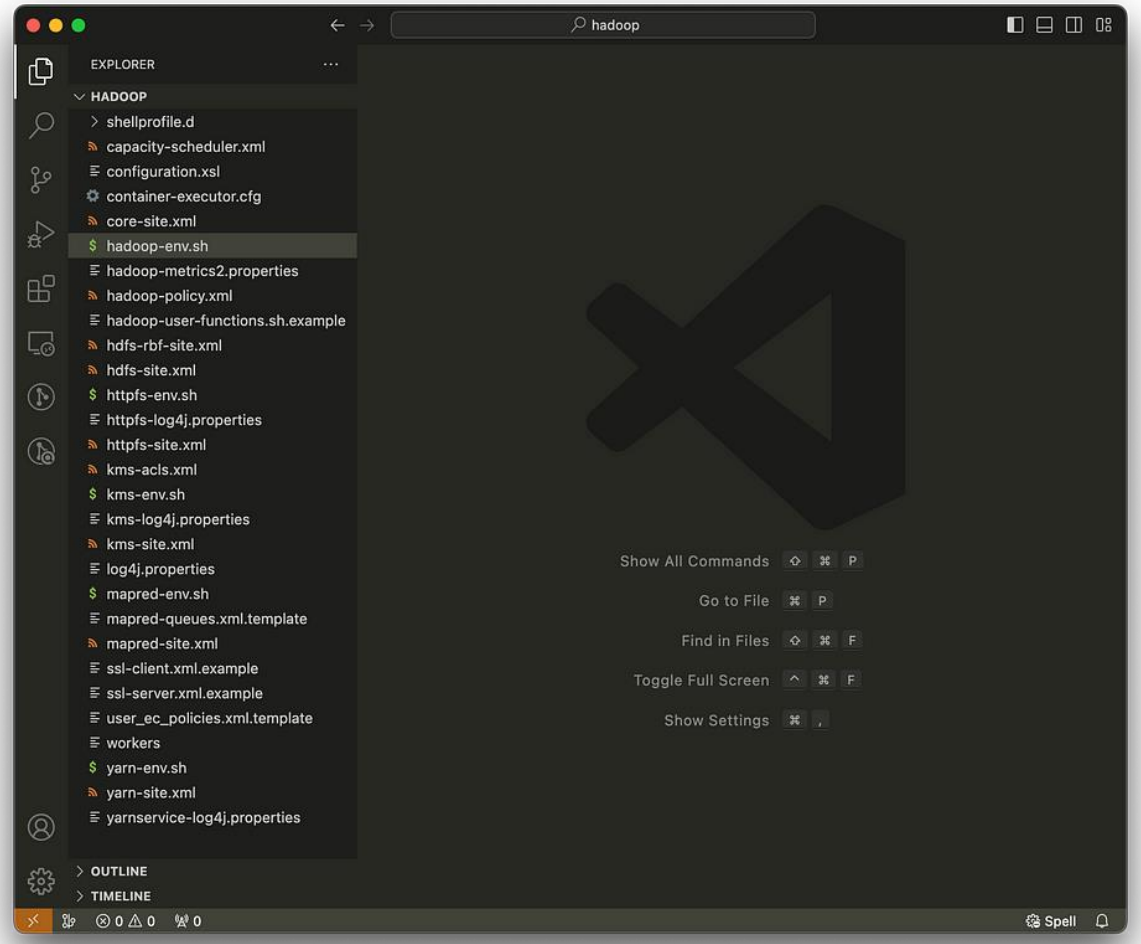
To install hadoop use brew command:
`brew install hadoop`

Check your installed hadoop version 3.3.6/3.4.0
`hadoop version`

Move to hadoop directory and verify the files/path :
`cd /opt/homebrew/Cellar/hadoop/3.4.0`
or cd /opt/homebrew/Cellar/hadoop/3.3.6 based on hadoop version

Step 4: Updated Hadoop config files -> 5 files

Go into the directory:
cd /opt/homebrew/Cellar/hadoop/3.4.0/libexec/etc/hadoop
Open this directory in any code editor of choice. E.g — vscode, and update following files with the given code add in the end.



Get Java home path using below command:
/usr/libexec/java_home

-> FILE 1: hadoop-env.sh

Add the code line to the end of the file and replace the path with your Java home path from your local system

```
export
JAVA_HOME=/Library/Java/JavaVirtualMachines/adoptopenjdk-8.jdk/Contents/Home
```

Following are file names and code to add to configure hadoop:

-> FILE 2: core-site.xml

```
<configuration>
<property>
<name>fs.defaultFS</name>
<value>hdfs://localhost:9000</value>
<final>true</final>
</property>
</configuration>
```

-> FILE 3: hdfs-site.xml

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

-> FILE 4: mapred-site.xml

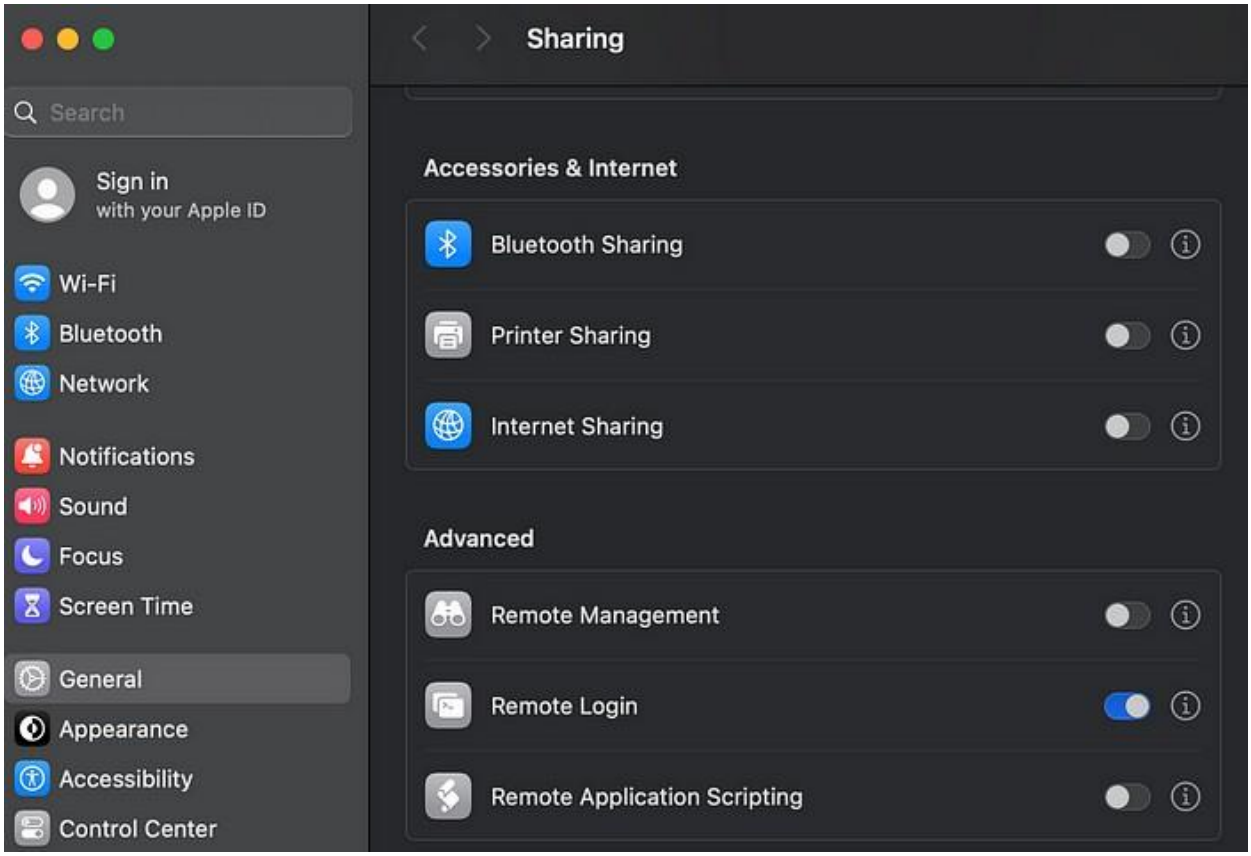
```
<configuration>
<property>
<name>yarn.app.mapreduce.am.env</name>
<value>HADOOP_MAPRED_HOME=/opt/homebrew/opt/hadoopp</value>
</property>
<property>
<name>mapreduce.map.env</name>
<value>HADOOP_MAPRED_HOME=/opt/homebrew/opt/hadoop</value>
</property>
<property>
```

```
<name>mapreduce.reduce.env</name>
<value>HADOOP_MAPRED_HOME=/opt/homebrew/opt/hadoop</value>
</property>
</configuration>
```

-> FILE 5: yarn-site.xml

```
<configuration>
<property>
<name>yarn.nodemanager.aux-services</name>
<value>mapreduce_shuffle</value>
</property>
</configuration>
```

Also, Now open **System Preferences** and open Sharing and select Remote Login



Step 5: Format HDFS and Start Hadoop

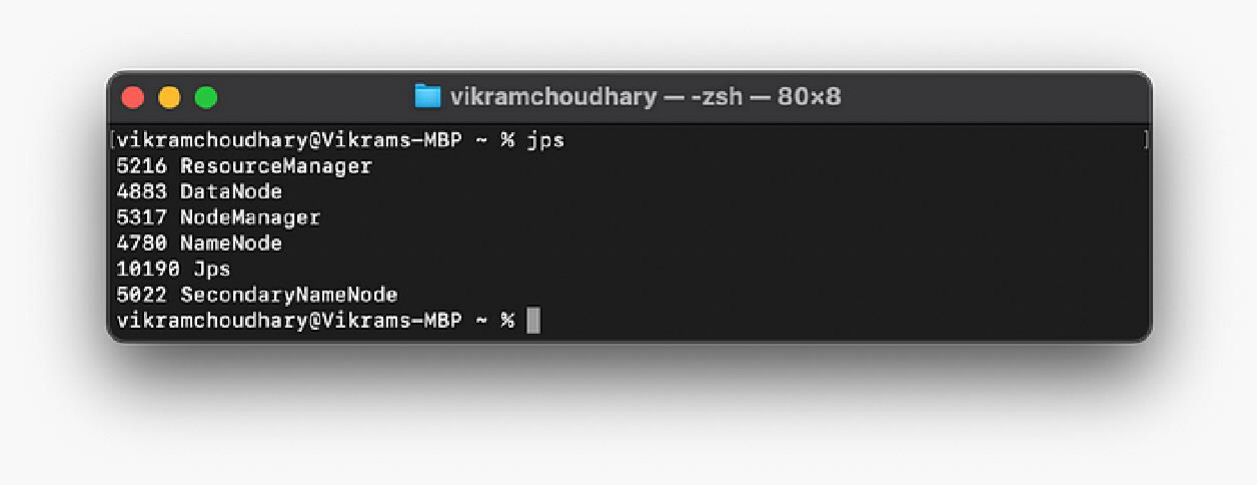
Run following commands to start hadoop:
hadoop namenode -format

Start hadoop:
start-all.sh

You can stop hadoop by using command : ***stop-all.sh***
If Error: Hadoop "Permission denied (publickey,password,keyboard-interactive)" warning

Proceed with the following steps:
1. Generate new keygen.
ssh-keygen -t rsa -P '' -f ~/.ssh/id_rsa
2. Register key gen:
cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys

You can start hadoop by using command : ***start-all.sh***
We can verify if all processes are running in machine using this command.
jps



http://localhost:9864/

http://localhost:8088/cluster

RESULT:

The step-by-step installation and configuration of Hadoop on Mac system have been successfully completed.