Ex.No – 1 Roll No – 210701149

Install, Configure and Run Hadoop and HDFS

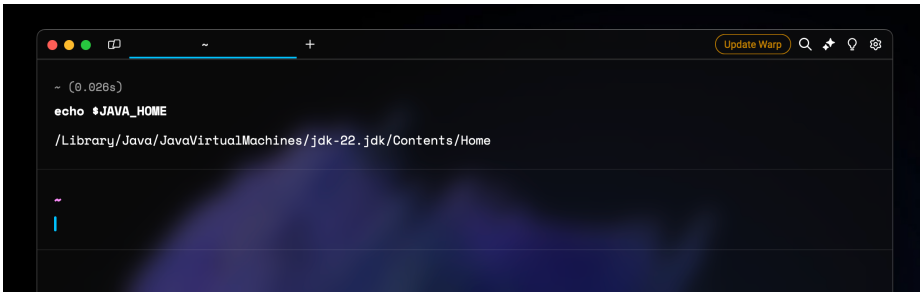
AIM:

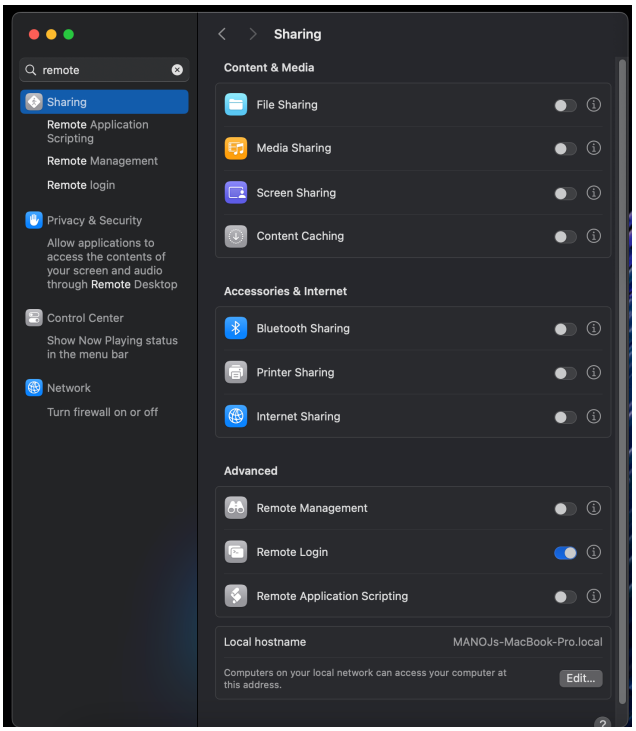
To install, configure and run Hadoop and HDFS on the system.

PROCEDURE:

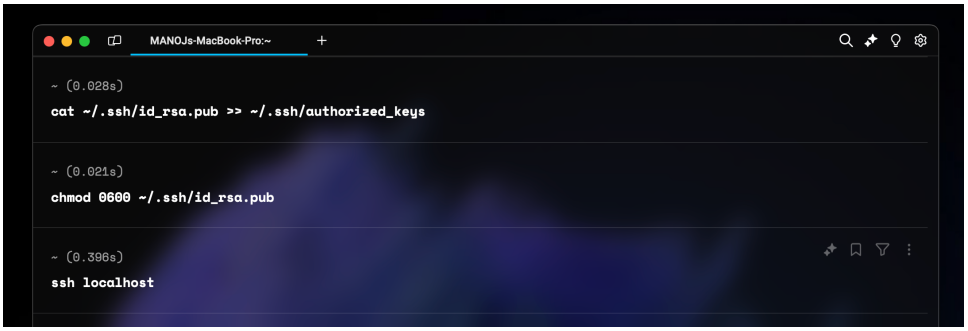
1. Install JDK 8 and set path in the environment variables.
2. Enable SSH for Local Host.
3. Download Hadoop and configure the Hadoop-env.sh file.
4. Set the path of Hadoop in the variables for global access.
5. Set the Java Path inside Hadoop.
6. Configure the core-site.xml, mapred-site.xml, hdfs-site.xml, yarn-site.xml
7. Start Hadoop using start-all.sh command.
8. Open the browser and go to the URL localhost:9870 and verify that Hadoop has been installed correctly.

OUTPUT:




Sharing
remote
Sharing
Remote Application Scripting
Remote Management
Remote login
Privacy & Security
Allow applications to access the contents of your screen and audio through Remote Desktop
Control Center
Show Now Playing status in the menu bar
Network
Turn firewall on or off
Content & Media
File Sharing
Media Sharing
Screen Sharing
Content Caching
Accessories & Internet
Bluetooth Sharing
Printer Sharing
Internet Sharing
Advanced
Remote Management
Remote Login
Remote Application Scripting
Local hostname
MANOJs-MacBook-Pro.local
Computers on your local network can access your computer at this address.
Edit...

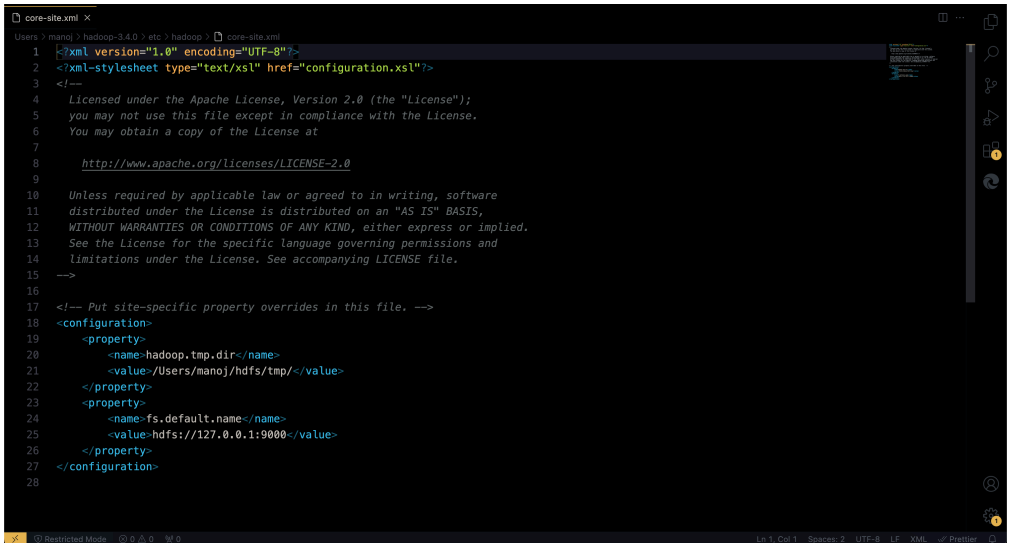

~ (3.81s)
ssh-keygen -t rsa -P '' -f ~/.ssh/id_rsa
Generating public/private rsa key pair.
/Users/manoj/.ssh/id_rsa already exists.
Overwrite (y/n)? y
Your identification has been saved in /Users/manoj/.ssh/id_rsa
Your public key has been saved in /Users/manoj/.ssh/id_rsa.pub
The key fingerprint is:
SHA256:4aX1tE27vidBGGwpMaeENvAD4cgEAVMA6ynSMU/61ug manoj@MANOJs-MacBook-Pro.local
The key's randomart image is:
+---[RSA 3072]----+
|*++o. +o .+o.. |
| o o o o+..+= |
|. o + ..+.+o.o. |
|...*   . * o.+.. |
|ooo .   S   o.o |
|o . o   .. |
| + .   .. |
| o   .. .|
| E   o+ |
+----[SHA256]-----+


MANOJs-MacBook-Pro:~
~ (0.028s)
cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
~ (0.021s)
chmod 0600 ~/.ssh/id_rsa.pub
~ (0.396s)
ssh localhost

Download

Hadoop is released as source code tarballs with corresponding binary tarballs for convenience. The downloads are distributed via mirror sites and should be checked for tampering using GPG or SHA-512.

Version	Release date	Source download	Binary download	Release notes
3.4.0	2024 Mar 17	source (checksum signature)	binary (checksum signature) binary-aarch64 (checksum signature)	Announcement
3.3.6	2023 Jun 23	source (checksum signature)	binary (checksum signature) binary-aarch64 (checksum signature)	Announcement
2.10.2	2022 May 31	source (checksum signature)	binary (checksum signature)	Announcement



hdfs-site.xml

Users > manoj > hadoop-3.4.0 > etc > hadoop > hdfs-site.xml

```xml
<configuration>
  <property>
      <name>dfs.data.dir</name>
      <value>/Users/manoj/hdfs/namenode</value>
  </property>
  <property>
      <name>dfs.data.dir</name>
      <value>/Users/manoj/hdfs/datanode</value>
  </property>
  <property>
      <name>dfs.replication</name>
      <value>1</value>
  </property>
</configuration>
```

mapred-site.xml

Users > manoj > hadoop-3.4.0 > etc > hadoop > mapred-site.xml

```xml
<!-- Put site-specific property overrides in this file. -->
<configuration>
  <property>
    <name>mapreduce.framework.name</name>
    <value>yarn</value>
  </property>
</configuration>
```

yarn-site.xml

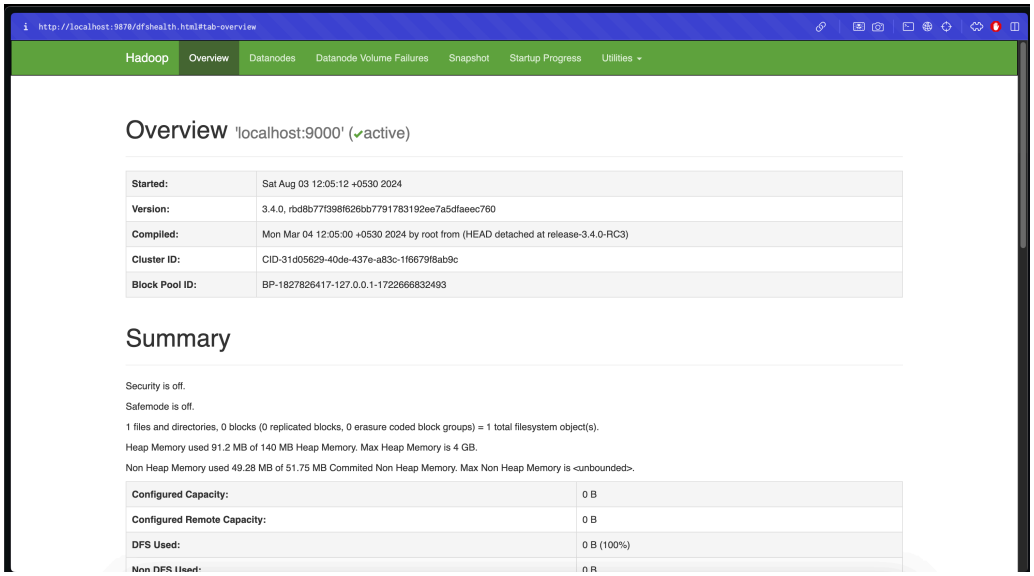
Users > manoj > hadoop-3.4.0 > etc > hadoop > yarn-site.xml

```
<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>
  <property>
    <name>yarn.resourcemanager.hostname</name>
    <value>127.0.0.1</value>
  </property>
  <property>
    <name>yarn.acl.enable</name>
    <value>0</value>
  </property>
  <property>
    <name>yarn.nodemanager.env-whitelist</name>
    <value>JAVA_HOME,HADOOP_COMMON_HOME,HADOOP_HDFS_HOME,HADOOP_CONF_DIR,CLASSPATH_PERPEND_DISTCACHE,HADOOP_YARN_HOME,HADOOP_MAPRE
  </property>
</configuration>
```

MANOJs-MacBook-Pro:~

~ (8.434s)

```
hdfs namenode -format
2024-08-03 12:03:45,604 INFO util.GSet: capacity      = 2^17 = 131072 entries
Re-format filesystem in Storage Directory root= /Users/manoj/hdfs/tmp/dfs/name; location= null ? (Y or N) y
2024-08-03 12:03:52,517 INFO namenode.FSImage: Allocated new BlockPoolId: BP-1827826417-127.0.0.1-17226668324
93
2024-08-03 12:03:52,518 INFO common.Storage: Will remove files: [/Users/manoj/hdfs/tmp/dfs/name/current/fsima
ge_0000000000000000000, /Users/manoj/hdfs/tmp/dfs/name/current/VERSION, /Users/manoj/hdfs/tmp/dfs/name/curren
t/fsimage_0000000000000000000.md5, /Users/manoj/hdfs/tmp/dfs/name/current/seen_txid, /Users/manoj/hdfs/tmp/df
s/name/current/edits_inprogress_0000000000000000001]
2024-08-03 12:03:52,545 INFO common.Storage: Storage directory /Users/manoj/hdfs/tmp/dfs/name has been succes
sfully formatted.
2024-08-03 12:03:52,564 INFO namenode.FSImageFormatProtobuf: Saving image file /Users/manoj/hdfs/tmp/dfs/name
/current/fsimage.ckpt_0000000000000000000 using no compression
2024-08-03 12:03:52,626 INFO namenode.FSImageFormatProtobuf: Image file /Users/manoj/hdfs/tmp/dfs/name/curren
t/fsimage.ckpt_0000000000000000000 of size 400 bytes saved in 0 seconds .
2024-08-03 12:03:52,635 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
2024-08-03 12:03:52,638 INFO blockmanagement.DatanodeManager: Slow peers collection thread shutdown
2024-08-03 12:03:52,653 INFO namenode.FSNamesystem: Stopping services started for active state
2024-08-03 12:03:52,653 INFO namenode.FSNamesystem: Stopping services started for standby state
2024-08-03 12:03:52,655 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid=0 when meet shutdown.
2024-08-03 12:03:52,656 INFO namenode.NameNode: SHUTDOWN_MSG:
/************************************************************
SHUTDOWN_MSG: Shutting down NameNode at MANOJs-MacBook-Pro.local/127.0.0.1
************************************************************/
```



RESULT:

Thus, to install, configure and to run Hadoop and HDFS has been done successfully.