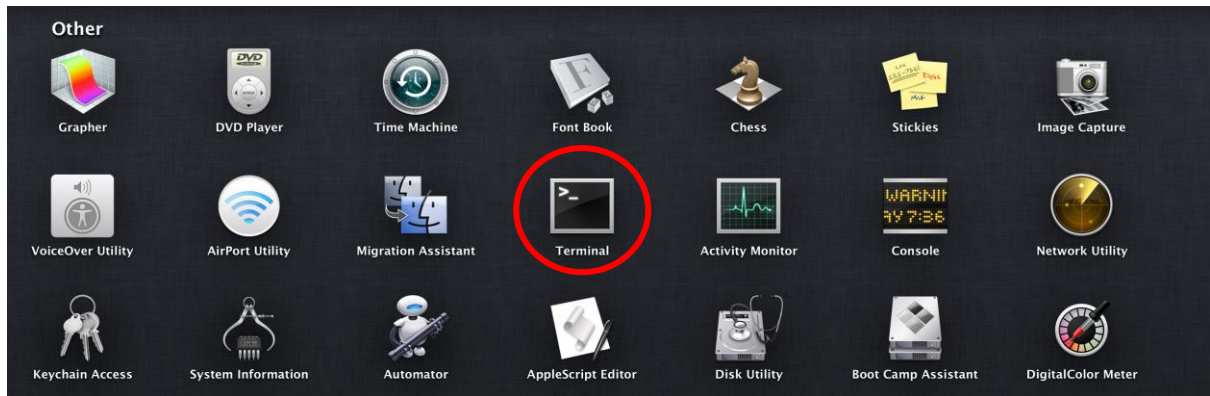


Hadoop Installation with Single DataNode on Mac



➤ **Open Terminal**



➤ **Check if Java is Installed :**

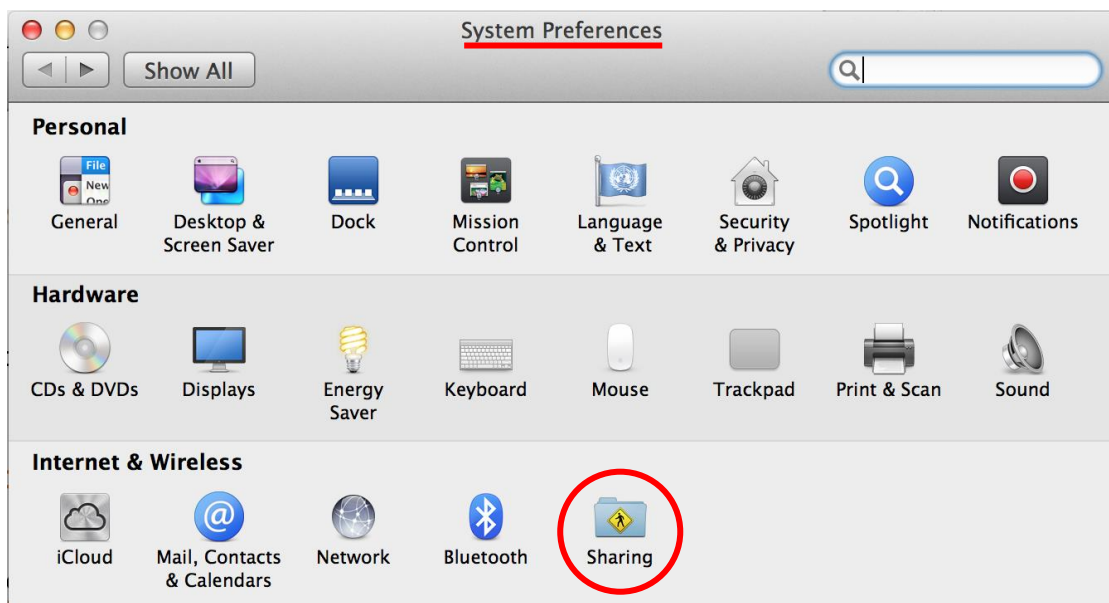
- **Command: java -version**

If you receive the below lines in the snap-shot then java is installed, else you will receive a pop-up requesting you **install Java** on your mac. Click on install and try the above command once again after the installation is done.

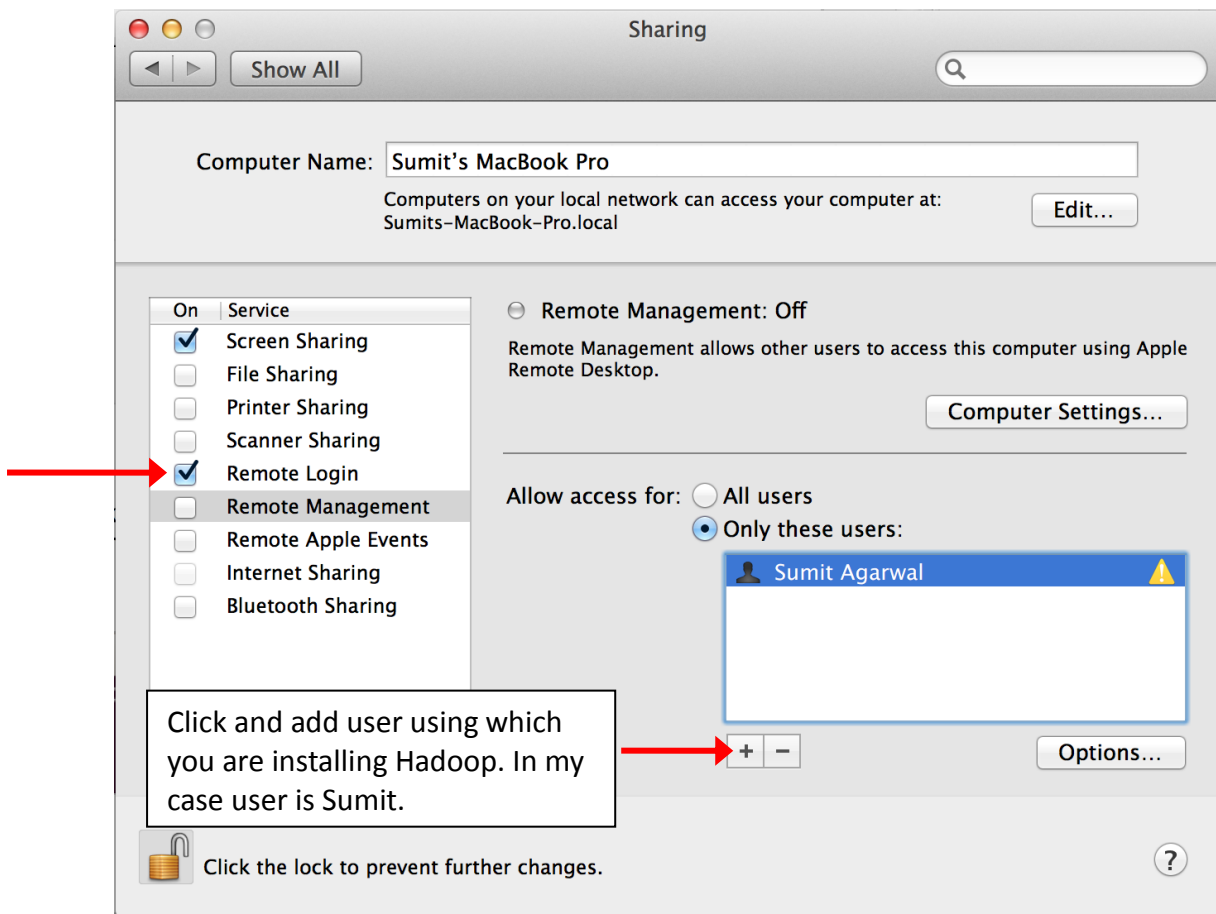
```
Sumits-MacBook-Pro:~ sumitagarwal$ java -version
java version "1.6.0_65"
Java(TM) SE Runtime Environment (build 1.6.0_65-b14-462-11M4609)
Java HotSpot(TM) 64-Bit Server VM (build 20.65-b04-462, mixed mode)
Sumits-MacBook-Pro:~ sumitagarwal$
```

➤ **Install SSH:** SSH comes installed on your Mac. However, you need to enable access to your own machine (so hadoop doesn't ask you for a password). To do this, go to

- **Go To: System Preferences > Sharing (under Internet & Network)**



Under the list of services, check “Remote Login”. For extra security, you can hit the radio button for “Only these Users”



➤ **Create a ssh key:**

- **Command:** `ssh-keygen -t rsa -P ""`

➤ **Moving the key to authorized key:**

- **Command:** `cat $HOME/.ssh/id_rsa.pub >> $HOME/.ssh/authorized_keys`

➤ **Now try to login to localhost using ssh:**

- **Command:** `ssh localhost`

➤ `exit`

➤ **Download Hadoop:**

- **Link:** <http://www.dsgnwrl.com/am/hadoop/common/hadoop-1.2.1/hadoop-1.2.1.tar.gz>

By default the file will be downloaded in Downloads folder

➤ **Extract Hadoop:**

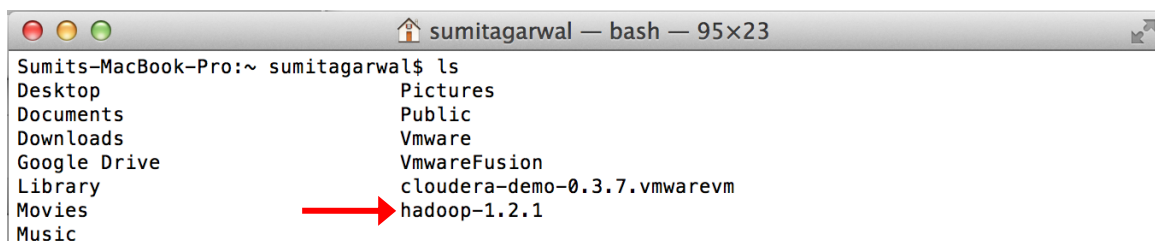
- **Command:** `tar -xzf /Users/Enter your username/Downloads/hadoop-1.2.1.tar.gz`



```
sumitagarwal — bash — 95x23
Sumits-MacBook-Pro:~ sumitagarwal$ tar -xzf /Users/sumitagarwal/Downloads/hadoop-1.2.1.tar.gz
```

➤ **Check whether hadoop has extracted or not:**

- **Command:** `ls`

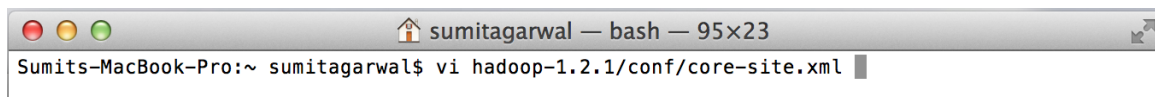


```
sumitagarwal — bash — 95x23
Sumits-MacBook-Pro:~ sumitagarwal$ ls
Desktop          Pictures
Documents        Public
Downloads        VMware
Google Drive     VMwareFusion
Library          cloudera-demo-0.3.7.vmwarevm
Movies           hadoop-1.2.1
Music
```

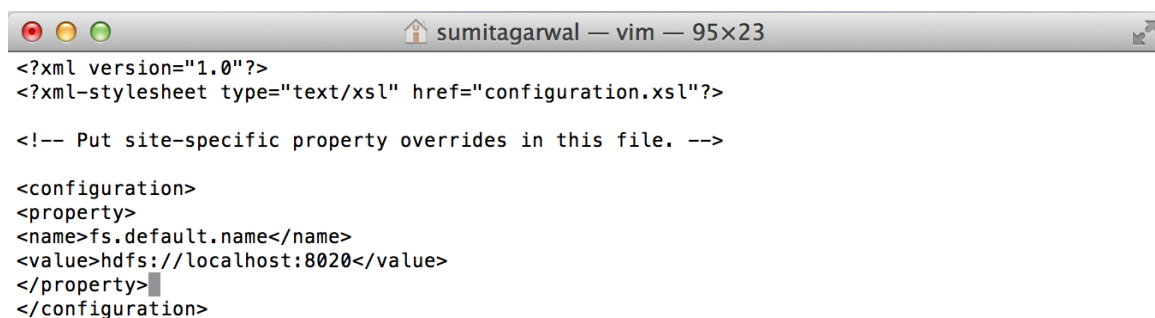
➤ **Edit core-site.xml:**

- **Command:** `vi hadoop-1.2.1/conf/core-site.xml`

```
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:8020</value>
</property>
```



```
sumitagarwal — bash — 95x23
Sumits-MacBook-Pro:~ sumitagarwal$ vi hadoop-1.2.1/conf/core-site.xml
```



```
sumitagarwal — vim — 95x23
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

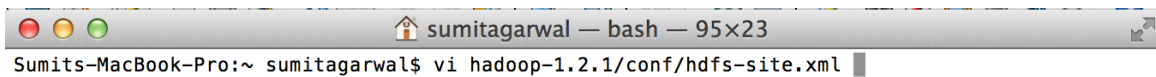
<!-- Put site-specific property overrides in this file. -->

<configuration>
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:8020</value>
</property>
</configuration>
```

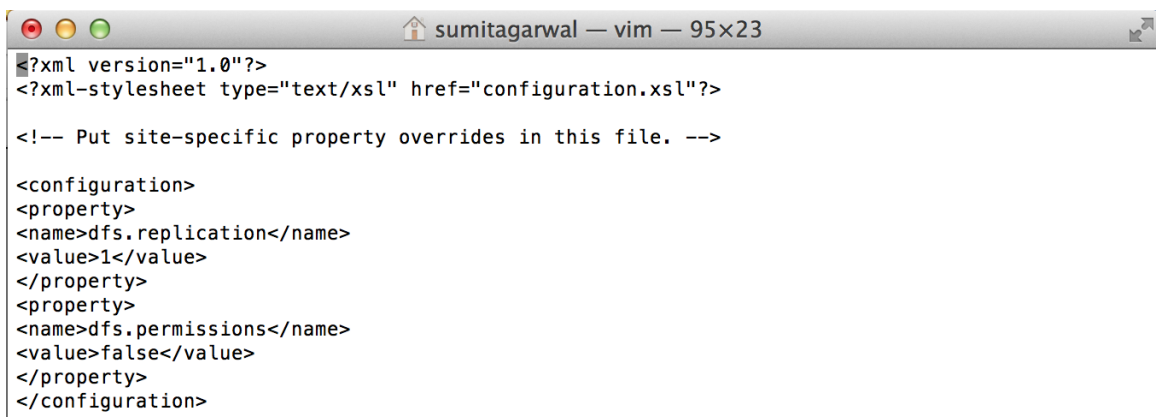
➤ **Edit hdfs-site.xml:**

- **Command: vi hadoop-1.2.1/conf/hdfs-site.xml**

```
<property>
<name>dfs.replication</name>
<value>1</value>
</property>
<property>
<name>dfs.permissions</name>
<value>>false</value>
</property>
```



A terminal window titled 'sumitagarwal — bash — 95x23'. The prompt is 'Sumits-MacBook-Pro:~ sumitagarwal\$'. The command entered is 'vi hadoop-1.2.1/conf/hdfs-site.xml'.



A vim editor window titled 'sumitagarwal — vim — 95x23'. The content of the file is as follows:

```
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<!-- Put site-specific property overrides in this file. -->

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

➤ **Edit mapred-site.xml:**

- **Command: vi hadoop-1.2.1/conf/mapred-site.xml**

```
<property>
<name>mapred.job.tracker</name>
<value>localhost:8021</value>
</property>
```

```
sumitagarwal — bash — 95x23
Sumits-MacBook-Pro:~ sumitagarwal$ vi hadoop-1.2.1/conf/mapred-site.xml
```

```
sumitagarwal — vim — 95x23
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

<!-- Put site-specific property overrides in this file. -->

<configuration>
<property>
<name>mapred.job.tracker</name>
<value>localhost:8021</value>
</property>
</configuration>
```

➤ **Edit hadoop-env.sh file:**

- **Command:** `vi hadoop-1.2.1/conf/hadoop-env.sh`

```
sumitagarwal — bash — 95x23
Sumits-MacBook-Pro:~ sumitagarwal$ vi hadoop-1.2.1/conf/hadoop-env.sh
```

Uncomment the below shown export and add the below the path to your JAVA_HOME & HADOOP_HEAPSIZE:

Type : `export JAVA_HOME=/Library/Java/Home`
`export HADOOP_HEAPSIZE=2000`

```
sumitagarwal — vim — 95x23
# Set Hadoop-specific environment variables here.

# The only required environment variable is JAVA_HOME. All others are
# optional. When running a distributed configuration it is best to
# set JAVA_HOME in this file, so that it is correctly defined on
# remote nodes.

# The java implementation to use. Required.
export JAVA_HOME=/Library/Java/Home

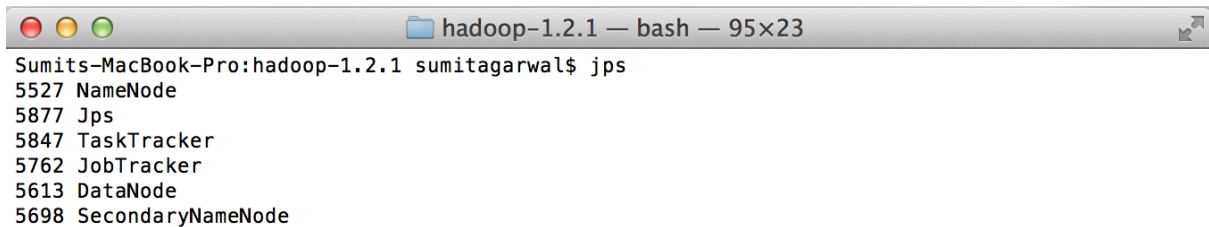
# Extra Java CLASSPATH elements. Optional.
# export HADOOP_CLASSPATH=

# The maximum amount of heap to use, in MB. Default is 1000.
export HADOOP_HEAPSIZE=2000

# Extra Java runtime options. Empty by default.
# export HADOOP_OPTS=-server

# Command specific options appended to HADOOP_OPTS when specified
export HADOOP_NAMENODE_OPTS="-Dcom.sun.management.jmxremote $HADOOP_NAMENODE_OPTS"
@
```

- **Change the directory where hadoop is installed.**
 - **Command:** `cd hadoop-1.2.1`
- **Format the name node**
 - **Command:** `bin/hadoop namenode -format`
- **Start the namenode, datanode, secondary namenode**
 - **Command:** `bin/start-dfs.sh`
- **Start the task tracker and job tracker**
 - **Command:** `bin/start-mapred.sh`
- **To check if Hadoop started correctly**
 - **Command:** `jps`



```
hadoop-1.2.1 — bash — 95x23
Sumits-MacBook-Pro:hadoop-1.2.1 sumitagarwal$ jps
5527 NameNode
5877 Jps
5847 TaskTracker
5762 JobTracker
5613 DataNode
5698 SecondaryNameNode
```