INPT
المعهد الوطني للبريد والمواصلات
ⵓⵔⴰⵣ ⵏⵏⵙⵔ ⵏ ⵜⵓⵍⵓⵙⵜ ⴰ ⵙⵎⵙⵓⵍⴻ
Institut National des Postes et Télécommunications

anrt
agence nationale de réglementation des télécommunications
الوكالة الوطنية لتقنين المواصلات
ⵜⴰⴳⵍⵙⵓⵙⵜ ⵜⵏⵏⵎⵙⵓⵜ ⵏ ⵙⴳⵣⵣⵏ ⵏ ⵜⵎⴰⵡⴰⵍⵉⵏ

Rapport de :

TP 1 : Installation et utilisation d'Apache Hbase

Réalisé par

Mouad Riali

Kamal Addi

Sous l'encadrement de :

Pr.Zaidouni Dounia

TP 1 : Installation et utilisation d'Apache Hbase

Installation et configuration d'Apache Hbase en mode « Standalone » et en mode « Pseudo-distribué » :

1- Préparation de l'environnement :

Installation de "Apache Hadoop" :

Etape 1 : Création d'un utilisateur hduser :

```
mouad-kamal@mouadkamal-VirtualBox:~$ sudo adduser hduser
Adding user `hduser' ...
Adding new group `hduser' (1002) ...
Adding new user `hduser' (1002) with group `hduser' ...
Creating home directory `/home/hduser' ...
Copying files from `/etc/skel' ...
Enter new UNIX password:
Retype new UNIX password:
Sorry, passwords do not match
passwd: Authentication token manipulation error
passwd: password unchanged
Try again? [y/N] Y
Enter new UNIX password:
Retype new UNIX password:
passwd: password updated successfully
Changing the user information for hduser
Enter the new value, or press ENTER for the default
    Full Name []:
    Room Number []:
    Work Phone []:
    Home Phone []:
    Other []:
Is the information correct? [Y/n]
```

```
mouad-kamal@mouadkamal-VirtualBox:~$ sudo adduser hduser sudo
Adding user `hduser' to group `sudo' ...
Adding user hduser to group sudo
Done.
```

D'abord on redémarre la machine virtuelle et on utilise le nouveau compte hduser.

Etape 2 : Mise en place de la clé ssh

On installe le paquet nécessaire pour ssh en tapant la commande :

```
hduser@mouadkamal-VirtualBox:~$ sudo apt-get install openssh-server
[sudo] password for hduser:
Reading package lists... Done
Building dependency tree
Reading state information... Done
openssh-server is already the newest version (1:7.6p1-4ubuntu0.3).
```

Maintenant, il faut mettre en place la clé ssh pour son propre compte. Pour cela,on exécute les commandes suivantes :

```
hduser@mouadkamal-VirtualBox:~$ ssh-keygen -t rsa -P ""
Generating public/private rsa key pair.
Enter file in which to save the key (/home/hduser/.ssh/id_rsa):
Created directory '/home/hduser/.ssh'.
Your identification has been saved in /home/hduser/.ssh/id_rsa.
Your public key has been saved in /home/hduser/.ssh/id_rsa.pub.
The key fingerprint is:
SHA256:69oKER0r+4CycvU+XmZjGZUN22+TxRGLVkXWBBf3m8I hduser@mouadkamal-VirtualBox
The key's randomart image is:
+---[RSA 2048]----+
|      .   .   .O%|
|     . o   *  o+*|
|    o o   + oo .+|
|   . +   .  o. oo|
|. . =   S    E=o |
| o . =   +   ... |
|o . . o O        |
|..   o.B .       |
|     .=+o        |
+----[SHA256]-----+
```

```
hduser@mouadkamal-VirtualBox:~$ cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
hduser@mouadkamal-VirtualBox:~$ chmod 0600 ~/.ssh/authorized_keys
```

On copie la clé public sur le serveur localhost :

```
hduser@mouadkamal-VirtualBox:~$ ssh-copy-id -i /home/hduser/.ssh/id_rsa.pub -f hduser@localhost
/usr/bin/ssh-copy-id: INFO: Source of key(s) to be installed: "/home/hduser/.ssh/id_rsa.pub"

Number of key(s) added: 1

Now try logging into the machine, with:   "ssh 'hduser@localhost'"
and check to make sure that only the key(s) you wanted were added.
```

On teste la connexion à localhost :

```
hduser@mouadkamal-VirtualBox:~$ ssh hduser@localhost
Welcome to Ubuntu 18.04.4 LTS (GNU/Linux 5.3.0-28-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/advantage


 * Canonical Livepatch is available for installation.
   - Reduce system reboots and improve kernel security. Activate at:
     https://ubuntu.com/livepatch

359 packages can be updated.
292 updates are security updates.

New release '20.04.2 LTS' available.
Run 'do-release-upgrade' to upgrade to it.

Your Hardware Enablement Stack (HWE) is supported until April 2023.
```

```
hduser@mouadkamal-VirtualBox:~$ exit
logout
Connection to localhost closed.
```

Etape 3 : Installation de JAVA 8

version 1.8.0

On installe ***JAVA 8*** dans le répertoire **/opt/java** :

```
hduser@mouadkamal-VirtualBox:~$ sudo mkdir /opt/java
[sudo] password for hduser:
hduser@mouadkamal-VirtualBox:~$ ls -R /opt
/opt:
java

/opt/java:
```

On extrait l'archive en utilisant la commande tar comme indiqué ci-dessous :

```
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ tar -zxvf jdk-8u71-linux-x64.tar.gz
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ sudo mv jdk1.8.0_71/ /opt/java/
```

On utilise la commande update-alternatives pour dire au système où java et ses exécutables sont installés.

```
mouad-kamal@mouadkamal-VirtualBox:/home/hduser$ sudo update-alternatives --install /usr/bin/java java /opt/java/jdk1.8.0_71/bin/java
update-alternatives: using /opt/java/jdk1.8.0_71/bin/java to provide /usr/bin/java (java) in auto mode
```

```
mouad-kamal@mouadkamal-VirtualBox:/home/hduser$ update-alternatives --config java
There is only one alternative in link group java (providing /usr/bin/java): /opt/java/jdk1.8.0_71/bin/java
Nothing to configure.
```

Maintenant,on met à jour aussi javac alternatives,et on ouvre le fichier **/etc/profile** en ecrivant :

```
export JAVA_HOME=/opt/java/jdk1.8.0_71/
export JRE_HOME=/opt/java/jdk1.8.0._71/jre
export PATH=$PATH:/opt/java/jdk1.8.0_71/bin:/opt/java/jdk1.8.0_71/jre/bin
```

Après avoir enregistré le fichier profile,on exécute la commande source pour recharger le fichier (en tant que root et avec l'utilisateur hadoop) :

```
~$ source /etc/profile
~$ source .bashrc
```

Etape 4 : Installation d'Apache Hadoop 2.7.4

```
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ tar -zxvf hadoop-2.7.4.tar.gz
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ mv hadoop-2.7.4 hadoop
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ sudo mv hadoop /usr/local/hadoop/
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ sudo chown -R hduser /usr/local/hadoop
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ sudo mkdir -p /usr/local/hadoop_store/hdfs/namenode
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ sudo mkdir -p /usr/local/hadoop_store/hdfs/datanode
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ sudo chown -R hduser /usr/local/hadoop_store
```

Etape 5 : Configuration d'Apache Hadoop 2.7.4

On modifie le fichier :**.bashrc** en ajoutant les lignes suivantes à la fin du fichier :

```
export JAVA_HOME=/opt/java/jdk1.8.0_71/
export JRE_HOME=/opt/java/jdk1.8.0._71/jre
export PATH=$PATH:/opt/java/jdk1.8.0_71/bin:/opt/java/jdk1.8.0_71/jre/bin
#HADOOP VARIABLES START
export JAVA_HOME=/opt/java/jdk1.8.0_71/
export HADOOP_INSTALL=/usr/local/hadoop
export PATH=$PATH:$HADOOP_INSTALL/bin
export PATH=$PATH:$HADOOP_INSTALL/sbin
export HADOOP_MAPRED_HOME=$HADOOP_INSTALL
export HADOOP_COMMON_HOME=$HADOOP_INSTALL
export HADOOP_HDFS_HOME=$HADOOP_INSTALL
export YARN_HOME=$HADOOP_INSTALL
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_INSTALL/lib/native
#export HADOOP_OPTS=" Djava.library.path=$HADOOP_INSTALL/lib"
#HADOOP VARIABLES END
```

Maintenant,on ouvre le fichier ***/usr/local/hadoop/etc/hadoop/hadoop-env.sh*** et on modifie la variable d'environnement JAVA_HOME :

```
# The java implementation to use. By default, this environment
# variable is REQUIRED on ALL platforms except OS X!
export JAVA_HOME=/opt/java/jdk1.8.0_71/
# Location of Hadoop.  By default, Hadoop will attempt to determine
# this location based upon its execution path.
# export HADOOP_HOME=
```

Et On crée le répertoire des fichiers temporaires de hadoop :

```
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ sudo mkdir -p /app/hadoop/tmp
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ sudo chown -R hduser /app/hadoop/tmp
```

> On modifie d'abord des fichiers pour la configuration de Hadoop Dans le rep ***/usr/local/hadoop/etc/hadoop/*** : On ouvre le fichier ***core-site.xml*** et on entre ce qui suit entre et </ configuration> :

```
<configuration>
<property>
<name>hadoop.tmp.dir</name>
<value>/app/hadoop/tmp</value>
</property>
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:54310</value>
</property>
</configuration>
```

le fichier ***hdfs-site.xml*** et on entre ce qui suit entre et </ configuration> :

```
<configuration>
<property>
<name>dfs.replication</name>
<value>1</value>
</property>
<property>
<name>dfs.namenode.name.dir</name>
<value>file:/usr/local/hadoop_store/hdfs/namenode</value>
</property>
<property>
<name>dfs.datanode.data.dir</name>
<value>file:/usr/local/hadoop_store/hdfs/datanode</value>
</property>
</configuration>
```

le fichier ***mapred-site.xml*** et on entre ce qui suit entre et </ configuration> :

```
<configuration>
<property>
<name>mapred.job.tracker</name>
<value>localhost:54311</value>
</property>
</configuration>
```

le fichier ***yarn-site.xml*** et on entre ce qui suit entre et </ configuration> :

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

Et on formate le Namenode :

```
hduser@mouadkamal-VirtualBox:/usr/local/hadoop/etc/hadoop$ hdfs namenode -format
WARNING: /usr/local/hadoop/logs does not exist. Creating.
2021-02-16 00:04:28,458 INFO namenode.NameNode: STARTUP_MSG:
/************************************************************
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = mouadkamal-VirtualBox/127.0.1.1
STARTUP_MSG:   args = [-format]
STARTUP_MSG:   version = 2.7.4
STARTUP_MSG:   classpath = /usr/local/hadoop/etc/hadoop:/usr/local/hadoop/share/hadoop/common/lib/jetty-server-9.3.24.v20180605.jar:
...etc
```

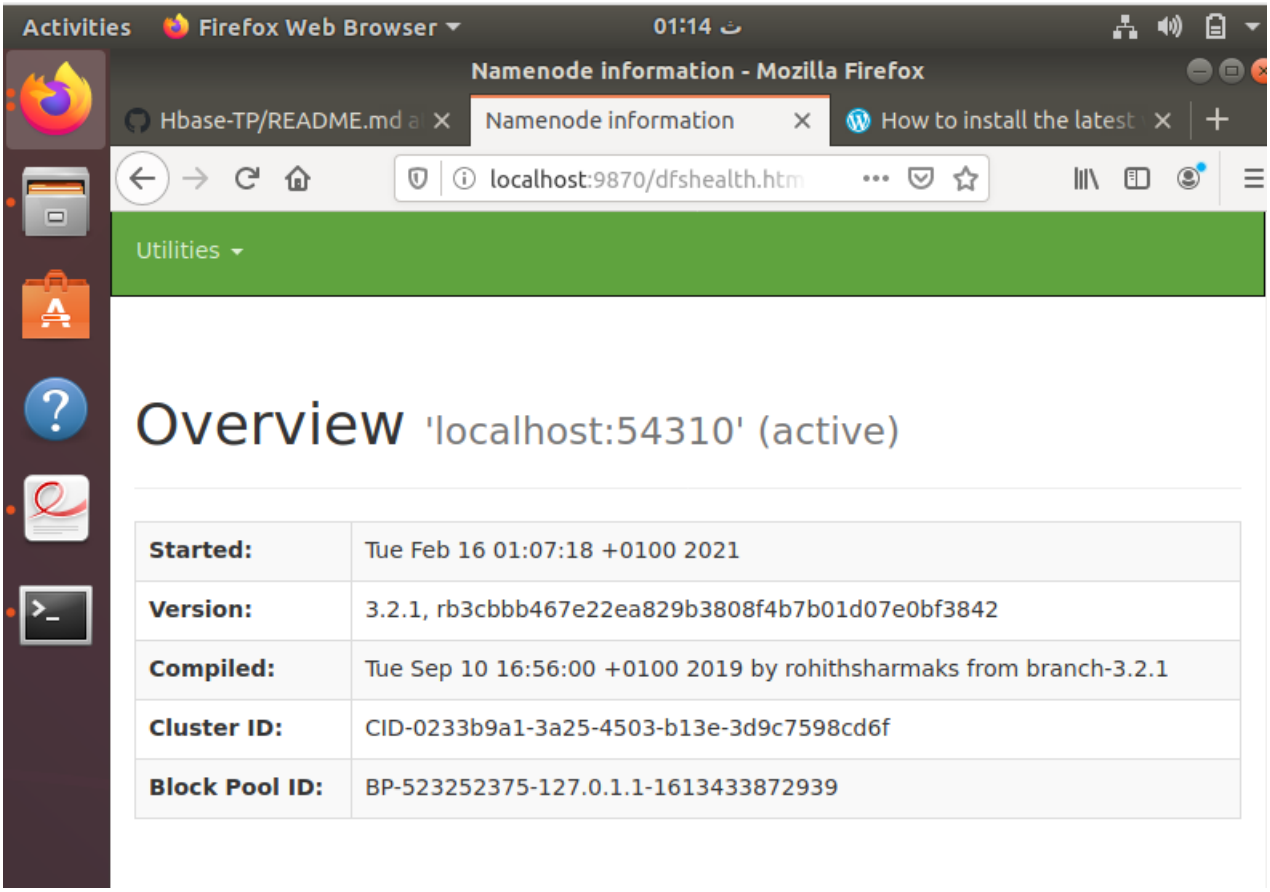
Maintenant, il est temps de démarrer le cluster à nœud unique nouvellement installé.

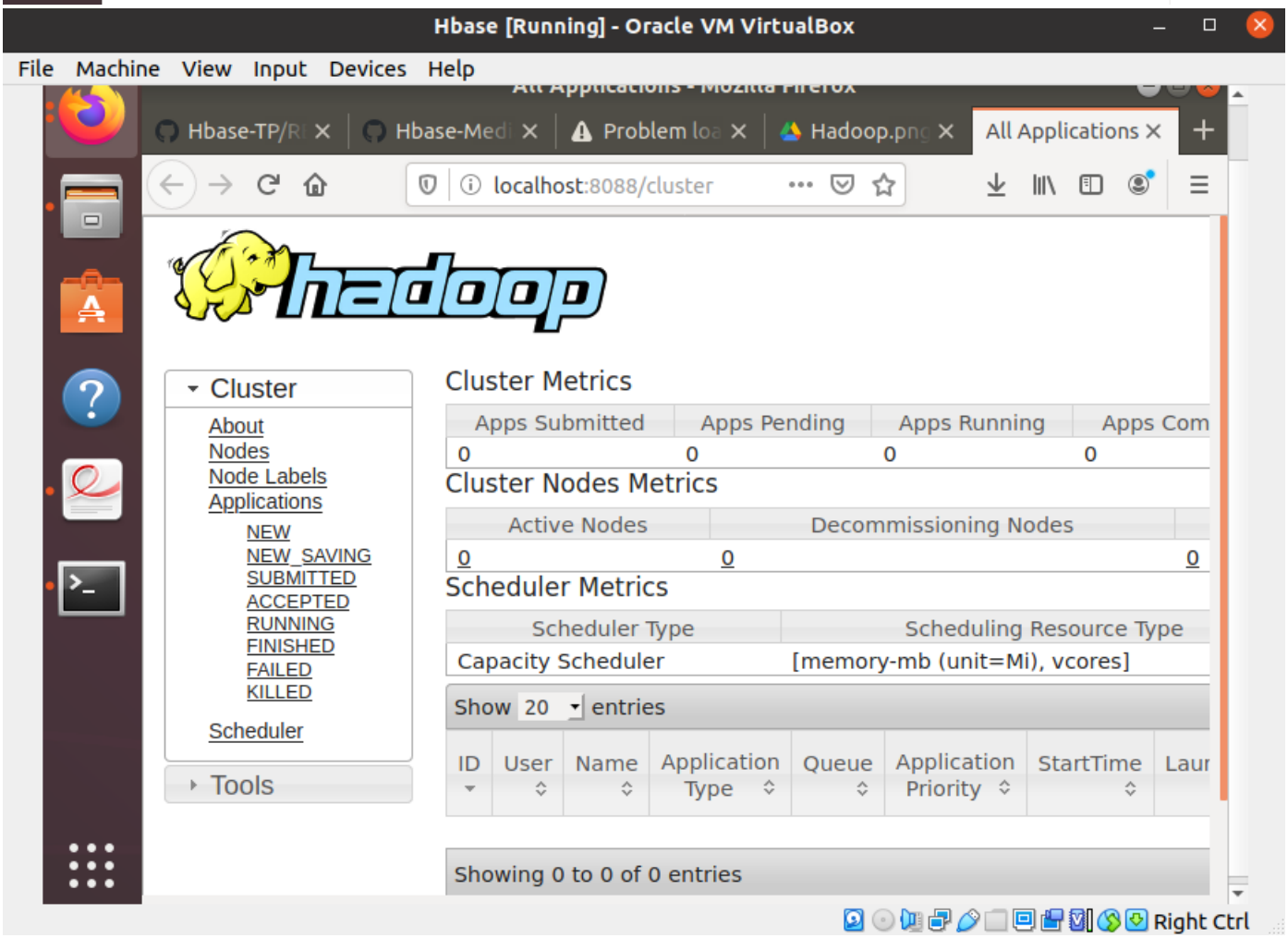
```
hduser@mouadkamal-VirtualBox:/usr/local/hadoop/etc/hadoop$ start-dfs.sh
Starting namenodes on [localhost]
localhost: starting namenode, logging to /usr/local/hadoop/logs/hadoop-hduser-namenode-mouadkamal-VirtualBox.out
localhost: starting datanode, logging to /usr/local/hadoop/logs/hadoop-hduser-datanode-mouadkamal-VirtualBox.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop/logs/hadoop-hduser-secondarynamenode-mouadkamal-VirtualBox.out
2021-02-16 00:07:35,016 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cla
```

```
hduser@mouadkamal-VirtualBox:/usr/local/hadoop/etc/hadoop$ start-yarn.sh
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-hduser-resourcemanager-mouadkamal-VirtualBox.out
localhost: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-hduser-nodemanager-mouadkamal-VirtualBox.out
```

```
hduser@mouadkamal-VirtualBox:/usr/local/hadoop/etc/hadoop$ jps
5345 DataNode
6005 NodeManager
5846 ResourceManager
5190 NameNode
5562 SecondaryNameNode
6381 Jps
```

> Voila !!





Installation et configuration de spark en local :

De même pour Apache Hadoop on va exécuter le code suivant :

```
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ tar -zxvf spark-2.4.3-bin-hadoop2.7.tgz
spark-2.4.3-bin-hadoop2.7/
spark-2.4.3-bin-hadoop2.7/python/
spark-2.4.3-bin-hadoop2.7/python/setup.cfg
spark-2.4.3-bin-hadoop2.7/python/pyspark/
spark-2.4.3-bin-hadoop2.7/python/pyspark/resultiterable.py
spark-2.4.3-bin-hadoop2.7/python/pyspark/python/
spark-2.4.3-bin-hadoop2.7/python/pyspark/python/pyspark/
...etc
```

```
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ mv spark-2.4.3-bin-hadoop2.7 spark
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ sudo mv spark /usr/local/
[sudo] password for hduser:
```

Ajouter les lignes suivantes au fichier ***.bashrc*** :

```
export SPARK_HOME=/usr/local/spark
export PATH=$PATH:$SPARK_HOME/bin
```

```
hduser@mouadkamal-VirtualBox:~$ source .bashrc
```

Installation Python :

```
hduser@mouadkamal-VirtualBox:~$ sudo apt-get install python
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following packages were automatically installed and are no longer required:
  fonts-liberation2 fonts-opensymbol gir1.2-gst-plugins-base-1.0
...etc
```

Apres, on aura le resultat suivant :

```
hduser@mouadkamal-VirtualBox:/usr/local/spark$ ./bin/pyspark
Python 2.7.17 (default, Sep 30 2020, 13:38:04)
[GCC 7.5.0] on linux2
Type "help", "copyright", "credits" or "license" for more information.
21/02/16 01:15:56 WARN Utils: Your hostname, mouadkamal-VirtualBox resolves to a loopback address: 127.0.1.1; using 10.0.2.15 instea
21/02/16 01:15:56 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
21/02/16 01:15:59 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
Welcome to
      ____              __
     / __/__  ___ _____/ /__
    _\ \/ _ \/ _ `/ __/  '_/
   /__ / .__/\_,_/_/ /_/\_\   version 2.4.3
      /_/

Using Python version 2.7.17 (default, Sep 30 2020 13:38:04)
SparkSession available as 'spark'.
>>>
```

et :

```
hduser@mouadkamal-VirtualBox:/usr/local/spark$ ./bin/spark-shell
21/02/16 01:18:09 WARN Utils: Your hostname, mouadkamal-VirtualBox resolves to a loopback address: 127.0.1.1; using 10.0.2.15 instea
21/02/16 01:18:09 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
21/02/16 01:18:16 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
Spark context Web UI available at http://10.0.2.15:4040
Spark context available as 'sc' (master = local[*], app id = local-1613438315750).
Spark session available as 'spark'.
Welcome to
      ____              __
     / __/__  ___ _____/ /__
    _\ \/ _ \/ _ `/ __/  '_/
   /___/ .__/\_,_/_/ /_/\_\   version 2.4.3
      /_/

Using Scala version 2.11.12 (Java HotSpot(TM) 64-Bit Server VM, Java 1.8.0_71)
Type in expressions to have them evaluated.
Type :help for more information.

scala>
```

Connexion de Spark à une distribution de Hadoop :

Pour utiliser ces packages de Hadoop, on doit modifier SPARK_DIST_CLASSPATH afin d'inclure les fichiers jar relatifs à ces packages. Pour ce faire, il est préférable d'ajouter une entrée dans ***conf/spark-env.sh*** :

```
hduser@mouadkamal-VirtualBox:/usr/local/spark$ cd conf
hduser@mouadkamal-VirtualBox:/usr/local/spark/conf$ cp spark-env.sh.template spark-env.sh
hduser@mouadkamal-VirtualBox:/usr/local/spark/conf$ sudo nano spark-env.sh
[sudo] password for hduser:
Sorry, try again.
[sudo] password for hduser:
```

On va insérer le contenu suivant dans spark-env.sh :

```
  GNU nano 2.9.3                  spark-env.sh                    Modified

# - SPARK_HISTORY_OPTS, to set config properties only for the history server ($
# - SPARK_SHUFFLE_OPTS, to set config properties only for the external shuffle$
# - SPARK_DAEMON_JAVA_OPTS, to set config properties for all daemons (e.g. "-D$
# - SPARK_DAEMON_CLASSPATH, to set the classpath for all daemons
# - SPARK_PUBLIC_DNS, to set the public dns name of the master or workers

# Generic options for the daemons used in the standalone deploy mode
# - SPARK_CONF_DIR      Alternate conf dir. (Default: ${SPARK_HOME}/conf)
# - SPARK_LOG_DIR       Where log files are stored.  (Default: ${SPARK_HOME}/l$
# - SPARK_PID_DIR       Where the pid file is stored. (Default: /tmp)
# - SPARK_IDENT_STRING  A string representing this instance of spark. (Default$
# - SPARK_NICENESS      The scheduling priority for daemons. (Default: 0)
# - SPARK_NO_DAEMONIZE  Run the proposed command in the foreground. It will no$
# Options for native BLAS, like Intel MKL, OpenBLAS, and so on.
# You might get better performance to enable these options if using native BLA$
# - MKL_NUM_THREADS=1        Disable multi-threading of Intel MKL
# - OPENBLAS_NUM_THREADS=1   Disable multi-threading of OpenBLAS
### in conf/spark-env.sh ###
# If 'hadoop' binary is on your PATH
export SPARK_DIST_CLASSPATH=/usr/local/hadoop
# With explicit path to 'hadoop' binary
export SPARK_DIST_CLASSPATH=/usr/local/hadoop/bin
# Passing a Hadoop configuration directory
export SPARK_DIST_CLASSPATH=/usr/local/hadoop/etc/hadoop


^G Get Help    ^O Write Out   ^W Where Is    ^K Cut Text    ^J Justify
^X Exit        ^R Read File   ^\ Replace     ^U Uncut Text  ^T To Linter
```

3-Installation et configuration d'Apache Hbase :

1- Récupérer les fichiers sources de Hbase :

> Voire version Badge!!

2- Décompresser le fichier récupéré dans le répertoire de votre choix :

on va repeter les memes etapes qu'on a deja fait avec **Apache Hadoop** et **Apache Spark** Alors on execute les commandes suivantes :

```
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ tar -zxvf hbase-1.4.7-bin.tar.gz
```

```
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ ls
hbase-1.4.7  hbase-1.4.7-bin.tar.gz  TP1_Hadoop.pdf  TP2.pdf
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ mv hbase-1.4.7 hbase
hduser@mouadkamal-VirtualBox:~/Desktop/BIG-DATA$ sudo mv hbase /usr/local/
[sudo] password for hduser:
```

3- Configurer le PATH dans le .bashrc :

on ajoute les lignes suivantes au fichier ***.bashrc*** afin de configurer le PATH vers ***Hbase*** repertoire :

```
export HBASE_HOME=/usr/local/hbase
export PATH=$PATH:$HBASE_HOME/bin
```

d'ou on aura :

```
  GNU nano 2.9.3                      .bashrc                       Modified


export JAVA_HOME=/opt/java/jdk1.8.0_71/
export JRE_HOME=/opt/java/jdk1.8.0._71/jre
export PATH=$PATH:/opt/java/jdk1.8.0_71/bin:/opt/java/jdk1.8.0_71/jre/bin
#HADOOP VARIABLES START
export JAVA_HOME=/opt/java/jdk1.8.0_71/
export HADOOP_INSTALL=/usr/local/hadoop
export PATH=$PATH:$HADOOP_INSTALL/bin
export PATH=$PATH:$HADOOP_INSTALL/sbin
export HADOOP_MAPRED_HOME=$HADOOP_INSTALL
export HADOOP_COMMON_HOME=$HADOOP_INSTALL
export HADOOP_HDFS_HOME=$HADOOP_INSTALL
export YARN_HOME=$HADOOP_INSTALL
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_INSTALL/lib/native
#export HADOOP_OPTS="-Djava.library.path=$HADOOP_INSTALL/lib"
#HADOOP VARIABLES END


export SPARK_HOME=/usr/local/spark
export PATH=$PATH:$SPARK_HOME/bin

export HBASE_HOME=/usr/local/hbase
export PATH=$PATH:$HBASE_HOME/bin




^G Get Help    ^O Write Out   ^W Where Is    ^K Cut Text    ^J Justify
^X Exit        ^R Read File   ^\ Replace     ^U Uncut Text  ^T To Spell
```

Ensuite, pour maintenir les modifications et mettre a jour le fichier.**bashrc** on tape:

```
hduser@mouadkamal-VirtualBox:~$ source .bashrc
```

4- Configuration du fichier « hbase-env.sh » :

D'abord, il faut rappeler que ***Apache Hbase*** necessite **JAVA 8** - heureusement, on l'a deja installe-, alors il faut ajouter le chemin vers le ***JDK 8*** au fichier ***« hbase-env.sh »*** , d'ou vient l'utlite des commandes suivantes :

```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ ls conf
hadoop-metrics2-hbase.properties  hbase-policy.xml  regionservers
hbase-env.cmd                     hbase-site.xml
hbase-env.sh                      log4j.properties
```

```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ sudo nano conf/hbase-env.sh
```

```
  GNU nano 2.9.3                 conf/hbase-env.sh                 Modified


# The directory where pid files are stored. /tmp by default.
# export HBASE_PID_DIR=/var/hadoop/pids

# Seconds to sleep between slave commands.  Unset by default.  This
# can be useful in large clusters, where, e.g., slave rsyncs can
# otherwise arrive faster than the master can service them.
# export HBASE_SLAVE_SLEEP=0.1

# Tell HBase whether it should manage it's own instance of Zookeeper or not.
# export HBASE_MANAGES_ZK=true

# The default log rolling policy is RFA, where the log file is rolled as per t$
# RFA appender. Please refer to the log4j.properties file to see more details $
# In case one needs to do log rolling on a date change, one should set the env$
# HBASE_ROOT_LOGGER to "<DESIRED_LOG LEVEL>,DRFA".
# For example:
# HBASE_ROOT_LOGGER=INFO,DRFA
# The reason for changing default to RFA is to avoid the boundary case of fill$
# DRFA doesn't put any cap on the log size. Please refer to HBase-5655 for mor$
export JAVA_HOME=/opt/java/jdk1.8.0_71/




^G Get Help    ^O Write Out   ^W Where Is    ^K Cut Text    ^J Justify
^X Exit        ^R Read File   ^\ Replace     ^U Uncut Text  ^T To Linter
```

5-Modification de /etc/hosts :

Puisque dans ce TP nous allons travaillé avec un seul serveur, il faut modifier le fichier /etc/hosts pour modifier l'adresse de notre serveur de 127.0.1.1 à l'adresse 127.0.0.1, comme suit :

```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ sudo nano /etc/hosts
```

```
  GNU nano 2.9.3                    /etc/hosts                     Modified

127.0.0.1       localhost
127.0.0.1       mouadkamal-VirtualBox

# The following lines are desirable for IPv6 capable hosts
::1     ip6-localhost ip6-loopback
fe00::0 ip6-localnet
ff00::0 ip6-mcastprefix
ff02::1 ip6-allnodes
ff02::2 ip6-allrouters
```

Pour prendre les modifications en compte, on redémarre notre machine.

3- Démarrage du cluster Hadoop configuré dans la machine "Hbase" :

on execute les commandes suivantes :

```
 hduser@mouadkamal-VirtualBox:/usr/local/hadoop$ cd ../hadoop_store
hduser@mouadkamal-VirtualBox:/usr/local/hadoop_store$ rm -rf *
hduser@mouadkamal-VirtualBox:/usr/local/hadoop_store$ mkdir -p /usr/local/hadoop_store/hdfs/namenode
hduser@mouadkamal-VirtualBox:/usr/local/hadoop_store$ mkdir -p /usr/local/hadoop_store/hdfs/datanode
hduser@mouadkamal-VirtualBox:/usr/local/hadoop_store$ chown -R hduser /usr/local/hadoop_store/hdfs/datanode
hduser@mouadkamal-VirtualBox:/usr/local/hadoop_store$ chown -R hduser /usr/local/hadoop_store/hdfs/namenode
```

```
hduser@mouadkamal-VirtualBox:/usr/local/hadoop_store$ cd ../hadoop
hduser@mouadkamal-VirtualBox:/usr/local/hadoop$ hdfs namenode -format
21/02/16 22:51:55 INFO namenode.NameNode: STARTUP_MSG:
/************************************************************
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = localhost/127.0.0.1
STARTUP_MSG:   args = [-format]
STARTUP_MSG:   version = 2.7.4
.
.
.
.
21/02/16 22:51:57 INFO namenode.FSImageFormatProtobuf: Image file /usr/local/hadoop_store/hdfs/namenode/current/fsimage.ckpt_0000000
21/02/16 22:51:57 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
21/02/16 22:51:57 INFO util.ExitUtil: Exiting with status 0
21/02/16 22:51:57 INFO namenode.NameNode: SHUTDOWN_MSG:
/************************************************************
SHUTDOWN_MSG: Shutting down NameNode at localhost/127.0.0.1
************************************************************/
```

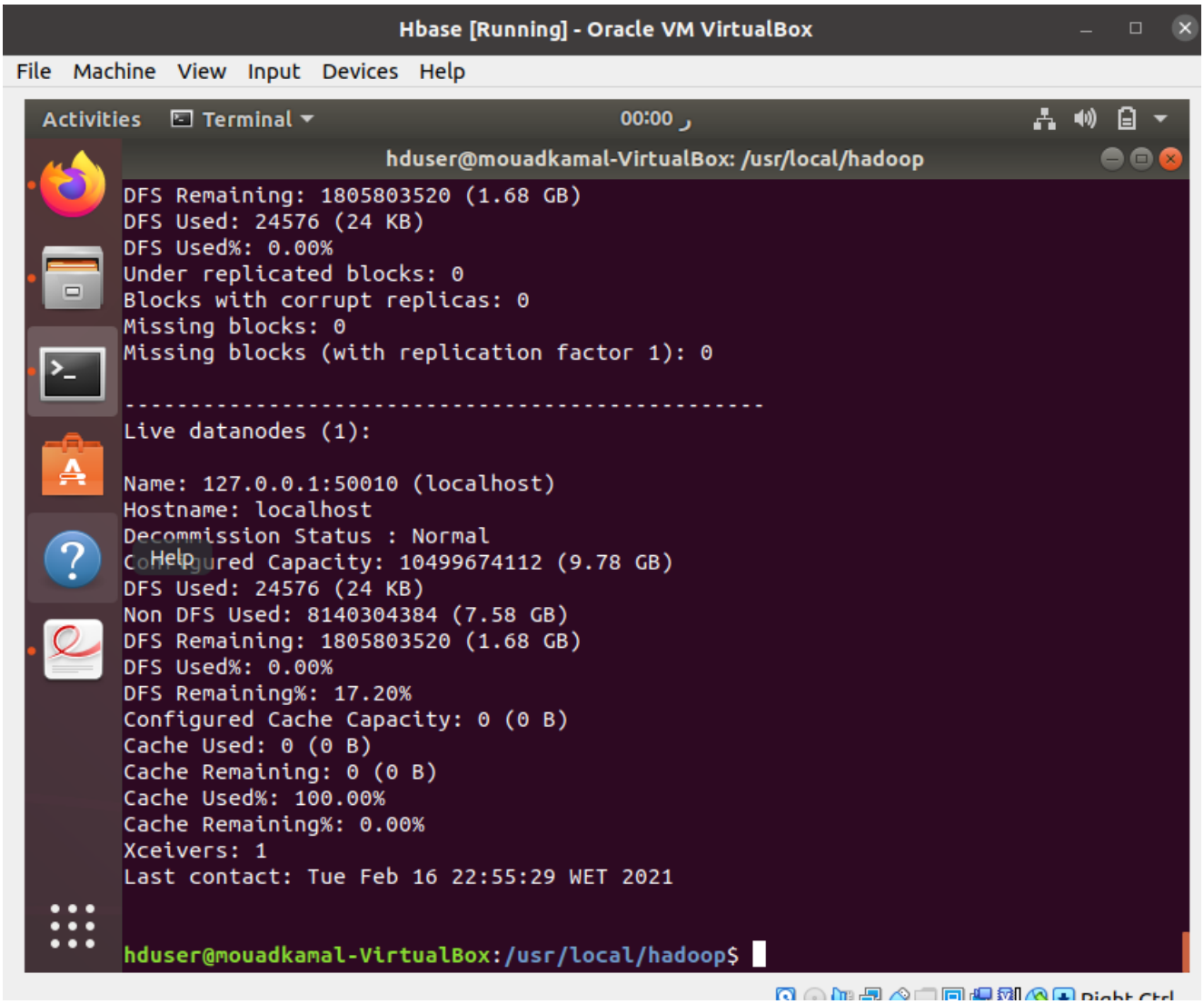
```
hduser@mouadkamal-VirtualBox:/usr/local/hadoop$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
21/02/16 22:54:36 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes w
Starting namenodes on [localhost]
localhost: starting namenode, logging to /usr/local/hadoop/logs/hadoop-hduser-namenode-mouadkamal-VirtualBox.out
localhost: starting datanode, logging to /usr/local/hadoop/logs/hadoop-hduser-datanode-mouadkamal-VirtualBox.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop/logs/hadoop-hduser-secondarynamenode-mouadkamal-VirtualBox.out
21/02/16 22:54:59 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes w
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-hduser-resourcemanager-mouadkamal-VirtualBox.out
localhost: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-hduser-nodemanager-mouadkamal-VirtualBox.out
```

et enfin, pour tester a quel point on a configure Hadoop, on execute :

```
hduser@mouadkamal-VirtualBox:/usr/local/hadoop$ hdfs dfsadmin -report
```

dans notre cas on a :

```
-------------------------------------------------
Live datanodes (1):

Name: 127.0.0.1:50010 (localhost)
Hostname: localhost
Decommission Status : Normal
Configured Capacity: 10499674112 (9.78 GB)
DFS Used: 24576 (24 KB)
Non DFS Used: 8140304384 (7.58 GB)
DFS Remaining: 1805803520 (1.68 GB)
DFS Used%: 0.00%
DFS Remaining%: 17.20%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 1
Last contact: Tue Feb 16 22:55:29 WET 2021
```



4- Configuration d'Apache Hbase pour un mode « Standalone » :

Apres s'assurer que Hadoop deamons marchent bien on poursuit les etapes suivants :

Configuration de « hbase-site.xml » pour un mode « Standalone » :

Pour installer HBase en mode Standalone, on ajoute les lignes suivantes dans le ***« hbase-site.xml »*** entre et :

```
 <property>
   <name>hbase.rootdir</name>
   <value>file:///home/mouad-kamal/hbase</value>
 </property>
 <property>
   <name>hbase.zookeeper.property.dataDir</name>
   <value>file:///home/mouad-kamal/zookeeper</value>
 </property>
 <property>
   <name>hbase.unsafe.stream.capability.enforce</name>
   <value>false</value>
 </property>
```

Alors le fichier ***hbase-site.xml*** devient :

```
  GNU nano 2.9.3                  hbase-site.xml                   Modified

 * "License"); you may not use this file except in compliance
 * with the License.  You may obtain a copy of the License at
 *
 *     http://www.apache.org/licenses/LICENSE-2.0
 *
 * Unless required by applicable law or agreed to in writing, software
 * distributed under the License is distributed on an "AS IS" BASIS,
 * WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
 * See the License for the specific language governing permissions and
 * limitations under the License.
 */
-->
<configuration>
 <property>
   <name>hbase.rootdir</name>
   <value>file:///home/mouad-kamal/hbase</value>
 </property>
 <property>
   <name>hbase.zookeeper.property.dataDir</name>
   <value>/home/mouad-kamal/zookeeper</value>
 </property>
 <property>
   <name>hbase.unsafe.stream.capability.enforce</name>
   <value>false</value>
 </property>
           [ line 17/37 (45%), col 2/69 (2%), char 644/1248 (51%) ]
^G Get Help    ^O Write Out   ^W Where Is    ^K Cut Text    ^J Justify
^X Exit        ^R Read File   ^\ Replace     ^U Uncut Text  ^T To Spell
```

Lancement de Hbase pour un mode « Standalone » :

on lance la commande suivante :

```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ ./bin/start-hbase.sh
running master, logging to /usr/local/hbase/logs/hbase-hduser-master-mouadkamal-VirtualBox.out
Java HotSpot(TM) 64-Bit Server VM warning: ignoring option PermSize=128m; support was removed in 8.0
Java HotSpot(TM) 64-Bit Server VM warning: ignoring option MaxPermSize=128m; support was removed in 8.0
```

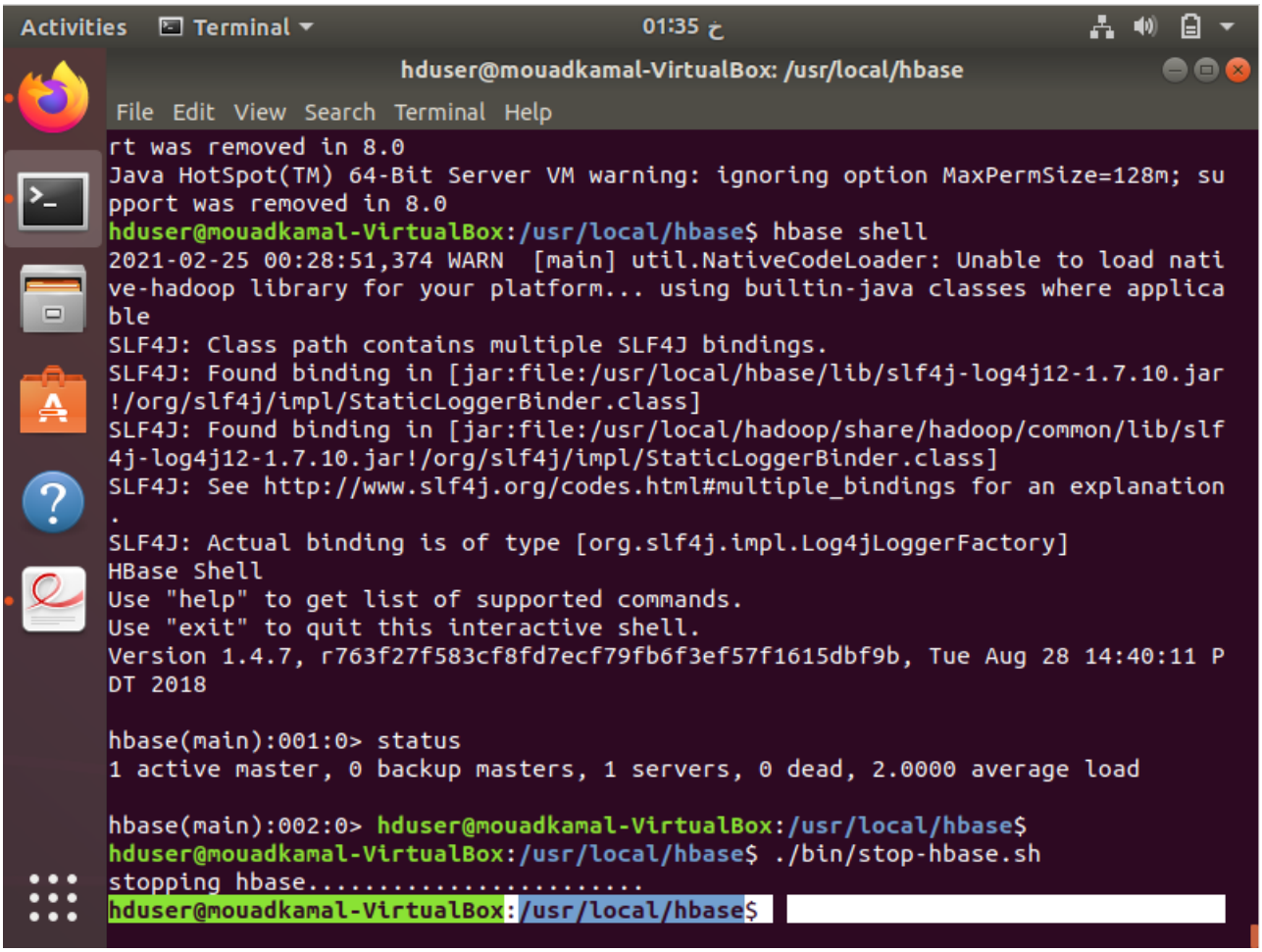
> Avec :

```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ jps
3521 ResourceManager
4785 HMaster
3682 NodeManager
2950 NameNode
5111 Jps
3115 DataNode
```

> Voila !!

```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ hbase shell
```

```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ hbase shell
2021-02-25 00:28:51,374 WARN  [main] util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBin
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
HBase Shell
Use "help" to get list of supported commands.
Use "exit" to quit this interactive shell.
Version 1.4.7, r763f27f583cf8fd7ecf79fb6f3ef57f1615dbf9b, Tue Aug 28 14:40:11 PDT 2018

hbase(main):001:0> status
1 active master, 0 backup masters, 1 servers, 0 dead, 2.0000 average load

hbase(main):002:0>
```



Then we stop Hbase Master :

```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ ./bin/stop-hbase.sh
stopping hbase.....................
hduser@mouadkamal-VirtualBox:/usr/local/hbase$
```

5- Configuration d'Apache Hbase en mode « Pseudo-distribué » :

1- Création des fichiers nécessaires pour Hbase dans le HDFS :

pour bien faire cette tache, il faut s'assurer que Hadoop est bien connecte :

```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ hadoop fs -mkdir -p /hbase
21/02/25 00:45:01 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes w
```

On vérifie avec la commande suivante que les répértoires sont bien créés :

```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ cd ../hadoop/etc/hadoop
hduser@mouadkamal-VirtualBox:/usr/local/hadoop/etc/hadoop$ hadoop fs -ls /
21/02/25 00:48:44 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes w
Found 2 items
drwxr-xr-x   - hduser supergroup          0 2021-02-25 00:45 /hbase
```

Pour installer HBase en mode pseudo-distribué, on ajoute les lignes suivantes dans le « hbase-site.xml » :

> On accede d'abord au rep /usr/local/hbase/conf :

```
hduser@mouadkamal-VirtualBox:/usr/local/hadoop/etc/hadoop$ cd ../../../hbase
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ cd conf
hduser@mouadkamal-VirtualBox:/usr/local/hbase/conf$ sudo nano hbase-site.xml
```

> Puis on ajoute les lignes suivantes au fichier ***hbase-site.xml***

```
<property>
<name>hbase.rootdir</name>
<value>hdfs://localhost:54310/hbase</value>
</property>
<property>
<name>hbase.zookeeper.quorum</name>
<value>localhost</value>
</property>
<property>
<name>hbase.cluster.distributed</name>
<value>true</value>
</property>
```

Et on aura le resultat suivant :

```
  GNU nano 2.9.3                  hbase-site.xml                   Modified

 *
 *     http://www.apache.org/licenses/LICENSE-2.0
 *
 * Unless required by applicable law or agreed to in writing, software
 * distributed under the License is distributed on an "AS IS" BASIS,
 * WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
 * See the License for the specific language governing permissions and
 * limitations under the License.
 */
-->
<configuration>
 <property>
  <name>hbase.rootdir</name>
  <value>hdfs://localhost:54310/hbase</value>
 </property>
 <property>
  <name>hbase.zookeeper.quorum</name>
  <value>localhost</value>
 </property>
 <property>
  <name>hbase.cluster.distributed</name>
  <value>true</value>
 </property>
</configuration>

^G Get Help    ^O Write Out   ^W Where Is    ^K Cut Text    ^J Justify
^X Exit        ^R Read File   ^\ Replace     ^U Uncut Text  ^T To Spell
```

> Attention, il faut utiliser dans hbase-site.xml la même adresse configurée dans le fichier de configuration de hadoop : core-site.xml. (Dans ce TP, nous avons utilisé localhost:54310 dans core-site.xml alors dans le fichier hbase-site.xml nous utilisons aussi la même adresse localhost:54310).
>
> On peut s'assurer de port convenable en tapant la commande suivante :

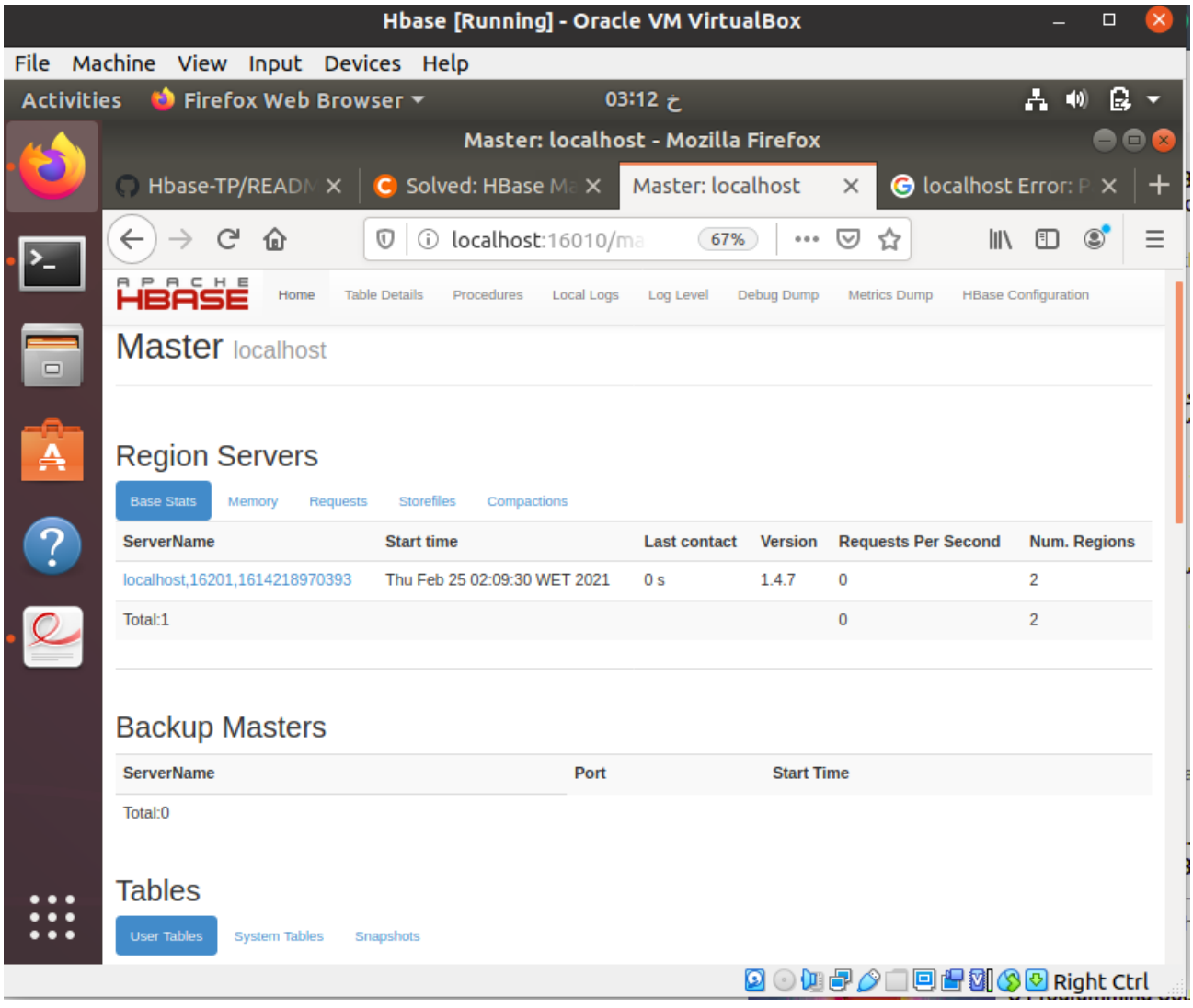
```
hduser@mouadkamal-VirtualBox:/usr/local/hbase/conf$ hdfs getconf -confKey fs.defaultFS
21/02/25 02:05:55 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes w
hdfs://localhost:54310
```

Enfin, on lance Hbase en executant la commande suivante :

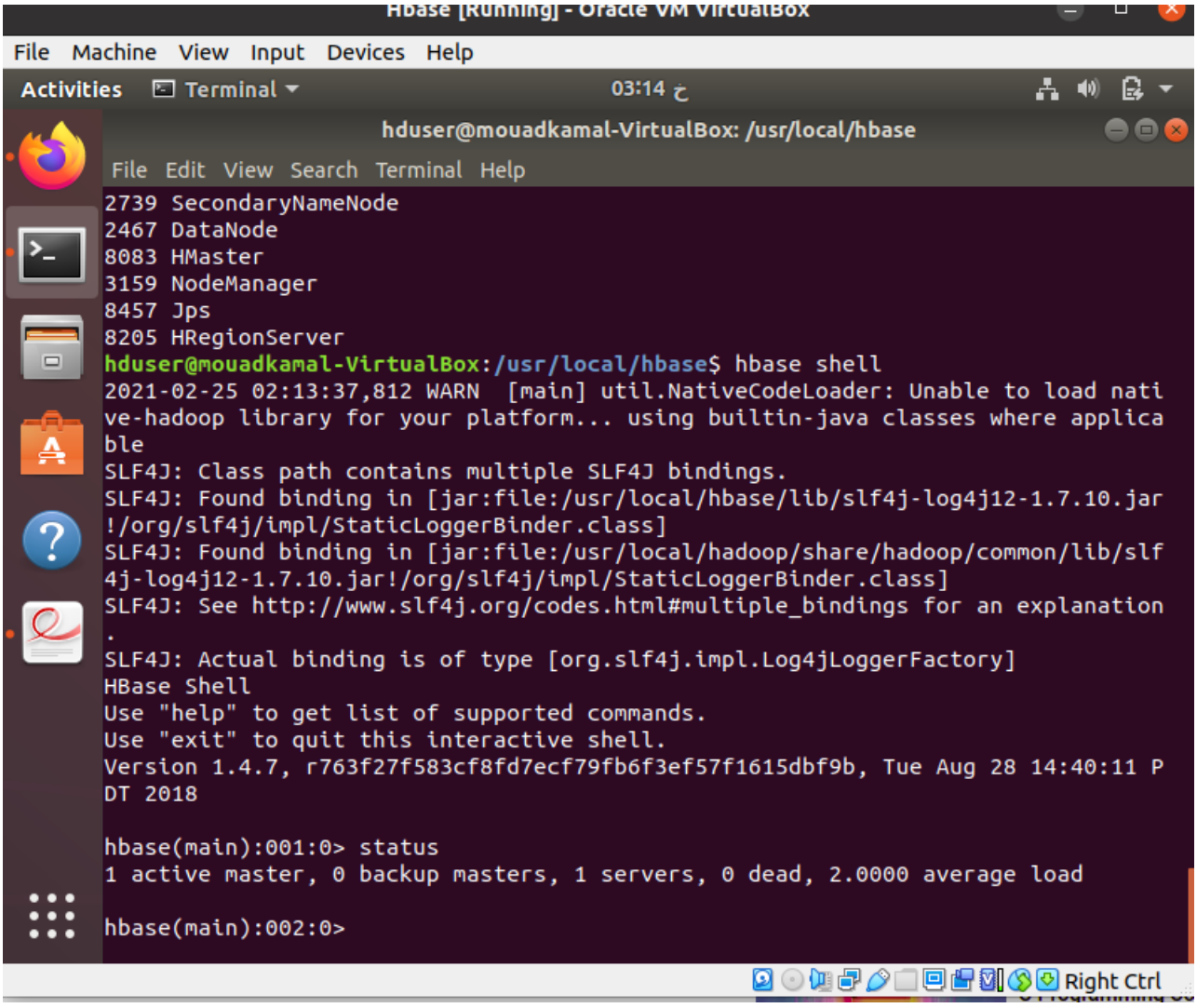
```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ ./bin/start-hbase.sh
localhost: running zookeeper, logging to /usr/local/hbase/bin/../logs/hbase-hduser-zookeeper-mouadkamal-VirtualBox.out
running master, logging to /usr/local/hbase/logs/hbase-hduser-master-mouadkamal-VirtualBox.out
: running regionserver, logging to /usr/local/hbase/logs/hbase-hduser-regionserver-mouadkamal-VirtualBox.out
```

> *Et Voila !!*

```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ jps
2992 ResourceManager
8017 HQuorumPeer
2307 NameNode
2739 SecondaryNameNode
2467 DataNode
8083 HMaster
3159 NodeManager
8457 Jps
8205 HRegionServer
hduser@mouadkamal-VirtualBox:/usr/local/hbase$
```



```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ hbase shell
2021-02-25 02:13:37,812 WARN  [main] util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBin
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
HBase Shell
Use "help" to get list of supported commands.
Use "exit" to quit this interactive shell.
Version 1.4.7, r763f27f583cf8fd7ecf79fb6f3ef57f1615dbf9b, Tue Aug 28 14:40:11 PDT 2018

hbase(main):001:0> status
1 active master, 0 backup masters, 1 servers, 0 dead, 2.0000 average load
```



2- Manipulation de « HBase » :

1-Création d'une BD :

Toujours en mode "pseudo-distribue" on lance le shell :

```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ hbase shell
```

1- On crée la table, ainsi que les familles de colonnes associées : create 'registre_ventes','client','ventes'

```
hbase(main):001:0> create 'registre_ventes','client','ventes'
0 row(s) in 10.1210 seconds

=> Hbase::Table - registre_ventes
hbase(main):002:0>
```

2- On vérifier que la table est bien créée:

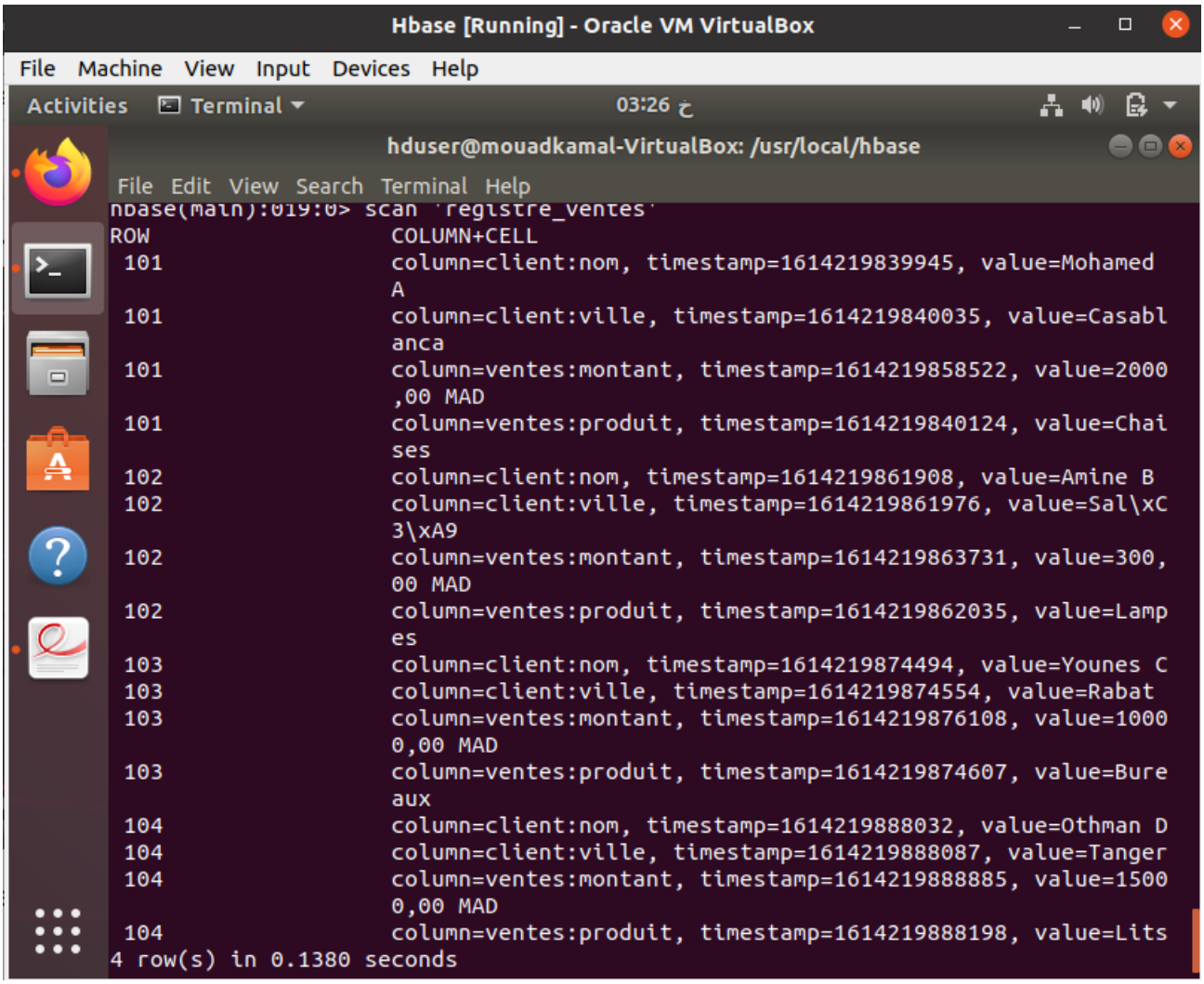
```
hbase(main):002:0> list
TABLE
registre_ventes
1 row(s) in 0.0560 seconds

=> ["registre_ventes"]
hbase(main):003:0>
```

3- On insére les différentes lignes de la table « registre_ventes » en tapant :

```
put 'registre_ventes', '101', 'client:nom', 'Mohamed A'
put 'registre_ventes', '101', 'client:ville', 'Casablanca'
put 'registre_ventes', '101', 'ventes:produit', 'Chaises'
put 'registre_ventes', '101', 'ventes:montant', '2000,00 MAD'
put 'registre_ventes', '102', 'client:nom', 'Amine B'
put 'registre_ventes', '102', 'client:ville', 'Salé'
put 'registre_ventes', '102', 'ventes:produit', 'Lampes'
put 'registre_ventes', '102', 'ventes:montant', '300,00 MAD'
put 'registre_ventes', '103', 'client:nom', 'Younes C'
put 'registre_ventes', '103', 'client:ville', 'Rabat'
put 'registre_ventes', '103', 'ventes:produit', 'Bureaux'
put 'registre_ventes', '103', 'ventes:montant', '10000,00 MAD'
put 'registre_ventes', '104', 'client:nom', 'Othman D'
put 'registre_ventes', '104', 'client:ville', 'Tanger'
put 'registre_ventes', '104', 'ventes:produit', 'Lits'
put 'registre_ventes', '104', 'ventes:montant', '15000,00 MAD'
```

3- On visualise le résultat de l'insertion :

```
hbase(main):019:0> scan 'registre_ventes'
ROW                   COLUMN+CELL
 101                  column=client:nom, timestamp=1614219839945, value=Mohamed
                      A
 101                  column=client:ville, timestamp=1614219840035, value=Casabl
                      anca
 101                  column=ventes:montant, timestamp=1614219858522, value=2000
                      ,00 MAD
 101                  column=ventes:produit, timestamp=1614219840124, value=Chai
                      ses
 102                  column=client:nom, timestamp=1614219861908, value=Amine B
 102                  column=client:ville, timestamp=1614219861976, value=Sal\xC
                      3\xA9
 102                  column=ventes:montant, timestamp=1614219863731, value=300,
                      00 MAD
 102                  column=ventes:produit, timestamp=1614219862035, value=Lamp
                      es
 103                  column=client:nom, timestamp=1614219874494, value=Younes C
 103                  column=client:ville, timestamp=1614219874554, value=Rabat
 103                  column=ventes:montant, timestamp=1614219876108, value=1000
                      0,00 MAD
 103                  column=ventes:produit, timestamp=1614219874607, value=Bure
                      aux
 104                  column=client:nom, timestamp=1614219888032, value=Othman D
 104                  column=client:ville, timestamp=1614219888087, value=Tanger
 104                  column=ventes:montant, timestamp=1614219888885, value=1500
                      0,00 MAD
 104                  column=ventes:produit, timestamp=1614219888198, value=Lits
4 row(s) in 0.1380 seconds
```



5- Affichons par exemple les valeurs de la colonne « produit » de la ligne 101 :

```
hbase(main):020:0> get 'registre_ventes','101',{COLUMN => 'ventes:produit'}
COLUMN               CELL
 ventes:produit      timestamp=1614219840124, value=Chaises
1 row(s) in 0.1380 seconds
```

6- Pour supprimer une table, il faut faire les étapes suivantes:

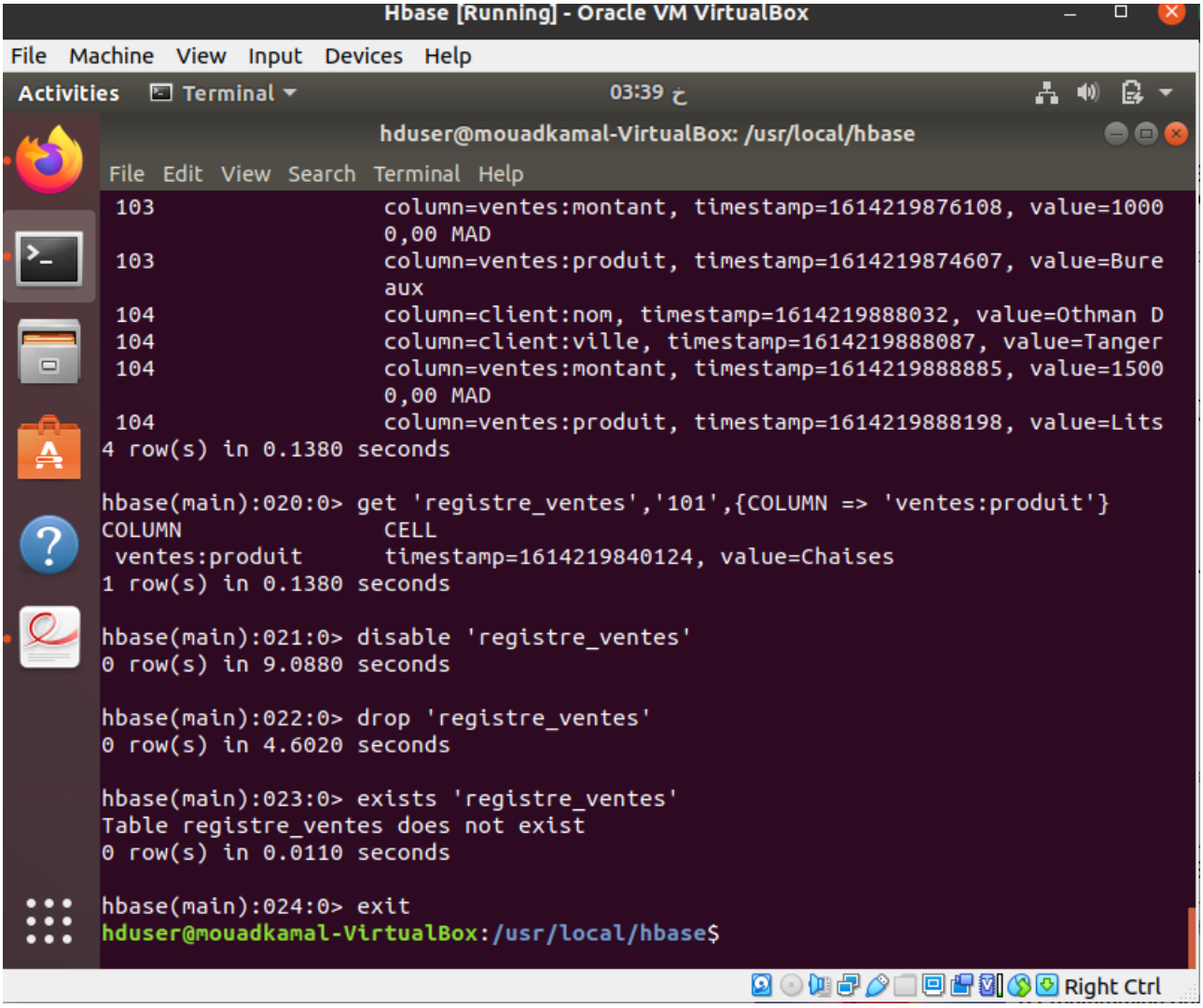
```
hbase(main):021:0> disable 'registre_ventes'
0 row(s) in 9.0880 seconds
```

puis :

```
hbase(main):022:0> drop 'registre_ventes'
0 row(s) in 4.6020 seconds
```

Et pour vérifier que la table n'existe plus, on tape :

```
hbase(main):023:0> exists 'registre_ventes'
Table registre_ventes does not exist
0 row(s) in 0.0110 seconds
```



3- L'utilisation de l'API Java de HBase :

Pour compiler sans problèmes notre code java, il faut inclure les librairies de Hbase le classpath par défaut, grâce à la variable d'environnement $CLASSPATH, pour cela, on l'ajoute dans le fichier **.bashrc** :

```
export CLASSPATH=/usr/local/hbase/lib/*:/usr/local/hadoop/share/hadoop/common/*
```

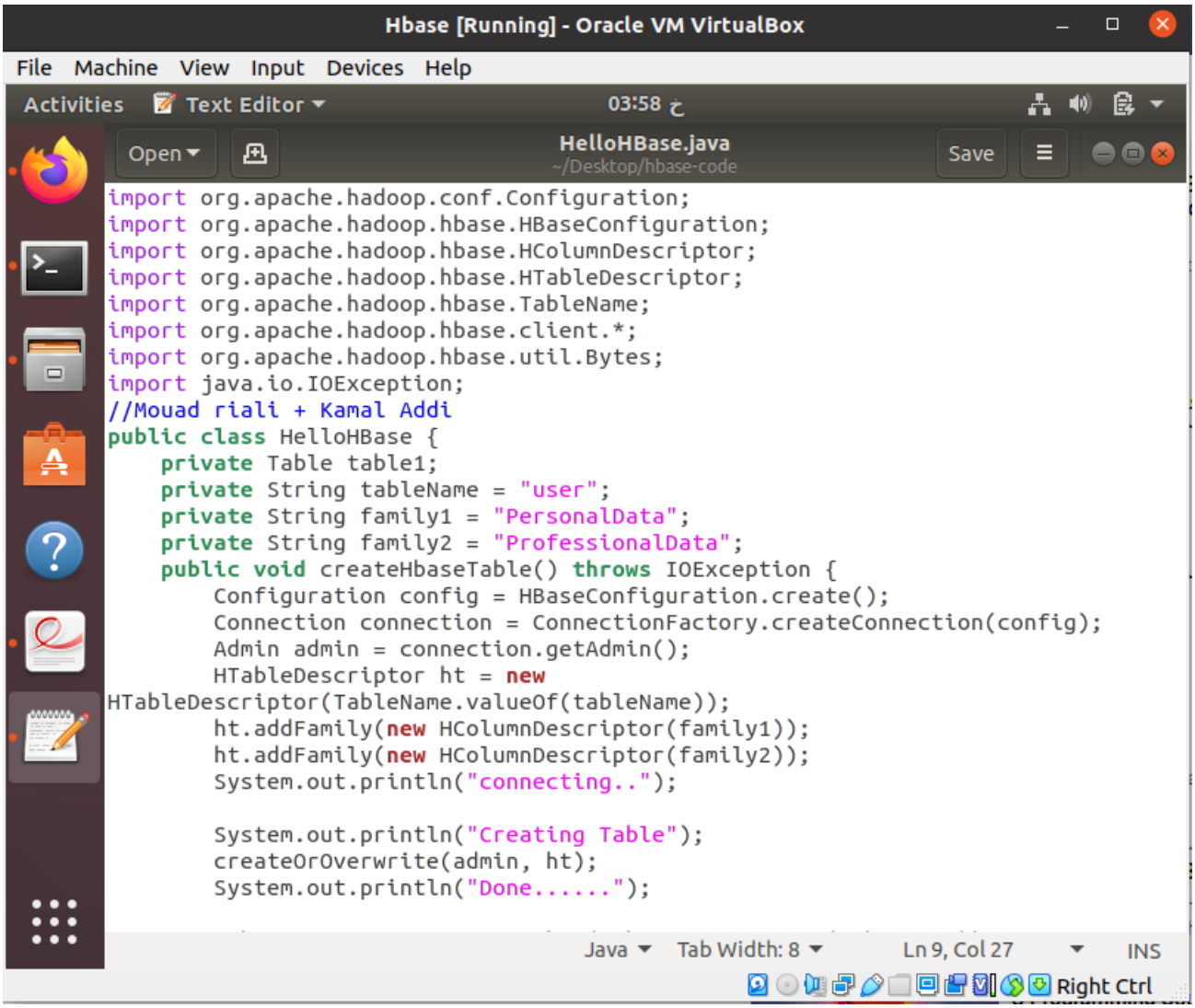
d'ou l'ensemble des lignes ajoutees est:

```
 export JAVA_HOME=/opt/java/jdk1.8.0_71/
export JRE_HOME=/opt/java/jdk1.8.0._71/jre
export PATH=$PATH:/opt/java/jdk1.8.0_71/bin:/opt/java/jdk1.8.0_71/jre/bin
#HADOOP VARIABLES START
export JAVA_HOME=/opt/java/jdk1.8.0_71/
export HADOOP_INSTALL=/usr/local/hadoop
export PATH=$PATH:$HADOOP_INSTALL/bin
export PATH=$PATH:$HADOOP_INSTALL/sbin
export HADOOP_MAPRED_HOME=$HADOOP_INSTALL
export HADOOP_COMMON_HOME=$HADOOP_INSTALL
export HADOOP_HDFS_HOME=$HADOOP_INSTALL
export YARN_HOME=$HADOOP_INSTALL
export HADOOP_COMMON_LIB_NATIVE_DIR=$HADOOP_INSTALL/lib/native
#export HADOOP_OPTS=" Djava.library.path=$HADOOP_INSTALL/lib"
#HADOOP VARIABLES END
export SPARK_HOME=/usr/local/spark
export PATH=$PATH:$SPARK_HOME/bin

export HBASE_HOME=/usr/local/hbase
export PATH=$PATH:$HBASE_HOME/bin

export CLASSPATH=/usr/local/hbase/lib/*:/usr/local/hadoop/share/hadoop/common/*




^G Get Help    ^O Write Out   ^W Where Is    ^K Cut Text    ^J Justify
^X Exit        ^R Read File   ^\ Replace     ^U Uncut Text  ^T To Spell
```

D'abords on cree un fichier java pour executer ce programme ci-dessous :

```
 import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.hbase.HBaseConfiguration;
import org.apache.hadoop.hbase.HColumnDescriptor;
import org.apache.hadoop.hbase.HTableDescriptor;
import org.apache.hadoop.hbase.TableName;
import org.apache.hadoop.hbase.client.*;
import org.apache.hadoop.hbase.util.Bytes;
import java.io.IOException;
//Mouad riali + Kamal Addi
public class HelloHBase {
    private Table table1;
    private String tableName = "user";
    private String family1 = "PersonalData";
    private String family2 = "ProfessionalData";
    public void createHbaseTable() throws IOException {
        Configuration config = HBaseConfiguration.create();
        Connection connection = ConnectionFactory.createConnection(config);
        Admin admin = connection.getAdmin();
        HTableDescriptor ht = new HTableDescriptor(TableName.valueOf(tableName));
        ht.addFamily(new HColumnDescriptor(family1));
        ht.addFamily(new HColumnDescriptor(family2));
        System.out.println("connecting..");

        System.out.println("Creating Table");
        createOrOverwrite(admin, ht);
        System.out.println("Done......");

        table1 = connection.getTable(TableName.valueOf(tableName));
        try {
            System.out.println("Adding user: user1");
            byte[] row1 = Bytes.toBytes("user1");
            Put p = new Put(row1);
            p.addColumn(family1.getBytes(), "name".getBytes(), Bytes.toBytes("mohamed"));
```

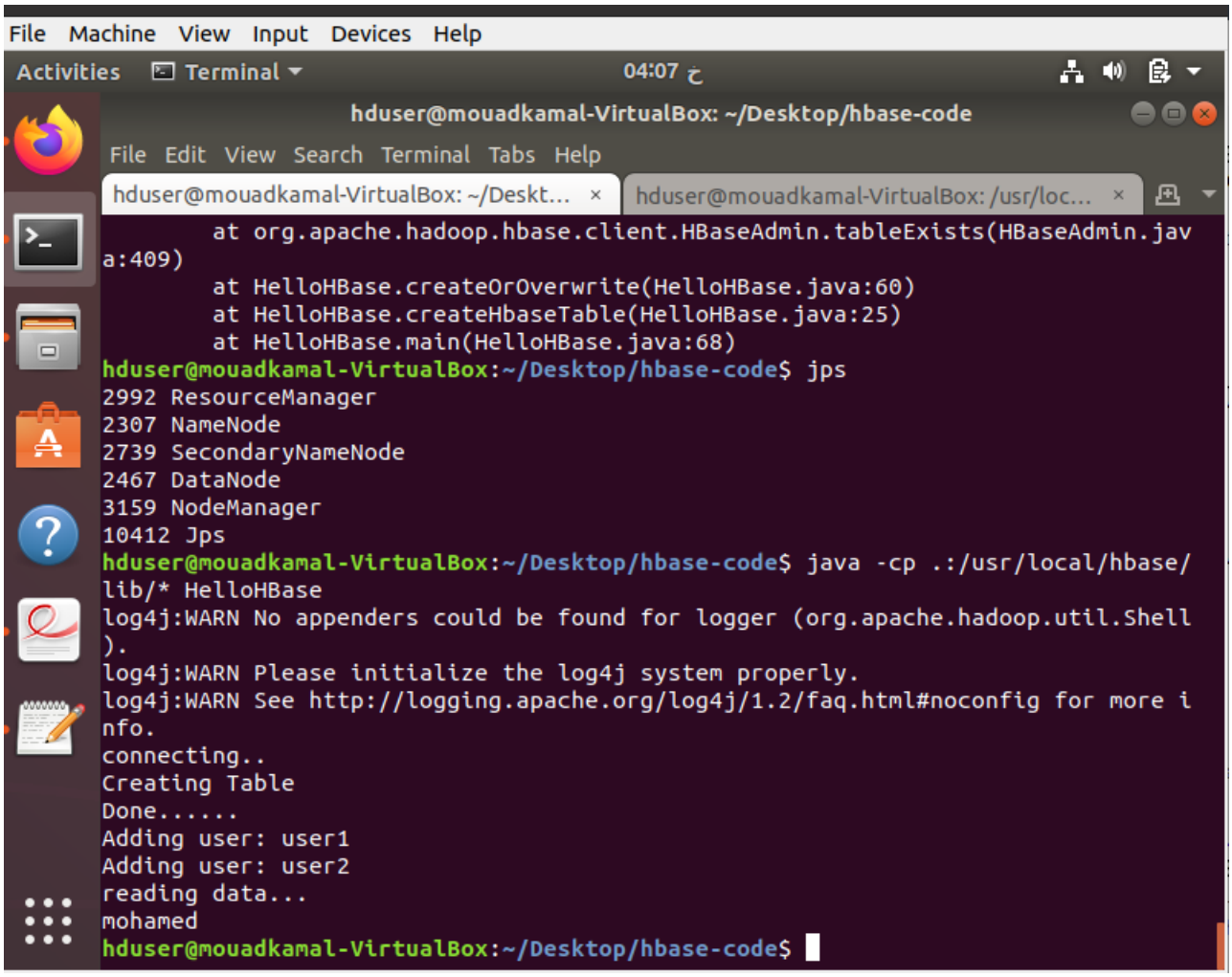
```
            p.addColumn(family1.getBytes(), "address".getBytes(), Bytes.toBytes("maroc"));
            p.addColumn(family2.getBytes(), "company".getBytes(), Bytes.toBytes("corp"));
            p.addColumn(family2.getBytes(), "salary".getBytes(), Bytes.toBytes("10000"));
            table1.put(p);
            System.out.println("Adding user: user2");
            byte[] row2 = Bytes.toBytes("user2");
            Put p2 = new Put(row2);
            p2.addColumn(family1.getBytes(), "name".getBytes(), Bytes.toBytes("younes"));
            p2.addColumn(family1.getBytes(), "tel".getBytes(), Bytes.toBytes("21212121"));
            p2.addColumn(family2.getBytes(), "profession".getBytes(), Bytes.toBytes("Engineer"));
            p2.addColumn(family2.getBytes(), "company".getBytes(), Bytes.toBytes("entrep"));
            table1.put(p2);
            System.out.println("reading data...");
            Get g = new Get(row1);

            Result r = table1.get(g);
            System.out.println(Bytes.toString(r.getValue(family1.getBytes(), "name".getBytes())));
        } catch (Exception e) {
            e.printStackTrace();
        } finally {
            table1.close();
            connection.close();
        }
    }
    public static void createOrOverwrite(Admin admin, HTableDescriptor table) throws
    IOException {
        if (admin.tableExists(table.getTableName())) {
            admin.disableTable(table.getTableName());
            admin.deleteTable(table.getTableName());
        }
        admin.createTable(table);
    }
    public static void main(String[] args) throws IOException {
        HelloHBase admin = new HelloHBase();
        admin.createHbaseTable();
    }
}
```



Et on compile cette classe ci dissus :

```
hduser@mouadkamal-VirtualBox:~/Desktop/hbase-code$ javac HelloHBase.java
```

```
hduser@mouadkamal-VirtualBox:~/Desktop/hbase-code$ java -cp .:/usr/local/hbase/lib/* HelloHBase
log4j:WARN No appenders could be found for logger (org.apache.hadoop.util.Shell).
log4j:WARN Please initialize the log4j system properly.
log4j:WARN See http://logging.apache.org/log4j/1.2/faq.html#noconfig for more info.
connecting..
Creating Table
Done.....
Adding user: user1
Adding user: user2
reading data...
mohamed
```



Chargement de fichiers :

Il est possible de charger des fichiers volumineux dans la base HBase, à partir de HDFS. Pour cela, on va télécharger le ficher sur le lien : https://www.dropbox.com/s/1aobaf5ibm5e7gm/purchases2.txt?dl=0 1- On commence par charger le fichier dans le répertoire input de HDFS (mais d'abord,on créer ce répertoire car) :

```
hduser@mouadkamal-VirtualBox:~/Desktop/hbase-code$ hdfs dfs -mkdir -p /input
21/02/25 03:15:18 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes w
```

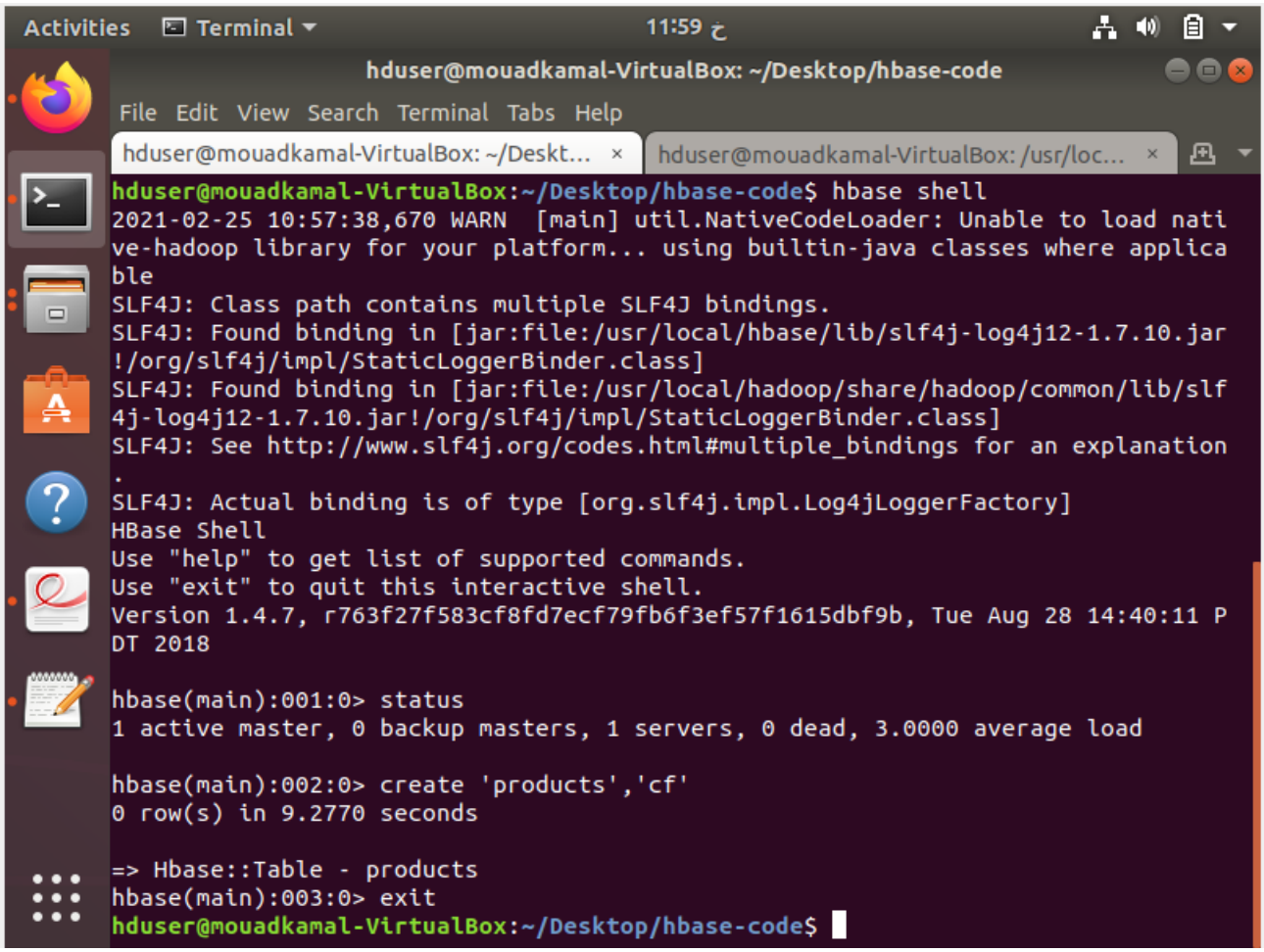
```
hduser@mouadkamal-VirtualBox:~/Desktop/hbase-code$ hdfs dfs -ls /
21/02/25 03:15:34 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes w
Found 2 items
drwxr-xr-x   - hduser supergroup          0 2021-02-25 03:05 /hbase
drwxr-xr-x   - hduser supergroup          0 2021-02-25 03:15 /input
```

```
hduser@mouadkamal-VirtualBox:~/Desktop/hbase-code$ hdfs dfs -put /home/hduser/Downloads/purchases2.txt /input
21/02/25 10:49:21 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes w
```

```
hduser@mouadkamal-VirtualBox:~/Desktop/hbase-code$ hdfs dfs -ls -R /input
21/02/25 10:49:51 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes w
-rw-r--r--   1 hduser supergroup  243309628 2021-02-25 10:49 /input/purchases2.txt
```

2- On Crée la base products avec une famille de colonnes 'cf'(Hbase shell) :

```
hbase(main):002:0> create 'products','cf'
0 row(s) in 9.2770 seconds

=> Hbase::Table - products
hbase(main):003:0> exit
```



3- On exécute la commande suivante. ImportTsv est une utilité qui permet de charger des données au format tsv dans HBase. Elle permet de déclencher une opération MapReduce sur le fichier principal stocké dans HDFS, pour lire les données puis les insérer via des put dans la base.

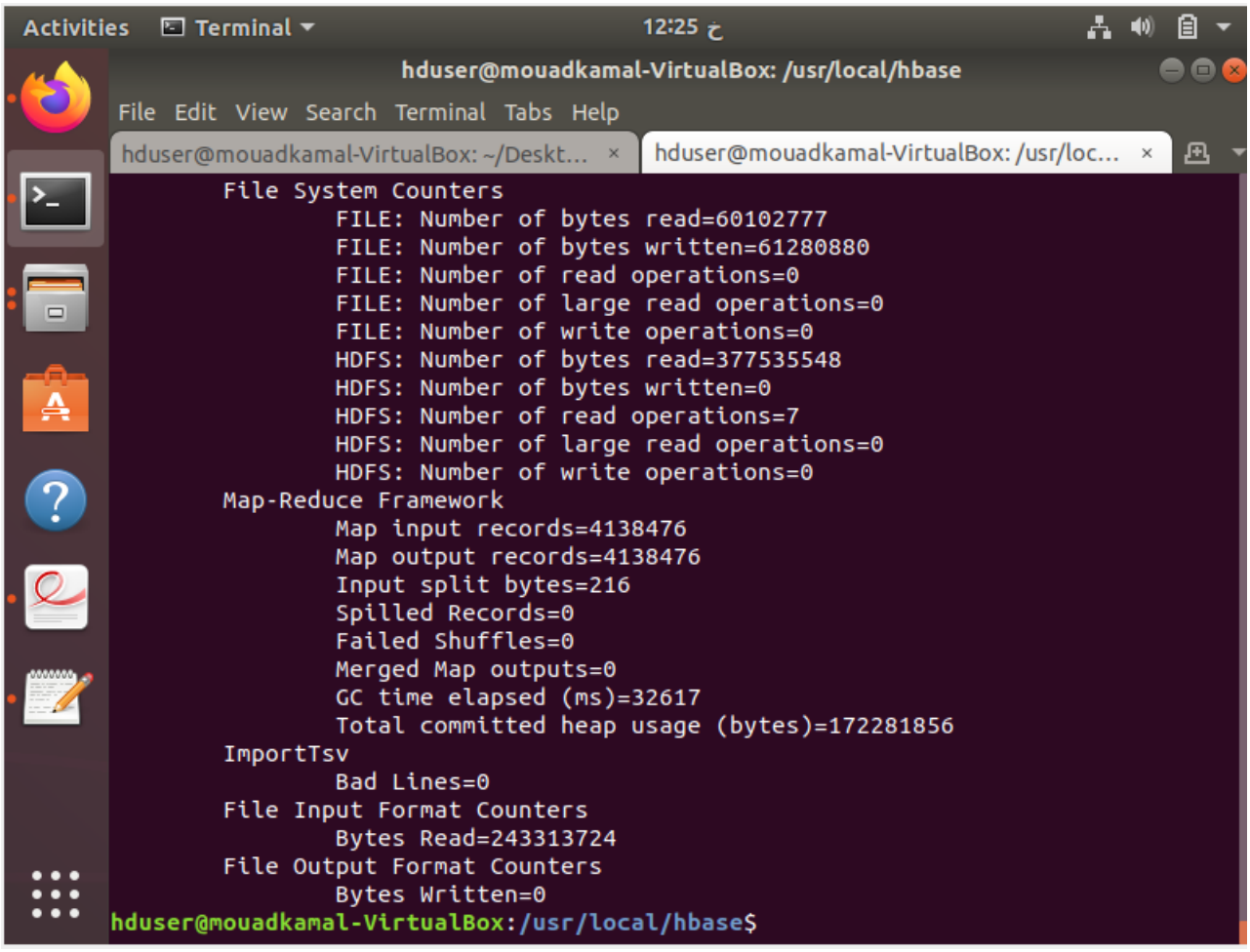
```
hduser@mouadkamal-VirtualBox:/usr/local/hbase$ hbase org.apache.hadoop.hbase.mapreduce.ImportTsv -Dimporttsv.separator=',' -Dimporti
```

output :

```
2021-02-25 11:14:41,118 WARN  [main] util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBin
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
2021-02-25 11:14:42,778 INFO  [main] zookeeper.RecoverableZooKeeper: Process identifier=hconnection-0x3fc2959f connecting to ZooKeep
2021-02-25 11:14:42,807 INFO  [main] zookeeper.ZooKeeper: Client environment:zookeeper.version=3.4.10-39d3a4f269333c922ed3db283be479
{...}
2021-02-25 11:14:54,838 INFO  [main] mapreduce.Job:  map 0% reduce 0%
2021-02-25 11:15:00,329 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:01,029 INFO  [main] mapreduce.Job:  map 1% reduce 0%
2021-02-25 11:15:03,330 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:04,154 INFO  [main] mapreduce.Job:  map 2% reduce 0%
2021-02-25 11:15:06,336 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:09,363 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:10,209 INFO  [main] mapreduce.Job:  map 3% reduce 0%
2021-02-25 11:15:12,375 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:13,226 INFO  [main] mapreduce.Job:  map 5% reduce 0%
2021-02-25 11:15:15,394 INFO  [communication thread] mapred.LocalJobRunner: map > map
```

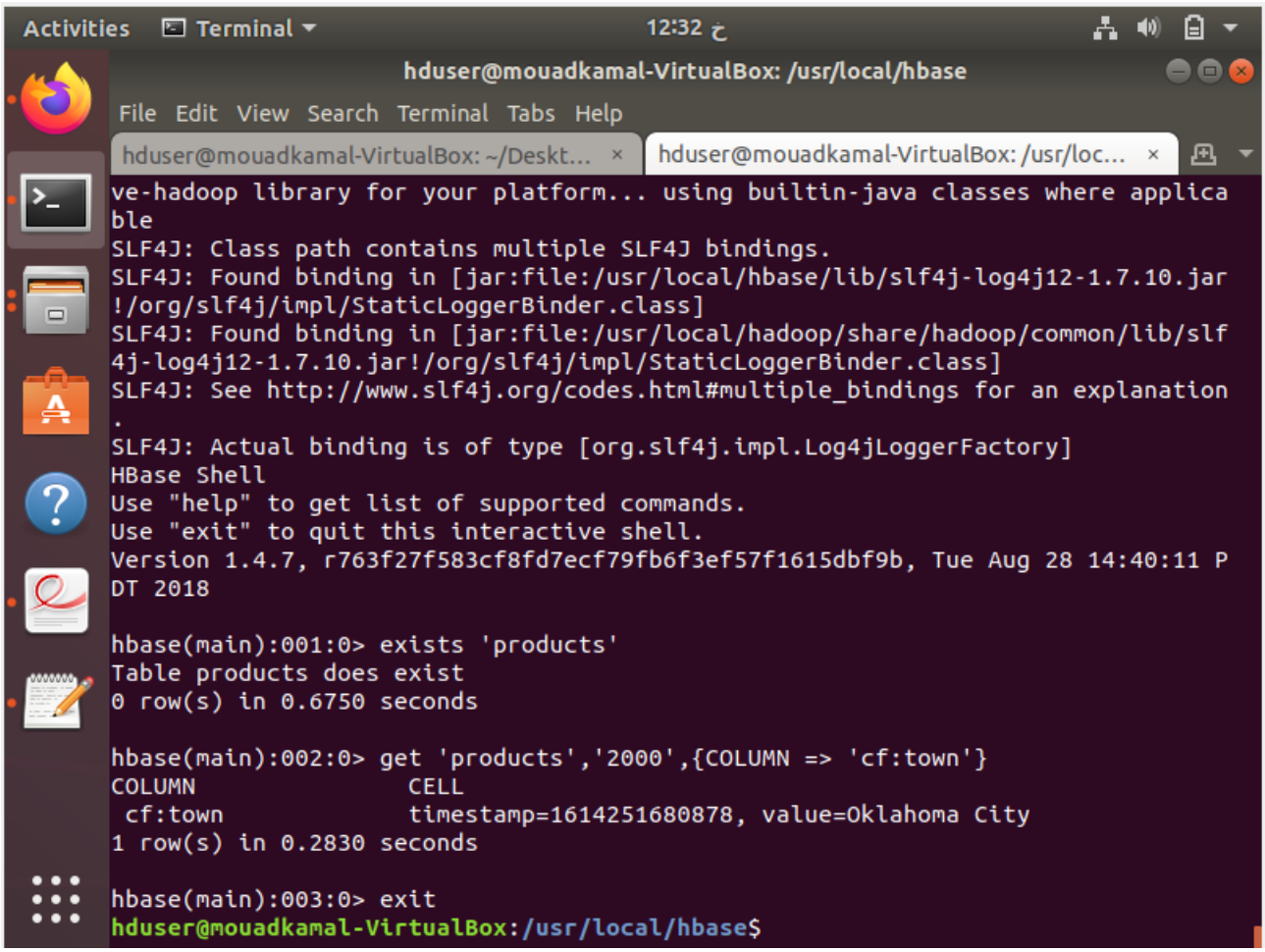
```
2021-02-25 11:15:18,402 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:19,318 INFO  [main] mapreduce.Job:  map 6% reduce 0%
2021-02-25 11:15:21,414 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:22,349 INFO  [main] mapreduce.Job:  map 7% reduce 0%
2021-02-25 11:15:24,418 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:25,393 INFO  [main] mapreduce.Job:  map 8% reduce 0%
2021-02-25 11:15:27,424 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:28,418 INFO  [main] mapreduce.Job:  map 10% reduce 0%
2021-02-25 11:15:30,434 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:33,449 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:34,457 INFO  [main] mapreduce.Job:  map 11% reduce 0%
2021-02-25 11:15:36,454 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:36,499 INFO  [main] mapreduce.Job:  map 13% reduce 0%
2021-02-25 11:15:39,496 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:39,574 INFO  [main] mapreduce.Job:  map 14% reduce 0%
2021-02-25 11:15:42,532 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:43,092 INFO  [main] mapreduce.Job:  map 15% reduce 0%
2021-02-25 11:15:45,549 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:46,141 INFO  [main] mapreduce.Job:  map 16% reduce 0%
2021-02-25 11:15:48,557 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:49,185 INFO  [main] mapreduce.Job:  map 18% reduce 0%
2021-02-25 11:15:51,567 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:15:52,206 INFO  [main] mapreduce.Job:  map 19% reduce 0%
2021-02-25 11:15:54,568 INFO  [communication thread] mapred.LocalJobRunner: map > map
{...}
 > map
2021-02-25 11:18:51,445 INFO  [main] mapreduce.Job:  map 70% reduce 0%
2021-02-25 11:18:53,661 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:18:54,469 INFO  [main] mapreduce.Job:  map 71% reduce 0%
2021-02-25 11:18:56,674 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:18:57,498 INFO  [main] mapreduce.Job:  map 72% reduce 0%
2021-02-25 11:18:59,681 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:19:00,620 INFO  [main] mapreduce.Job:  map 74% reduce 0%
2021-02-25 11:19:02,702 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:19:03,642 INFO  [main] mapreduce.Job:  map 75% reduce 0%
2021-02-25 11:19:05,703 INFO  [communication thread] mapred.LocalJobRunner: map > map
2021-02-25 11:19:06,684 INFO  [main] mapreduce.Job:  map 76% reduce 0%
2021-02-25 11:19:08,714 INFO  [communication thread] mapred.LocalJobRunner: map > map
{...}
2021-02-25 11:19:57,758 INFO  [LocalJobRunner Map Task Executor #0] mapred.LocalJobRunner: map
2021-02-25 11:19:57,758 INFO  [LocalJobRunner Map Task Executor #0] mapred.Task: Task 'attempt_local449278986_0001_m_000001_0' done.
2021-02-25 11:19:57,758 INFO  [LocalJobRunner Map Task Executor #0] mapred.LocalJobRunner: Finishing task: attempt_local449278986_00
2021-02-25 11:19:57,848 INFO  [Thread-46] mapred.LocalJobRunner: map task executor complete.
2021-02-25 11:19:58,450 INFO  [main] mapreduce.Job: Job job_local449278986_0001 completed successfully
2021-02-25 11:20:02,147 INFO  [main] mapreduce.Job: Counters: 21
    File System Counters
        FILE: Number of bytes read=60102777
        FILE: Number of bytes written=61280880
        FILE: Number of read operations=0
        FILE: Number of large read operations=0
        FILE: Number of write operations=0
        HDFS: Number of bytes read=377535548
        HDFS: Number of bytes written=0
        HDFS: Number of read operations=7
        HDFS: Number of large read operations=0
        HDFS: Number of write operations=0
    Map-Reduce Framework
        Map input records=4138476
        Map output records=4138476
        Input split bytes=216
        Spilled Records=0
        Failed Shuffles=0
        Merged Map outputs=0
        GC time elapsed (ms)=32617
        Total committed heap usage (bytes)=172281856
    ImportTsv
```

```
        Bad Lines=0
    File Input Format Counters
        Bytes Read=243313724
    File Output Format Counters
        Bytes Written=0
```



On vérifie que la base a bien été créée en consultant la ville de l'enregistrement numéro 2000:

```
hbase(main):001:0> exists 'products'
Table products does exist
0 row(s) in 0.6750 seconds

hbase(main):002:0> get 'products','2000',{COLUMN => 'cf:town'}
COLUMN                CELL
 cf:town              timestamp=1614251680878, value=Oklahoma City
1 row(s) in 0.2830 seconds
```



4- Traitement de données avec Spark :

Préparation de l'environnement :

1- On télécharge maven-3.5.0 :

Maven™ version 3.0.5 size 8.1MB

```
hduser@mouadkamal-VirtualBox:~/Downloads$ tar -zxvf apache-maven-3.5.0-bin.tar.gz
hduser@mouadkamal-VirtualBox:~/Downloads$ sudo mv apache-maven-3.5.0 /opt/
```

Pour mettre en place de manière permanente la variable d'environnement PATH pour tous les utilisateurs : On ouvre le fichier /etc/profile et modifie le PATH en ajoutant le chemin où se trouve le bin de maven dans export PATH :

```
export PATH=$PATH:/opt/java/jdk1.8.0_71/bin:/opt/java/jdk1.8.0_71/jre/bin::/opt/apache-maven-3.5.0/bin
```

Alors on aura :

```
  GNU nano 2.9.3                  /etc/profile                    Modified

      PS1='# '
    else
      PS1='$ '
    fi
  fi
fi

if [ -d /etc/profile.d ]; then
  for i in /etc/profile.d/*.sh; do
    if [ -r $i ]; then
      . $i
    fi
  done
  unset i
fi

export JAVA_HOME=/opt/java/jdk1.8.0_71/
export JRE_HOME=/opt/java/jdk1.8.0._71/jre
export PATH=$PATH:/opt/java/jdk1.8.0_71/bin:/opt/java/jdk1.8.0_71/jre/bin
export PATH=$PATH:/opt/java/jdk1.8.0_71/bin:/opt/java/jdk1.8.0_71/jre/bin::/op$




^G Get Help    ^O Write Out   ^W Where Is    ^K Cut Text    ^J Justify
^X Exit        ^R Read File   ^\ Replace     ^U Uncut Text  ^T To Spell
```

```
hduser@mouadkamal-VirtualBox:~/Downloads$ sudo nano /etc/profile
hduser@mouadkamal-VirtualBox:~/Downloads$ source /etc/profile
```

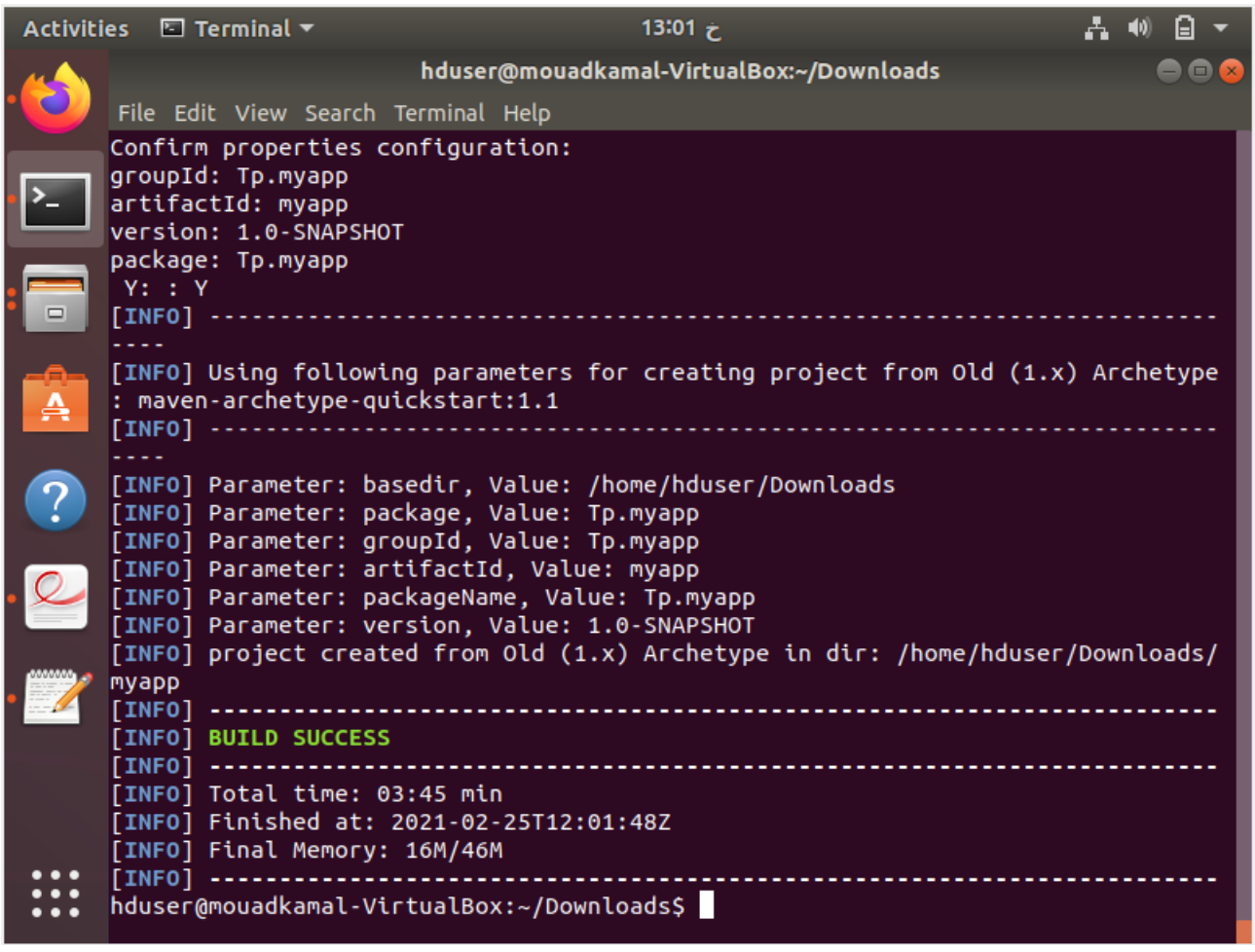
```
hduser@mouadkamal-VirtualBox:~/Downloads$ mvn -v
Apache Maven 3.5.0 (ff8f5e7444045639af65f6095c62210b5713f426; 2017-04-03T20:39:06+01:00)
Maven home: /opt/apache-maven-3.5.0
Java version: 1.8.0_71, vendor: Oracle Corporation
Java home: /opt/java/jdk1.8.0_71/jre
Default locale: en_US, platform encoding: UTF-8
OS name: "linux", version: "5.3.0-28-generic", arch: "amd64", family: "unix"
hduser@mouadkamal-VirtualBox:~/Downloads$
```

2- On crée un projet Maven avec la commande suivante :

```
hduser@mouadkamal-VirtualBox:~/Downloads$ mvn archetype:generate -DarchetypeArtifactId=maven-archetype-quickstart -DarchetypeVersion
```

output :

```
[INFO] Scanning for projects...
Downloading: https://repo.maven.apache.org/maven2/org/apache/maven/plugins/maven-clean-plugin/2.5/maven-clean-plugin-2.5.pom
Downloaded: https://repo.maven.apache.org/maven2/org/apache/maven/plugins/maven-clean-plugin/2.5/maven-clean-plugin-2.5.pom (3.9 kB
Downloading: https://repo.maven.apache.org/maven2/org/apache/maven/plugins/maven-plugins/22/maven-plugins-22.pom
Downloaded: https://repo.maven.apache.org/maven2/org/apache/maven/plugins/maven-plugins/22/maven-plugins-22.pom (13 kB at 44 kB/s)
Downloading: https://repo.maven.apache.org/maven2/org/apache/maven/maven-parent/21/maven-parent-21.pom
Downloaded: https://repo.maven.apache.org/maven2/org/apache/maven/maven-parent/21/maven-parent-21.pom (26 kB at 89 kB/s)
Downloading: https://repo.maven.apache.org/maven2/org/apache/apache/10/apache-10.pom
{...}
groupId: Tp.myapp
artifactId: myapp
version: 1.0-SNAPSHOT
package: Tp.myapp
 Y: : Y
[INFO] ----------------------------------------------------------------------------
[INFO] Using following parameters for creating project from Old (1.x) Archetype: maven-archetype-quickstart:1.1
[INFO] ----------------------------------------------------------------------------
[INFO] Parameter: basedir, Value: /home/hduser/Downloads
[INFO] Parameter: package, Value: Tp.myapp
[INFO] Parameter: groupId, Value: Tp.myapp
[INFO] Parameter: artifactId, Value: myapp
[INFO] Parameter: packageName, Value: Tp.myapp
[INFO] Parameter: version, Value: 1.0-SNAPSHOT
[INFO] project created from Old (1.x) Archetype in dir: /home/hduser/Downloads/myapp
[INFO] ------------------------------------------------------------------------
[INFO] BUILD SUCCESS
[INFO] ------------------------------------------------------------------------
[INFO] Total time: 03:45 min
[INFO] Finished at: 2021-02-25T12:01:48Z
[INFO] Final Memory: 16M/46M
[INFO] ------------------------------------------------------------------------
hduser@mouadkamal-VirtualBox:~/Downloads$
```



```
hduser@mouadkamal-VirtualBox:~/Downloads$ cd myapp
hduser@mouadkamal-VirtualBox:~/Downloads/myapp$ mvn package
[INFO] Scanning for projects...
[INFO]
[INFO] ------------------------------------------------------------------------
[INFO] Building myapp 1.0-SNAPSHOT
[INFO] ------------------------------------------------------------------------
Downloading: https://repo.maven.apache.org/maven2/org/apache/maven/plugins/maven-resources-plugin/2.6/maven-resources-plugin-2.6.pom
Downloaded: https://repo.maven.apache.org/maven2/org/apache/maven/plugins/maven-resources-plugin/2.6/maven-resources-plugin-2.6.pom
Downloading: https://repo.maven.apache.org/maven2/org/apache/maven/plugins/mave
{...}
Downloaded: https://repo.maven.apache.org/maven2/org/codehaus/plexus/plexus-archiver/2.1/plexus-archiver-2.1.jar (184 kB at 137 kB/s
Downloaded: https://repo.maven.apache.org/maven2/commons-lang/commons-lang/2.1/commons-lang-2.1.jar (208 kB at 133 kB/s)
Downloaded: https://repo.maven.apache.org/maven2/org/codehaus/plexus/plexus-utils/3.0/plexus-utils-3.0.jar (226 kB at 87 kB/s)
[INFO] Building jar: /home/hduser/Downloads/myapp/target/myapp-1.0-SNAPSHOT.jar
[INFO] ------------------------------------------------------------------------
[INFO] BUILD SUCCESS
[INFO] ------------------------------------------------------------------------
[INFO] Total time: 01:14 min
[INFO] Finished at: 2021-02-25T12:07:55Z
[INFO] Final Memory: 18M/46M
[INFO] ------------------------------------------------------------------------
```

![](Dans le fichier ***pom.xml*** et apres tout les ajouts, on va trouver:

```
<project
    xmlns="http://maven.apache.org/POM/4.0.0"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://maven.apache.org/POM/4.0.0 http://maven.apache.org/xsd/maven-4.0.0.xsd">
    <modelVersion>4.0.0</modelVersion>
    <groupId>Tp.myapp</groupId>
    <artifactId>myapp</artifactId>
    <version>1.0-SNAPSHOT</version>
    <packaging>jar</packaging>
    <name>myapp</name>
    <url>http://maven.apache.org</url>
    <properties>
    <project.build.sourceEncoding>UTF-8</project.build.sourceEncoding>
    </properties>
        <build>
    <plugins>
        <plugin>
            <groupId>org.apache.maven.plugins</groupId>
            <artifactId>maven-compiler-plugin</artifactId>
            <configuration>
                <source>1.8</source>
                <target>1.8</target>
            </configuration>
        </plugin>
    </plugins>
    </build>
    <dependencies>
    <dependency>
        <groupId>junit</groupId>
        <artifactId>junit</artifactId>
        <version>3.8.1</version>
        <scope>test</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.hbase</groupId>
        <artifactId>hbase</artifactId>
        <version>2.1.3</version>
        <type>pom</type>
    </dependency>
    <dependency>
        <groupId>org.apache.hbase</groupId>
        <artifactId>hbase-spark</artifactId>
        <version>2.0.0-alpha4</version>
    </dependency>
    <dependency>
        <groupId>org.apache.spark</groupId>
        <artifactId>spark-core_2.11</artifactId>
        <version>2.2.1</version>
    </dependency>
    </dependencies>
</project>
```

On renomme ***App.java*** en ***HbaseSparkProcess.java*** :

```
hduser@mouadkamal-VirtualBox:~/Downloads/myapp/src/main/java/Tp/myapp$ mv App.java HbaseSparkProcess.java
```

```
hduser@mouadkamal-VirtualBox:~/Downloads/myapp/src/main/java/Tp/myapp$ mv App.java HbaseSparkProcess.java
hduser@mouadkamal-VirtualBox:~/Downloads/myapp/src/main/java/Tp/myapp$ sudo nano HbaseSparkProcess.java
[sudo] password for hduser:
hduser@mouadkamal-VirtualBox:~/Downloads/myapp/src/main/java/Tp/myapp$
```

```
package Tp.myapp;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.hbase.HBaseConfiguration;
import org.apache.hadoop.hbase.client.Result;
import org.apache.hadoop.hbase.io.ImmutableBytesWritable;
import org.apache.hadoop.hbase.mapreduce.TableInputFormat;
import org.apache.spark.SparkConf;
import org.apache.spark.api.java.JavaPairRDD;
import org.apache.spark.api.java.JavaSparkContext;
public class HbaseSparkProcess {
  public void createHbaseTable() {
    Configuration config = HBaseConfiguration.create();
    SparkConf sparkConf = new
    SparkConf().setAppName("SparkHBaseTest").setMaster("local[4]");
    JavaSparkContext jsc = new JavaSparkContext(sparkConf);
    config.set(TableInputFormat.INPUT_TABLE, "products");
    JavaPairRDD < ImmutableBytesWritable,
    Result > hBaseRDD = jsc.newAPIHadoopRDD(config, TableInputFormat.class, ImmutableBytesWritable.class, Result.class);
    System.out.println("nombre d'enregistrements: " + hBaseRDD.count());
  }
  public static void main(String[] args) {
    HbaseSparkProcess admin = new HbaseSparkProcess();
    admin.createHbaseTable();
  }
}
```

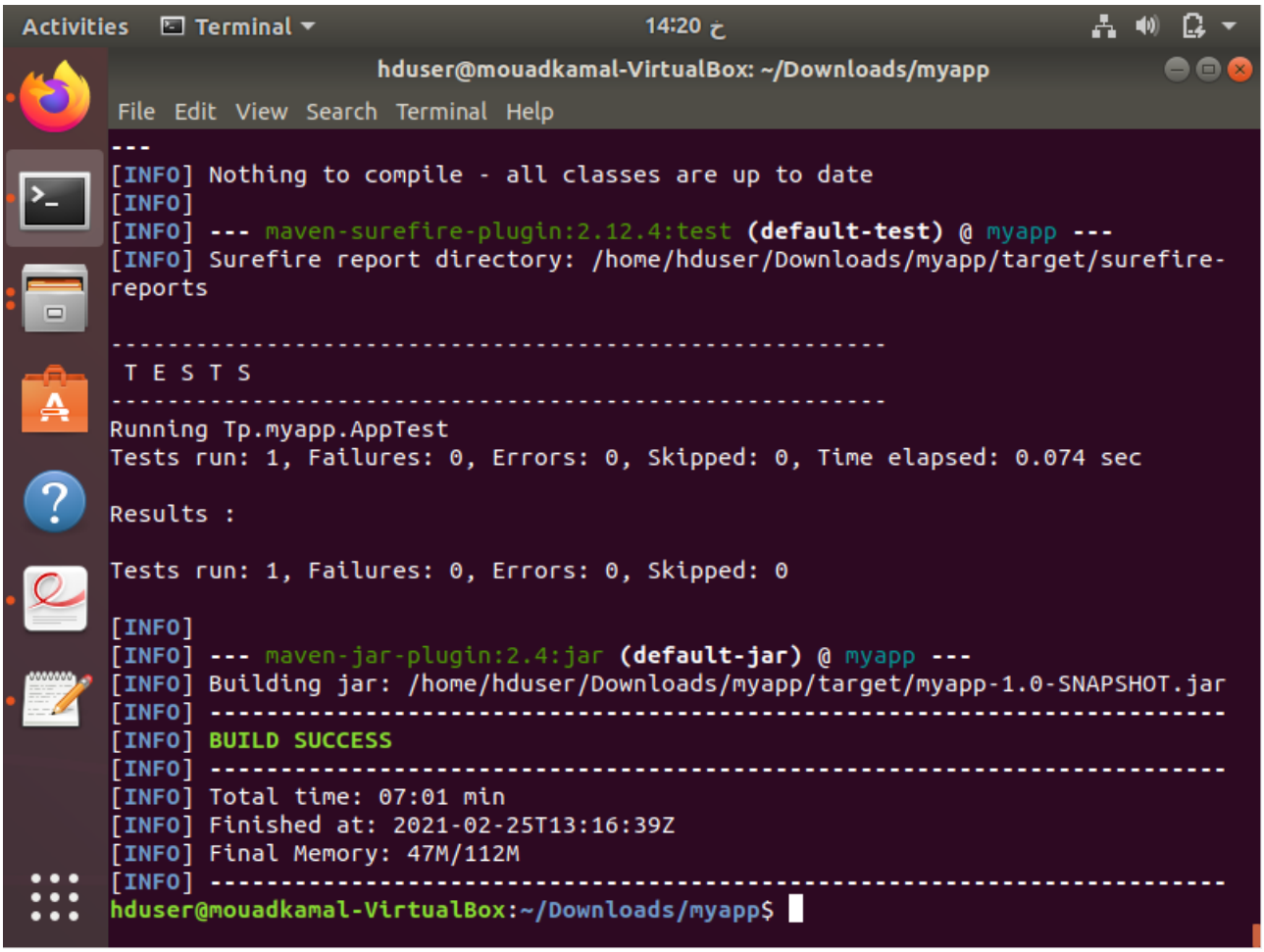
On enregistre HbaseSparkProcess.java, et on lance la commande :

```
hduser@mouadkamal-VirtualBox:~/Downloads/myapp$ mvn package
```

output:

```
[INFO] Scanning for projects...
[WARNING]
[WARNING] Some problems were encountered while building the effective model for Tp.myapp:myapp:jar:1.0-SNAPSHOT
[WARNING] 'build.plugins.plugin.version' for org.apache.maven.plugins:maven-compiler-plugin is missing. @ line 17, column 11
[WARNING]
[WARNING] It is highly recommended to fix these problems because they threaten the stability of your build.
[WARNING]
[WARNING] For this reason, future Maven versions might no longer support building such malformed projects.
[WARNING]
[INFO]
[INFO] ------------------------------------------------------------------------
[INFO] Building myapp 1.0-SNAPSHOT
[INFO] ------------------------------------------------------------------------
Downloading: https://repo.maven.apache.org/maven2/org/apache/hbase/hbase/2.1.3/hbase-2.1.3.pom
Downloaded: https://repo.maven.apache.org/maven2/org/apache/hbase/hbase/2.1.3/hbase-2.1.3.pom (150 kB at 89 kB/s)
Downloading: https://repo.maven.apache.org/maven2/junit/junit/4.12/junit-4.12.pom
Downloaded: https://repo.maven.apache.org/maven2/junit/junit/4.12/junit-4.12.pom (24 kB at 82 kB/s)
Downloading: https://repo.maven.apache.org/maven2/org/apache/hbase/hbase-spark/2.0.0-alpha4/hbase-spark-2.0.0-alpha4.pom
Downloaded: https://repo.maven.apache.org/maven2/org/apache/hbase/hbase-spark/2.0.0-alpha4/hbase-spark-2.0.0-alpha4.pom (23 kB at 10
Downloading: https://repo.maven.apache.org/maven2/org/apache/hbase/hbase-build-configuration/2.0.0-alpha4/hbase-build-configuration-
Downloaded: https://repo.maven.apache.org/maven2/org/apache/hbase/hbase-build-configuration/2.0.0-alpha4/hbase-build-configuration-2
Downloading: https://repo.maven.apache.org/maven2/org/apache/hbase/hbase/2.0.0-alpha4/hbase-2.0.0-alpha4.pom
Downloaded: https://repo.maven.apache.org/maven2/org/apache/hbase/hbase/2.0.0-alpha4/hbase-2.0.0-alpha4.pom (139 kB at 188 kB/s)
Downloading: https://repo.maven.apache.org/maven2/org/apache/hbase/thirdparty/hbase-shaded-miscellaneous/1.0.1/hbase-shaded-miscella
Downloaded: https://repo.maven.apache.org/maven2/org/apache/hbase/thirdparty/hbase-shaded-miscellaneous/1.0.1/hbase-shaded-miscellan
Downloading: https://repo.maven.apache.org/maven2/org/apache/hbase/thirdparty/hbase-thirdparty/1.0.1/hbase-thirdparty-1.0.1.pom
Downloaded: https://repo.maven.apache.org/maven2/org/apache/hbase/thirdparty/hbase-thirdparty/1.0.1/hbase-thirdparty-1.0.1.pom (13 k
Downloading: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/module/jackson-module-scala_2.10/2.9.1/jackson-module-scala_
Downloaded: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/module/jackson-module-scala_2.10/2.9.1/jackson-module-scala_2
Downloading: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/core/jackson-core/2.9.1/jackson-core-2.9.1.pom
Downloaded: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/core/jackson-core/2.9.1/jackson-core-2.9.1.pom (6.1 kB at 30
Downloading: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/jackson-parent/2.9.1/jackson-parent-2.9.1.pom
Downloaded: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/jackson-parent/2.9.1/jackson-parent-2.9.1.pom (8.0 kB at 38 k
```

```
Downloaded: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/jackson-parent/2.9.1/jackson-parent-2.9.1.pom (6.8 kB at 38 k
Downloading: https://repo.maven.apache.org/maven2/com/fasterxml/oss-parent/30/oss-parent-30.pom
Downloaded: https://repo.maven.apache.org/maven2/com/fasterxml/oss-parent/30/oss-parent-30.pom (21 kB at 80 kB/s)
Downloading: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/core/jackson-annotations/2.9.1/jackson-annotations-2.9.1.pom
Downloaded: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/core/jackson-annotations/2.9.1/jackson-annotations-2.9.1.pom
Downloading: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/core/jackson-databind/2.9.1/jackson-databind-2.9.1.pom
Downloaded: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/core/jackson-databind/2.9.1/jackson-databind-2.9.1.pom (6.8 k
Downloading: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/jackson-bom/2.9.1/jackson-bom-2.9.1.pom
Downloaded: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/jackson-bom/2.9.1/jackson-bom-2.9.1.pom (12 kB at 54 kB/s)
Downloading: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/core/jackson-annotations/2.9.0/jackson-annotations-2.9.0.pom
Downloaded: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/core/jackson-annotations/2.9.0/jackson-annotations-2.9.0.pom
Downloading: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/jackson-parent/2.9.0/jackson-parent-2.9.0.pom
Downloaded: https://repo.maven.apache.org/maven2/com/fasterxml/jackson/jackson-parent/2.9.0/jackson-parent-2.9.0.pom (7.8 kB at 32 k
Downloading: https://repo.maven.apache.org/maven2/com/fasterxml/oss-parent/28/oss-parent-28.pom
Downloaded: https://repo.maven.apache.org/maven2/com/fasterxml/oss-parent/28/os
{...}

-------------------------------------------------------
 T E S T S
-------------------------------------------------------
Running Tp.myapp.AppTest
Tests run: 1, Failures: 0, Errors: 0, Skipped: 0, Time elapsed: 0.074 sec

Results :

Tests run: 1, Failures: 0, Errors: 0, Skipped: 0

[INFO]
[INFO] --- maven-jar-plugin:2.4:jar (default-jar) @ myapp ---
[INFO] Building jar: /home/hduser/Downloads/myapp/target/myapp-1.0-SNAPSHOT.jar
[INFO] ------------------------------------------------------------------------
[INFO] BUILD SUCCESS
[INFO] ------------------------------------------------------------------------
[INFO] Total time: 07:01 min
[INFO] Finished at: 2021-02-25T13:16:39Z
[INFO] Final Memory: 47M/112M
[INFO] ------------------------------------------------------------------------
```



On copie le fichier **myapp-1.0-SNAPSHOT.jar** dans le répertoir **/usr/local/spark** :

```
hduser@mouadkamal-VirtualBox:~/Downloads/myapp/target$ ls
classes            maven-archiver  myapp-1.0-SNAPSHOT.jar  test-classes
generated-sources  maven-status    surefire-reports
hduser@mouadkamal-VirtualBox:~/Downloads/myapp/target$ cp myapp-1.0-SNAPSHOT.jar /usr/local/spark/
```

On copie tous les fichiers de la bibliothèque hbase dans le répertoire jars de spark:

```
hduser@mouadkamal-VirtualBox:~/Downloads/myapp/target$ cp -r /usr/local/hbase/lib/* /usr/local/spark/jars
```

On exécute ce fichier grâce à spark-submit comme suit :

```
hduser@mouadkamal-VirtualBox:/usr/local/spark$ spark-submit --class TaseSparkProcess myapp-1.0-SNAPSHOT.jar
```

output:

```
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/spark/jars/slf4j-log4j12-1.7.16.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/spark/jars/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
21/02/25 13:30:16 WARN util.Utils: Your hostname, mouadkamal-VirtualBox resolves to a loopback address: 127.0.0.1; using 10.0.2.15 i
21/02/25 13:30:16 WARN util.Utils: Set SPARK_LOCAL_IP if you need to bind to another address
21/02/25 13:30:23 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes w
21/02/25 13:30:25 INFO spark.SparkContext: Running Spark version 2.4.3
21/02/25 13:30:25 INFO spark.SparkContext: Submitted application: SparkHBaseTest
{...}
nombre d'enregistrements: 4138476
21/02/25 13:32:55 INFO spark.SparkContext: Invoking stop() from shutdown hook
21/02/25 13:32:55 INFO server.AbstractConnector: Stopped Spark@2780f7da{HTTP/1.1,[http/1.1]}{0.0.0.0:4040}
21/02/25 13:32:55 INFO ui.SparkUI: Stopped Spark web UI at http://10.0.2.15:4040
21/02/25 13:32:57 INFO spark.MapOutputTrackerMasterEndpoint: MapOutputTrackerMasterEndpoint stopped!
21/02/25 13:32:58 INFO memory.MemoryStore: MemoryStore cleared
21/02/25 13:32:58 INFO storage.BlockManager: BlockManager stopped
21/02/25 13:32:58 INFO storage.BlockManagerMaster: BlockManagerMaster stopped
21/02/25 13:32:58 INFO scheduler.OutputCommitCoordinator$OutputCommitCoordinatorEndpoint: OutputCommitCoordinator stopped!
21/02/25 13:32:59 INFO spark.SparkContext: Successfully stopped SparkContext
21/02/25 13:32:59 INFO util.ShutdownHookManager: Shutdown hook called
21/02/25 13:32:59 INFO util.ShutdownHookManager: Deleting directory /tmp/spark-94d4ec4a-b523-4d13-bf56-5538a7fcb82f
21/02/25 13:32:59 INFO util.ShutdownHookManager: Deleting directory /tmp/spark-ef905603-8c75-4828-a9e9-5f84ffbec045
```