

République Tunisienne
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE DE CARTHAGE

INSTITUT DES HAUTE S ETUDES COMMERCIALES



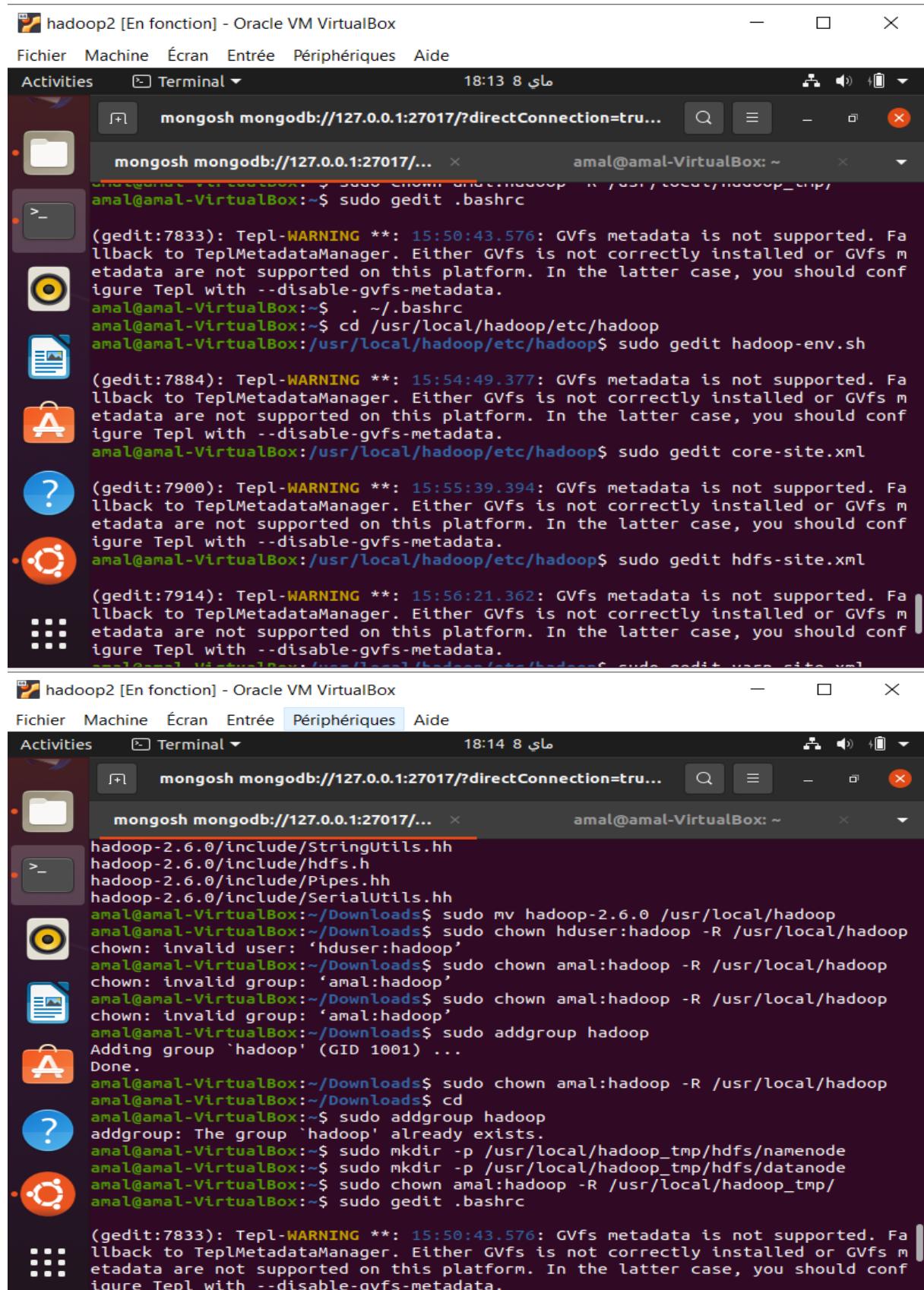
Projet MSI

Enseignante : Fayech Sondes

*Elaboré par : Rahma Ben Saber
Khayati Amal
Latifa Mankai*

Année universitaire 2021-2022

1/ Installation Hadoop sur une machine virtuel



The screenshot shows two terminal windows in an Oracle VM VirtualBox interface. Both windows are titled "hadoop2 [En fonction] - Oracle VM VirtualBox". The top window has its tabs set to "Fichier", "Machine", "Écran", "Entrée", "Périphériques", and "Aide". The bottom window has its tabs set to "Activities", "Terminal", and "Périphériques". Both windows show a terminal session with the user "amal" at "amal@amal-VirtualBox".

Top Terminal Window (Périphériques tab selected):

```
mongosh mongodb://127.0.0.1:27017/?directConnection=true... amal@amal-VirtualBox:~$ sudo gedit .bashrc
(gedit:7833): Tepl-WARNING **: 15:50:43.576: GVfs metadata is not supported. Fallback to TeplMetadataManager. Either GVfs is not correctly installed or GVfs metadata are not supported on this platform. In the latter case, you should configure Tepl with --disable-gvfs-metadata.
amal@amal-VirtualBox:~$ . ~/.bashrc
amal@amal-VirtualBox:~$ cd /usr/local/hadoop/etc/hadoop
amal@amal-VirtualBox:/usr/local/hadoop/etc/hadoop$ sudo gedit hadoop-env.sh
(gedit:7884): Tepl-WARNING **: 15:54:49.377: GVfs metadata is not supported. Fallback to TeplMetadataManager. Either GVfs is not correctly installed or GVfs metadata are not supported on this platform. In the latter case, you should configure Tepl with --disable-gvfs-metadata.
amal@amal-VirtualBox:/usr/local/hadoop/etc/hadoop$ sudo gedit core-site.xml
(gedit:7900): Tepl-WARNING **: 15:55:39.394: GVfs metadata is not supported. Fallback to TeplMetadataManager. Either GVfs is not correctly installed or GVfs metadata are not supported on this platform. In the latter case, you should configure Tepl with --disable-gvfs-metadata.
amal@amal-VirtualBox:/usr/local/hadoop/etc/hadoop$ sudo gedit hdfs-site.xml
(gedit:7914): Tepl-WARNING **: 15:56:21.362: GVfs metadata is not supported. Fallback to TeplMetadataManager. Either GVfs is not correctly installed or GVfs metadata are not supported on this platform. In the latter case, you should configure Tepl with --disable-gvfs-metadata.
amal@amal-VirtualBox:/usr/local/hadoop/etc/hadoop$ sudo gedit mapred-site.xml
```

Bottom Terminal Window (Activities tab selected):

```
mongosh mongodb://127.0.0.1:27017/?directConnection=true... amal@amal-VirtualBox:~$ ls -l /usr/local/hadoop/include/
total 0
amal@amal-VirtualBox:~/Downloads$ sudo mv hadoop-2.6.0 /usr/local/hadoop
amal@amal-VirtualBox:~/Downloads$ sudo chown hduser:hadoop -R /usr/local/hadoop
chown: invalid user: 'hduser:hadoop'
amal@amal-VirtualBox:~/Downloads$ sudo chown amal:hadoop -R /usr/local/hadoop
chown: invalid group: 'amal:hadoop'
amal@amal-VirtualBox:~/Downloads$ sudo chown amal:hadoop -R /usr/local/hadoop
chown: invalid group: 'amal:hadoop'
amal@amal-VirtualBox:~/Downloads$ sudo addgroup hadoop
Adding group `hadoop' (GID 1001) ...
Done.
amal@amal-VirtualBox:~/Downloads$ sudo chown amal:hadoop -R /usr/local/hadoop
amal@amal-VirtualBox:~/Downloads$ cd
amal@amal-VirtualBox:~$ sudo addgroup hadoop
addgroup: The group `hadoop' already exists.
amal@amal-VirtualBox:~$ sudo mkdir -p /usr/local/hadoop_tmp/hdfs/namenode
amal@amal-VirtualBox:~$ sudo mkdir -p /usr/local/hadoop_tmp/hdfs/datanode
amal@amal-VirtualBox:~$ sudo chown amal:hadoop -R /usr/local/hadoop_tmp/
amal@amal-VirtualBox:~$ sudo gedit .bashrc
(gedit:7833): Tepl-WARNING **: 15:50:43.576: GVfs metadata is not supported. Fallback to TeplMetadataManager. Either GVfs is not correctly installed or GVfs metadata are not supported on this platform. In the latter case, you should configure Tepl with --disable-gvfs-metadata.
```

hadoop2 [En fonction] - Oracle VM VirtualBox

Fichier Machine Écran Entrée Périphériques Aide

Activities Terminal 18:15 8 مای

mongosh mongodb://127.0.0.1:27017/?directConnection=true... amal@amal-VirtualBox: ~

```
etadatas are not supported on this platform. In the latter case, you should configure Tepl with --disable-gvfs-metadata.
amal@amal-VirtualBox:/usr/local/hadoop/etc/hadoop$ cd
amal@amal-VirtualBox:~$ hdfs namenode -format
22/05/08 15:58:40 INFO namenode.NameNode: STARTUP_MSG:
/*****
STARTUP_MSG: Starting NameNode
STARTUP_MSG: host = amal-VirtualBox/127.0.1.1
STARTUP_MSG: args = [-format]
STARTUP_MSG: version = 2.6.0
STARTUP_MSG: classpath = /usr/local/hadoop/etc/hadoop:/usr/local/hadoop/share/hadoop/common/lib/httpclient-4.2.5.jar:/usr/local/hadoop/share/hadoop/common/lib/asm-3.2.jar:/usr/local/hadoop/share/hadoop/common/lib/curator-framework-2.6.0.jar:/usr/local/hadoop/share/hadoop/common/lib/junit-4.11.jar:/usr/local/hadoop/share/hadoop/common/lib/commons-compress-1.4.1.jar:/usr/local/hadoop/share/hadoop/common/lib/jetty-6.1.26.jar:/usr/local/hadoop/share/hadoop/common/lib/commons-beanutils-core-1.8.0.jar:/usr/local/hadoop/share/hadoop/common/lib/jackson-mapper-asl-1.9.13.jar:/usr/local/hadoop/share/hadoop/common/lib/jsr305-1.3.9.jar:/usr/local/hadoop/share/hadoop/common/lib/jersey-json-1.9.jar:/usr/local/hadoop/share/hadoop/common/lib/hamcrest-core-1.3.jar:/usr/local/hadoop/share/hadoop/common/lib/jsp-api-2.1.jar:/usr/local/hadoop/share/hadoop/common/lib/xz-1.0.jar:/usr/local/hadoop/share/hadoop/common/lib/gson-2.2.4.jar:/usr/local/hadoop/share/hadoop/common/lib/mockito-all-1.8.5.jar:/usr/local/hadoop/share/hadoop/common/lib/commons-el-1.0.jar:/usr/local/hadoop/share/hadoop/common/lib/java-xmlbundler-0.4.jar:/usr/local/hadoop/share/hadoop/common/lib/jackson-xc-1.9.13.jar:/usr/local/hadoop/share/hadoop/common/lib/sl4j-log4j12-1.7.5.jar:/usr/local/hadoop/share/hadoop/common
```

hadoop2 [En fonction] - Oracle VM VirtualBox

Fichier Machine Écran Entrée Périphériques Aide

Activities Terminal 18:17 8 مای

mongosh mongodb://127.0.0.1:27017/?directConnection=true... amal@amal-VirtualBox: ~

```
amal@amal-VirtualBox:~$ start-dfs.sh
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.6.0.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/05/08 16:08:24 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting namenodes on [localhost]
localhost: namenode running as process 8177. Stop it first.
localhost: datanode running as process 8328. Stop it first.
Starting secondary namenodes [0.0.0.0]
0.0.0.0: secondarynamenode running as process 8504. Stop it first.
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.6.0.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/05/08 16:08:31 WARN util.NativeCodeLoader: Unable to load native-hadoop libr
```

hadoop2 [En fonction] - Oracle VM VirtualBox

Fichier Machine Écran Entrée Périphériques Aide

Activities Terminal amal@amal-VirtualBox: ~ 18:17 8 مای

```
mongosh mongodb://127.0.0.1:27017/... amal@amal-VirtualBox: ~
WARNING: All illegal access operations will be denied in a future release
22/05/08 16:08:31 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
amal@amal-VirtualBox:~$ start-yarn.sh
starting yarn daemons
resourcemanager running as process 8624. Stop it first.
localhost: nodemanager running as process 8747. Stop it first.
amal@amal-VirtualBox:~$ hdfs dfs -mkdir /amal
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.6.0.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/05/08 16:10:07 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
amal@amal-VirtualBox:~$ hdfs dfs -ls /
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.6.0.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
```

Activities Terminal amal@amal-VirtualBox: ~ 10:41 13 مای

```
localhost: starting datanode, logging to /usr/local/hadoop/logs/hadoop-amal-datanode-amal-VirtualBox.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop/logs/hadoop-amal-secondarynamenode-amal-VirtualBox.out
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.6.0.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/05/13 10:41:12 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-amal-resourcemanager-amal-VirtualBox.out
localhost: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-amal-nodemanager-amal-VirtualBox.out
amal@amal-VirtualBox:~$ jps
3380 NameNode
3510 DataNode
3720 SecondaryNameNode
3977 NodeManager
4267 Jps
3853 ResourceManager
amal@amal-VirtualBox:~$
```

hadoop2 [En fonction] - Oracle VM VirtualBox

Fichier Machine Écran Entrée Périphériques Aide

Activities Terminal 18:17 ٨ ماء amal@amal-VirtualBox: ~

mongosh mongodb://127.0.0.1:27017/... amal@amal-VirtualBox: ~

```
localhost: nodemanager running as process 8747. Stop it first.
amal@amal-VirtualBox:~$ hdfs dfs -mkdir /amal
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.6.0.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/05/08 16:10:07 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
amal@amal-VirtualBox:~$ hdfs dfs -ls /
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.6.0.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/05/08 16:10:19 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
drwxr-xr-x - amal supergroup 0 2022-05-08 16:10:19 /amal
```

hadoop2 [En fonction] - Oracle VM VirtualBox

Fichier Machine Écran Entrée Périphériques Aide

Activities Firefox Web Browser 10:54 ٩ ماء

All Applications localhost:46213 X Getting Started X NiFi Flow X traduction - C Solved: Nifi 1.14 X GitHub - apache/ X What's New in Ni X +

localhost:8088/cluster

hadoop All Applications Logged in as: dr.who

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	Vcores Used	Vcores Total	Vcores Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes
0	0	0	0	0 B	8 GB	0 B	0	8	0	1	0	0	0	0	0

Show 20 entries Search:

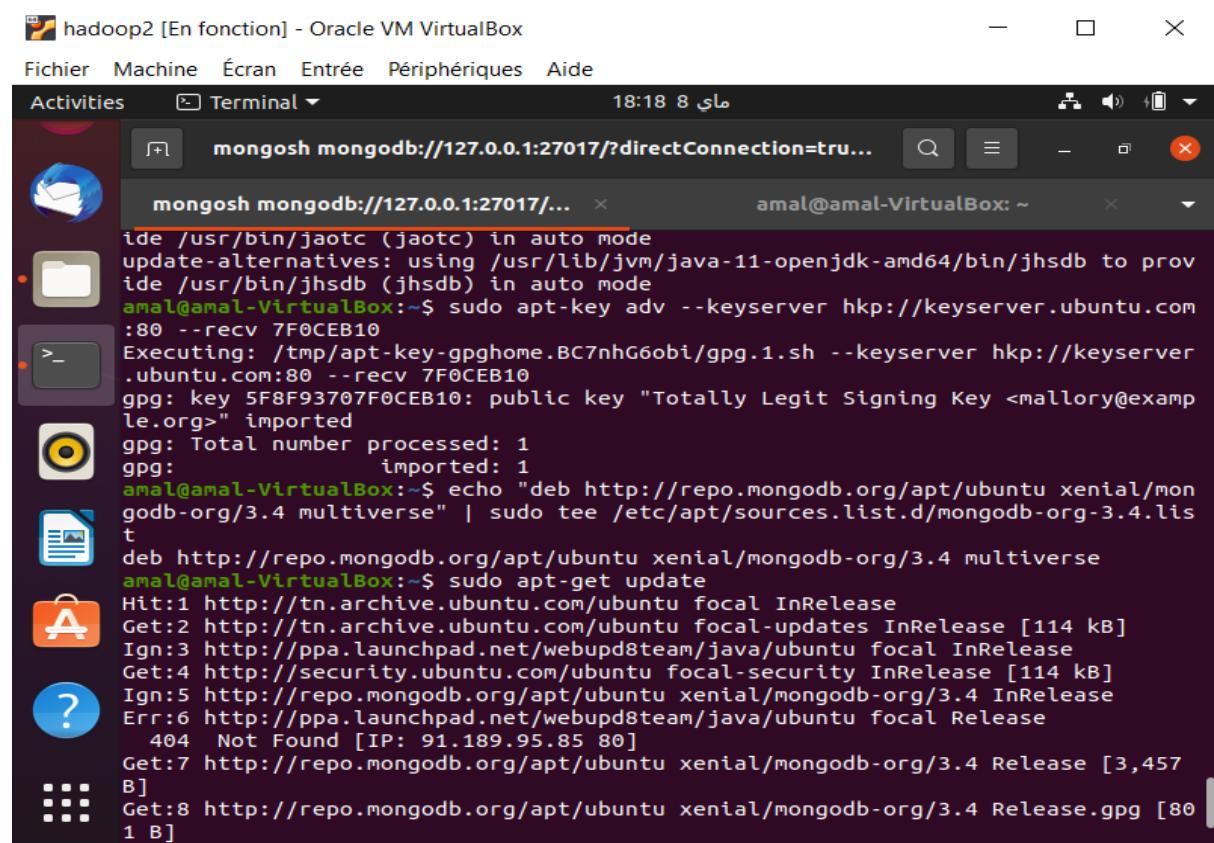
ID User Name Application Type Queue StartTime FinishTime State FinalStatus Progress Tracking UI

No data available in table

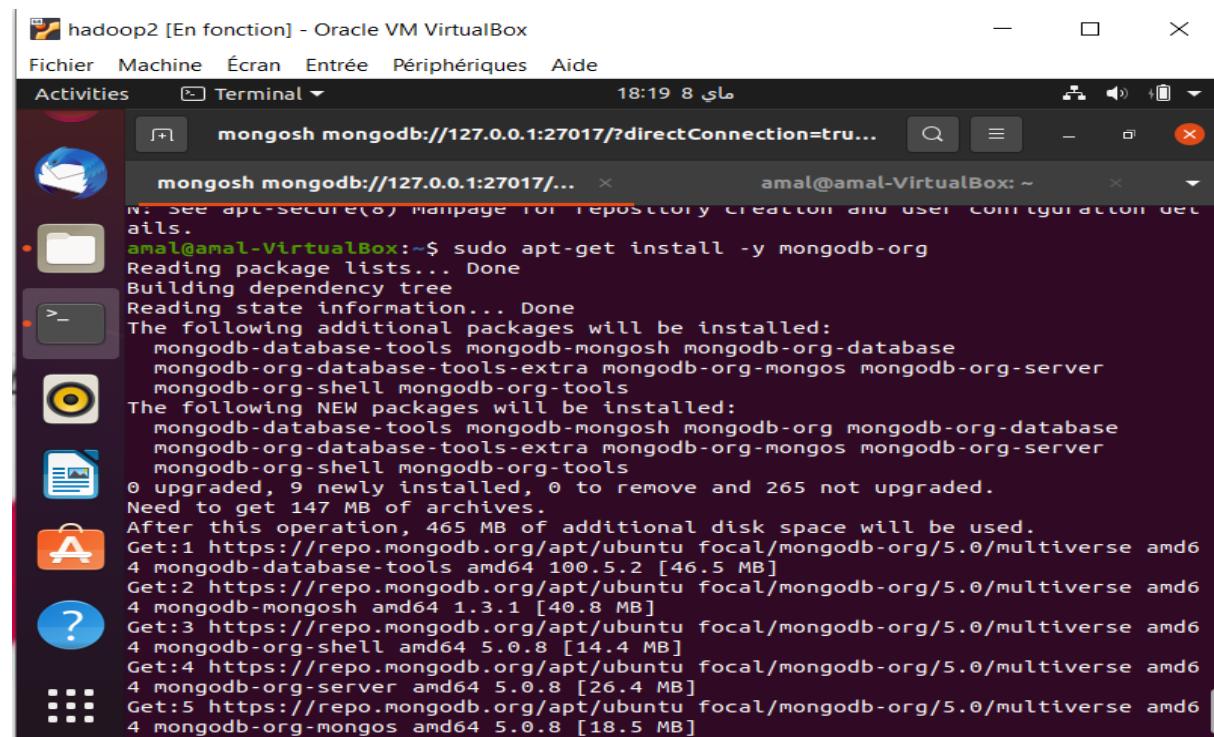
Showing 0 to 0 of 0 entries First Previous Next Last

Scheduler Tools

2/ Installation MongoDB sous la VM :



```
hadoop2 [En fonction] - Oracle VM VirtualBox
Fichier Machine Écran Entrée Périphériques Aide
Activities Terminal 18:18 8 مای
mongosh mongodb://127.0.0.1:27017/?directConnection=true...
mongosh mongodb://127.0.0.1:27017/...
ide /usr/bin/jaotc (jaotc) in auto mode
update-alternatives: using /usr/lib/jvm/java-11-openjdk-amd64/bin/jhsdb to provide jhsdb
ide /usr/bin/jhsdb (jhsdb) in auto mode
amal@amal-VirtualBox:~$ sudo apt-key adv --keyserver hkp://keyserver.ubuntu.com:80 --recv 7F0CEB10
Executing: /tmp/apt-key-gpghome.BC7nhG6obi/gpg.1.sh --keyserver hkp://keyserver.ubuntu.com:80 --recv 7F0CEB10
gpg: key 5F8F93707F0CEB10: public key "Totally Legit Signing Key <mallory@example.org>" imported
gpg: Total number processed: 1
gpg: imported: 1
amal@amal-VirtualBox:~$ echo "deb http://repo.mongodb.org/apt/ubuntu xenial/mongodb-org/3.4 multiverse" | sudo tee /etc/apt/sources.list.d/mongodb-org-3.4.list
deb http://repo.mongodb.org/apt/ubuntu xenial/mongodb-org/3.4 multiverse
amal@amal-VirtualBox:~$ sudo apt-get update
Hit:1 http://tn.archive.ubuntu.com/ubuntu focal InRelease
Get:2 http://tn.archive.ubuntu.com/ubuntu focal-updates InRelease [114 kB]
Ign:3 http://ppa.launchpad.net/webupd8team/java/ubuntu focal InRelease
Get:4 http://security.ubuntu.com/ubuntu focal-security InRelease [114 kB]
Ign:5 http://repo.mongodb.org/apt/ubuntu xenial/mongodb-org/3.4 InRelease
Err:6 http://ppa.launchpad.net/webupd8team/java/ubuntu focal Release
  404 Not Found [IP: 91.189.95.85 80]
Get:7 http://repo.mongodb.org/apt/ubuntu xenial/mongodb-org/3.4 Release [3,457 B]
Get:8 http://repo.mongodb.org/apt/ubuntu xenial/mongodb-org/3.4 Release.gpg [80 1 B]
```



```
hadoop2 [En fonction] - Oracle VM VirtualBox
Fichier Machine Écran Entrée Périphériques Aide
Activities Terminal 18:19 8 مای
mongosh mongodb://127.0.0.1:27017/?directConnection=true...
mongosh mongodb://127.0.0.1:27017/...
W: See apt-secure(8) manpage for repository creation and user configuration details.
amal@amal-VirtualBox:~$ sudo apt-get install -y mongodb-org
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  mongodb-database-tools mongodb-mongosh mongodb-org-database
  mongodb-org-database-tools-extra mongodb-org-mongos mongodb-org-server
  mongodb-org-shell mongodb-org-tools
The following NEW packages will be installed:
  mongodb-database-tools mongodb-mongosh mongodb-org mongodb-org-database
  mongodb-org-database-tools-extra mongodb-org-mongos mongodb-org-server
  mongodb-org-shell mongodb-org-tools
0 upgraded, 9 newly installed, 0 to remove and 265 not upgraded.
Need to get 147 MB of archives.
After this operation, 465 MB of additional disk space will be used.
Get:1 https://repo.mongodb.org/apt/ubuntu focal/mongodb-org/5.0/multiverse amd64 4 mongodb-database-tools amd64 100.5.2 [46.5 MB]
Get:2 https://repo.mongodb.org/apt/ubuntu focal/mongodb-org/5.0/multiverse amd64 4 mongodb-mongosh amd64 1.3.1 [40.8 MB]
Get:3 https://repo.mongodb.org/apt/ubuntu focal/mongodb-org/5.0/multiverse amd64 4 mongodb-org-shell amd64 5.0.8 [14.4 MB]
Get:4 https://repo.mongodb.org/apt/ubuntu focal/mongodb-org/5.0/multiverse amd64 4 mongodb-org-server amd64 5.0.8 [26.4 MB]
Get:5 https://repo.mongodb.org/apt/ubuntu focal/mongodb-org/5.0/multiverse amd64 4 mongodb-org-mongos amd64 5.0.8 [18.5 MB]
```

- Démarrer le serveur : mongod

```

mongosh mongodb://127.0.0.1:27017/?directConnection=true...
Setting up mongodb-org (5.0.8) ...
Processing triggers for man-db (2.9.1-1) ...
amal@amal-VirtualBox:~$ sudo systemctl start mongod
amal@amal-VirtualBox:~$ mongod
{"t":{"$date":"2022-05-08T16:25:38.328+01:00"}, "s":"I", "c":"CONTROL", "id":23285, "ctx":"thread1", "msg":"Automatically disabling TLS 1.0, to force-enable TLS 1.0 specify --sslDisabledProtocols 'none'"}
{"t":{"$date":"2022-05-08T16:25:38.328+01:00"}, "s":"I", "c":"NETWORK", "id":4915701, "ctx":"thread1", "msg":"Initialized wire specification", "attr":{"spec":{"incomingExternalClient":{"minWireVersion":0,"maxWireVersion":13}, "incomingInternalClient":{"minWireVersion":0,"maxWireVersion":13}, "outgoing":{"minWireVersion":0, "maxWireVersion":13}, "isInternalClient":true}}}
{"t":{"$date":"2022-05-08T16:25:38.328+01:00"}, "s":"W", "c":"ASIO", "id":2601, "ctx":"thread1", "msg":"No TransportLayer configured during NetworkInterface startup"}
{"t":{"$date":"2022-05-08T16:25:38.329+01:00"}, "s":"I", "c":"NETWORK", "id":4648601, "ctx":"thread1", "msg":"Implicit TCP FastOpen unavailable. If TCP FastOpen is required, set tcpFastOpenServer, tcpFastOpenClient, and tcpFastOpenQueueSize."}
{"t":{"$date":"2022-05-08T16:25:38.329+01:00"}, "s":"W", "c":"ASIO", "id":2601, "ctx":"thread1", "msg":"No TransportLayer configured during NetworkInterface startup"}
{"t":{"$date":"2022-05-08T16:25:38.329+01:00"}, "s":"I", "c":"REPL", "id":5123008, "ctx":"thread1", "msg":"Successfully registered PrimaryOnlyService", "attr":{"service":"TenantMigrationDonorService", "ns":"config.tenantMigrationDonors"}}
}

```

- Lancer le terminal : mongosh

```

mongosh mongodb://127.0.0.1:27017/?directConnection=true...
3092, "ctx":"initandlisten", "msg":"\n\n***aborted after fassert() failure\n\n"
amal@amal-VirtualBox:~$ mongosh
Current Mongosh Log ID: 6277e0f8e4109bc27dfc3e39
Connecting to: mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+1.3.1
Using MongoDB: 5.0.8
Using Mongosh: 1.3.1

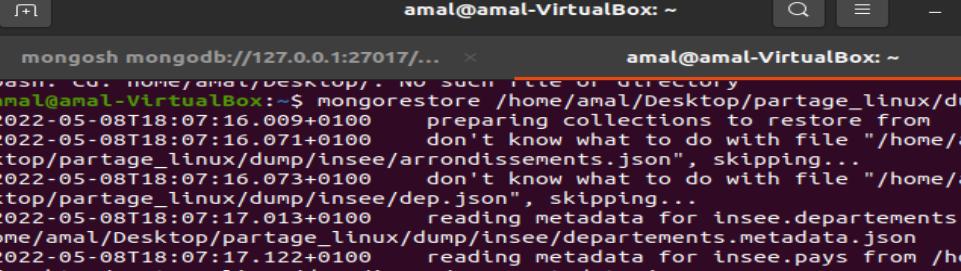
For mongosh info see: https://docs.mongodb.com/mongodb-shell/

To help improve our products, anonymous usage data is collected and sent to MongoDB periodically (https://www.mongodb.com/legal/privacy-policy).
You can opt-out by running the disableTelemetry() command.

-----
The server generated these startup warnings when booting:
2022-05-08T16:25:25.754+01:00: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem
2022-05-08T16:25:25.992+01:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted
-----
test>

```

3/ Dump la base de données insee :



The screenshot shows a Linux desktop environment with a terminal window open. The terminal window title is "mongosh mongodb://127.0.0.1:27017/...". The command run is "mongorestore /home/amal/Desktop/partage_linux/dump/". The output of the command is as follows:

```
amal@amal-VirtualBox:~$ mongorestore /home/amal/Desktop/partage_linux/dump/
2022-05-08T18:07:16.009+0100  preparing collections to restore from
2022-05-08T18:07:16.071+0100  don't know what to do with file "/home/amal/Desktop/partage_linux/dump/insee/arrondissements.json", skipping...
2022-05-08T18:07:16.073+0100  don't know what to do with file "/home/amal/Desktop/partage_linux/dump/insee/dep.json", skipping...
2022-05-08T18:07:17.013+0100  reading metadata for insee.departements from /home/amal/Desktop/partage_linux/dump/insee/departements.metadata.json
2022-05-08T18:07:17.122+0100  reading metadata for insee.pays from /home/amal/Desktop/partage_linux/dump/insee/pays.metadata.json
2022-05-08T18:07:17.187+0100  reading metadata for insee.regions from /home/amal/Desktop/partage_linux/dump/insee/regions.metadata.json
2022-05-08T18:07:17.330+0100  reading metadata for insee.arrondissements from /home/amal/Desktop/partage_linux/dump/insee/arrondissements.metadata.json
2022-05-08T18:07:17.458+0100  reading metadata for insee.cantons from /home/amal/Desktop/partage_linux/dump/insee/cantons.metadata.json
2022-05-08T18:07:18.399+0100  reading metadata for insee.communes from /home/amal/Desktop/partage_linux/dump/insee/communes.metadata.json
2022-05-08T18:07:18.520+0100  reading metadata for insee.dep from /home/amal/Desktop/partage_linux/dump/insee/dep.metadata.json
2022-05-08T18:07:18.558+0100  restoring insee.dep from /home/amal/Desktop/partage_linux/dump/insee/dep.bson
2022-05-08T18:07:18.577+0100  restoring insee.arrondissements from /home/amal/Desktop/partage_linux/dump/insee/arrondissements.bson
2022-05-08T18:07:18.615+0100  restoring insee.communes from /home/amal/Desktop/partage_linux/dump/insee/communes.bson
```

```
hadoop2 [En fonction] - Oracle VM VirtualBox
Fichier Machine Écran Entrée Périphériques Aide
Activities Terminal 18:21 8 مای
amal@amal-VirtualBox: ~
mongosh mongodb://127.0.0.1:27017/...
amal@amal-VirtualBox: ~
terExpression:primitive.D(nil)}
2022-05-08T18:09:56.141+0100      restoring indexes for collection insee.communes
from metadata
2022-05-08T18:09:56.142+0100      index: &idx.IndexDocument{Options:primitive.M{"name": "DEP_1_COM_1", "ns": "communes.communes", "v": 2}, Key:primitive.D{primitive.E{Key: "DEP", Value: 1}, primitive.E{Key: "COM", Value: 1}}, PartialFilterExpression:primitive.D(nil)}
2022-05-08T18:09:56.142+0100      index: &idx.IndexDocument{Options:primitive.M{"2dsphereIndexVersion": 3, "name": "contour_2dsphere", "ns": "communes.communes", "v": 2}, Key:primitive.D{primitive.E{Key: "contour", Value: "2dsphere"}}, PartialFilterExpression:primitive.D(nil)}
2022-05-08T18:09:56.142+0100      index: &idx.IndexDocument{Options:primitive.M{"2dsphereIndexVersion": 3, "name": "localisation_2dsphere", "ns": "communes.communes", "v": 2}, Key:primitive.D{primitive.E{Key: "localisation", Value: "2dsphere"}}, PartialFilterExpression:primitive.D(nil)}
2022-05-08T18:10:02.665+0100      no indexes to restore for collection insee.dep
2022-05-08T18:10:02.665+0100      restoring indexes for collection insee.departements
from metadata
2022-05-08T18:10:02.665+0100      index: &idx.IndexDocument{Options:primitive.M{"2dsphereIndexVersion": 3, "name": "contour_2dsphere", "ns": "communes.departements", "v": 2}, Key:primitive.D{primitive.E{Key: "contour", Value: "2dsphere"}}, PartialFilterExpression:primitive.D(nil)}
2022-05-08T18:10:03.974+0100      no indexes to restore for collection insee.pays
2022-05-08T18:10:32.763+0100      3443985 document(s) restored successfully. 0 document(s) failed to restore.
amal@amal-VirtualBox: ~$
```

The screenshot shows a Linux desktop environment with a terminal window open. The terminal window title is "mongosh mongodb://127.0.0.1:27017/?directConnection=true...". The terminal content displays startup warnings from MongoDB and a list of databases and collections in the "insee" database.

```
-----  
The server generated these startup warnings when booting:  
2022-05-08T16:25:25.754+01:00: Using the XFS filesystem is strongly recommended with the WiredTiger storage engine. See http://dochub.mongodb.org/core/prodnotes-filesystem  
2022-05-08T16:25:25.992+01:00: Access control is not enabled for the database. Read and write access to data and configuration is unrestricted  
-----  
test> show databases  
admin 41 kB  
config 111 kB  
insee 227 MB  
local 41 kB  
test> use insee  
switched to db insee  
insee> show collections  
arrondissements  
cantons  
communes  
dep  
departements  
pays  
regions  
insee> ■
```

4/ Exporter les collections de la base INSEE en format .csv

5/ Déplacer les données sous HDFS

The screenshot shows a Linux desktop environment with a terminal window open. The terminal window title is "amal@amal-VirtualBox: ~". The terminal content shows the user mounting a shared folder from their desktop to HDFS and then listing the contents of the "/amal" directory.

```
amal@amal-VirtualBox:~$ sudo mount -t vboxsf partage_db /home/amal/Desktop/linux_partage  
amal@amal-VirtualBox:~$ hdfs dfs -put /home/amal/Desktop/linux_partage /amal  
WARNING: An illegal reflective access operation has occurred  
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.6.0.jar) to method sun.security.krb5.Config.newInstance()  
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil  
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations  
WARNING: All illegal access operations will be denied in a future release  
22/05/11 16:04:57 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
amal@amal-VirtualBox:~$ hdfs dfs -ls /amal  
WARNING: An illegal reflective access operation has occurred  
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.6.0.jar) to method sun.security.krb5.Config.newInstance()  
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil  
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations  
WARNING: All illegal access operations will be denied in a future release  
22/05/11 16:05:26 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
Found 8 items
```

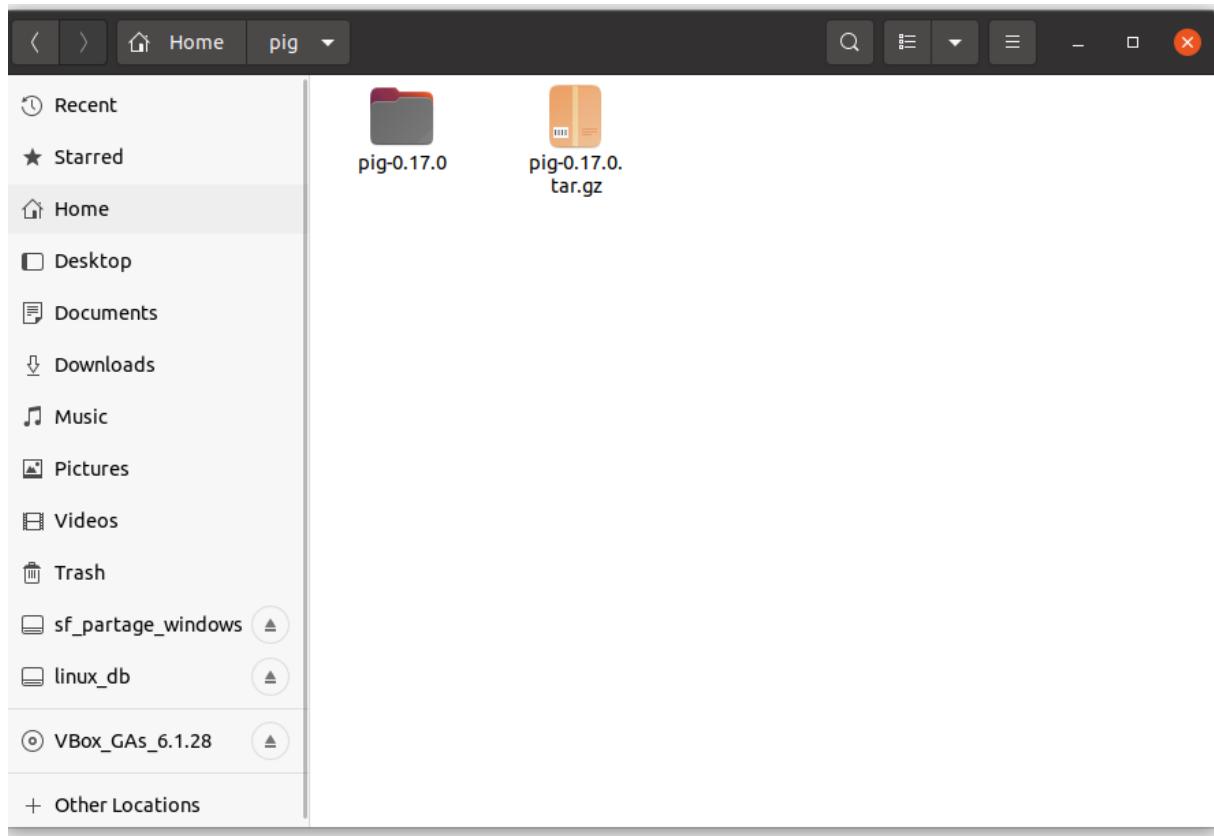
hadoop3 [En fonction] - Oracle VM VirtualBox

Fichier Machine Écran Entrée Périphériques Aide

Activities Terminal amal@amal-VirtualBox: ~ 16:09 11 ماء amal@amal-VirtualBox: ~/pig

```
22/05/11 16:04:57 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
amal@amal-VirtualBox:~$ hdfs dfs -ls /amal
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.6.0.jar) to method sun.security.krb5.Config.newInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
22/05/11 16:05:26 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 8 items
-rw-r--r-- 1 amal supergroup 20293905 2022-05-11 16:05 /amal/arrondissement.s.csv
-rw-r--r-- 1 amal supergroup 8380599 2022-05-11 16:05 /amal/cantons.csv
-rw-r--r-- 1 amal supergroup 102521893 2022-05-11 16:05 /amal/communes.csv
-rw-r--r-- 1 amal supergroup 164055671 2022-05-11 16:05 /amal/dep.csv
-rw-r--r-- 1 amal supergroup 14562792 2022-05-11 16:05 /amal/departements.csv
-rw-r--r-- 1 amal supergroup 25598 2022-05-11 16:05 /amal/pays.csv
-rw-r--r-- 1 amal supergroup 494 2022-05-11 16:05 /amal/plantes.csv
-rw-r--r-- 1 amal supergroup 6059116 2022-05-11 16:05 /amal/regions.csv
amal@amal-VirtualBox:~$
```

6/ Installation Apache Pig sous la VM :



Activities Terminal 10:42 13 مای amal@amal-VirtualBox: ~

```
2022-05-13 10:41:59,155 [main] INFO org.apache.pig.Main - Apache Pig version 0.17.0 (r1797386) compiled Jun 02 2017, 15:41:58
2022-05-13 10:41:59,156 [main] INFO org.apache.pig.Main - Logging error messages to: /home/amal/pig_1652434919155.log
2022-05-13 10:41:59,249 [main] INFO org.apache.pig.impl.util.Utils - Default bootup file /home/amal/.pigbootup not found
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.6.0.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
2022-05-13 10:42:00,086 [main] WARN org.apache.hadoop.util.NativeCodeLoader - Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2022-05-13 10:42:00,098 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
2022-05-13 10:42:00,098 [main] INFO org.apache.pig.backend.hadoop.executionengine.HExecutionEngine - Connecting to hadoop file system at: hdfs://localhost:9000
2022-05-13 10:42:00,587 [main] INFO org.apache.pig.PigServer - Pig Script ID for the session: PIG-default-d6311a12-a81f-401e-b13c-d01c4ddb7fb7
2022-05-13 10:42:00,592 [main] WARN org.apache.pig.PigServer - ATS is disabled since yarn.timeline-service.enabled set to false
grunt> pig
```

7/ Charger les données d'un fichier HDFS sous Pig sans schéma :

```
Activities Terminal 18:32 11 مای amal@amal-VirtualBox: ~
amal@amal-VirtualB... amal@amal-VirtualB... amal@amal-VirtualB...
.....,58b83d0c8aed46744b17e4d,58b84
21185aed46744b21437)
grunt> departements = LOAD '/amal/departements.csv' USING PigStorage(',');
```

```

Activities Terminal 18:45 11 مای amal@amal-VirtualBox: ~
amal@amal-VirtualB... amal@amal-VirtualB... amal@amal-VirtualB...
5148,45.287173396448,-12.799517581784,45.287100676524,-12.799580960285,45.28696
5389485,-12.79972307885,45.286906512107,-12.799805432003,45.286802573328,-12.79
9965601857,45.285633634881,-12.801902020471,45.284470493494,-12.804915419317,45
.284272213305,-12.80805075457,45.28423014486,-12.809156728048,45.284225785915,-
12.809378280066,45.284290811373,-12.810667705586,45.28431042649,-12.81090370008
8,45.284338263757,-12.811076386481,45.284378993315,-12.811241824748,45.28443627
869,-12.811380116883,45.284517557485,-12.811550032686,45.284679120852,-12.81182
4757513,45.286390753641,-12.814702110691,45.286061222587,-12.815832811766,45.28
488044409,-12.815501319473,45.280957062192,-12.813225860166,45.280308783502,-12
.81278255231,45.280175068812,-12.812689554341,45.2800404008,-12.812567620521,45
.279951817861,-12.812456488927,45.279857686284,-12.812327277701,45.279693403271
,-12.812088723713,45.279317736331,-12.811514003678,45.279274321337,-12.81141729
2312,45.279217956216,-12.811275380923,45.279181905971,-12.811172331804,45.27901
4650199,-12.809905619779,45.279095516828,-12.808867426265,45.27912217998,-12.80
7976686967,45.279082746084,-12.807310277919,45.279075266115,-12.807211719843,45
Rhythmbox 08,-12.807106833799,45.278673252047,-12.805292368954,45.278050685692
,868231,45.276652922938,-12.801412443065,45.276177989429,-12.80104309
2473,45.275711386395,-12.800707190481,45.275335170643,-12.800451673763,45.27508
3442737,-12.800284645612,45.274354115315,-12.799834172207,45.272885179719,-12.7
97076750529,45.271367318378,-12.793340942459,45.271347010101,-12.793305696569
,,)
grunt> communes = LOAD '/amal/communes.csv' USING PigStorage(',');

```

8/ Traitement des données sous Pig :

- Départements :
- Générer une transformation dans l'affichage (dans ce cas on va afficher une seule colonne REGION (\$1) à partir du variable existante départements).

```

Activities Terminal 19:44 11 مای amal@amal-VirtualBox: ~
amal@amal-VirtualB... amal@amal-VirtualB... amal@amal-VirtualB...
(28)
(11)
(75)
(32)
(76)
(76)
(93)
(93)
(52)
(75)
(75)
(44)
(27)
(27)
(11)
(11)
(11)
(11)
(01)
(02)
(03)
(04)
(06)
grunt> C = FOREACH departements GENERATE ( $1 ) ;

```

- Afficher les données dont l'attribut REGION (\$1) > 11 => condition

The screenshot shows a Linux desktop environment with a dark theme. On the left, there's a vertical dock containing icons for various applications: a browser (Firefox), a terminal (Terminal), a code editor (Sublime Text), a file manager (Nautilus), a system monitor (System Monitor), a file viewer (File Viewer), and a terminal (Terminal). The main area features a terminal window titled "amal@amal-VirtualBox: ~". The terminal is displaying a command-line interface with the following text:

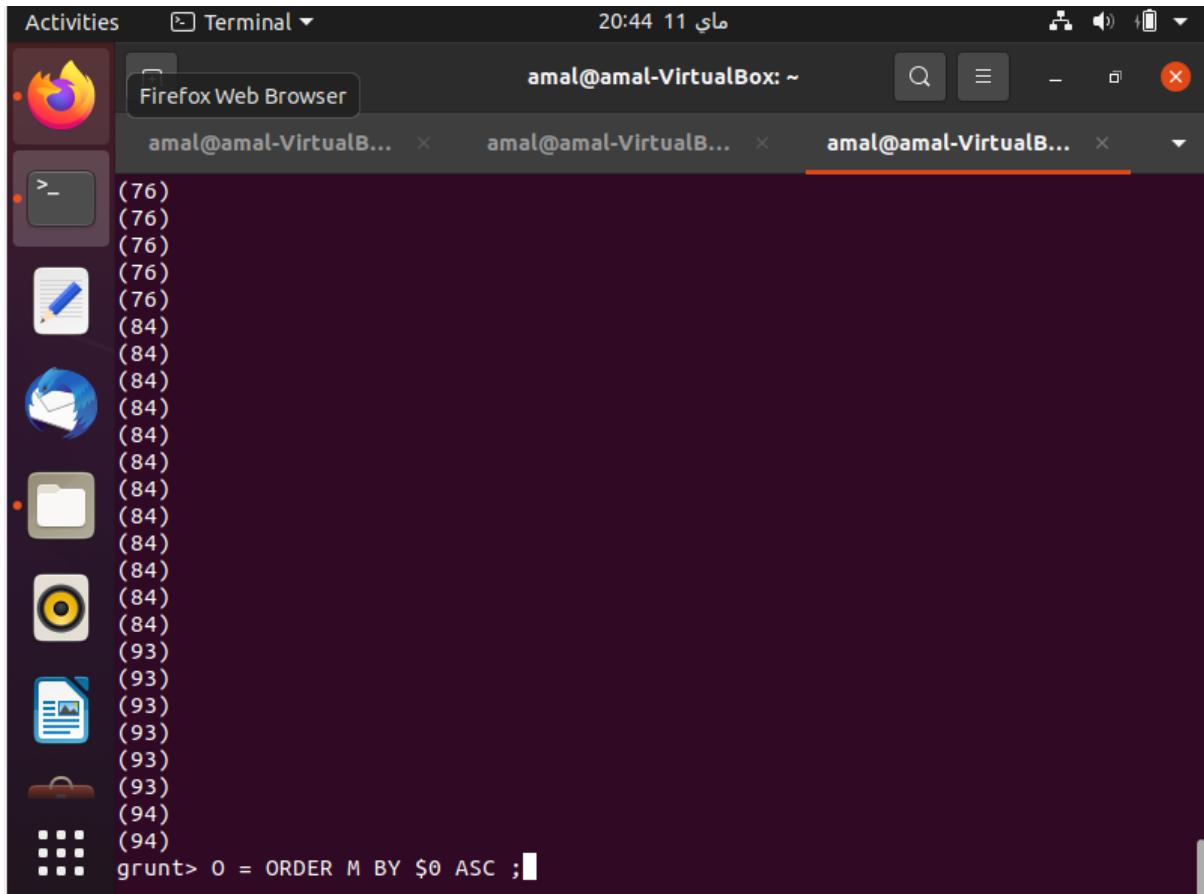
```
.....  
.....,58b83d0c85aed46744b17e50,58b8421085aed46744b2116b)  
grunt> N = filter departements by ( $1 > 11 ) ;
```

- D'après l'affichage donnée à partir de la condition du filtre, on génère un nouveau affichage dont l'attribut choisi est REGION (\$1) .

The screenshot shows a Linux desktop interface with a terminal window open. The terminal title is "amal@amal-VirtualBox: ~". The output of the command "grunt> M = FOREACH N GENERATE (\$1) ;" is displayed, showing the following results:

- (84)
- (75)
- (76)
- (76)
- (44)
- (44)
- (84)
- (27)
- (27)
- (52)
- (84)
- (84)
- (28)
- (75)
- (32)
- (76)
- (76)
- (93)
- (93)
- (52)
- (75)
- (75)
- (44)
- (27)
- (27)

- Trier la première colonne dans un ordre croissant à partir du résultat donné du FOREACH .

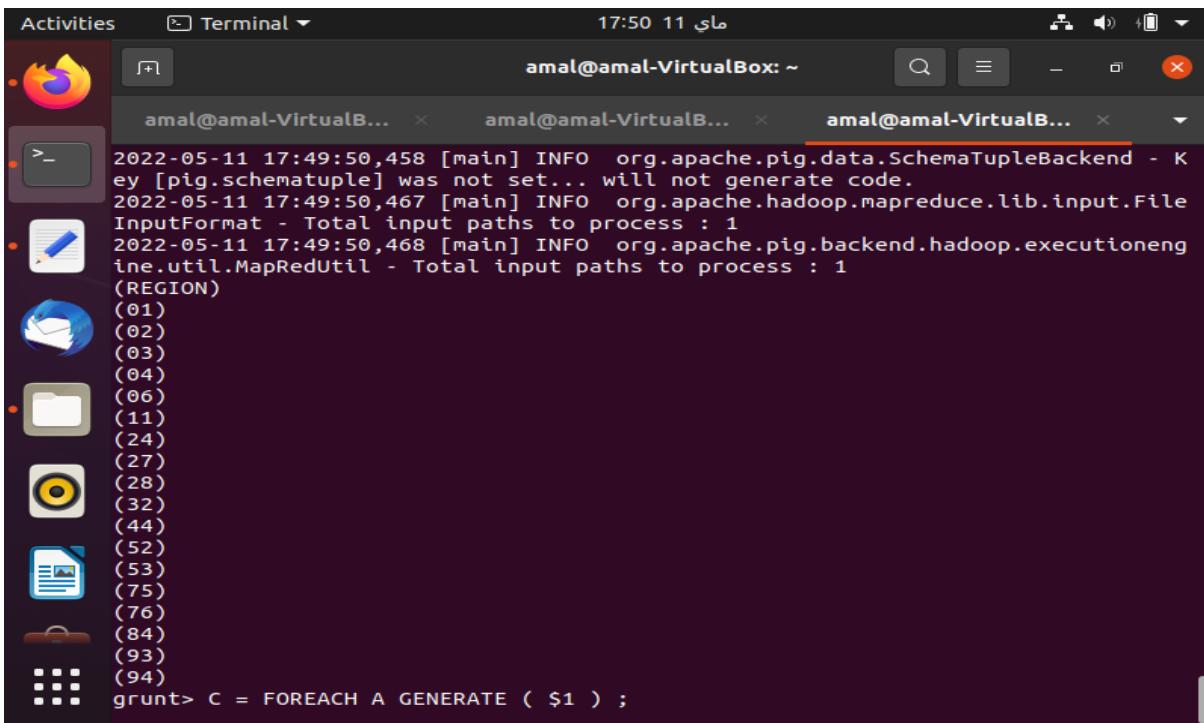


A screenshot of a Linux desktop environment showing a terminal window. The terminal window title is "Terminal" and the session name is "amal@amal-VirtualBox: ~". The terminal content shows the following command being run:

```
grunt> 0 = ORDER M BY $0 ASC ;
```

The terminal window is part of a desktop interface with other application icons visible in the dock.

- Regions :
- Générer une transformation dans l'affichage (dans ce cas on va afficher une seule colonne REGION (\$1) à partir du variable existante région).



A screenshot of a Linux desktop environment showing a terminal window. The terminal window title is "Terminal" and the session name is "amal@amal-VirtualBox: ~". The terminal content shows the following command being run:

```
2022-05-11 17:49:50,458 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig.schematuple] was not set... will not generate code.  
2022-05-11 17:49:50,467 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1  
2022-05-11 17:49:50,468 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1  
(REGION)  
(01)  
(02)  
(03)  
(04)  
(06)  
(11)  
(24)  
(27)  
(28)  
(32)  
(44)  
(52)  
(53)  
(75)  
(76)  
(84)  
(93)  
(94)  
grunt> C = FOREACH A GENERATE ( $1 ) ;
```

The terminal window is part of a desktop interface with other application icons visible in the dock.

- Afficher les données dont l'attribut REGION (\$1) == 11 => condition

```

Activities Terminal 17:43 11 ماء amal@amal-VirtualBox: ~
amal@amal-VirtualBox: ~ x amal@amal-VirtualBox: ~ x amal@amal-VirtualBox: ~ Capture d'écran
,,,,,,,
grunt> B = FILTER A BY ($1 == 11);
2022-05-11 17:32:59,214 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_CHARARRAY 2 time(s).
2022-05-11 17:32:59,219 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_INT 1 time(s).
grunt> dump B;
2022-05-11 17:33:03,937 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_INT 1 time(s).
2022-05-11 17:33:03,950 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: FILTER
2022-05-11 17:33:03,983 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2022-05-11 17:33:03,984 [main] INFO org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer - {RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, NestedLimitOptimizer, PartitionFilterOptimizer, PredicatePushdownOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2022-05-11 17:33:03,989 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File concatenation threshold: 100 optimistic? false
2022-05-11 17:33:03,989 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR plan size before optimization: 1
2022-05-11 17:33:03,990 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR plan size after optimization: 1
2022-05-11 17:33:04,007 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Co

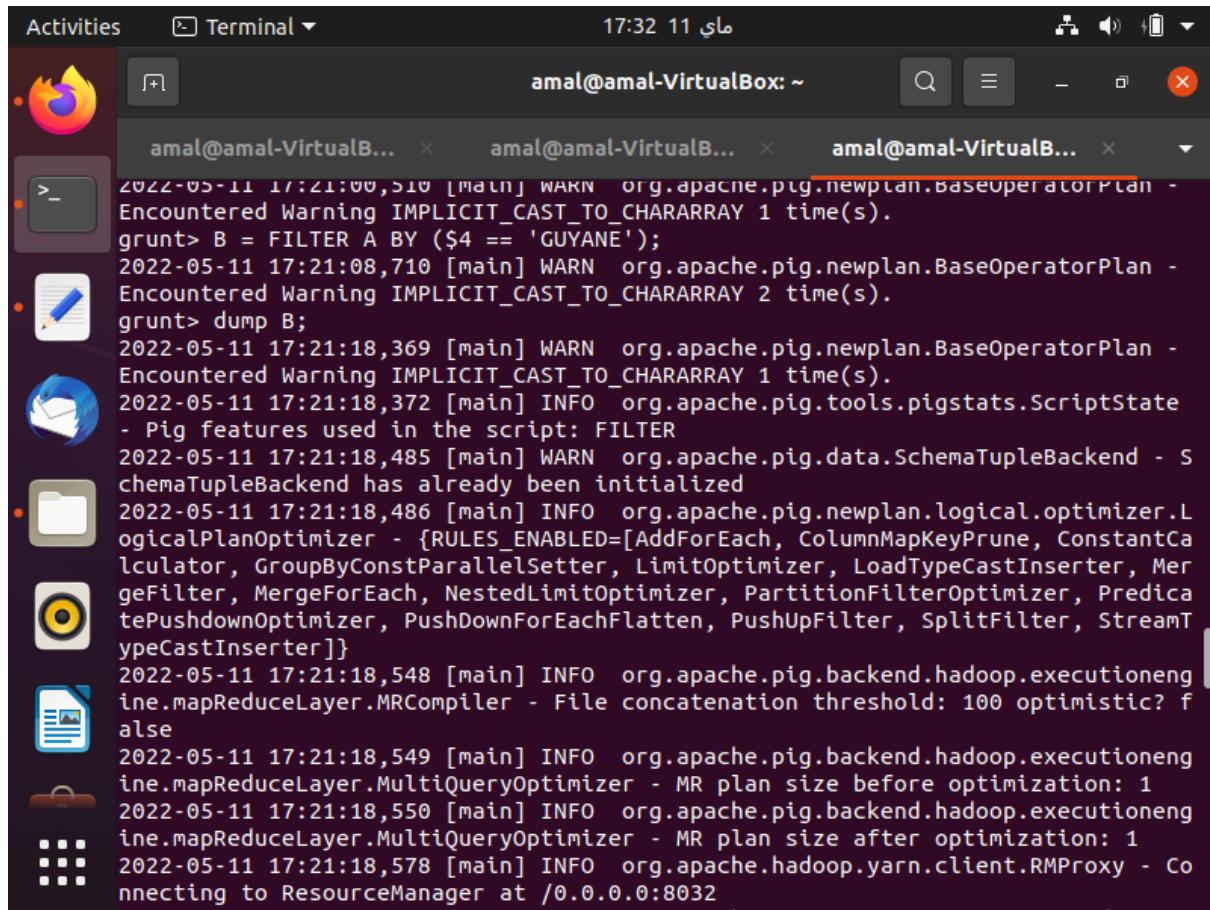
```

```

Activities Terminal 17:45 11 ماء amal@amal-VirtualBox: ~
amal@amal-VirtualBox: ~ x amal@amal-VirtualBox: ~ x amal@amal-VirtualBox: ~ Capture d'écran
connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 7 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISCONDS)
2022-05-11 17:35:50,673 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 8 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISCONDS)
2022-05-11 17:35:51,673 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 9 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISCONDS)
2022-05-11 17:35:51,775 [main] WARN org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Unable to retrieve job to compute warning aggregation.
2022-05-11 17:35:51,775 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2022-05-11 17:35:51,776 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2022-05-11 17:35:51,794 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2022-05-11 17:35:51,794 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(58b83d0c85aed46744b17e4e,11,75056,1,ILE-DE-FRANCE,Île-de-France,25,155,1278,12027565,12179602,17.9,28.6,15.8,11.6,11.9,15.9,24.4,15.0,11.5,10.5,19.4,31.2,16.6,11.7,12.2,10.4,17.7,21.8,10.7,18.8,24.7,Polygon,,,,,,,,,,,

```

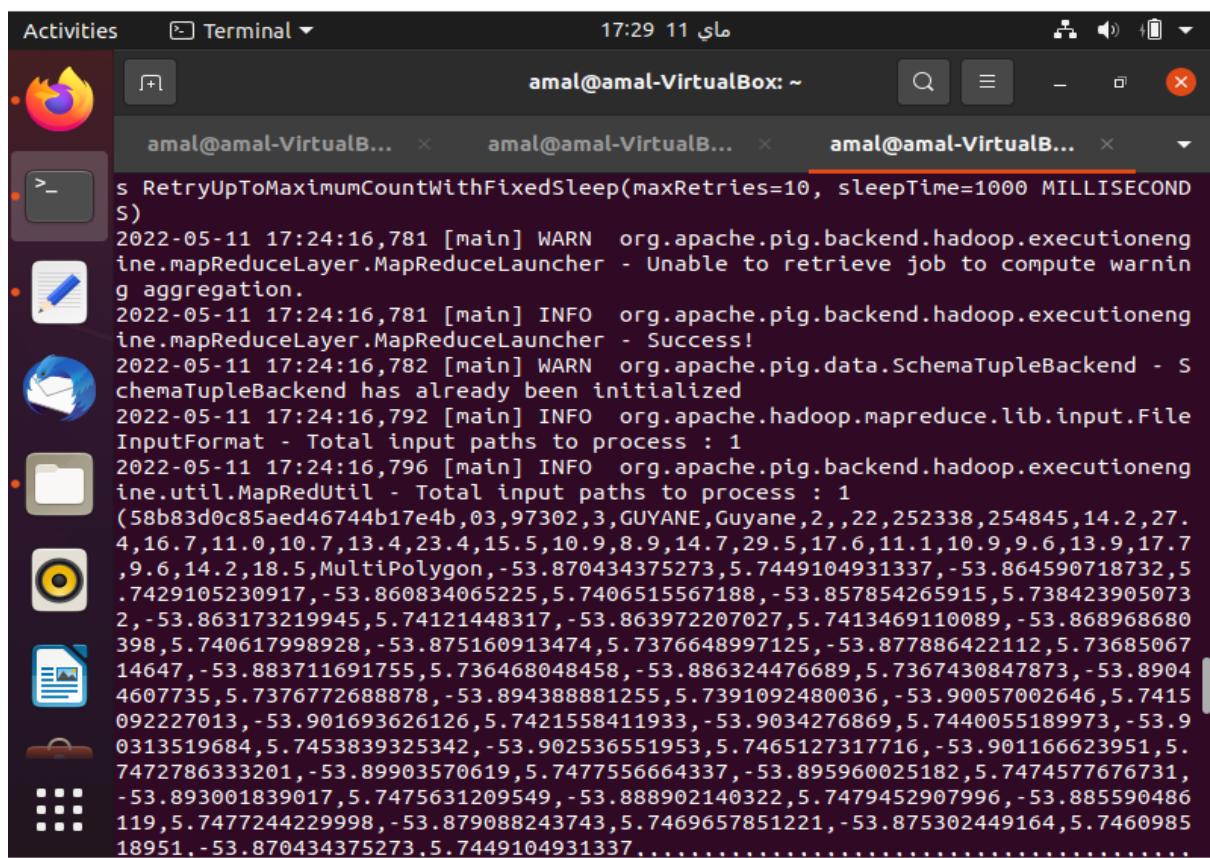
- Afficher les données dont l'attribut NCC (\$4) == 'GUYANE' => condition



```

Activities Terminal 17:32 11 ماء amal@amal-VirtualBox: ~
amal@amal-VirtualB... x amal@amal-VirtualB... x amal@amal-VirtualB... x
2022-05-11 17:21:00,510 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_CHARARRAY 1 time(s).
grunt> B = FILTER A BY ($4 == 'GUYANE');
2022-05-11 17:21:08,710 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_CHARARRAY 2 time(s).
grunt> dump B;
2022-05-11 17:21:18,369 [main] WARN org.apache.pig.newplan.BaseOperatorPlan - Encountered Warning IMPLICIT_CAST_TO_CHARARRAY 1 time(s).
2022-05-11 17:21:18,372 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: FILTER
2022-05-11 17:21:18,485 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2022-05-11 17:21:18,486 [main] INFO org.apache.pig.newplan.logical.optimizer.LogicalPlanOptimizer - {RULES_ENABLED=[AddForEach, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, MergeFilter, MergeForEach, NestedLimitOptimizer, PartitionFilterOptimizer, PredicatePushdownOptimizer, PushDownForEachFlatten, PushUpFilter, SplitFilter, StreamTypeCastInserter]}
2022-05-11 17:21:18,548 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MRCompiler - File concatenation threshold: 100 optimistic? false
2022-05-11 17:21:18,549 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR plan size before optimization: 1
2022-05-11 17:21:18,550 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MultiQueryOptimizer - MR plan size after optimization: 1
2022-05-11 17:21:18,578 [main] INFO org.apache.hadoop.yarn.client.RMProxy - Connecting to ResourceManager at /0.0.0.0:8032

```



```

Activities Terminal 17:29 11 ماء amal@amal-VirtualBox: ~
amal@amal-VirtualB... x amal@amal-VirtualB... x amal@amal-VirtualB... x
s RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISCONDS)
2022-05-11 17:24:16,781 [main] WARN org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Unable to retrieve job to compute warning aggregation.
2022-05-11 17:24:16,781 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2022-05-11 17:24:16,782 [main] WARN org.apache.pig.data.SchemaTupleBackend - SchemaTupleBackend has already been initialized
2022-05-11 17:24:16,792 [main] INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat - Total input paths to process : 1
2022-05-11 17:24:16,796 [main] INFO org.apache.pig.backend.hadoop.executionengine.util.MapRedUtil - Total input paths to process : 1
(58b83d0c85aed46744b17e4b,03,97302,3,GUYANE,Guyane,,2,,22,252338,254845,14.2,27.4,16.7,11.0,10.7,13.4,23.4,15.5,10.9,8.9,14.7,29.5,17.6,11.1,10.9,9.6,13.9,17.7,9.6,14.2,18.5,MultiPolygon,-53.870434375273,5.7449104931337,-53.864590718732,5.7429105230917,-53.860834065225,5.7406515567188,-53.857854265915,5.7384239050732,-53.863173219945,5.74121448317,-53.863972207027,5.7413469110089,-53.868968680398,5.740617998928,-53.875160913474,5.7376648997125,-53.87786422112,5.7368506714647,-53.883711691755,5.736468048458,-53.886324476689,5.7367430847873,-53.89044607735,5.7376772688878,-53.894388881255,5.7391092480036,-53.90057002646,5.7415092227013,-53.901693626126,5.7421558411933,-53.9034276869,5.7440055189973,-53.90313519684,5.7453839325342,-53.902536551953,5.7465127317716,-53.901166623951,5.7472786333201,-53.89903570619,5.7477556664337,-53.895960025182,5.7474577676731,-53.893001839017,5.7475631209549,-53.888902140322,5.7479452907996,-53.885590486119,5.7477244229998,-53.879088243743,5.7469657851221,-53.875302449164,5.746098518951,-53.870434375273,5.7449104931337.....
```

- Regrouper les données par REGION (\$1)

```
Activities Terminal 17:59 11 ماي amal@amal-VirtualBox: ~ amal@amal-VirtualBox: ~ amal@amal-VirtualBox: ~ Util. Capture d'écran  
+  
amal@amal-VirtualB... x amal@amal-VirtualB... x amal@amal-VirtualB...  
nates.11.0.52.0,contour.coordinates.11.0.52.1,contour.coordinates.11.0.53.0,con-  
tour.coordinates.11.0.53.1,contour.coordinates.11.0.54.0,contour.coordinates.11.  
.0.54.1,contour.coordinates.11.0.55.0,contour.coordinates.11.0.55.1,contour.co-  
ordinates.11.0.56.0,contour.coordinates.11.0.56.1,contour.coordinates.11.0.57.0,  
contour.coordinates.11.0.57.1,contour.coordinates.11.0.58.0,contour.coordinates.  
.11.0.58.1,contour.coordinates.12.0.0.0,contour.coordinates.12.0.0.1,contour.co-  
ordinates.12.0.1.0,contour.coordinates.12.0.1.1,contour.coordinates.12.0.2.0,co-  
ntour.coordinates.12.0.2.1,contour.coordinates.12.0.3.0,contour.coordinates.12.  
0.3.1,contour.coordinates.12.0.4.0,contour.coordinates.12.0.4.1,contour.co-  
ordinates.12.0.5.0,contour.coordinates.12.0.5.1,contour.coordinates.12.0.6.0,contour  
.coordinates.12.0.6.1,contour.coordinates.12.0.7.0,contour.coordinates.12.0.7.1  
,contour.coordinates.12.0.8.0,contour.coordinates.12.0.8.1,contour.coordinates.  
.12.0.9.0,contour.coordinates.12.0.9.1,contour.coordinates.12.0.10.0,contour.co-  
ordinates.12.0.10.1,contour.coordinates.12.0.11.0,contour.coordinates.12.0.11.1,  
contour.coordinates.12.0.12.0,contour.coordinates.12.0.12.1,contour.coordinates.  
.12.0.13.0,contour.coordinates.12.0.13.1,contour.coordinates.12.0.14.0,contour.  
coordinates.12.0.14.1,contour.coordinates.12.0.15.0,contour.coordinates.12.0.15  
.1,contour.coordinates.12.0.16.0,contour.coordinates.12.0.16.1,contour.co-  
ordinates.12.0.17.0,contour.coordinates.12.0.17.1,contour.coordinates.12.0.18.0,co-  
ntur.coordinates.12.0.18.1,contour.coordinates.12.0.19.0,contour.coordinates.12.0  
.19.1,contour.coordinates.12.0.20.0,contour.coordinates.12.0.20.1,contour.co-  
ordinates.12.0.21.0,contour.coordinates.12.0.21.1,contour.coordinates.12.0.22.0,co-  
ntour.coordinates.12.0.22.1,contour.coordinates.12.0.23.0,contour.coordinates.12.  
0.23.1,contour.coordinates.12.0.24.0,contour.coordinates.12.0.24.1,id_pays,id  
_cheflieu}))  
grunt> F = GROUP A BY ($1) ;
```

- Communes
 - Générer une transformation dans l'affichage (dans ce cas on va afficher une seule colonne DEP (\$4) à partir du variable existante communes).

- Regrouper les données par CDC (\$1)

```

Activities Terminal 19:24 11 مای amal@amal-VirtualBox: ~
amal@amal-VirtualB... amal@amal-VirtualB... amal@amal-VirtualB...
5.7,10.9,11.9,15.5,22.9,14.1,10.7,12.7,19.8,28.1,17.1,11.4,11.6,9.9,17.6,20.7,1
0.2,19.0,24.1),(58b8420a85aed46744b1f483,,0,84,69,044,1,,0,,CHARBONNIERES-LES-B
AINS,,Charbonnières-les-Bains,4988,5146,658,1,21,59,500,93,77,128,37,4,2193,204
7,35,111,1452,2924,190,2073,2057,5155.5,33312.0,80.5,,,,,,,73.5,66.7,6.8,28.
4,22.9,1.8,1.1,0.3,0.4,-26.6,69260,3875,460,11.87,3415,88.13,118,3.05,3.46,3297
,85,08,96.54,2466,831,21.45,25.2,63.64,74.8,3864,472,12.22,3392,87.78,39,1.01,1
.15,3353,86.78,98.85,80,2.07,2.39,310,8.02,9.25,1890,137,3.55,4.09,13,0.34,0.39
,2,0.05,0.06,10,0.26,0.3,352,9.11,10.5,61,1.58,1.82,498,12.89,14.85,48.91,56.37
,Polygon,4.7268163204804,45.776866805752,4.7278098682301,45.777519918224,4.7291
471180432,45.7778707278011,4.7292853708944,45.778929434821,4.7283451595197,45.77
9488767448,4.7292323997338,45.783611212889,4.7319689702072,45.787377429364,4.73
18471237961,45.787515306024,4.7332119120229,45.787846529222,4.7431764188504,45.
790648275078,4.7525641049488,45.790537107674,4.7531927772012,45.788723217239,4.
7551305841346,45.783055538406,4.7556737762537,45.777868894441,4.7579523149683,4
5.77439449597,4.7607739380817,45.771007948911,4.7612101999448,45.770683187324,4
.760757168063,45.770602889826,4.7491471544029,45.771108653124,4.7392212320088,4
5.771516112584,4.734633702732,45.773594417155,4.7277646788328,45.776486600999,4
.7268163204804,45.776866805752,,,,,,,,,,,
,,,,,,,,,,,
,,,,,,,,,,,
,,,,,,,,,,,
,,,,,,,,,,,
,,,,,,,,,,,
Point,4.73333,45.7833,211.0,295.0,4.13,58b
83d0c85aed46744b17e58,58b83dbb85aed46744b17ea0,58b83eca85aed46744b17fb2,,24.5,3
3.6,23.2,11.8,14.6,16.9,25.3,14.4,11.6,12.2,30.3,37.4,34.0,12.5,15.1,10.0,24.1,
28.9,10.5,30.1,35.0})
grunt> F = GROUP communes BY ($1) ;

```

8/ Enregistrer la résultat dans un fichier

```

Activities Terminal 17:42 11 مای amal@amal-VirtualBox: ~
FireFox Web Browser amal@amal-VirtualB... amal@amal-VirtualB... amal@amal-VirtualB...
2022-05-11 17:41:36,225 [main] INFO org.apache.hadoop.ipc.Client - Retrying co
nnect to server: 0.0.0.0/0.0.0.0:10020. Already tried 5 time(s); retry policy i
s RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECOND
S)
2022-05-11 17:41:37,226 [main] INFO org.apache.hadoop.ipc.Client - Retrying co
nnect to server: 0.0.0.0/0.0.0.0:10020. Already tried 6 time(s); retry policy i
s RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECOND
S)
2022-05-11 17:41:38,227 [main] INFO org.apache.hadoop.ipc.Client - Retrying co
nnect to server: 0.0.0.0/0.0.0.0:10020. Already tried 7 time(s); retry policy i
s RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECOND
S)
2022-05-11 17:41:39,228 [main] INFO org.apache.hadoop.ipc.Client - Retrying co
nnect to server: 0.0.0.0/0.0.0.0:10020. Already tried 8 time(s); retry policy i
s RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECOND
S)
2022-05-11 17:41:40,229 [main] INFO org.apache.hadoop.ipc.Client - Retrying co
nnect to server: 0.0.0.0/0.0.0.0:10020. Already tried 9 time(s); retry policy i
s RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECOND
S)
2022-05-11 17:41:40,331 [main] WARN org.apache.pig.backend.hadoop.executioneng
ine.mapReduceLayer.MapReduceLauncher - Unable to retrieve job to compute warnin
g aggregation.
2022-05-11 17:41:40,331 [main] INFO org.apache.pig.backend.hadoop.executioneng
ine.mapReduceLayer.MapReduceLauncher - Success!
grunt> store B into '/resultfiltreregion/';

```

Activities Firefox Web Browser 20:12 12 آگو

The screenshot shows a Linux desktop interface with a dock on the left containing various application icons. A Firefox browser window is open, displaying a table of file metadata. The table has columns for Permission, Owner, Group, Size, Replication, Block Size, and Name. All files listed are owned by 'amal' and belong to the 'supergroup'. The 'Size' column shows '0 B' for most files, except for one which is '10.89 MB'. The 'Replication' column shows values of 0 or 1. The 'Block Size' column shows '128 MB' for all files. The 'Name' column lists several files, including 'RSTFOREACHFILTREDEPART', 'amal', 'region', 'resultatFILTERDEPART', 'resultatdumpcommunes', 'resultatdumpdepartements', 'resultatfiltrecommunes', 'resultatfiltregyrane', 'resultatforeachcommunes', 'resultatforeachdep', and 'resultatforeachregion'.

Permission	Owner	Group	Size	Replication	Block Size	Name
drwxr-xr-x	amal	supergroup	0 B	0	0 B	RSTFOREACHFILTREDEPART
drwxr-xr-x	amal	supergroup	0 B	0	0 B	amal
drwxr-xr-x	amal	supergroup	0 B	0	0 B	region
drwxr-xr-x	amal	supergroup	0 B	0	0 B	resultatFILTERDEPART
drwxr-xr-x	amal	supergroup	0 B	0	0 B	resultatdumpcommunes
drwxr-xr-x	amal	supergroup	0 B	0	0 B	resultatdumpdepartements
drwxr-xr-x	amal	supergroup	0 B	0	0 B	resultatfiltrecommunes
drwxr-xr-x	amal	supergroup	0 B	0	0 B	resultatfiltregyrane
drwxr-xr-x	amal	supergroup	0 B	0	0 B	resultatforeachcommunes
drwxr-xr-x	amal	supergroup	0 B	0	0 B	resultatforeachdep
drwxr-xr-x	amal	supergroup	0 B	0	0 B	resultatforeachregion

Activities Firefox Web Browser 20:15 12 آگو

The screenshot shows a Linux desktop interface with a dock on the left containing various application icons. A Firefox browser window is open, displaying a 'Startup Progress' bar at the top. Below it, a section titled 'Utilities' is visible. The main content area is titled 'Browse Directory' and shows a table of files in the '/resultatFILTERDEPART' directory. The table has columns for Permission, Owner, Group, Size, Replication, Block Size, and Name. The first file is a symbolic link named '_SUCCESS' with size '0 B' and replication '1'. The second file is a regular file named 'part-m-00000' with size '10.89 MB' and replication '1'. The 'Block Size' column shows '128 MB' for both files. At the bottom of the page, there is a note: 'Hadoop, 2014.'

Startup Progress

Utilities

Browse Directory

/resultatFILTERDEPART

Permission	Owner	Group	Size	Replication	Block Size	Name
-rw-r--r--	amal	supergroup	0 B	1	128 MB	_SUCCESS
-rw-r--r--	amal	supergroup	10.89 MB	1	128 MB	part-m-00000

Hadoop, 2014.

Activities Firefox Web Browser 20:55 11 مای

PigSt All App X Browsin debu apad Apac > + - □ ×

localhost:8088/cluster/apps

All Applications

Containers	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioned Nodes	Last Node
0 B	8 GB	0 B	0	8	0	0	1	0	0

Search:

Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus
amal:DefaultJobName	MAPREDUCE	default	Wed, 11 May 2022 19:48:37 GMT	Wed, 11 May 2022 19:48:52 GMT	FINISHED	SUCCEEDED
amal:DefaultJobName	MAPREDUCE	default	Wed, 11 May 2022 19:46:45 GMT	Wed, 11 May 2022 19:46:59 GMT	FINISHED	SUCCEEDED
amal:DefaultJobName	MAPREDUCE	default	Wed, 11 May 2022 19:44:58	Wed, 11 May 2022 19:45:09	FINISHED	SUCCEEDED

Activities Firefox Web Browser 20:56 11 مای

PigSt All App X Browsin debu apad Apac > + - □ ×

localhost:8088/cluster/apps

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total
33	0	0	33	0	0 B	8 GB

Show 20 entries

ID	User	Name	App Type
application_1652281007914_0033	amal	PigLatin:DefaultJobName	MAPR
application_1652281007914_0032	amal	PigLatin:DefaultJobName	MAPR
application_1652281007914_0031	amal	PigLatin:DefaultJobName	MAPR
application_1652281007914_0030	amal	PigLatin:DefaultJobName	MAPR

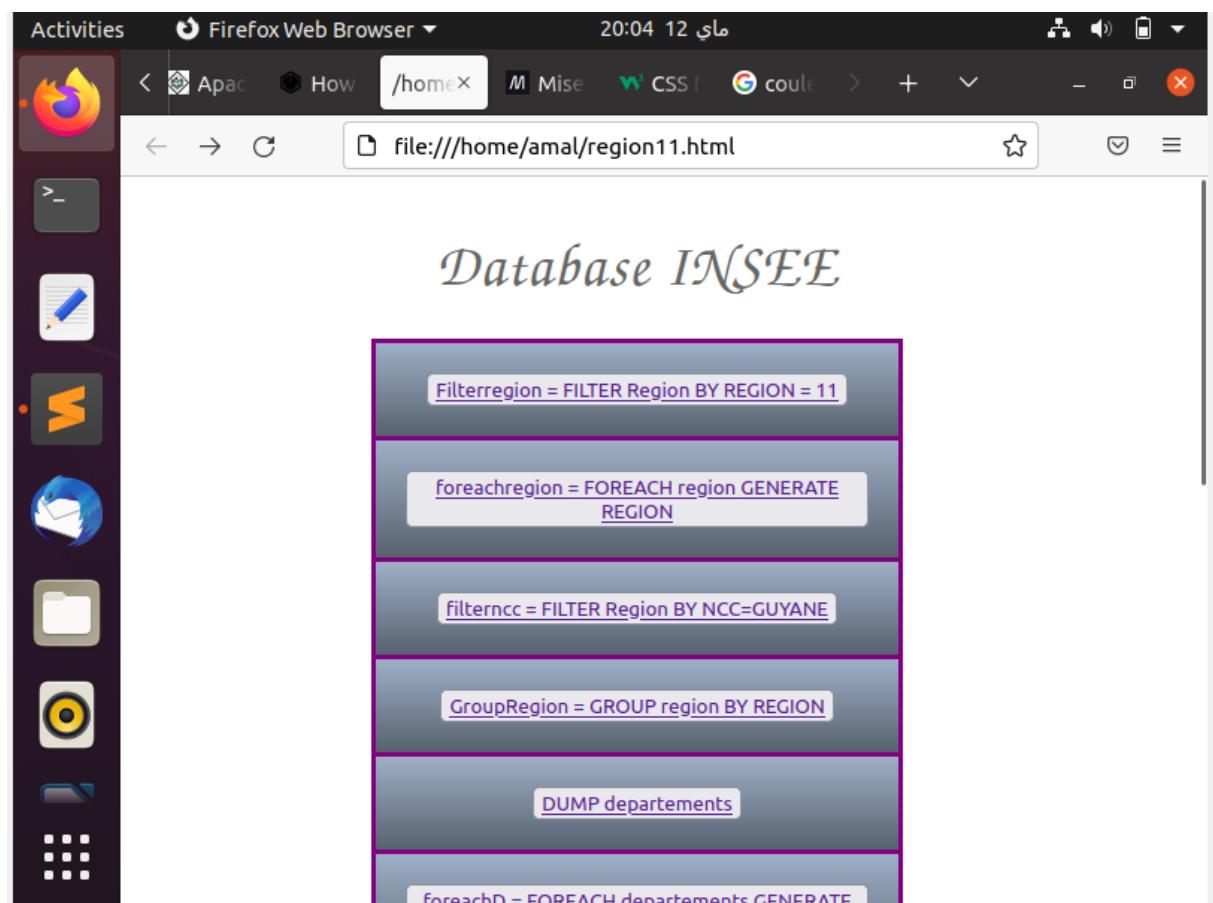
9/ Interface

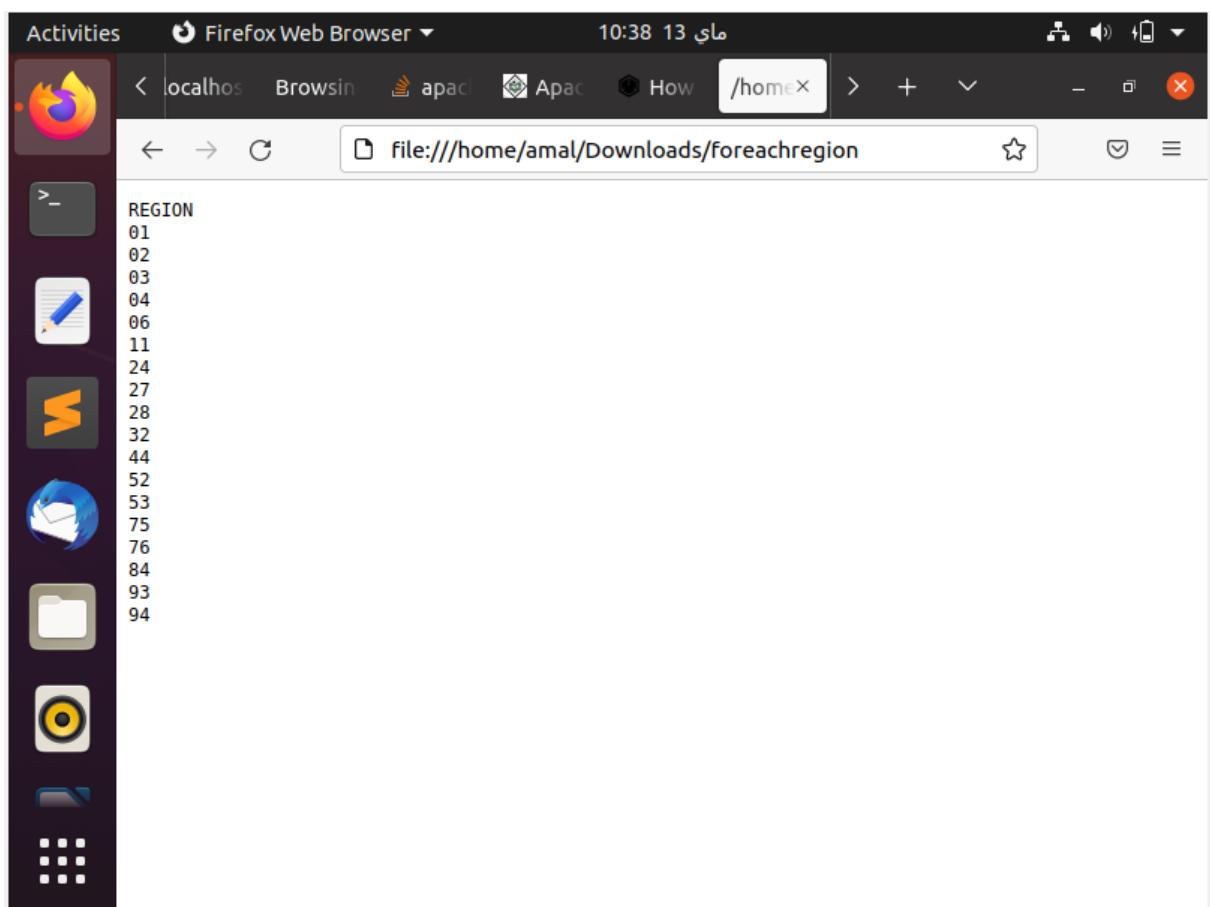
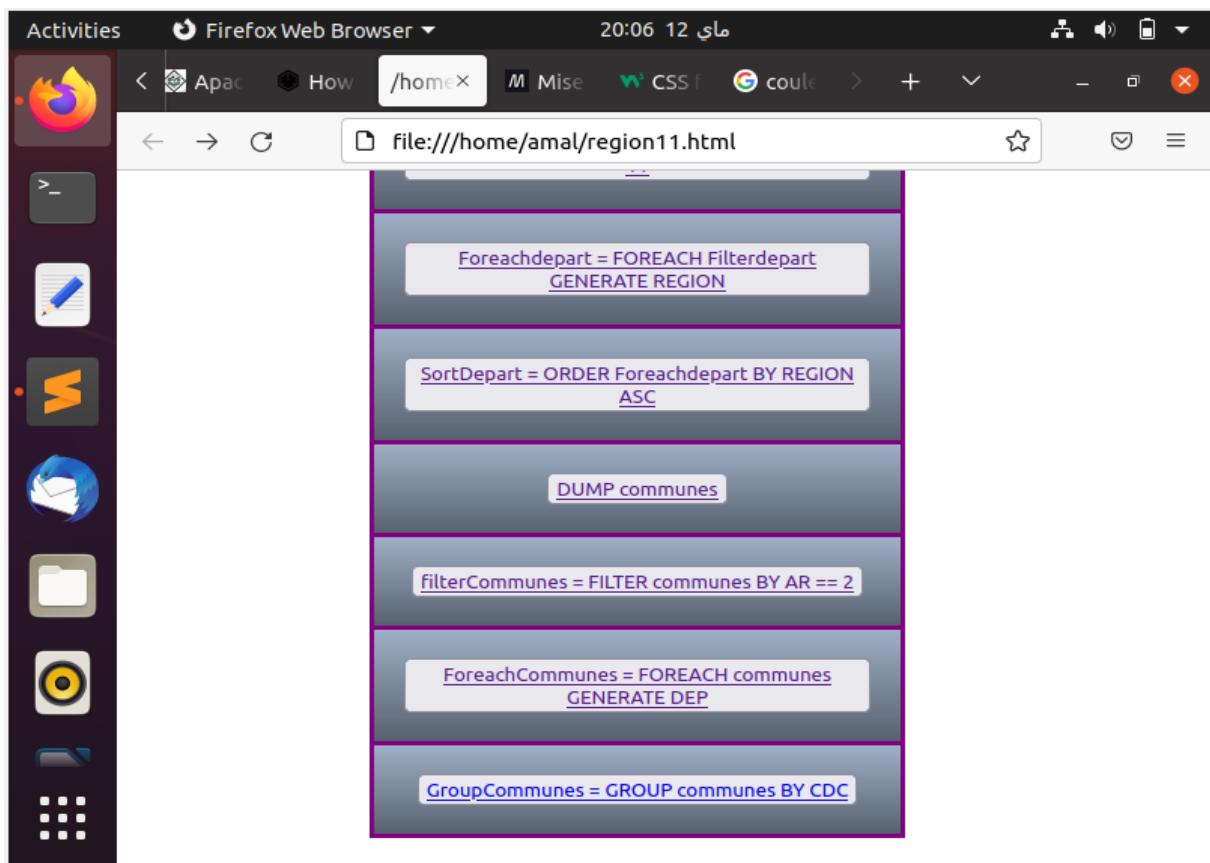
- Code :

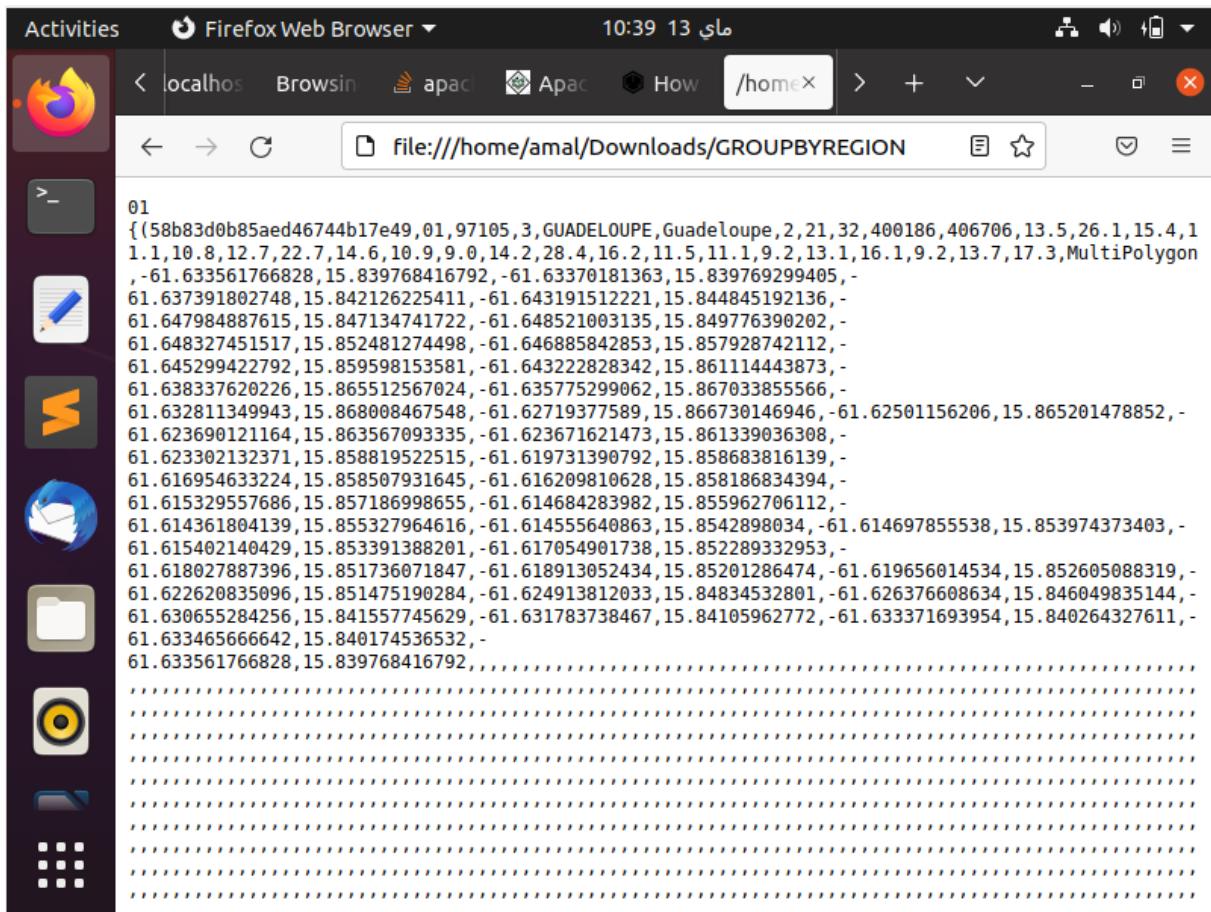
The screenshot shows the Sublime Text editor interface. The title bar displays "Activities" and "Sublime Text". The status bar shows the time as "20:06 12 مارس". The main window has two tabs: "region11.html" and "style.css". The "region11.html" tab contains the following code:

```
<head>
<link href="style.css" rel="stylesheet" type="text/css">
</head>
<body>
<table>
<caption>Database INSEE</caption>
<tr>
<td> <button> <a rel="stylesheet" type="text/css" href="/home/amal/Downloads/filterR11" > Filterregion = FILTER Region BY REGION = 11 </a> </button></td> </tr>
<tr>
<td><button> <a rel="stylesheet" type="text/css" href="/home/amal/Downloads/foreachregion" > foreachregion = FOREACH region GENERATE REGION </a> </button></td> </tr>
<tr>
<td><button> <a rel="stylesheet" type="text/css" href="/home/amal/Downloads/filtreguyaneregion" > filterncc = FILTER Region BY NCC=GUYANE</a> </button></td>
</tr>
<tr>
<td>
<button> <a rel="stylesheet" type="text/css" href="/home/amal/Downloads/GROUPBYREGION" > GroupRegion = GROUP region BY REGION </a> </button></td>
</tr>
<tr>
<td><button> <a rel="stylesheet" type="text/css" href="/home/amal/Downloads/dumpdepart" > DUMP departements</a> </button></td>
</tr>
<tr>
```

- Affichage :







Activities Firefox Web Browser 10:39 13 مای 13:00

localhos browsin apac Apac How /home/x

file:///home/amal/Downloads/GROUPBYREGION

```
01
{(58b83d0b85aed46744b17e49,01,97105,3,GUADELOUPE,Guadeloupe,2,21,32,400186,406706,13.5,26.1,15.4,1
1.1,10.8,12.7,22.7,14.6,10.9,9.0,14.2,28.4,16.2,11.5,11.1,9.2,13.1,16.1,9.2,13.7,17.3,Multipolygon
,-61.633561766828,15.839768416792,-61.63370181363,15.839769299405,-
61.637391802748,15.842126225411,-61.643191512221,15.844845192136,-
61.647984887615,15.847134741722,-61.648521003135,15.849776390202,-
61.648327451517,15.852481274498,-61.646885842853,15.857928742112,-
61.645299422792,15.859598153581,-61.643222828342,15.861114443873,-
61.638337620226,15.865512567024,-61.635775299062,15.867033855566,-
61.632811349943,15.868008467548,-61.62719377589,15.866730146946,-61.62501156206,15.865201478852,-
61.623690121164,15.863567093335,-61.623671621473,15.861339036308,-
61.623302132371,15.858819522515,-61.619731390792,15.858683816139,-
61.616954633224,15.858507931645,-61.616209810628,15.858186834394,-
61.615329557686,15.857186998655,-61.614684283982,15.855962706112,-
61.614361804139,15.855327964616,-61.614555640863,15.8542898034,-61.614697855538,15.853974373403,-
61.615402140429,15.853391388201,-61.617054901738,15.852289332953,-
61.618027887396,15.851736071847,-61.618913052434,15.85201286474,-61.619656014534,15.852605088319,-
61.622620835096,15.851475190284,-61.624913812033,15.84834532801,-61.626376608634,15.846049835144,-
61.630655284256,15.841557745629,-61.631783738467,15.84105962772,-61.633371693954,15.840264327611,-
61.633465666642,15.840174536532,-
61.633561766828,15.839768416792,.....
```

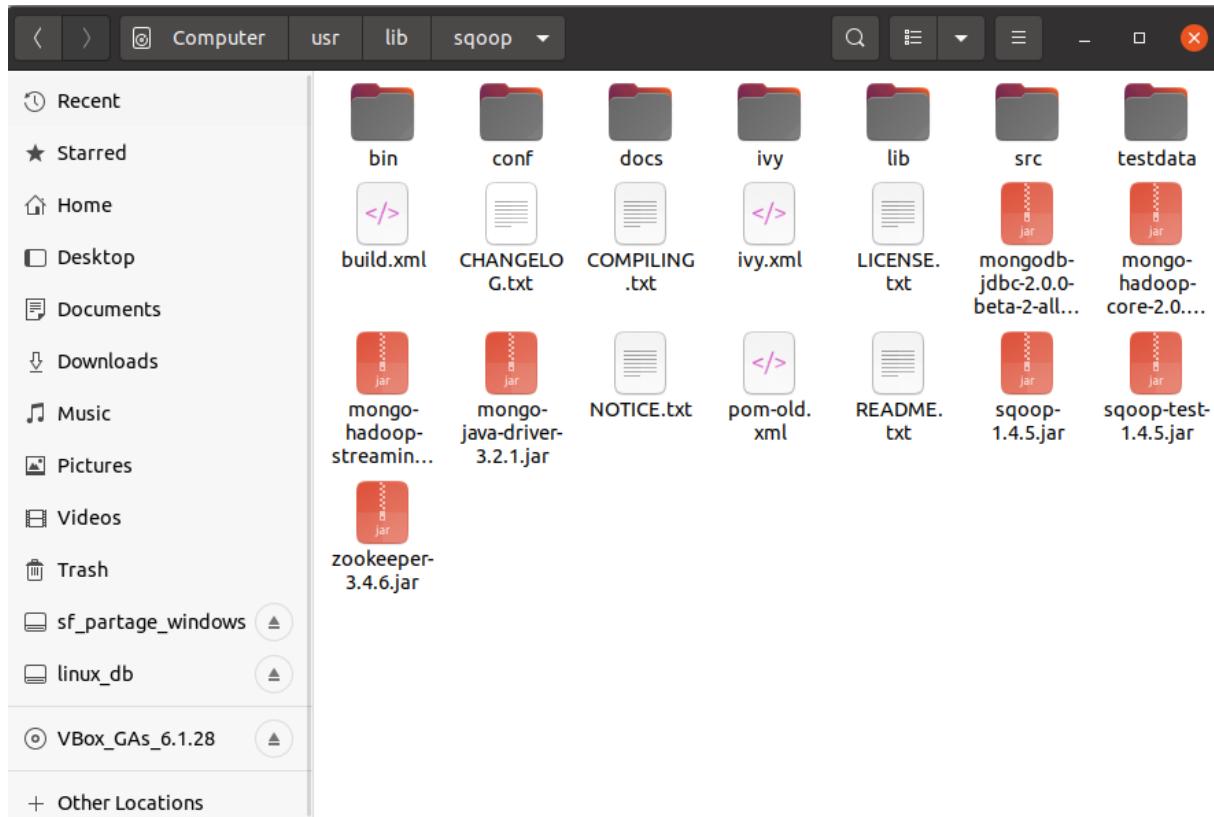
❖ Essai :

- Sqoop

```
mongosh mongodb://... × amal@amal-VirtualB... × amal@amal-VirtualB...
r/lib/sqoop': Permission denied
amal@amal-VirtualBox:~$ sudo mv /home/amal/Desktop/sqoop-1.4.5-bin__hadoop-2.0.4-alpha /usr/lib/sqoop
[sudo] password for amal:
amal@amal-VirtualBox:~$ sudo gedit .bashrc

(gedit:5587): Tepl-WARNING **: 11:48:47.940: GVfs metadata is not supported. Fallback to TeplMetadataManager. Either GVfs is not correctly installed or GVfs metadata are not supported on this platform. In the latter case, you should configure Tepl with --disable-gvfs-metadata.
amal@amal-VirtualBox:~$ sudo source ~/.bashrc
sudo: source: command not found
amal@amal-VirtualBox:~$ source ~/.bashrc
amal@amal-VirtualBox:~$ sudo gedit /usr/lib/sqoop/conf/sqoop-env-template.sh

(gedit:5671): Tepl-WARNING **: 11:55:51.941: GVfs metadata is not supported. Fallback to TeplMetadataManager. Either GVfs is not correctly installed or GVfs metadata are not supported on this platform. In the latter case, you should configure Tepl with --disable-gvfs-metadata.
amal@amal-VirtualBox:~$ cd $SQOOP_HOME/conf
amal@amal-VirtualBox:/usr/lib/sqoop/conf$ sudo mv sqoop-env-template.sh sqoop-env.sh
amal@amal-VirtualBox:/usr/lib/sqoop/conf$ cd
amal@amal-VirtualBox:~$
```



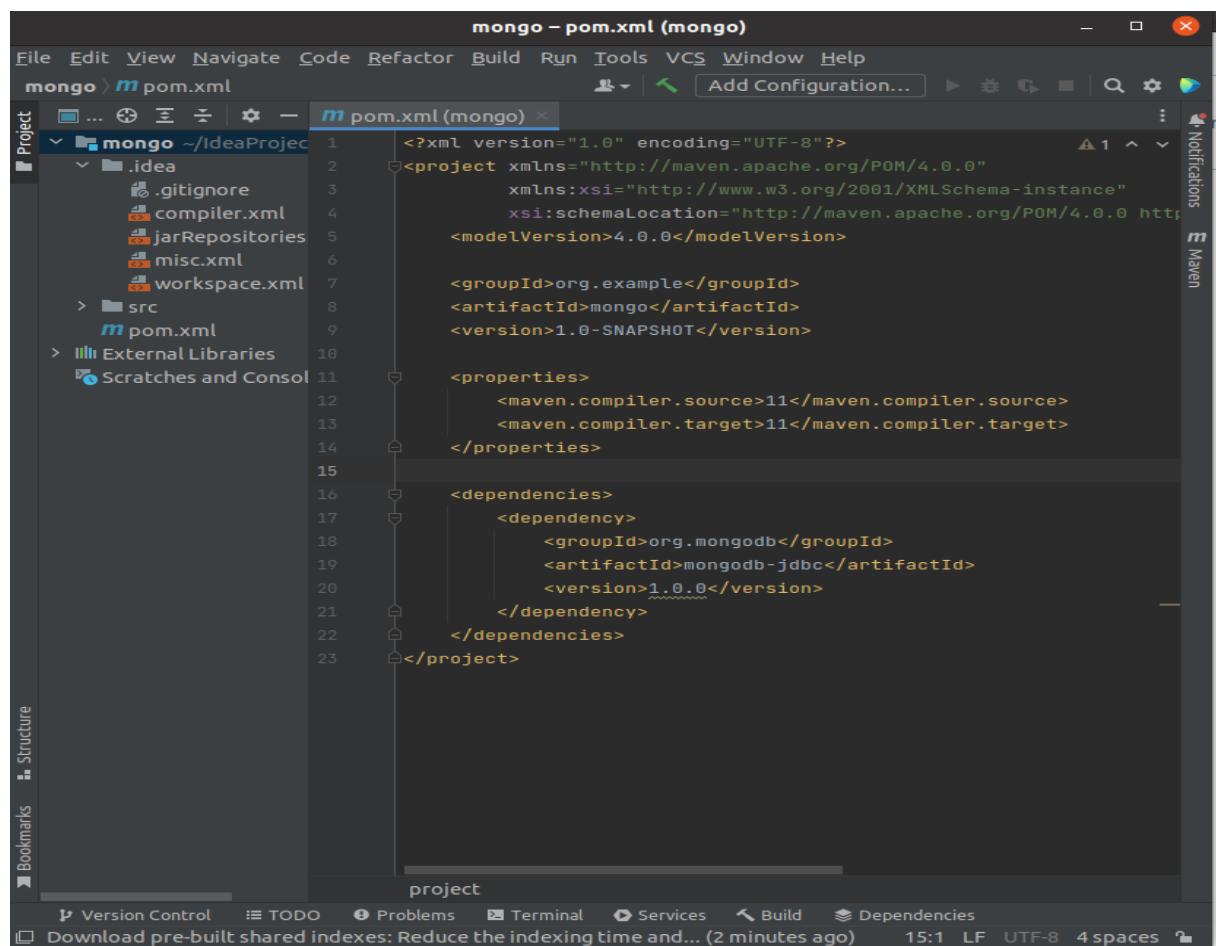
```
mongosh mo...      amal@amal-VirtualBox: /usr/lib/sqoop/bin      amal@amal-...      amal@amal-...
amal@amal-VirtualBox:/usr/lib/sqoop/bin$ sqoop import-all-tables --connect jdbc:mongodb://localhost:27017/insee --warehouse-dir /amal1
Warning: /usr/lib/sqoop/../hbase does not exist! HBase imports will fail.
Please set $HBASE_HOME to the root of your HBase installation.
Warning: /usr/lib/sqoop/../hcatalog does not exist! HCatalog jobs will fail.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /usr/lib/sqoop/../accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
Warning: /usr/lib/sqoop/../zookeeper does not exist! Accumulo imports will fail.
Please set $ZOOKEEPER_HOME to the root of your Zookeeper installation.
22/05/10 16:03:21 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5
22/05/10 16:03:21 ERROR tool.BaseSqoopTool: Got error creating database manager:
java.io.IOException: No manager for connect string: jdbc:mongodb://localhost:27017/insee
        at org.apache.sqoop.ConnFactory.getManager(ConnFactory.java:191)
        at org.apache.sqoop.tool.BaseSqoopTool.init(BaseSqoopTool.java:247)
        at org.apache.sqoop.tool.ImportTool.init(ImportTool.java:89)
        at org.apache.sqoop.tool.ImportAllTablesTool.run(ImportAllTablesTool.java:82)
        at org.apache.sqoop.Sqoop.run(Sqoop.java:143)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:70)
        at org.apache.sqoop.Sqoop.runSqoop(Sqoop.java:179)
        at org.apache.sqoop.Sqoop.runTool(Sqoop.java:218)
        at org.apache.sqoop.Sqoop.runTool(Sqoop.java:227)
```

```

amal@amal-VirtualBox: /usr/lib/sqoop/bin$ sqoop import-all-tables --connect jdbc
c:mongodb://127.0.0.1:27017/insee --warehouse-dir /amal1
Warning: /usr/lib/sqoop/..hbbase does not exist! HBase imports will fail.
Please set $HBASE_HOME to the root of your HBase installation.
Warning: /usr/lib/sqoop/..hcatalog does not exist! HCatalog jobs will fail.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /usr/lib/sqoop/..accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
Warning: /usr/lib/sqoop/..zookeeper does not exist! Accumulo imports will fail.
Please set $ZOOKEEPER_HOME to the root of your Zookeeper installation.

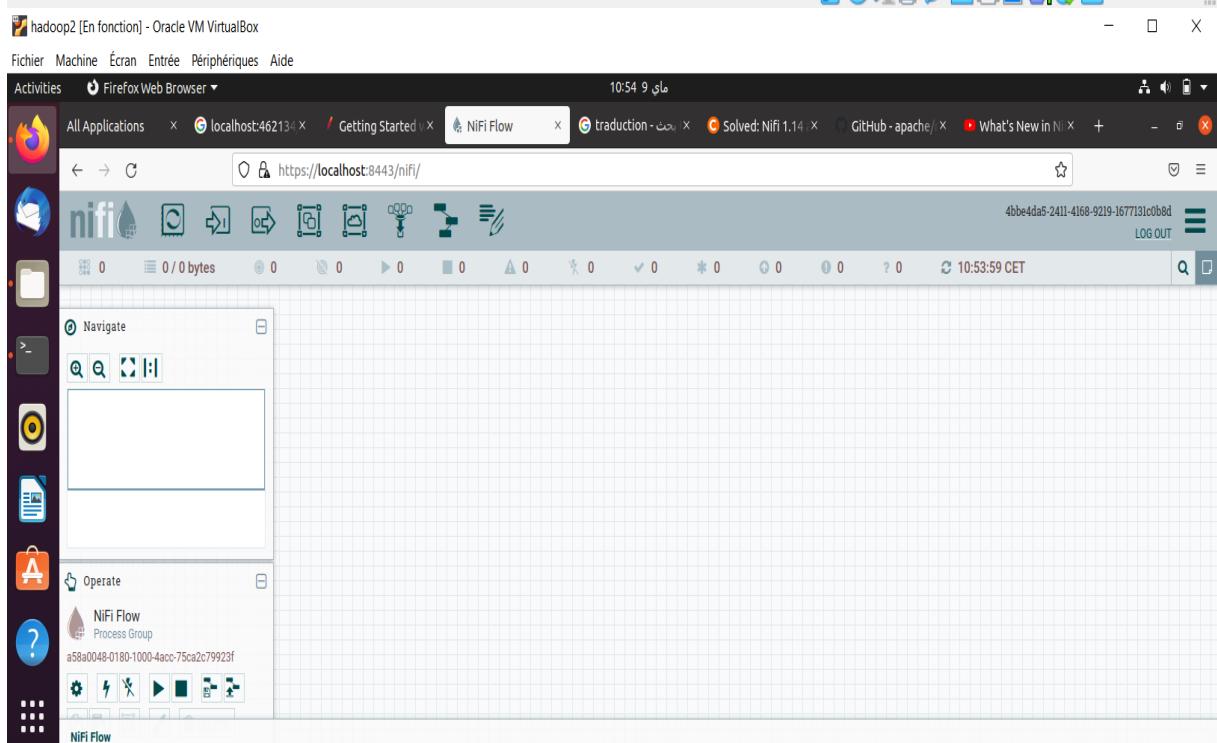
22/05/10 15:54:38 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5
22/05/10 15:54:39 ERROR tool.BaseSqoopTool: Got error creating database manager:
java.io.IOException: No manager for connect string: jdbc:mongodb://127.0.0.1:27
017/insee
        at org.apache.sqoop.ConnFactory.getManager(ConnFactory.java:191)
        at org.apache.sqoop.tool.BaseSqoopTool.init(BaseSqoopTool.java:247)
        at org.apache.sqoop.tool.ImportTool.init(ImportTool.java:89)
        at org.apache.sqoop.tool.ImportAllTablesTool.run(ImportAllTablesTool.jav
a:82)
        at org.apache.sqoop.Sqoop.run(Sqoop.java:143)
        at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:70)
        at org.apache.sqoop.Sqoop.runSqoop(Sqoop.java:179)
        at org.apache.sqoop.Sqoop.runTool(Sqoop.java:218)

```



- Nifi :

```
hadoop2 [En fonction] - Oracle VM VirtualBox
Fichier Machine Écran Entrée Périphériques Aide
Activities Terminal amal@amal-VirtualBox: ~/Desktop/nifi-1.16.1
amal@amal-VirtualBox:~$ sudo /home/amal/Desktop/nifi-1.16.1/bin/nifi.sh start
[sudo] password for amal:
Java home: /usr/lib/jvm/default-java
NiFi home: /home/amal/Desktop/nifi-1.16.1
Bootstrap Config File: /home/amal/Desktop/nifi-1.16.1/conf/bootstrap.conf
amal@amal-VirtualBox:~$ sudo /home/amal/Desktop/nifi-1.16.1/bin/nifi.sh status
Java home: /usr/lib/jvm/default-java
NiFi home: /home/amal/Desktop/nifi-1.16.1
Bootstrap Config File: /home/amal/Desktop/nifi-1.16.1/conf/bootstrap.conf
2022-05-09 10:33:19,186 INFO [main] org.apache.nifi.bootstrap.Command Apache NiFi is currently running, listening to Bootstrap on port 46213, PID=3498
amal@amal-VirtualBox:~$ laptop:nifi-1.0.0-SNAPSHOT myuser$ grep Generated logs/nifi-app*log
laptop:nifi-1.0.0-SNAPSHOT: command not found
amal@amal-VirtualBox:~$ sudo laptop:nifi-1.0.0-SNAPSHOT myuser$ grep Generated logs/nifi-app*log
sudo: laptop:nifi-1.0.0-SNAPSHOT: command not found
amal@amal-VirtualBox:~$ cd /home/amal/Desktop/nifi-1.16.1/logs
amal@amal-VirtualBox:~/Desktop/nifi-1.16.1/logs$ laptop:nifi-1.0.0-SNAPSHOT myuser$ grep Generated logs/nifi-app*log
laptop:nifi-1.0.0-SNAPSHOT: command not found
amal@amal-VirtualBox:~/Desktop/nifi-1.16.1/logs$ grep Generated logs/nifi-app*log
amal@amal-VirtualBox:~$ cd /home/amal/Desktop/nifi-1.16.1/
amal@amal-VirtualBox:~/Desktop/nifi-1.16.1$ grep Generated logs/nifi-app*log
logs/nifi-app_2022-05-08_22.0.log:Generated Username [4bbe4da5-2411-4168-9219-1677131c0b8d]
logs/nifi-app_2022-05-08_22.0.log:Generated Password [g9AbdAv6r1RC8Jrvu9m002CKLm5PuGoR]
amal@amal-VirtualBox:~/Desktop/nifi-1.16.1$
```



- Flume :

```

mongosh mongodb://... amal@amal-VirtualB... amal@amal-VirtualB...
amal@amal-VirtualBox:~$ flume-ng version
Flume 1.9.0
Source code repository: https://git-wip-us.apache.org/repos/asf/flume.git
Revision: d4fcab4f501d41597bc616921329a4339f73585e
Compiled by fszabo on Mon Dec 17 20:45:25 CET 2018
From source with checksum 35db629a3bda49d23e9b3690c80737f9
amal@amal-VirtualBox:~$ sudo gedit /flume.conf
[sudo] password for amal:
amal@amal-VirtualBox:~$ sudo gedit /home/amal/Desktop/apache-flume-1.9.0-bin/conf/flume.conf

(gedit:4620): Tepl-WARNING **: 11:06:28.047: GVfs metadata is not supported. Fallback to TeplMetadataManager. Either GVfs is not correctly installed or GVfs metadata are not supported on this platform. In the latter case, you should configure Tepl with --disable-gvfs-metadata.
amal@amal-VirtualBox:~$ flume-ng agent --conf conf --conf-file /home/amal/Desktop/apache-flume-1.9.0-bin/conf/flume.conf --name agent2
Info: Including Hadoop libraries found via (/usr/local/hadoop/bin/hadoop) for HDFS access
Info: Including Hive libraries found via () for Hive access
+ exec /usr/lib/jvm/default-java/bin/java -Xmx20m -cp 'conf:/home/amal/Desktop/apache-flume-1.9.0-bin/lib/*:/usr/local/hadoop/etc/hadoop:/usr/local/hadoop/share/hadoop/common/lib/*:/usr/local/hadoop/share/hadoop/common/*:/usr/local/hadoop/share/hadoop/hdfs:/usr/local/hadoop/share/hadoop/lib/*:/usr/local/hadoop/sha

```

```

Open flume.conf
/home/amal/Desktop/apache-flume-1.9.0-bin/conf

1 agent2.sources = source2
2 agent2.channels = channel2
3 agent2.sinks = sink2
4
5 agent2.sources.source2.type = jms
6 agent2.sources.source2.initialContextFactory = org.apache.activemq.jndi.ActiveMQInitialContextFactory
7 agent2.sources.source2.connectionFactory = GenericConnectionFactory
8 agent2.sources.source2.providerURL = mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+1.3.1
9 agent2.sources.source2.port = 27017
10 agent2.sources.source2.destinationName = insee
11 agent2.sources.source2.destinationType = QUEUE
12
13 agent2.sources.source2.channels = channel2
14
15 agent2.sinks.sink2.type = hdfs
16 agent2.sinks.sink2.path = /amal1
17 agent2.sinks.sink2.fileSuffix = .csv
18
19
20
21 agent2.sinks.sink2.channel = channel2
22
23 agent2.channels.channel2.type = memory
24 agent2.channels.channel2.capacity = 1000000
25
26

```

```

mongosh mongodb://... amal@amal-VirtualB... amal@amal-VirtualB...
22/05/10 11:13:43 INFO conf.FlumeConfiguration: Processing:source2
22/05/10 11:13:43 INFO conf.FlumeConfiguration: Processing:source2
22/05/10 11:13:43 WARN conf.FlumeConfiguration: Agent configuration for 'agent2' has no configfilters.
22/05/10 11:13:43 INFO conf.FlumeConfiguration: Post-validation flume configuration contains configuration for agents: [agent2]
22/05/10 11:13:43 INFO node.AbstractConfigurationProvider: Creating channels
22/05/10 11:13:43 INFO channel.DefaultChannelFactory: Creating instance of channel channel2 type memory
22/05/10 11:13:43 INFO node.AbstractConfigurationProvider: Created channel channel2
22/05/10 11:13:43 INFO source.DefaultSourceFactory: Creating instance of source source2, type jms
22/05/10 11:13:44 ERROR node.AbstractConfigurationProvider: Source source2 has been removed due to an error during configuration
org.apache.flume.FlumeException: Could not create initial context org.apache.activemq.jndi.ActiveMQInitialContextFactory provider mongodb://127.0.0.1:27017/?directConnection=true&serverSelectionTimeoutMS=2000&appName=mongosh+1.3.1
        at org.apache.flume.source.jms.JMSSource.doConfigure(JMSSource.java:259)
        at org.apache.flume.source.BasicSourceSemantics.configure(BasicSourceSemantics.java:65)
        at org.apache.flume.source.AbstractPollableSource.configure(AbstractPollableSource.java:65)
        at org.apache.flume.conf.Configurables.configure(Configurables.java:41)

```