TP4 :HBASE

```
PS C:\Users\DELL\Downloads\HBase> docker start hadoop-master hadoop-slave1 hadoop-slave2
hadoop-master
hadoop-slave1
hadoop-slave2
PS C:\Users\DELL\Downloads\HBase> docker exec -it hadoop-master bash
root@hadoop-master:~# ./start-hadoop.sh



Starting namenodes on [hadoop-master]
hadoop-master: Warning: Permanently added 'hadoop-master,172.21.0.2' (ECDSA) to the list of known hosts.
```

```
root@hadoop-master:~# /usr/local/hbase/bin/start-hbase.sh
hadoop-master: Warning: Permanently added 'hadoop-master,172.21.0.2' (ECDSA) to the list of known hosts.
hadoop-master: running zookeeper, logging to /usr/local/hbase/bin/../logs/hbase-root-zookeeper-hadoop-master.out
root@hadoop-master:~# jps
1536 HQuorumPeer
1786 HRegionServer
251 NameNode
1611 HMaster
539 SecondaryNameNode
844 ResourceManager
2220 Jps
```

```
root@hadoop-master:~# hbase shell
SLF4J: Class path contains multiple SLF4J bindings.
```

- Nous allons créer une base de données qui contient les données suivantes:

Row Key	customer		sales	
ROW_ID	name	city	product	amount
101	John White	Los Angeles, CA	Chairs	$400.00
102	Jane Brown	Atlanta, GA	Lamps	$200.00
103	Bill Green	Pittsburgh, PA	Desk	$500.00
104	Jack Black	St. Louis, MO	Bed	$1,600.00

```
hbase(main):001:0> create 'sales_ledger','customer','sales'
0 row(s) in 1.9120 seconds

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

```
hbase(main):001:0> create 'sales_ledger','customer','sales'
0 row(s) in 1.9120 seconds

=> Hbase::Table - sales_ledger
hbase(main):002:0> list
TABLE
sales_ledger
1 row(s) in 0.0310 seconds

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

Chargement :

```
PS C:\Users\DELL\Downloads\HBase> docker exec hadoop-master bash -c "hbase org.apache.hadoop.hbase.ma
preduce.ImportTsv -libjars /usr/local/hbase/lib/commons-lang-2.6.jar -Dimporttsv.separator=',' -Dimpo
rttsv.columns=HBASE_ROW_KEY,cf:date,cf:time,cf:town,cf:product,cf:price,cf:payment products input/pur
chases2.txt"
```

Insérer les différentes lignes:

```
hbase(main):012:0> put 'sales_ledger','103','customer:name','Bill Green'
sburgh, PA'
put 'sales_ledger','103','sales:produc0 row(s) in 0.0150 seconds

t','Desk'
put 'sales_ledger','103','sales:amount',hbase(main):013:0> put 'sales_ledger','103','customer:city','Pitts
0 row(s) in 0.0080 seconds

hbase(main):014:0> put 'sales_ledger','103','sales:product','Desk'
0 row(s) in 0.0070 seconds
```

Visualiser le résultat de l'insertion, en tapant:

```
hbase(main):020:0> scan 'sales_ledger'
ROW                     COLUMN+CELL
 101                    column=customer:city, timestamp=1763028623521, value=Los Angeles, CA
 101                    column=customer:name, timestamp=1763028623458, value=John White
 101                    column=sales:amount, timestamp=1763028625894, value=$400.00
 101                    column=sales:product, timestamp=1763028623558, value=Chairs
 102                    column=customer:city, timestamp=1763028660071, value=Atlanta, GA
 102                    column=customer:name, timestamp=1763028659969, value=Jane Brown
 102                    column=sales:amount, timestamp=1763028661478, value=$200.00
 102                    column=sales:product, timestamp=1763028660172, value=Lamps
 103                    column=customer:city, timestamp=1763028671158, value=Pittsburgh, PA
```

Afficher les valeurs de la colonne product de la ligne 102

```
hbase(main):021:0> get 'sales_ledger','102',{COLUMN => 'sales:product'}
COLUMN                     CELL
 sales:product             timestamp=1763028660172, value=Lamps
1 row(s) in 0.0390 seconds
```

Explorer les filtres dans HBase :

Filter – RowFilter

```
hbase(main):006:0> scan 'sales_ledger', {FILTER => "RowFilter(>, 'binary:102')"}
ROW                      COLUMN+CELL
 103                     column=customer:city, timestamp=1763028671158, value=Pittsburgh, PA
 103                     column=customer:name, timestamp=1763028671118, value=Bill Green
 103                     column=sales:amount, timestamp=1763028673600, value=$500.00
 103                     column=sales:product, timestamp=1763028671188, value=Desk
 104                     column=customer:city, timestamp=1763028694679, value=St. Louis, MO
 104                     column=customer:name, timestamp=1763028694615, value=Jack Black
 104                     column=sales:amount, timestamp=1763028705259, value=$1,600.00
 104                     column=sales:product, timestamp=1763028694713, value=Bed
2 row(s) in 0.0670 seconds
```

```
hbase(main):007:0> scan 'sales_ledger', {FILTER => "FamilyFilter(=, 'binary:customer')"}
ROW                      COLUMN+CELL
 101                     column=customer:city, timestamp=1763028623521, value=Los Angeles, CA
 101                     column=customer:name, timestamp=1763028623458, value=John White
 102                     column=customer:city, timestamp=1763028660071, value=Atlanta, GA
 102                     column=customer:name, timestamp=1763028659969, value=Jane Brown
 103                     column=customer:city, timestamp=1763028671158, value=Pittsburgh, PA
 103                     column=customer:name, timestamp=1763028671118, value=Bill Green
 104                     column=customer:city, timestamp=1763028694679, value=St. Louis, MO
 104                     column=customer:name, timestamp=1763028694615, value=Jack Black
4 row(s) in 0.0310 seconds
```

Filter – FamilyFilter

```
hbase(main):008:0> scan 'sales_ledger', {FILTER => "FamilyFilter(=, 'binary:customer')"}
ROW                      COLUMN+CELL
 101                     column=customer:city, timestamp=1763028623521, value=Los Angeles, CA
 101                     column=customer:name, timestamp=1763028623458, value=John White
 102                     column=customer:city, timestamp=1763028660071, value=Atlanta, GA
 102                     column=customer:name, timestamp=1763028659969, value=Jane Brown
 103                     column=customer:city, timestamp=1763028671158, value=Pittsburgh, PA
 103                     column=customer:name, timestamp=1763028671118, value=Bill Green
 104                     column=customer:city, timestamp=1763028694679, value=St. Louis, MO
 104                     column=customer:name, timestamp=1763028694615, value=Jack Black
4 row(s) in 0.0320 seconds
```

Filter – QualifierFilter

```
hbase(main):009:0> scan 'sales_ledger', {FILTER => "QualifierFilter(=, 'binary:amount')"}
ROW                      COLUMN+CELL
 101                     column=sales:amount, timestamp=1763028625894, value=$400.00
 102                     column=sales:amount, timestamp=1763028661478, value=$200.00
 103                     column=sales:amount, timestamp=1763028673600, value=$500.00
 104                     column=sales:amount, timestamp=1763028705259, value=$1,600.00
4 row(s) in 0.0280 seconds
```

Filter – SingleColumnValueFilter

```
hbase(main):010:0> scan 'sales_ledger', {FILTER => "SingleColumnValueFilter('sales', 'amount', =, 'binary:
hbase(main):011:1" $450.00')"}
ROW                      COLUMN+CELL
0 row(s) in 0.0290 seconds
```

Filter – PrefixFilter

```
hbase(main):001:0> scan 'sales_ledger', {FILTER => "PrefixFilter('10')"}
ROW                      COLUMN+CELL
 101                     column=customer:city, timestamp=1763028623521, value=Los Angeles, CA
 101                     column=customer:name, timestamp=1763028623458, value=John White
 101                     column=sales:amount, timestamp=1763028625894, value=$400.00
 101                     column=sales:product, timestamp=1763028623558, value=Chairs
 102                     column=customer:city, timestamp=1763028660071, value=Atlanta, GA
 102                     column=customer:name, timestamp=1763028659969, value=Jane Brown
 102                     column=sales:amount, timestamp=1763028661478, value=$200.00
 102                     column=sales:product, timestamp=1763028660172, value=Lamps
```

Filter – PageFilter

```
hbase(main):002:0> scan 'sales_ledger', {FILTER => "PageFilter(2)"}
ROW                         COLUMN+CELL
 101                         column=customer:city, timestamp=1763028623521, value=Los Angeles, CA
 101                         column=customer:name, timestamp=1763028623458, value=John White
 101                         column=sales:amount, timestamp=1763028625894, value=$400.00
 101                         column=sales:product, timestamp=1763028623558, value=Chairs
 102                         column=customer:city, timestamp=1763028660071, value=Atlanta, GA
 102                         column=customer:name, timestamp=1763028659969, value=Jane Brown
 102                         column=sales:amount, timestamp=1763028661478, value=$200.00
 102                         column=sales:product, timestamp=1763028660172, value=Lamps
2 row(s) in 0.0440 seconds
```

Chargement donne :

```
PS C:\Users\DELL\Downloads\HBase> docker exec hadoop-master bash -c 'export HADOOP_CLASSPATH=$(hbase
classpath) && hbase org.apache.hadoop.hbase.mapreduce.ImportTsv -Dimporttsv.separator="," -Dimporttsv
.columns=HBASE_ROW_KEY,cf:date,cf:time,cf:town,cf:product,cf:price,cf:payment products /user/root/inp
ut/purchases2.txt'
oducts /user/root/input/purchases2.txt';d79dace4-461b-421c-9a6a-7c72ef73000aSLF4J: Class path contain
s multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/Stati
cLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!
```

```
PS C:\Users\DELL\Downloads\HBase> docker cp "C:\Users\DELL\Downloads\HBase\hbase-code\scan.hbase" had
oop-master:/tmp/scan.hbase; docker exec hadoop-master bash -c "hbase shell /tmp/scan.hbase 2>&1" | Se
lect-String -Pattern "ROW|column=|row\(s\)"
461b-421c-9a6a-7c72ef73000aSuccessfully copied 2.05kB to hadoop-master:/tmp/scan.hbase

ROW  COLUMN+CELL
 1 column=cf:date, timestamp=1763033080722, value=2012-01-01
 1 column=cf:payment, timestamp=1763033080722, value=Amex
 1 column=cf:price, timestamp=1763033080722, value=214.05
 1 column=cf:product, timestamp=1763033080722, value=Men's Clothing
 1 column=cf:time, timestamp=1763033080722, value=09:00
```