UNIVERSITE GASTON BERGER

L'excellence au service du développement

INSTITUT POLYTECHNIQUE DE SAINT-LOUIS

Evaluation **: TP1 HDFS Prise en main**

Michel Nassalang

1. Prise en main de HDFS

Pour démarrer hdfs on a la commande : start-all.sh

Pour voir les commandes possibles avec hdfs : hdfs dfs

Pour voir ce que fait une commande avec hdfs : hdfs dfs -usage COMMANDE

2. Création des repertoires et copies de fichiers

- Pour la création de dossier dans hdfs, on a la commande : hdfs dfs -mkdir path

Pour le dossier nassalang créé, on a : hdfs dfs -mkdir /user/vagrant/nassalang

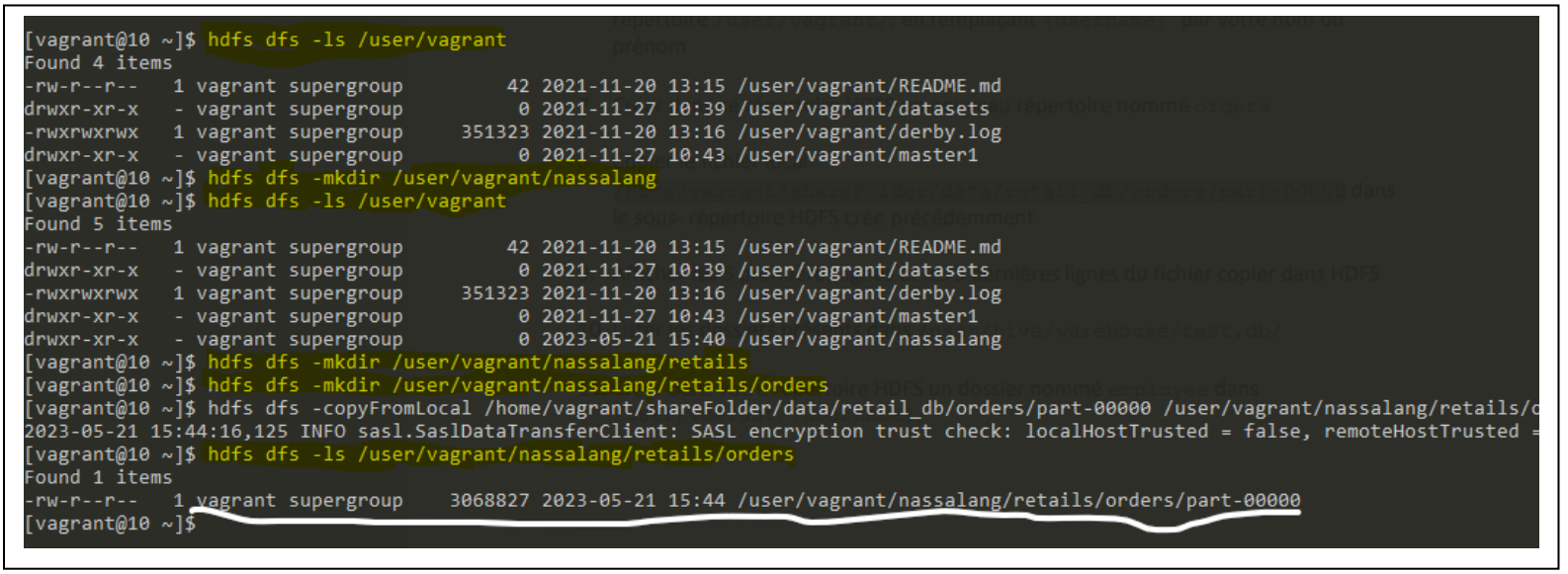
Pour le dossier retails : hdfs dfs -mkdir /user/vagrant/nassalang/retails

Pour le dossier orders : hdfs dfs -mkdir /user/vagrant/nassalang/retails/orders

- Pour la copie du fichier dans hdfs, on a : hdfs dfs -copyFromLocal localdst dst

Pour la copie du fichier /home/vagrant/shareFolder/data/retail_db/orders/part-00000 dans le repertoire hdfs créé précedemment on a comme commande :

hdfs dfs -copyFromLocal /home/vagrant/shareFolder/data/retail_db/orders/part-00000 /user/vagrant/nassalang/retails/orders

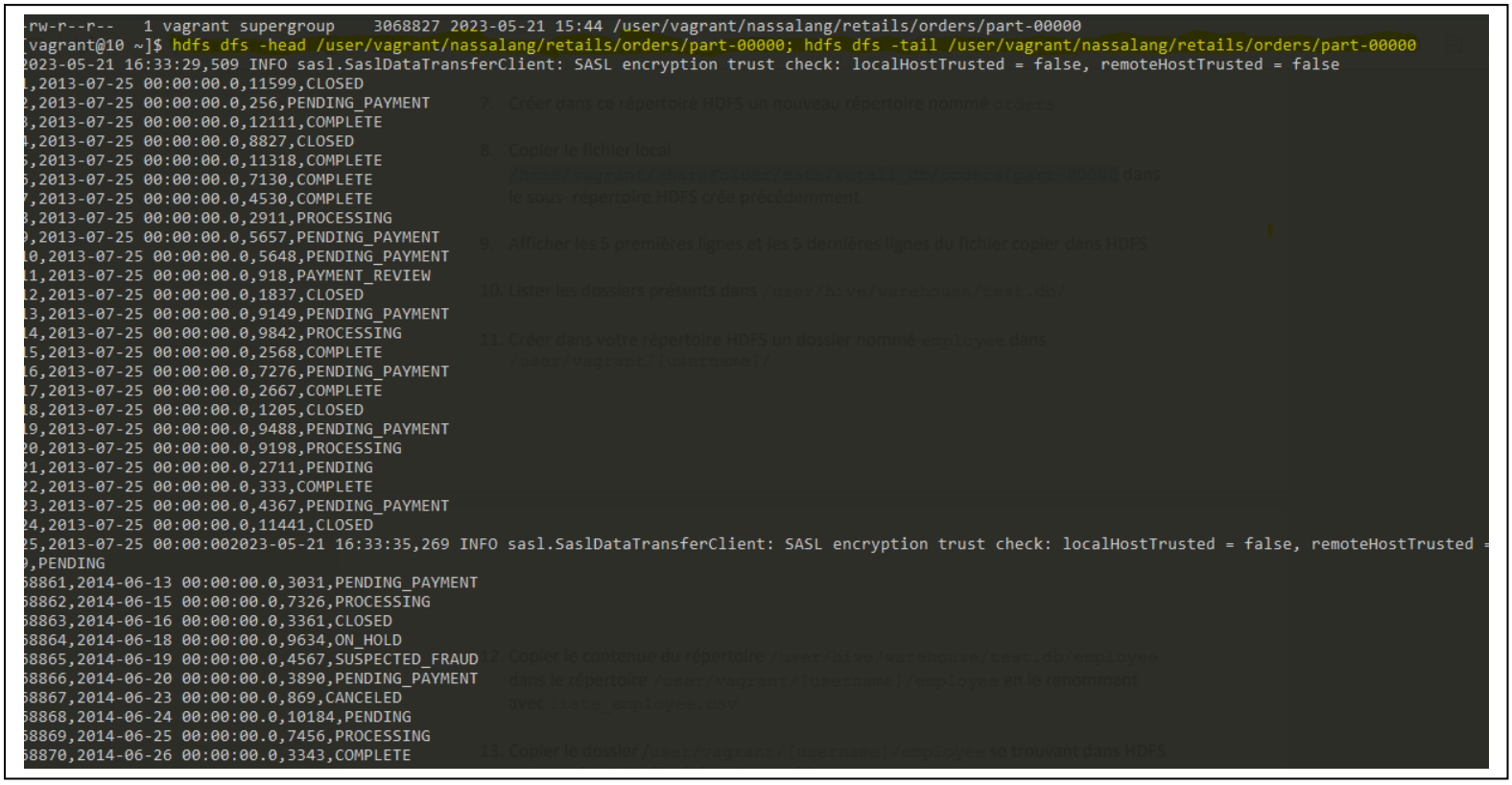


3. Visualisation des données par hdfs

Pour afficher les lignes d'un fichier, on utilise les commandes head et tail :

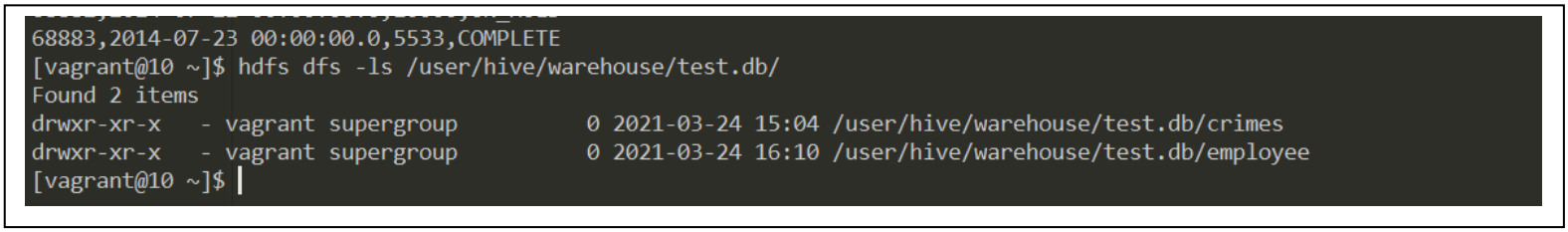
hdfs dfs -head file ; hdfs dfs -tail file

Pour le fichier part-00000: on a : hdfs dfs -head /user/vagrant/nassalang/retails/orders/part-00000; hdfs dfs -tail /user/vagrant/nassalang/retails/orders/part-00000



Pour lister les fichiers présents dans un dossier on utilise : hdfs dfs -ls path

Pour le repertoire /user/hive/warehouse/test.db/, on a : hdfs dfs -ls /user/hive/warehouse/test.db/



4. Création des repertoires et copies de cotenu

Pour le dossier employee créé, on a : hdfs dfs -mkdir /user/vagrant/nassalang/employee

Pour la copie du contenu de dans ce repertoire créé précedemment , on a : hdfs dfs -cp /user/hive/warehouse/test.db/employee/000000_1 /user/vagrant/nassalang/employee

```
[vagrant@10 ~]$ hdfs dfs -ls /user/hive/warehouse/test.db/
Found 2 items
drwxr-xr-x   - vagrant supergroup          0 2021-03-24 15:04 /user/hive/warehouse/test.db/crimes
drwxr-xr-x   - vagrant supergroup          0 2021-03-24 16:10 /user/hive/warehouse/test.db/employee
[vagrant@10 ~]$ hdfs dfs -mkdir /user/vagrant/nassalang/employee
[vagrant@10 ~]$ hdfs dfs -cp /user/hive/warehouse/test.db/employee /user/vagrant/nassalang/employee
2023-05-21 17:05:48,536 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2023-05-21 17:05:49,061 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
[vagrant@10 ~]$ hdfs dfs -ls /user/vagrant/nassalang/employee
Found 1 items
drwxr-xr-x   - vagrant supergroup          0 2023-05-21 17:05 /user/vagrant/nassalang/employee/employee
[vagrant@10 ~]$ hdfs dfs -ls /user/vagrant/nassalang/employee/employee
Found 1 items
-rw-r--r--   1 vagrant supergroup         23 2023-05-21 17:05 /user/vagrant/nassalang/employee/employee/000000_1
[vagrant@10 ~]$ hdfs dfs -cp /user/vagrant/nassalang/employee/employee/000000_1 /user/vagrant/nassalang/employee
2023-05-21 17:07:35,382 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2023-05-21 17:07:35,695 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
[vagrant@10 ~]$ hdfs dfs -ls /user/vagrant/nassalang/employee
Found 2 items
-rw-r--r--   1 vagrant supergroup         23 2023-05-21 17:07 /user/vagrant/nassalang/employee/000000_1
drwxr-xr-x   - vagrant supergroup          0 2023-05-21 17:05 /user/vagrant/nassalang/employee/employee
```

Pour le renommage du fichier, on a : hdfs dfs -mv /user/vagrant/nassalang/employee/000000_1 /user/vagrant/nassalang/employee/liste_employee.csv

```
[vagrant@10 ~]$ hdfs dfs -mv /user/vagrant/nassalang/employee/000000_1 /user/vagrant/nassalang/employee/liste_employee.csv
[vagrant@10 ~]$ hdfs dfs -ls /user/vagrant/nassalang/employee
Found 1 items
-rw-r--r--   1 vagrant supergroup         23 2023-05-21 17:07 /user/vagrant/nassalang/employee/liste_employee.csv
[vagrant@10 ~]$
```

Pour la copie du dossier /user/vagrant/[username]/employee se trouvant dans HDFS vers un répertoire local de votre machine, on a : hdfs dfs -copyToLocal /user/vagrant/nassalang/employee /home/vagrant/nassalang

```
[vagrant@10 ~]$ hdfs dfs -copyToLocal /user/vagrant/nassalang/employee /home/vagrant/nassalang
2023-05-21 17:38:18,197 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
[vagrant@10 ~]$ ls /home/vagrant/nassalang
employee  hbase-3.0.0-alpha-3-bin.tar.gz  retails
[vagrant@10 ~]$ ls /home/vagrant/nassalang/employee/
liste_employee.csv
[vagrant@10 ~]$
```