

## TP Hive

Création de la base de données

### Création de table externe

- 1- Créer une table externe comptes, cette dernière pointe sur le fichier comptes de hdfs.

```
CREATE TABLE external comptes(numero STRING,type STRING,cledeux STRING, typecollectif  
STRING, intitule STRING, debit STRING,credit STRING)ROW FORMAT DELIMITED FIELDS  
TERMINATED BY ','  
LOCATION '/mydata/input/hive'  
TBLPROPERTIES ("skip.header.line.count"="1")
```

- 2- Vérifier que tout a été créé comme attendu :

```
SHOW TABLES;  
DESCRIBE Comptes;  
SELECT * FROM Comptes limit 10;
```

- 3- Vérifier que les champs de la table sont correctement mappés en faisant exécuter à Hive un job MapReduce simple :

```
SELECT DISTINCT numero FROM comptes;
```

Vérifier le type et l'emplacement de la nouvelle table 'comptes':

- 4- DESCRIBE FORMATTED Comptes;

Notez que dans ce mode d'utilisation, la commande `DROP TABLE` supprimera la table dans Hive, mais pas le fichier de données. Vous pouvez le vérifier en supprimant puis recréant la table 'records'.

### Création des tables interne

- 5- Créer une table interne shd\_comptes avec les mêmes informations de la table externe que vous venez de créer
- 6- Charger les données à partir de la table externe
- 7- Afficher les comptes qui commence par 7
- 8- Ajouter une partition sur le numéro de compte comptable
- 9- Vérifier si vous avez bien ajouté la partition
- 10- Après la partition afficher les comptes qui commence par 7
- 11- Comparer le temps d'exécution sur Yarn