

Lab Sqoop

Objectif : Utiliser Sqoop pour importer/exporter des données entre MySQL et HDFS, en manipulant un fichier de consommation électrique.

1. Préparation de l'environnement Docker

Cloner le dépôt et démarrer les services :

```
git clone https://github.com/YASSINEABN/sqoop-lab.git
cd sqoop-lab
docker-compose up -d
```

2. Création de la table et insertion des données dans MySQL

Connexion au conteneur MySQL :

```
docker exec -it mysql bash
mysql -u root -proot
```

Créer la base `energie` et la table `consommation_energie` :

```
CREATE DATABASE IF NOT EXISTS energie;
USE energie;

CREATE TABLE consommation_energie (
    id_client VARCHAR(10),
    ville VARCHAR(50),
    date_consommation DATE,
    consommation_kwh FLOAT,
    type_compteur VARCHAR(20)
);

INSERT INTO consommation_energie VALUES
('C001', 'Casablanca', '2025-03-01', 42.5, 'Monophasé'),
('C002', 'Rabat', '2025-03-01', 38.2, 'Triphasé'),
('C003', 'Fès', '2025-03-01', 55.0, 'Monophasé'),
('C001', 'Casablanca', '2025-03-02', 41.7, 'Monophasé'),
('C002', 'Rabat', '2025-03-02', 39.1, 'Triphasé'),
('C004', 'Marrakech', '2025-03-01', 47.3, 'Monophasé'),
('C005', 'Tanger', '2025-03-01', 36.0, 'Monophasé'),
('C006', 'Agadir', '2025-03-01', 60.5, 'Triphasé'),
```

```
('C001', 'Casablanca', '2025-03-03', 43.0, 'Monophasé');
```

Quitter :

```
exit
```

3. Importer les données de MySQL vers HDFS

Connexion au conteneur **sqoop** :

```
docker exec -it sqoop bash
```

Commande d'import :

```
sqoop import \  
--connect jdbc:mysql://mysql:3306/energie \  
--username root \  
--password root \  
--table consommation_energie \  
--target-dir /user/root/consommation_import \  
--m 1
```

4. Vérification dans HDFS

```
hdfs dfs -ls /user/root/consommation_import  
hdfs dfs -cat /user/root/consommation_import/part-m-00000
```

5. Export de HDFS vers MySQL

5.1 Créer une nouvelle table dans MySQL

```
docker exec -it mysql bash  
mysql -u root -proot  
  
USE energie;  
  
CREATE TABLE consommation_export (  
    id_client VARCHAR(10),  
    ville VARCHAR(50),  
    date_consommation DATE,  
    consommation_kwh FLOAT,  
    type_compteur VARCHAR(20)  
);  
exit
```

5.2 Export depuis HDFS

Toujours depuis le conteneur sqoop :

```
hdfs dfs -cp /user/root/consommation_import/part-m-00000  
/user/root/consommation_export.csv
```

Puis :

```
sqoop export \  
--connect jdbc:mysql://mysql:3306/energie \  
--username root \  
--password root \  
--table consommation_export \  
--export-dir /user/root/consommation_export.csv \  
--input-fields-terminated-by ',' \  
--m 1
```

6. Vérification finale côté MySQL

```
docker exec -it mysql bash  
mysql -u root -proot  
  
USE energie;  
SELECT * FROM consommation_export;  
exit
```

7. Nettoyage (facultatif)

```
docker exec -it sqoop bash  
hdfs dfs -rm -r /user/root/consommation_import  
hdfs dfs -rm /user/root/consommation_export.csv  
  
docker exec -it mysql bash  
mysql -u root -proot  
  
USE energie;  
DROP TABLE consommation_export;  
DROP TABLE consommation_energie;
```