

## Peuplement de la base noSQL ScyllaDB

### Prérequis :

-Installer ScyllaDB dans vagrant :

```
sudo yum install -y scylla
```

-Démarrer le service Scylla :

```
sudo systemctl start scylla-server
```

-Se connecter à Scylla et créer un keyspace nommé tpaprojet :

```
cqlsh
```

```
CREATE KEYSPACE tpaprojet WITH replication = {'class': 'NetworkTopologyStrategy',  
'replication_factor' : 1};
```

-S'assurer que les fichiers csv soient bien dans le répertoire `../data`

### Scripts de peuplement :

Le peuplement de la base se fait à l'aide de 4 scripts python :

-peuplerScyllaClients.py

-peuplerScyllaCatalogue.py

-peuplerScyllaMarketing.py

-peuplerScyllaImmatriculations.py

Il y a également un script principal (**peuplerScylla.py**) qui va exécuter ces 4 scripts un à un. Nous conseillons d'exécuter celui-la qui s'occupera de tout.

### Verification :

Si les scripts se sont exécutés sans erreur, vous devriez avoir votre base Scylla peuplé.

Vous pouvez verifier de cette façon :

Connecter vous à Scylla :

```
cqlsh
```

Utiliser le keyspace tpaprojet :

```
USE tpaprojet;
```

Verifier l'existence des tables Clients,Immatriculations,Marketing et Catalogue :

```
DESCRIBE TABLES;
```

Faire une requête sur l'une des tables :

```
SELECT * FROM clients;
```

## Peuplement de la base relationnelle mySQL

### Prérequis :

- S'assurer d'avoir mySQL
- S'assurer que le service mySQL est lancé
- S'assurer que les fichiers csv soient bien dans le répertoire `../data`
- S'assurer que iconv est également installé

### Scripts de peuplement :

Le peuplement de la base se fait à l'aide d'1 script en bash:

- `insert_csv_in_mysql.sh`

### Vérification :

Si les scripts se sont exécutés sans erreur, vous devriez avoir votre base MySQL peuplée.

Vous pouvez vérifier de cette façon :

Connecter vous à MySQL :

`mysql`

Puis placez-vous sur votre collection

`use TPA`

Vérifier l'existence des tables Clients\_N, Immatriculations, Marketing et Catalogue :

`show tables;`

Faire une requête sur l'une des tables :

`SELECT * FROM catalogues;`

## Peuplement de la base de données nosql MongoDB

### Prérequis :

- S'assurer d'avoir mongodb
- S'assurer que le service de mongodb est lancé (mongod)
- S'assurer que les fichiers csv soient bien dans le répertoire `../data`

### Scripts de peuplement :

Le peuplement de la base se fait à l'aide d'1 script en bash:

- `insert_csv_in_mongodb.sh`

### Vérification :

Si les scripts se sont exécutés sans erreur, vous devriez avoir votre base MySQL peuplée.

Vous pouvez vérifier de cette façon :

Connecter vous à MongoDB :

`mongo`

Puis placez-vous sur votre collection

`use TPA`

Vérifier l'existence des tables Clients\_N, Immatriculations, Marketing et Catalogue :

`show collections`

Faire une requête sur l'une des tables :

`db.catalogues.find()`

## MongoDB vers Hive

### Prérequis :

- S'assurer d'avoir Hive
- Lancer les lignes suivantes dans vagrant  
`start-dfs.sh`  
`start-yarn.sh`
- S'assurer que le service de hive est lancé  
`nohup hive --service metastore > /dev/null &`  
`nohup hiveserver2 > /dev/null &`
- S'assurer que les fichiers csv soient bien dans le répertoire `../data`

### Scripts de peuplement :

Le peuplement de la base se fait à l'aide d'1 script en bash:

- `mongodb_to_hive_ext.sh`

### Vérification :

Si les scripts se sont exécutés sans erreur, vous devriez avoir votre DATALAKE peuplée.

Vous pouvez vérifier de cette façon :

Connecter vous à Hive :

```
beeline -u jdbc:hive2://localhost:10000/default --username root
```

Puis placez-vous sur votre collection

```
use default;
```

Vérifier l'existence des tables Clients\_N, Immatriculations, Marketing et Catalogue :

```
show tables;
```

Faire une requête sur l'une des tables :

```
SELECT * FROM catalogues;
```

## MYSQL vers Hive

### Prérequis :

- S'assurer d'avoir Hive
- Lancer les lignes suivantes dans vagrant  
`start-dfs.sh`  
`start-yarn.sh`
- S'assurer que le service de hive est lancé  
`nohup hive --service metastore > /dev/null &`  
`nohup hiveserver2 > /dev/null &`
- S'assurer que les fichiers csv soient bien dans le répertoire `../data`

### Scripts de peuplement :

Le peuplement de la base se fait à l'aide d'1 script en bash:

- `mysql_to_hive.sh`

### Vérification :

Si les scripts se sont exécutés sans erreur, vous devriez avoir votre DATALAKE peuplée.

Vous pouvez vérifier de cette façon :

Connecter vous à Hive :

```
beeline -u jdbc:hive2://localhost:10000/default --username root
```

Puis placez-vous sur votre collection

```
use TPA;
```

Vérifier l'existence des tables Clients\_N, Immatriculations, Marketing et Catalogue :

```
show tables;
```

Faire une requête sur l'une des tables :

```
SELECT * FROM catalogues;
```