

BIG DATA

TP - usage avancé des bases de données et NoSQL

Nicolas LAGAILLARDIE

16 octobre 2018

1 Objectif

Suite du TP de gestion de bases de données : installation et utilisation d'outils de gestion de bases NoSQL.

1.1 Définitions

NoSQL Ce terme désigne une famille de systèmes de gestion de base de données (SGBD) qui s'écarte du paradigme classique des bases relationnelles.

2 Installation de Cassandra

J'ai préféré installer Cassandra SQL plutôt que Oracle NoSQL, car plus simple. J'ai suivi la documentation officielle pour l'installation, puis j'ai exécuté quelques requêtes. Voici les commandes entrées dans le terminal :

- a. Démarrage de Cassandra

```
$ sudo service cassandra start
```

- b. Vérification du bon démarrage

```
$ nodetool status
```

- c. Démarrage de *cqlsh*

```
$ cqlsh
```

d. Création d'un *keyspace*

```
CREATE KEYSPACE user WITH REPLICATION = { 'class' : 'SimpleStrategy', '  
  ↪ replication_factor' : 1 } ;
```

e. Création d'un table

```
USE demo;  
CREATE TABLE eleves (idEleve text, nom text, prenom text, mail text, telephone text,  
  ↪ PRIMARY KEY (idEleve));
```

f. Ajouter des colonnes *cours* et *notes*

```
ALTER TABLE eleves ADD cours list<textless text>textgreater ;  
ALTER TABLE eleves ADD notes list<textless int>textgreater ;
```

g. Sortir de cqlsh puis lancer Spyder, driver pour Python

```
EXIT ;
```

```
$ spyder
```

3 Peuplement de la table

J'ai utilisé Python pour peupler la table avec beaucoup d'informations. Puis, toujours avec Python, j'ai lancé des requêtes SQL vers Cassandra afin d'afficher des résultats précis. Vous pouvez observer le code Python dans le fichier *first.py*.