

Projet Bases NoSQL pour le Big Data

M2 GIL

Lina Soualmia

9 janvier 2018

Sujet

Le projet décrit ci-après peut être réalisé avec les mêmes groupes que pour le projet Fouille de Textes.

L'objectif est d'étudier et de comparer deux système NoSQL que vous aurez choisis (p. ex. parmi la liste suivante : MongoDB, CouchDB, RethinkDB, Hadoop/HBase Cassandra, ou d'autres systèmes <http://nosql-database.org/>) :

Les systèmes NoSQL serviront à stocker des collections de données au format JSON (issus des corpus de textes constitués pour le projet Fouille de Textes).

Parmi les aspects à étudier, on citera :

- les modalités d'installation,
- les méthodes d'insertion de données,
- le langage de recherche,
- l'indexation interne,
- le support de la concurrence d'accès,
- l'architecture du système,
- les techniques de distribution...etc.

Le rapport, au format PDF, doit refléter votre propre expérience d'utilisation des deux systèmes (tous les points doivent être illustrés p.ex. par des copies d'écran commentées), les difficultés rencontrées, les avantages/inconvénients. Il est également attendu une comparaison des deux systèmes retenus sur leurs points forts et points faibles.

Modalités

Le rapport est à mettre à disposition (lien FileX) au plus tard le **19 février 2018** en laissant actif le lien 3 semaines.