

UN MODÈLE DE COÛT POUR LE NoSQL

Hao ZHANG

Parcours Recherche 2018--2019

6 novembre 2018



ÉCOLE
D'INGÉNIEURS
PARIS-LA DÉFENSE

Sommaire

① Sujet d'Étude

Contexte de la Recherche
Objectif de Projet

② Évaluation de la Base de Données NoSQL

Évaluation des Attributs de Qualité
Aperçu de la Performance

③ Transformation et Migration vers NoSQL

Transformation d'un Modèle Relationnels
Transformation d'un Modèle conceptuel

Contexte de la Recherche

Dans le BigData, les bases NoSQL répondent au problème de stockage et de distribution des données à large échelle pour souvent répondre à des problèmes de performances en centralisé.

De nombreuses solutions existent, et faire le bon choix lors de la définition des besoins d'un Système d'Information devient un choix crucial, car il est difficile et risqué de faire un changement technologique.

Objectif de Projet

Le but de ce projet de recherche est de définir un modèle de coût d'évaluation de requêtes sur des solutions NoSQL.

Ainsi, ce modèle de coût (accès réseaux et accès locaux) sera capable d'estimer pour chaque type de requête ce qu'il en coutera, et ainsi de choisir la solution NoSQL qui minimise ce coût.

L'objectif de ce projet est donc de raffiner les modèles de coût existant dans la littérature pour avoir un modèle plus global, le plus précis possible.

Sommaire

① Sujet d'Étude

Contexte de la Recherche
Objectif de Projet

② Évaluation de la Base de Données NoSQL

Évaluation des Attributs de Qualité
Aperçu de la Performance

③ Transformation et Migration vers NoSQL

Transformation d'un Modèle Relationnels
Transformation d'un Modèle conceptuel

Évaluation des Attributs de Qualité

Conception de la recherche et méthodologie :

- Identifier plusieurs attributs de qualité souhaitables à évaluer dans des bases de données NoSQL.
 - disponibilité, cohérence, durabilité, maintenabilité, performance, fiabilité, robustesse, évolutivité, délai de stabilisation et de rétablissement
- Identifier les systèmes NoSQL les plus populaires et les plus utilisés.
 - Aerospike, Cassandra, Couchbase, CouchDB, HBase, MongoDB and Voldemort
- Explorer la littérature pour évaluer les attributs de qualité sélectionnés sur les bases de données susmentionnées.

Aperçu de la Performance

Conception de la recherche et méthodologie :

- Comparez les cinq bases de données NoSQL les plus courantes en termes de performances des requêtes, basées sur les lectures et les mises à jour, en prenant en compte les charges de travail typiques, telles que représentées par Yahoo! Benchmark Serving Cloud.
 - Cassandra et HBase : bases de données de familles de colonnes.
 - MongoDB et OrientDB : bases de données de familles de document.
 - Redis : base de données de familles de clé-valeur.

Sommaire

① Sujet d'Étude

Contexte de la Recherche
Objectif de Projet

② Évaluation de la Base de Données NoSQL

Évaluation des Attributs de Qualité
Aperçu de la Performance

③ Transformation et Migration vers NoSQL

Transformation d'un Modèle Relationnels
Transformation d'un Modèle conceptuel

Transformation d'un Modèle relationnels

Mappage de Schéma

Procédure générale :

- Phase I, Définir la base de données source.
- Phase II, Classification des tables.
- Phase III, Définition de la base de données cible.
- Phase IV, Execute Data Conversion.
 - denormalization
 - migration

Cas de recherche : SQL à MongoDB, SQL à HBase, SQL à NoSQL,
SQL(CMS, Content Management System) à NoSQL

Cadre de Migration

Cadre capable de prendre en charge facilement la migration d'une base de données relationnelle (par exemple, MySQL) vers une base de données NoSQL (par exemple, MongoDB).

Data Migration Module

Un ensemble de méthodes permettant une migration transparente entre les SGBD (par exemple, de MySQL vers MongoDB).

Module de Mappage de Données

Fournir une couche de persistance pour traiter les requêtes de base de données, tout en renvoyant les données dans un format approprié.

Transformation d'un Modèle conceptuel

Basé sur Model Driven Architecture (MDA)

Procédure générale :

- Phase I, construire les métamodèles du diagramme de classes UML et de la base de données NoSQL.
- Phase II, proposer les règles de mappage entre les deux métamodèles.
- Phase III, construire le diagramme de classes UML, et générer le modèle de base de données NoSQL par transformation.

Cas de recherche : UML à NoSQL, UML à HBase, UML à GraphDB