



Rapport de projet NoSQL

Partie II - Évaluation et Analyse des Performances

El Houiti Chakib
Kezzoul Massili

2 décembre 2021

Introduction

Ce rapport fait suite à sa première partie qui concerne l'implémentation d'un moteur de requête *SPARQL* en étoiles en utilisant l'approche *hexastore*. Dans cette partie, nous allons analyser les performances de notre implémentation par rapport à d'autres implémentation. Notamment celle de Jena.

Dans un premier temps, nous allons préparer et analyser des bancs d'essais en utilisant *WatDiv*¹. *WatDiv* est système développé afin de mesurer les performances d'un moteur de requête *SPARQL*. Il consiste en la génération de jeux de données ainsi que des jeux de requêtes.

Dans un second temps, nous allons définir et comparer plusieurs plans de tests afin d'en trouver un (ou plusieurs) qui donne des résultats correctes, significatifs et interprétables.

Enfin, viendra la partie concrète d'évaluation des performances. On exécutera les plans de tests précédemment réalisés. Suivant les résultats obtenus, nous les présenterons selon des représentations graphiques que nous allons analyser. Nous expliquerons aussi les raisons à la base de ces résultats.

1 Bancs d'essais

1.1 Préparation des bancs d'essais

La première partie est de générer les jeux de données sur lesquelles les tests vont être exécutés. Pour cela, on utilise le programme *WatDiv*.

1.1.1 Génération des données

La génération d'un *dataset* se fait en utilisant la commande suivante :

```
0 ./watdiv -d <model-file> <scale-factor>
```

Pour un *scaleFactor* = 1, on obtient environ 100K triplets. On peut donc augmenter le nombre de triplets en augmentant la valeur de *scaleFactor*.

1.1.2 Génération des requêtes

1.2 Analyse des bancs d'essais

2 Plan des tests

3 Évaluation des performances

Cette partie est en cours de réalisation.

4 Conclusion

4.1 Apport des évaluations

4.2 Perspectives d'amélioration

1. Waterloo SPARQL Diversity Test Suite.