

الجمهورية الشعبية الديمقراطية الجزائرية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
المدرسة العليا للإعلام الآلي 08 ماي 1945 بسيدي بلعباس
École Supérieure en Informatique
-08 Mai 1945- Sidi Bel Abbès



Mémoire de Fin d'étude

Pour l'obtention du diplôme de **Master**

Filière: **Informatique**

Spécialité : **Systèmes d'Information et Web (SIW)**

Thème

**Etude et Comparaison des Techniques d'Optimisation
de Requêtes: Cas de l'Ordre de Jointure**

Présenté par :

- Melle ZEBLAH Maroua

Soutenu le : **16 Juillet 2022** Devant le jury composé de

Pr. MALKI Mimoun	President
Pr. BENSABER Amar Djamel	Examineur
Dr. KECHAR Mohamed	Encadrant
Pr. BELLATRECHE Ladjel	Co-Encadrant

Année Universitaire : 2021/2022

Abstract

The world today is going through the appearance of multiple events that it has never experienced simultaneously in the past, namely: the covid-19 pandemic, the war in Ukraine, and energy conflicts. This has prompted governments, states, users and researchers to propose initiatives to reduce the energy consumption of digital equipment used in work and at home.

As scientists, we are obliged to study the energy efficiency of equipment and their hardware, software, and the applications they use. Database management systems (DBMSs) are becoming energy sinks because of the massive explosion of data they must collect, process and store. The query processor is one of the most energy-consuming components of DBMSs. Initially, its role is to process requests efficiently. Given the volume of data and the complexity of the analysis queries, the study of the energy efficiency of this component undoubtedly becomes a crucial and urgent question. Most current query optimizers are designed to minimize I/O operations and try to use RAM as much as possible. As a result, they generally ignore the energetic aspects. By reviewing the literature, we identified the lack of work on the energy efficiency of the join order. The latter has been widely studied under the performance constraint. In order to address this problem from an energy perspective, we conducted a state-of-the-art work on the work done on query performance oriented join order. Basically, the optimization of join orders followed several approaches: deterministic approach, random search, genetic algorithm, hybrid strategy and optimization technique based on reinforcement learning and deep learning. Each one is different from the other in terms of operating principle, search space and search strategy, etc.

In this master's thesis, we review several join order algorithms of the different approaches mentioned. We also present the experiments of the previous works, make a comparative study and interpret the obtained results.

Key Words: Join Order, Optimizer, solution Space, Execution plan, Query evaluation.

Résumé

Le monde d'aujourd'hui passe par l'apparition des événements multiples qu'il n'a jamais connu simultanément par le passé, à savoir : la pandémie du Covid-19, la guerre en Ukraine, et les conflits énergétiques. Cela a poussé les gouvernements, les états, les usagers et les chercheurs à proposer des initiatives pour réduire la consommation énergétique des équipements numériques utilisés dans le travail et à la maison.

En tant que scientifiques, nous sommes dans l'obligation de mettre la main à la patte en étudiant l'efficacité énergétique des équipements et leurs matériels, logiciels, ainsi que les applications qu'ils utilisent. Les systèmes de gestion de bases de données (SGBDs) deviennent des gouffres énergétiques à cause de l'explosion massive des données qu'ils doivent collecter, traiter et stocker. Le processeur des requêtes constitue l'un des composants le plus énergivore des SGBDs. Initialement, il a pour rôle de traiter d'une manière efficace les requêtes. Compte tenu du volume des données et la complexité des requêtes d'analyse, l'étude de l'efficacité énergétique de ce composant devient sans aucun doute une question cruciale et urgente. La majorité des optimiseurs des requêtes actuels sont conçus pour minimiser les opérations d'entrées-sorties et essaient d'exploiter la RAM autant que possible. En conséquence, ils ignorent généralement les aspects énergétiques. Dans ce travail, nous nous concentrons sur l'ordre de la jointure afin d'étudier son impact énergétique. En examinant la littérature, nous avons identifié l'absence des travaux sur l'efficacité énergétique de l'ordre de jointure. Ce dernier a été largement étudié sous la contrainte de performance. Afin d'aborder ce problème dans une perspective énergétique, nous avons mené un travail d'état de l'art sur les travaux réalisés sur l'ordre de jointure orienté performance de requêtes. Fondamentalement, l'optimisation des ordres de jointure a suivi plusieurs approches : approche déterministe, recherche aléatoire, algorithme génétique, stratégie hybride et technique d'optimisation basée sur l'apprentissage par renforcement et l'apprentissage en profondeur. Chacune est différente de l'autre en termes de principe de fonctionnement, d'espace de recherche et de stratégie de recherche, etc.

Dans ce mémoire de master, nous passons en revue plusieurs algorithmes d'ordre de jointure des différentes approches mentionnées. Nous présentons également les expériences des travaux précédents, faisons une étude comparative et interprétons les résultats obtenus.

Mots clés: Ordre de jointure, Optimiseur, Espace de solution, Plan d'exécution, Evaluation des requêtes.