

Exploitation de l'Interaction des Requêtes OLAP pour la Gestion de Cache et l'Ordonnancement de Traitements

Amira Kerkad*, Ladjel Bellatreche*, Dominique Geniet*

*LIAS/ENSMA- Université de Poitiers
Futuroscope, 86960 - France
(amira.kerkad, bellatreche, ageniet)@ensma.fr

Résumé. Le cache est l'une des composantes principales d'un système de gestion de bases de données (SGBD). Les SGBD manipulant des bases de données volumineuses comme les entrepôts de données stockent souvent les données sur le disque. En conséquence, l'interrogation nécessite un transfert des données du disque vers la mémoire centrale via le tampon. Un nombre important de travaux sur la gestion de tampon ont été proposés. Malheureusement, ils supposent que les requêtes soient ordonnées. Dans le contexte des entrepôts de données, les requêtes interagissent du fait qu'elles utilisent la table des faits. Cette interaction pourrait impacter la gestion de cache et offrir un bon ordonnancement de requêtes. Dans cet article, nous proposons d'étudier conjointement le problème de gestion de tampon (BMP) et le problème d'ordonnancement de requêtes (QSP). Trois algorithmes sont proposés pour résoudre le problème conjoint. Finalement, un simulateur et une validation sous Oracle11G sont proposés.

1 Introduction

La majorité des SGBD exploite la mémoire cache pour réduire le coût des requêtes. Le gestionnaire du cache est l'une des principales composantes d'un SGBD (Effelsberg et Härder (1984)), surtout avec un grand volume de données comme dans les Entrepôts de Données (ED)¹. La technologie ED est le noyau des applications de BI (Business Intelligence), qui offre une capacité de stockage des données de sources hétérogènes, et des solutions d'accès efficaces par des requêtes OLAP (On-Line Analytical Processing). Actuellement, les applications de BI gèrent d'immenses volumes de données, ainsi les requêtes deviennent très gourmandes en temps de réponse. Les applications des ED Relationnels sont souvent modélisées par des schémas en étoile². La Figure 1 montre un exemple de schéma en étoile pour le Star Schema Benchmark (O'Neil et al. (2009)) contenant une table de faits et des tables de dimension. Les requêtes qui interrogent ce genre de schéma sont dites *requêtes de jointure en étoile*. Dans les EDR, les jointures portent sur la table des faits, ainsi la présence de résultats communs entre les requêtes est fréquente. Ce phénomène est connu dans d'optimisation multi-requêtes (Sellis (1988)) et il a été largement considéré dans la sélection de structure d'optimisation comme

1. Par exemple, l'ED eBay contient 2 petaoctets de données.

2. Le schéma en flocon de neige est moins utilisé que le schéma en étoile