

Résumé

Nous visons à travers cette thèse à proposer un ensemble d'approches permettant d'optimiser les entrepôts de données et d'aider l'AED à bien mener cette optimisation. Nos approches d'optimisation reposent sur l'utilisation de trois techniques d'optimisation : la fragmentation horizontale primaire, dérivée et les index de jointure binaires (IJB). Nous commençons par proposer une approche de fragmentation qui prend en considération à la fois la performance (réduction du coût d'exécution) et la manageabilité (contrôle du nombre de fragments générés). Nous avons ensuite proposé une approche gloutonne de sélection d'IJB. Cette approche commence par trouver une solution initiale ensuite l'améliorer en utilisant plusieurs stratégies. L'utilisation séparée de la fragmentation horizontale (FH) et des IJB ne permet pas d'exploiter les similarités existantes entre ces deux techniques. Nous avons proposé une approche de sélection conjointe de la FH et des IJB. Cette approche peut être utilisée pour le tuning de l'entrepôt. Nous avons mené plusieurs expérimentations en utilisant des modèles de coût mathématiques suivies d'autres effectuées sur un benchmark sous Oracle 10g pour valider nos différentes approches.

Mots-clés: Conception Physique, Tuning, Techniques d'Optimisation, Fragmentation Horizontale, Index de Jointure Binaires.

Abstract

The main goal of this thesis is to propose a set of approaches to optimize performance of data warehouses and assist the data warehouse administrator (DWA) to well perform this optimization. Our approaches use three optimization techniques : primary and derived horizontal partitioning and bitmap join indexes (BJI). First, we propose a partitioning approach, which takes into account both the performance (reducing the response time) and manageability (controlling the number of fragments). Then, we propose a greedy approach that selects a set of BJI. This approach starts with an initial solution then improves it using multiple strategies. The use of FH and BJI independently cannot exploit the various similarities existing between these two techniques. We propose an approach for multiple selection of FH and BJI. This approach can be used for tuning data warehouses. We conducted several experiments using mathematics cost models then using a benchmark dataset under Oracle 10g to validate our approaches.

Keywords: Physical Design, Tuning, Optimization Techniques, Horizontal Partitioning, Bitmap join index