

Maintenance de charge pour l'optimisation des entrepôts de données évolutifs : aide à l'administrateur

Cécile Favre, Fadila Bentayeb, Omar Boussaid

Université de Lyon (Laboratoire ERIC - Lyon 2)
5 av. Pierre Mendès-France, 69676 Bron Cedex
{cfavre|bentayeb}@eric.univ-lyon2.fr, omar.boussaid@univ-lyon2.fr
<http://eric.univ-lyon2.fr>

Résumé. Dans un contexte où les entrepôts de données sont amenés à subir des évolutions, nous proposons d'aider l'administrateur à la maintenance de la charge (ensemble de requêtes) qui sert à l'évaluation des performances. En répercutant les évolutions de l'entrepôt de données sur la charge, il est alors possible pour l'administrateur d'avoir une gestion pro-active des performances, évitant d'attendre que l'entrepôt mis à jour soit interrogé par les utilisateurs pour construire une charge valide pour optimiser les performances.

Mots-clés : Entrepôt de données, modèle, évolution, optimisation, maintenance de charge, requête décisionnelle, administrateur.

1 Introduction

Les entrepôts de données présentent une modélisation multidimensionnelle définie en fonction des sources de données et des besoins d'analyse. L'évolution de ces deux derniers «paramètres» peut alors nécessiter une évolution du modèle multidimensionnel, définissant ainsi un contexte évolutif.

Parallèlement, un des objectifs principaux de la constitution des entrepôts de données est l'analyse dite «en ligne» (OLAP : On-Line Analytical Processing). Or, Krompass et al. (2007) ont précisé que, dans le contexte des entrepôts de données, les requêtes peuvent être très rapides (quelques minutes) mais peuvent également prendre plusieurs heures. Il est alors crucial pour l'administrateur de pouvoir gérer la performance de l'entrepôt, en particulier dans le contexte évolutif. Se pose alors le problème de l'évaluation de ces performances dans ce contexte évolutif. Classiquement la gestion des performances doit se faire en fonction de l'utilisation de l'entrepôt. Ainsi, l'évaluation de performances se base sur un ensemble de requêtes utilisateurs, appelé charge.

Dans un contexte évolutif, nous pensons qu'il est alors intéressant pour l'administrateur d'avoir une démarche pro-active en matière de gestion de performances, et de ne pas attendre que l'entrepôt de données mis à jour soit utilisé par les utilisateurs pour obtenir une nouvelle charge de requêtes cohérente avec le modèle courant. Précédemment, nous avons proposé une première approche simpliste d'évolution de charge automatique (Favre et al., 2007). Dans cet article, nous nous focalisons davantage sur l'aide à l'administration que nous pouvons apporter pour la maintenance de la charge. L'objectif est d'aider l'administrateur d'une part à maintenir