

Une approche de répartition des données d'un entrepôt basée sur l'Optimisation par Essaim Particulaire

Hacène Derrar*, Omar Boussaid**, Mohamed Ahmed-Nacer*

* Laboratoire LSI, Département Informatique, USTHB Alger
Bp 32 El Alia, bab Ezzouar/ Ager

Hderrar@yahoo.fr, Anacer@mail.cerist.dz

** Laboratoire ERIC, Université Lyon2
Campus Porte des Alpes, 69676 Bron Cedex
Omar.Boussaid@univ-lyon2.fr

Résumé. Dans le contexte des entrepôts de données, le partitionnement des tables, des index et des vues matérialisées en fragments stockés et consultés séparément apporte des améliorations considérables en terme de gestion des données et de coût de traitement. Lors de leurs conceptions, ces techniques se basent sur l'analyse d'informations statistiques recueillies à partir des requêtes les plus fréquentes. Cependant, en raison de leurs caractéristiques les requêtes décisionnelles rendent ce contexte d'analyse très variable. Ceci rend avec le temps, les schémas de fragmentation réalisés inappropriés et par conséquent une dégradation des performances du système. A partir de ce constat, et considérant que la validité d'une approche de partitionnement est soumise à l'épreuve du temps et dépend des besoins et de l'environnement, on propose dans ce papier une approche de répartition des données de l'entrepôt basée sur une méthode issue des approches biomimétiques.

1 Introduction

Dans le contexte des entrepôts de données, le choix de la méthode et du schéma de fragmentation conjugué avec l'évolution perpétuelle des systèmes, des applications et enfin des données constituent les véritables problèmes qui entravent une meilleure exploitation de cette technique. En effet, une approche de fragmentation inappropriée et mal conçue influe considérablement sur les performances du système et plus particulièrement lors de l'exécution des opérations coûteuses telles que la jointure et la multi-jointure qui caractérisent les requêtes décisionnelles.

L'approche présentée dans ce papier s'inscrit dans le cadre d'une solution aux problèmes induits par l'application d'une stratégie de fragmentation inadaptée aux changements perpétuels du contexte dans lequel est exploité l'entrepôt de données. Cette solution consiste à mettre en œuvre un processus d'automatisation visant la redistribution des données à l'issue d'une dégradation des performances observée lors de l'exploitation de l'entrepôt. Nous proposons dans cet article, une approche de répartition des données basée sur une technique d'optimisation issue des approches biomimétiques. Celle-ci permet, à l'issue d'une phase d'évaluation, de déterminer un schéma de fragmentation optimisant les performances. Cet article est organisé comme suit, dans la première section, nous présentons un état de l'art sur