

TITRE: Optimisation de requêtes par matérialisation dans un environnement distribué.

EQUIPE/THEME: Mabiovis (Visu) & CombAlgo (Algo Distribuée)

DIRECTEURS: N. hanusse & S. maabout

COURRIELS: {hanusse, [maabout](mailto:maabout@labri.fr)}@labri.fr

MOTS-CLES: Matérialisation, indexation, structures de données, requêtes, systèmes distribués, optimisation, algorithmique, algorithmes approchés,

DIRECTEURS HABILITES: N. Hanusse

DESCRIPTION du SUJET:

La matérialisation (ou pré-calcul de résultats) est une des techniques permettant d'optimiser les requêtes. Le problème qui se pose alors peut être décrit comme suit: étant donné un ensemble de requêtes cibles (workload), quelles sont les "meilleures" parties de la base de données qu'on doit matérialiser afin d'optimiser le workload ? La définition de "meilleure" peut prendre plusieurs dimensions selon la fonction de coût que l'on veut minimiser.

De premiers résultats ont été obtenus dans le cadre des cubes de données en tenant compte d'une famille restreinte de requêtes. Aussi, la dimension dynamique (des données et des requêtes) n'a pas été complètement prise en compte. Le travail demandé consiste à étendre les solutions au cas de requêtes plus complexes telles que les Skylines, Top-K, ou plus généralement les requêtes avec préférence. Dans un contexte interactif, l'utilisateur peut se contenter d'un « résumé » du résultat suffisamment significatif pour l'aider à préciser sa requête. La matérialisation peut être utilisée afin trouver « rapidement » les résumés. Enfin, on envisage d'étudier les problèmes traités dans un contexte distribué : non seulement on cherche les requêtes à pré-calculer mais aussi (i) par qui les calculer, (ii) où les stocker et (iii) comment les exploiter ? Le travail demande une étude théorique et une capacité à développer des programmes pour des expérimentations.