

Bases de données avancées

Exercices dirigés : optimisation

Une BD relationnelle contient les tables

Commande(*numéro*, idClient, produit, date)

Client(*idClient*, nom, destination)

La table Commande contient 100 000 lignes et la table Client 5000 lignes. Les tailles des champs sont (en octets) : numéro(10), idClient(20), produit(40), date(10), nom(80) et destination(100).

On considère que chaque produit apparaît dans environ 3% des commandes et que le nombre de commandes est en moyenne le même pour chaque client de la table Client.

Une page disque a une taille utile de 2000 octets. Une adresse de page/article occupe 10 octets. Une lecture/écriture de page prend en moyenne 10 ms en accès aléatoire et 1 ms en accès séquentiel.

- A. Soit la requête qui demande les dates où l'on a commandé du toner.
1. Exprimer la requête en SQL.
 2. Quel est le plan d'exécution s'il n'y a pas d'index ? Quel sera le temps d'exécution ?
 3. Même question s'il existe un index arbre B+ sur *Commande.produit*. On suppose que l'index a 3 niveaux.
- B. Même questions pour la requête qui demande les noms des clients qui ont commandé du toner. Considérer les cas suivants :
1. Aucun index. Comparez les plans avec boucles imbriquées et avec tri-fusion, en calculant le temps d'exécution dans chaque cas. On suppose qu'on dispose de M=11 pages mémoire pour la jointure (pour la lecture des tables par boucles imbriquées / pour le tri).
 2. Plan d'exécution qui utilise un index sur *Client.idClient* (arbre B+ à 3 niveaux).
 3. Index sur *Commande.idClient*.
 4. Index sur *Client.idClient* et sur *Commande.produit*.
 5. Index sur *Commande.idClient* et sur *Commande.produit*.