

Université Paris Nanterre
Master 1 Miage - BDA
TD Optimisation de requêtes

Exercice 1 :

Soit la base de données suivante :

Client(idClient, nom, prénom, adresse, tél)

Produit(idProduit, libelle, prix)

Commande(idCommande, idClient, idProduit, date, quantité)

Pour les requêtes ci-dessous, construisez l'arbre algébrique optimal :

a)

```
SELECT c.idClient, p.libelle
```

```
FROM Client c, Produit p, Commande co
```

```
WHERE c.idClient=co.idClient and p.idProduit= co.idProduit and co.date= '15/02/2017'
```

b)

```
SELECT p.libelle, co.quantité
```

```
FROM Produit p, Commande co
```

```
WHERE p.idProduit= co.idProduit and prix > 300
```

c)

```
SELECT c.nom, co.date, p.libelle
```

```
FROM Client c, Produit p, Commande co
```

```
WHERE c.idClient=co.idClient and p.idProduit= co.idProduit and co.quantité>40
```

Exercice 2 :

Soit la base STATION DE SKI de schéma :

- station (noms, gare, type, telephone,...)
- activité (type_activité, noms_a, contraintes,...)

Pour la requête suivante, indiquez :

- ✓ l'arbre algébrique de la requête
- ✓ l'arbre algébrique optimal

Requête : Quels sont les noms des stations proposant une activité 'tennis' et quelles sont les gares où elles se situent ?

```
SELECT noms, gare
```

```
FROM station, activité
```

```
WHERE type_activité = 'tennis'
```

```
AND station.noms=activité.noms_a
```

Exercice3 :

Soit la base de données suivante :

- Editeur (Id-editeur, nom, rue, ville, pays)
- Livre (Id-livre, titre, Id-editeur)

L'attribut Id-editeur est codé sur 10 octets et titre sur 100 octets.

Les pages de la base occupent 4000 octets. La mémoire peut contenir 101 pages. Chaque relation (Editeur et Livre) a un index arbre B+ sur les attributs clés (Id_editeur et Id_livre respectivement).

Les statistiques contenues dans le catalogue sont :

- NTuples(Editeur) = 15000
- BFactor(Editeur) = 20 \Rightarrow NBlocks (Editeur) = 15000/20 = 750
- NLevelId-editeur(I) = 3
- NTuples(Livre) = 400000
- BFactor(Livre) = 10 \Rightarrow NBlocks(Livre) = 400000/10 = 40000
- NDistincPays(Editeur) = 4 \Rightarrow SCPays(Editeur) = 15000/4 = 3750
- NLevelId-livre(I) = 4

Pour la requête « quels sont les titres des livres édités par les éditeurs du pays 'P' »

SELECT Titre

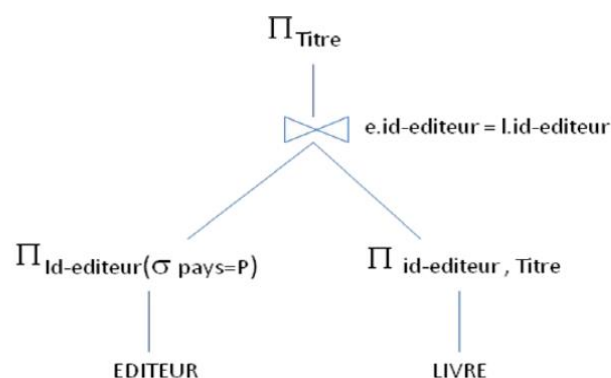
FROM Editeur e, Livre l

WHERE e.Id.editeur = l.Id-editeur

AND e.Pays = 'P'

On suppose que le résultat donne $\frac{1}{4}$ des livres.

Supposons qu'on exécute la requête en utilisant le plan d'exécution représenté par l'arbre algébrique ci-dessous et l'algorithme des boucles imbriquées en utilisant au maximum l'espace du buffer de mémoire. Évaluez le coût de la requête.



Exercice4:

Soit la base de données suivante :

Client(idClient, nom, prénom, adresse, tél)

Produit(idProduit, libelle, prix)

Commande(idCommande, idClient, idProduit, date, quantité)

L'attribut IdClient est codé sur 10 octets et nom sur 100 octets.

L'attribut idProduit est codé sur 20 octets.

Les pages de la base occupent 5000 octets. Chaque relation a un index arbre B+ sur les attributs clés (idClient, idProduit et idCommande).

Les statistiques contenues dans le catalogue sont :

- NTuples(Client) = 10.000
- BFactor(Client) = 20 => NBlocks (Client) = $10000/20 = 500$
- NLevelIdClient(I) = 3
- NTuples(Produit) = 50.000
- BFactor(Produit) = 50 => NBlocks(Produit) = $50000/50 = 1000$
- NLevelIdProduit(I) = 4
- NTuples (Commande) = 200.000
- BFactor(Commande) = 40 => NBlocks (Commande) = $200.000/40 = 5000$
- NLevelIdCommande(I) = 5

On suppose que 20% des produits uniquement ont un prix supérieur à 1000 euros.

Soit la requête « Donner le nom des clients ayant commandé au moins une fois un produit dont le prix est supérieur à 1000 euros »

L'algorithme de jointure utilisé est celui des boucles imbriquées indexées.

Questions :

- 1- Ecrivez la requête en SQL
- 2- Dessinez l'arbre algébrique optimal
- 3- Évaluez le coût d'exécution de la requête