

Exercice n°1 : Les différentes exécutions de la jointure

Supposons que les n-uplets des relations sont stockés dans des fichiers dans l'ordre ci-dessus et que chaque page de fichier (pour R comme pour S) contient au plus deux enregistrements. Pour chacune des méthodes : (1) listez les couples de n-uplets de R et S dans l'ordre de traitement ; (2) donnez une estimation du coût de la méthode.

R	
A	C
a2	c4
a1	c3
a3	c5
a1	c1
a1	c2
S	
A	B
a5	b6
a6	b7
a1	b1
a2	b3
a4	b2
a2	b4

Q1 : Jointure itérative “brute” (parcours des n-uplets)

Q2 : Jointure itérative page à page

Q3 : Jointure itérative par paquet (en supposant que le buffer est de taille 4)

Q4 : Jointure itérative avec index B+ ($h = 1$) non groupant de type 2 sur A pour S

Q5 : Jointure par tri-fusion (**Partie traitement seulement**)

Q6 : Jointure par hachage en supposant que la fonction de hachage appliquée est définie comme suit : $h(a1)=h(a4)=1$ $h(a2)=h(a3)=2$ $h(a5)=h(a6)=3$

Exercice n°2 : Exercice 2 : Évaluation du coût de la jointure

On considère la jointure $R \bowtie_{R.A=S.B} S$ et les éléments suivantes :

- La relation R contient 10 000 enregistrements et 10 enregistrements par page
- La relation S contient 2 000 enregistrements et 10 enregistrements par page
- Une page issue de la jointure ne peut contenir que 5 éléments
- L'attribut B de S est une clé primaire
- La taille du buffer est de 52 pages

Dans cet exercice, on considère pour le coût le nombre de lectures/écritures sur le disque.

Q1 : Quel est le coût de la jointure de R et S pour la méthode par itération page à page ? Quel est le nombre minimum de pages du buffer requises pour maintenir ce coût ?

Q2 : Quel est le coût minimal de la jointure de R et S pour la méthode itérative par paquets ? Quel est le nombre minimum de pages du buffer requises pour maintenir ce coût ?

Q3 : Quel est le coût de la jointure de R et S pour la méthode de tri-fusion ? Quel est le nombre minimum de pages du buffer requises pour maintenir ce coût ?

Q4 : Quel est le coût de la jointure de R et S pour la méthode de hachage ? Quel est le nombre minimum de pages du buffer requises pour maintenir ce coût ?

Q5 : Quel serait le coût minimum pour la jointure de R et S et avec quel algorithme ? Combien d'espace buffer cela nécessite-t-il ?

Q6 : Combien de pages au maximum sont produites par la jointure de R et S ?

Exercice n°3 : Exercice 3 : Jointure et tri fusion

Q1 : Considérez la jointure $Film \bowtie_{id_rel=id_art} Artiste$ en supposant que la table film occupe 1000 pages et la table Artiste 10 000 pages. On suppose que la mémoire disponible est de 252 pages.

- Quel est le coût de l'évaluation de cette jointure par l'algorithme par itération page à page ? par itération par paquet (utilisation du maximum de la mémoire disponible) ?
- Pour chacune des deux méthodes comparez entre le cas où Film est choisie en relation externe et le cas où Artiste est choisie comme relation externe ?

Q2 : Coût du tri fusion Supposons que le fichier que nous souhaitons trier occupe 75 000 pages de 4000 octets. Quel est le nombre d'étapes de fusion (on utilisera du tri multi-fusion) pour les différentes tailles de buffer suivantes : 400Mo, 2Mo, 1Mo ?