

Examen – Bases de Données Avancées – L3 Informatique – 2018-2019

pages : 1-2

Exercice 1 (4 points)

Soit un hash index extensible dont la taille maximale d'une bucket est 2, initialement vide, qui utilise la fonction de hachage (hash function) suivante :

x	h(x)
16	10101011
22	11110001
17	11010110
28	00000110
50	11110010
24	01000101

Insérez dans cet index les entrées de données suivantes : 24*, 22*, 16*, 28*, 50*, 17*.
Montrez l'état de l'index après chaque insertion.

Exercice 2 (6 points)

Soit un index B+Tree d'ordre 1, initialement vide.

2.1. (4 points)

Insérez une à une dans cet index les entrées de données suivantes : 7*, 29*, 9*, 24*, 14*, 32*.
Montrez l'état de l'index après chaque insertion.

2.2. (2 points)

Insérez (en partant de l'index vide) les mêmes entrées que ci-dessus, mais en utilisant un chargement par lot (bulk loading). Détaillez les étapes du bulk loading.

Exercice 3 (5 points)

Soit les relations A et B aux caractéristiques suivantes :

- A : schéma A(A1, A2, A3), 200 tuples / page

- B : schéma B(B1, B2), 40 tuples / page.

Supposons une instance de A avec 2000 tuples, dont 600 ont la valeur 4 sur A1, 600 ont la valeur 8 sur A1, et le reste ont la valeur 12 sur A1.

Supposons une instance de B avec 120400 tuples, dont 80000 ont la valeur 17 sur B1, et le reste ont la valeur 10 sur B1.

Supposons une jointure de A et B sur A1=B1.

Rappel : nous calculons le coût d'un algorithme de jointure en I/Os et *en ignorant le coût de l'écriture du résultat final de la jointure*.

3.1. (1 point)

Supposons l'usage d'un algorithme simple de jointure à boucles imbriquées (simple nested loops join). Quel choix de relation externe engendre le coût le plus bas de la jointure ? Justifiez.

3.2. (4 points)

Supposons maintenant l'usage d'un algorithme de jointure à base de hachage (hash join) qui dispose d'un total de 6 buffers (rappel : n'oubliez pas le buffer réservé pour le résultat) et utilise A comme relation «externe».

Laquelle parmi les deux fonctions de hachage ci-dessous mène au coût le plus bas de la jointure? Justifiez.

a. $h(x) = x \bmod 4$

b. $h(x) = x \bmod 3$

Exercice 4 (2 points)

Soient les relations A, B, C, D de schémas suivants (nous omettons le type des attributs):

A(A1, A2, A3, A4, A5)

B(B1, B2, B3, B4)

C(C1, C2, C3, C4, C5, C6)

Soit l'expression A.R. (Algèbre Relationnelle) suivante :

$\Pi_{A2} (\sigma_{C1 > 12 \wedge B3 = 2} (B \bowtie_{B2 = A3 \wedge B1 = C3} (C \bowtie_{A5 = C5} A)))$

Transformez cette expression en une expression A.R. équivalente, en poussant les sélections (σ) et projections (Π) « en bas » autant que possible (appliquer donc les sélections et projections le plus tôt possible). Montrez chaque étape de la transformation (1 étape = application d'une règle d'équivalence).

Exercice 5 (3 points)

Soit un Buffer Manager (BM) qui dispose de 4 cases (frames) numérotées de 1 à 4. Supposons des PageId de la forme PageN, avec $N \geq 1$ entier. Le BM traite donc des demandes de la forme :

- Get (PageN)

- Free (PageN, dirty) avec dirty = 1 si la page a été modifiée, 0 autrement.

Si, lors d'une demande Get nécessitant le chargement d'une page depuis le disque, le BM dispose de cases vides, le BM utilise la première de ces cases (celle du plus petit indice).

Soit la série de demandes :

Get(Page7)

Get(Page2)

Free(Page2,1)

Get(Page4)

Get(Page7)

Free(Page4,0)

Get(Page19)

Get(Page2)

Free(Page7,1)

Get(Page4)

Free(Page19,0)

Free(Page2,0)

Get(Page3)

Free(Page4,0)

Quelle est la première case dont le contenu est remplacé ? Quel est son contenu ? Est-ce que ce contenu est écrit sur le disque ? Justifiez :

a. Dans le cas d'une politique de remplacement LRU (« le moins récemment utilisé »).

b. Dans le cas d'une politique de remplacement Clock (« horloge »).