

Concepts Avancés de Bases de données

Joaquim LEFRANC et Jérôme SKODA

November 19, 2017

1 Nouveauté depuis le TP5

- bufferExtended supporte maintenant les nombres
- Nouvelle structure: DiskOutput (gestion des ecriture sur disque)
- Nouvelle structure: Table (table de Hash sur disque)
- Nouvelle structure: Bucket (bucket de table)
- Ajout de fonction de benchmark Lecture/ecriture dans Buffer et Table
- Nouveau tests
- Nouvelle démo
- Nouvelle commande make gen-tp5 et gen-tp6 (pour generer R et S)

2 Comparaison de Disque Nested loop join avec Disk Block Hash Join

La comparaison s'effectue sur les même disque R et S contenant chacun 256 valeurs. Les buffers sont de taille 10 supportant des données de taille décimales sur 2 octets. La table de hash a 10 bucket avec un hash linaire.

2.1 Hash join

- Création de table:
 - Disque R: Lecture de 256 lignes sur 26 fichiers
 - Disque S: Lecture de 256 lignes sur 26 fichiers
 - Table R: Ecriture de 256 lines sur 29 fichiers
 - Table S: Ecriture de 256 lines sur 32 fichiers
 - Total opération: 1024 ligne sur 113 fichiers
- Jointure:
 - Table R: Lecture de 256 lignes sur 29 fichiers
 - Table S: Lecture de 594 lignes sur 73 fichiers

- Disque RS: Ecriture de 32 lignes sur 4 fichiers
- Total opération: 882 ligne sur 106 fichiers
- Total opération: 1906 ligne sur 219 fichiers

Remarque: Les opérations de lecture sur Table de S peuvent être minimisées jusqu'à 256 lectures s'il n'y a pas de collision.

2.2 Nested Loop join

- Disque R: Lecture de 256 lignes sur 26 fichiers
- Disque S: Lecture de 6406 lignes sur 651 fichiers
- Disque RS: Ecriture de 32 lignes sur 4 fichiers
- Total d'opération: 6694 ligne sur 681 fichier

3 Comment compiler le projet

3.1 Avec le terminal

- make all : Compile tout les fichiers
- make test : Lancement de la série de tests automatiques
- make doc : Génération de la documentation (doxygen)
- make rapport : Génération du rapport (latex)
- make clean : Nettoyage du projet (suppression des objets et binaires)
- make demo-tp1 : Lancer la démo tp1
- make demo-tp2 : Lancer la démo tp2
- make demo-tp3 : Lancer la démo tp3
- make demo-tp4 : Lancer la démo tp4
- make demo-tp5 : Lancer la démo tp5
- make demo-tp5 : Lancer la démo tp6 (Nouveau)
- make rm-rs : Supprime le fichier res/RS.txt et res/disk/RS.txt

4 Arborescence

- bin : Binaire exécutable
 - demo : Exécutable de démonstration
 - test : Exécutable de test
- doc : Documentation doxygen sous différents formats

- rapport : Source du rapport
- res : Ressources necessaire au projet (fichier de bdd)
- script : Script utilisé pour les test
- src : Source du projet
 - bdd : Source de la bibliothèque
 - demo : Sources des differentes démonstrations d’utilisation
 - test : Sources des dufferents tests
- sujet.pdf : Sujet du projet
- README.md : Le readme du projet
- rappot.pdf : C’est moi
- refman.pdf : Documentation format pdf

5 Caracteristiques

- Le code est organisé
- Il y a des code des tests
- Il y a la doc
- Il y a un rapport
- Et il y a pleins d’autre chose

6 Démonstration

Les sources de demosntration sont diponible dans: src/demo Les exécutable de test sont généré dans: bin/demo La commande make pour lancer les demo sont: make demo-tp1, make demo-tp2, make demo-tp3 etc...

- tp1-natural-join: Natural join R et S
- tp2-merge-join-without-duplicate: Merge join sans duplication
- tp3-merge-join-with-duplicate: Merge join avec duplication
- tp4-hash-join : Hash join
- tp5-nested-loop-disk: Nested loop sur disque (Nouveau)

7 Test unitaire

Les sources de test sont disponibles dans: `src/test` Les exécutable de test sont générés dans: `bin/test` Le script de test est dans `script/test.sh` La commande `make` pour lancer les tests est: `make test`

- 00-storeFileBuffer: Ecriture d'un buffer dans un fichier
- 01-natural-join-1: Natural join R et S
- 02-natural-join-2: Natural join S et R
- 03-buf-quick-sort: Fonction de tri d'un buffer
- 04-merge-join-without-duplicate-1: Merge join sans duplication R et S
- 05-merge-join-without-duplicate-2: Merge join sans duplication S et R
- 06-merge-join-with-duplicate-1: Merge join avec duplication R et S
- 07-merge-join-with-duplicate-2: Merge join avec duplication S et R
- 08-hash-put-equilibre : Test ajout équilibré dans une table de hash
- 09-hash-put-desequilibre : Test ajout déséquilibré dans une table de hash
- 10-hash-full : Test remplissage complet dans une table de hash
- 11-hash-get : Test récupération d'une entrée dans la table de hash
- 12-hash-remove : Test suppression / rehash dans une table de hash
- 13-hash-join : Test hash join
- 14-buffer-read-file: test la lecture avec un buffer extended (Nouveau)
- 15-buffer-read-file-2: test la lecture avec un buffer extended (Nouveau)
- 16-disk-buffer-dump: test la lecture d'un disque (Nouveau)
- 17-disk-nested-loop-r-to-s-test: test le nested loop sur disque r to s (Nouveau)
- 18-disk-nested-loop-s-to-r-test: test le nested loop sur disque s to r (Nouveau)

```
.../Documents/Master2/CavBDM2/tp2
> make test
Script de test automatisé
[PASS] buf-quick-sort
[PASS] merge-join-1
[PASS] merge-join-2
[PASS] natural-join-1
[PASS] natural-join-2
[PASS] storeFileBuffer

Script de test terminé
.../Documents/Master2/CavBDM2/tp2
```