

Rapport d'OS

Andrius Ezerskis & Moïra Vanderslagmolen & Hasan Yildirim

November 1, 2022

1 Introduction

2 Stockage de la base de données

Afin de stocker les données de milliers d'étudiants, plusieurs choix se sont offerts à nous. Tout d'abord, nous avons utilisé `std::map`, comparable à un dictionnaire en python, dont les clés ont été l'identifiant de l'élève et les valeurs la struct `student`. Seulement, cela prenait beaucoup de place en mémoire et nous avons donc délaissé l'idée. Nous avons donc décidé de faire un tableau.

3 Checker l'identifiant

Pour voir si l'identifiant est déjà présent dans la base de données, nous avons pensé à mettre chaque étudiant dans son emplacement "correspondant" dans le tableau (étudiant `id=500` va à l'emplacement `db[500]`). La recherche par `id` et le `delete` auraient été rapides. Seulement, si l'utilisateur introduit un `id` très grand, le stockage de la data base sera très grand, pour un seul étudiant. Nous avons donc essayé d'implémenter une boucle `for`, qui vérifiait si l'`id` de l'étudiant n'était pas encore présent en itérant dans toute la base de données. Cela prenait beaucoup trop de temps, et nous avons donc décidé de partir sur une recherche dichotomique et de trier les étudiants par `id`.

4 Fin d'une query

Pour signaler la fin d'une query, nous avons implémenté une message queue grâce à une pipe. Lorsqu'on envoie une query à un processus, nous incrémentons une variable globale `OperationInProgress` qui indique le nombre d'opérations en cours. Lorsque qu'un processus termine sa query, il écrit dans sa pipe "SUCCESS". Le processus principal va lire cette pipe et va décrémenter la variable globale au fur et à mesure. Le pipe est non-bloquant, ce qui permet au processus principal de continuer même si il ne lit rien dans la pipe. Lors d'une transaction, on limite `OperationInProgress` à 1.

5 Gestion de signaux

Plusieurs choix se sont offerts à nous afin de terminer un programme. Nous avons d'abord pensé à 4 pipes qui indiquent au programme principal la fin d'une requête, mais ce n'était pas optimal. Nous avons repris la message queue qui traite les query, et dans cette message queue, le processus principal écrit un message, "kill". Les processus enfants le liront après avoir terminé leur requêtes, et le kill arrêtera la boucle et exit le programme.

6 Classes

Nous avons mis la `database_t` et le `query_result_t` dans une classe, afin de rendre le code plus lisible. Par conséquent, nous avons implémenté un constructeur dans `query_result_t`, au lieu du `query_result_init`, pour que le code crée directement le `query_result_t`, sans devoir appeler l'initialisation à chaque création d'objet. Dans `query result t`, nous avons aussi ajouté `log query`, pour pouvoir mettre les attributs en privés. Nous n'avons pas mis `student_t` en classe, parce que les attributs d'un student sont souvent accédés par la classe `query result t`. et nous n'aurions donc pas su mettre les attributs en privés.

7 Améliorations possibles

7.1 `std::map`

Nous avons comme idée de stocker, dans une structure de données `std::map` (équivalent d'un dictionnaire en Python), les identifiants des étudiants en clés ainsi que leur indice de position dans la base de données en valeurs de ces clés.

Selon nous, cela aurait permis de faciliter et d'améliorer l'efficacité de la vérification de si un étudiant est déjà dans la base de données ou non.

En effet, lors de l'ajout d'un étudiant, nous aurions juste à vérifier si l'identifiant est présent dans le map pour éviter d'avoir des doublons dans la base de données.

7.2 Threads

Au lieu d'utiliser des processus, nous aurions pu utiliser des threads. Seulement, si le processus main crashe, alors tous les threads crashent aussi. L'avantage des processus est leur indépendance, ce qui les rend plus stable que des threads.