

PCO Laboratoire 6

Producteur-Consommateur pour calcul différé (Hoare)

Benoît Delay, Eva Ray

January 28, 2024

1 Etape 1

Dans cette étape, nous avons mis en place la distribution des calculs grâce aux méthodes `requestComputation` et `getWork`.

1.1 Choix d'implémentation

Pour représenter le buffer contenant les calculs à réaliser, nous avons choisi d'utiliser une `std::map` ayant comme clé le type du calcul et comme valeur une `std::list` contenant les calculs sous forme de `Request` à réaliser de ce type. Notre technique consiste simplement à ajouter les calculs au début de la liste correspondante dans la map dès que `requestComputation` est appelée. Cependant, chaque liste a une taille maximale définie par la constante `MAX_TOLERATED_QUEUE_SIZE`. Pour modéliser cette contrainte, nous avons créé un tableau nommé `fullQueuePerType` contenant une `Condition` pour chaque type de calcul. Avant d'ajouter le calcul à la liste, on vérifie donc que la taille de la liste est inférieure à `MAX_TOLERATED_QUEUE_SIZE`. Si ce n'est pas le cas, on appelle la fonction `wait` sur la condition correspondante dans `fullQueuePerType` et on attend qu'un calcul du bon type arrive dans le buffer. La méthode `requestComputation` est aussi en charge de donner un indice unique à chaque calcul. Pour cela, nous avons utilisé un attribut statique `nextId` qui est incrémenté à chaque fois qu'un calcul est ajouté au buffer.

La méthode `getWork`, quant à elle, demande du travail d'un certain type. Pour cela, elle vérifie que la liste correspondante dans la map n'est pas vide. Si c'est le cas, elle appelle la méthode `pop_back` sur la liste afin de récupérer le dernier élément de la liste, qui sera donc le plus ancien. Si la liste est vide, elle appelle la méthode `wait` sur la condition correspondante dans `emptyQueuePerType` qui est un autre tableau contenant une `Condition` pour chaque type de calcul qui représente le fait qu'il n'y a pas de calcul de ce type dans le buffer. Dans ce cas, on attend donc qu'un calcul de ce type soit ajouté au buffer avant de pouvoir continuer.

Lorsqu'un calcul est ajouté au buffer dans la méthode `requestComputation`, on signale sur la condition correspondante dans `emptyQueuePerType` que le buffer contient maintenant un calcul de ce type. De même, lorsqu'un calcul est retiré du buffer dans la méthode `getWork`, on signale sur la condition correspondante dans `fullQueuePerType` que le buffer contient maintenant un calcul de ce type en moins.

D'une manière générale, nous appelons `monitorIn` et `monitorOut` dans toutes les méthodes, afin de s'assurer que les accès aux ressources partagées sont bien exclusifs.

1.2 Tests

L'environnement de test comporte 3 types de calculs différents, 2 calculateurs de type A, un calculateur de type B et un calculateur de type C. Les calculs de type A sont plus longs que les calculs de type B qui sont eux même plus longs que les calculs de type C.

- Les Google tests globaux et ceux concernant l'étape 1 passent tous.
- Si on envoie une demande de calcul de type A, un calculateur de type A prend le calcul. C'est le comportement attendu.
- Si on envoie 3 demandes de calcul de type A, les 2 calculateurs de type A prennent 2 calculs, le troisième est en attente jusqu'à ce qu'un calculateur de type A ait fini son calcul, celui-ci s'occupe alors du troisième calcul. C'est le comportement attendu.
- Si on envoie une demande de calcul de type A, une demande de calcul de type B et une demande de calcul de type C, 3 calculateurs du bon type prennent chacun un calcul. C'est le comportement attendu.
- Si on envoie une demande de calcul de type A alors que la queue du type A est pleine, il faut attendre qu'une place se libère dans la queue pour pouvoir ajouter le calcul et que celui-ci ait un identifiant. C'est le comportement attendu.
- Lorsqu'on envoie des demandes de calculs, un id leur est attribué dans leur ordre d'arrivée en partant de 0. Ceci peut être vérifié en regardant les logs. C'est le comportement attendu.

Des tests similaires ont aussi été effectués en changeant le nombre d'instances de calculateurs de chaque type et leur résultats a été concluant.

2 Etape 2

Dans cette étape nous avons mis en place la gestion des résultats grâce aux méthodes `getNextResult` et `provideResult`.

2.1 Choix d'implémentation

Pour représenter la structure contenant les résultats, nous avons choisi d'utiliser une `std::list` contenant les résultats sous forme de `ResultWithId`. `ResultWithId` est une structure que nous avons rajoutée contenant un id et un résultat qui est optionel (grâce à `std::optional`). Ainis, nous pouvons ajouter le résultat dans la liste directement après avoir reçu la demande de calcul, ce qui nous permet de connaître l'ordre des calculs en cours ou terminés à tout moment. Dans les faits, un résultat est ajouté à la liste dès que `requestComputation` est appelée.

La méthode `getNextResult` est en charge de récupérer le résultat le plus ancien dans la liste. Pour cela, elle n'as donc qu'à accéder au dernier élément contenu dans la liste. Si la partie optionelle du résultat est vide, cela signifie que le résultat n'est pas encore disponible et donc que le calcul est toujours en cours (ou pas commencé). Dans ce cas, on appelle la méthode `wait` sur la condition `notExpectedResult`. Cette dernière est une condition qui représente le fait que le prochain résultat n'est pas encore disponible. On attend donc que le résultat attendu soit disponible dans la liste avant de pouvoir continuer. Lorsque c'est le cas, on retire le résultat de la liste et on le retourne.

La méthode `provideResult` permet au calculateur de retourner le résultat du calcul. Pour cela, elle cherche l'id du résultat donné dans la liste `results` et met à jour la partie optionelle du résultat avec le résultat donné. Ensuite, elle signale sur la condition `notExpectedResult` que le prochain résultat est disponible.

2.2 Tests

L'environnement de test comporte 3 types de calculs différents, 2 calculateurs de type A, un calculateur de type B et un calculateur de type C. Les calculs de type A sont plus longs que les calculs de type B qui sont eux même plus longs que les calculs de type C.

- Les Google tests concernant l'étape 2 passent tous.

- Si on envoie une demande de calcul de type A, un calculateur de type A prend le calcul puis renvoie le résultat lorsqu'il est calculé. C'est le comportement attendu.
- Si on envoie 3 demandes de calcul de type B, le calculateur de type B s'occupe des calculs dans l'ordre d'arrivée puis renvoie le résultat à chaque fois qu'il est calculé. (Voir figure 1.a.)
- Si on envoie à la suite un calcul de type A, un calcul de type B et un calcul de type C, les résultats sont renvoyés dans l'ordre A, B, C, bien que le calcul de type C se soit terminé avant le calcul de type B, qui s'est terminé avant le calcul de type A. C'est le comportement attendu. (Voir figure 1.b.)
- Si on envoie à la suite un calcul de type A et 4 calculs de type C, les calculs de type C se terminent avant le calcul de type A mais le résultat du calcul de type A est renvoyé avant. C'est le comportement attendu. (Voir figure 1.c.)

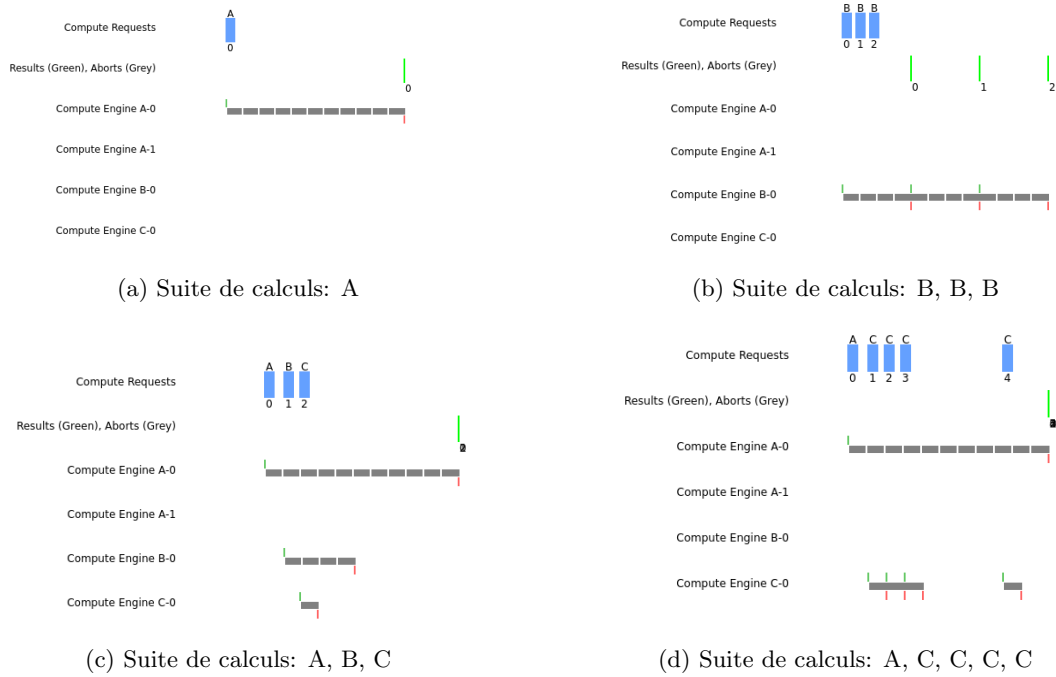


Figure 1: Exemple d'exécution pour l'étape 2

Des tests similaires ont aussi été effectués en changeant le nombre d'instances de calculateurs de chaque type et leur résultats a été concluant.

3 Etape 3

Dans cette étape, nous avons mis en place la possibilité d'annuler des calculs demandés par le client grâce aux méthodes `abortComputation` et `continueWork`.

3.1 Choix d'implémentation

La méthode `abortComputation` permet d'annuler un calcul en cours grâce à son identifiant. Il y a plusieurs cas à gérer. Si le calcul n'est pas encore en cours, il faut simplement le retirer de la map `buffer`. Dans ce cas, il faut signaler sur la condition `fullQueuePerType` que la queue pour ce type de calcul contient un calcul en moins. Si le calcul est en cours, il faut le retirer de la liste `results` et signaler sur la condition `notExpectedResult` pour potentiellement débloquent un thread qui attend sur ce résultat. Si le calcul est terminé, il faut simplement le retirer de la liste `results`. On notera qu'au vu de notre implémentation de `requestComputation`, qui ajoute directement un calcul demandé dans la liste `results`, il faut aussi retirer le calcul de cette liste s'il est demandé mais pas encore en cours.

La méthode `continueWork` est quant à elle assez simple, puisqu'elle va simplement chercher l'id passé en paramètre dans la liste `results`. Si l'id est présent, cela signifie que le calcul doit continuer, sinon, cela signifie que le calcul doit être annulé.

3.2 Tests

L'environnement de test comporte 3 types de calculs différents, 2 calculateurs de type A, un calculateur de type B et un calculateur de type C. Les calculs de type A sont plus longs que les calculs de type B qui sont eux même plus longs que les calculs de type C.

- Les Google tests concernant l'étape 3 passent tous.
- Si on envoie une demande de calcul de type A et qu'on l'annule avant la fin, le calculateur s'arrête et le résultat n'est pas renvoyé. C'est le comportement attendu. (Voir figure 2.a.)
- Si on envoie une demande de calcul de type A et qu'on l'annule après la fin alors que le résultat est déjà renvoyé, cela n'a pas de conséquence et on peut relancer un nouveau calcul. C'est le comportement attendu.
- Si on envoie des demandes de calcul de type A et B et qu'on annule le calcul de type B alors qu'il est déjà calculé mais avant la fin du calcul de type A, seul le calculateur de type A renvoie son résultat. C'est le comportement attendu. (Voir figure 2.b.)
- Si on envoie des demandes de calcul de type A, B et C et qu'on annule les calculs de type B et C alors qu'ils sont déjà calculés mais avant la fin du calcul de type A, seul le calculateur de type A renvoie son résultat. C'est le comportement attendu. (Voir figure 2.c.)
- Lorsqu'on a annulé un calcul, on peut relancer un nouveau calcul sans problème. C'est le comportement attendu. (Voir figure 2.d.)

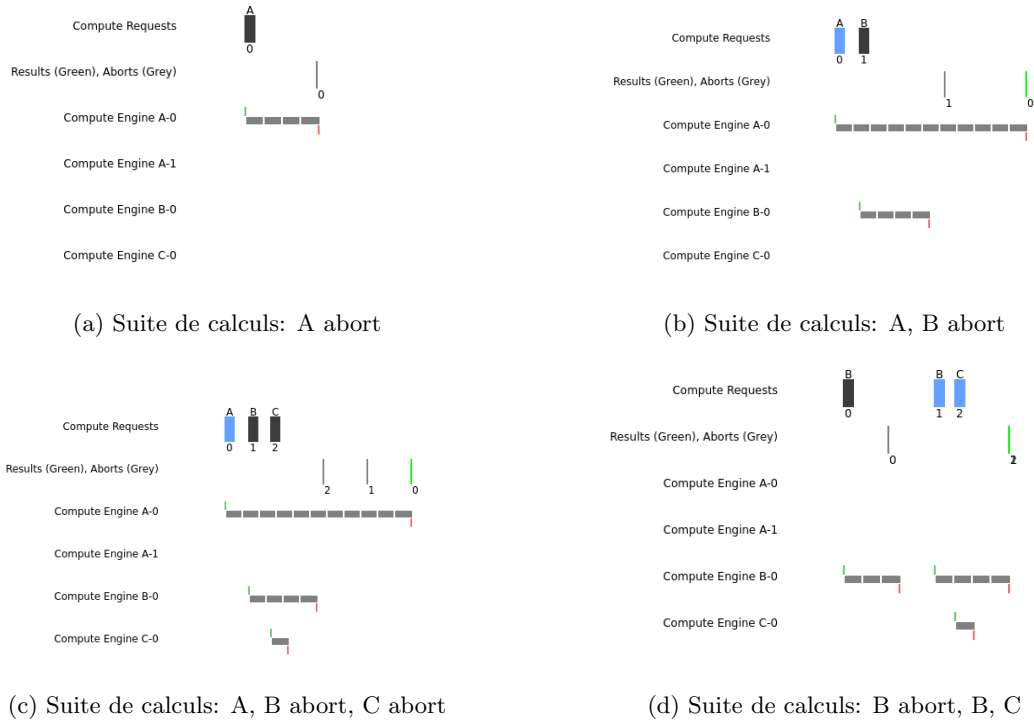


Figure 2: Exemple d'exécution pour l'étape 3

4 Etape 4

Dans cette étape, nous avons mis en place la gestion de la terminaison du buffer grâce aux méthodes `stop` et `throwStopException`.

4.1 Choix d'implémentation

La méthode `stop` permet de demander l'arrêt du buffer. Afin de modéliser le fait que le buffer est arrêté, nous avons créé un attribut booléen `stopped` qui est initialisé à `false` au lancement du programme. Lorsque la méthode `stop` est appelée, on met cet attribut à `true`. De plus, on signale sur toutes les conditions existantes dans le buffer afin de débloquer tous les threads qui attendent sur ces dernières, afin qu'ils puissent s'arrêter.

Pour que l'arrêt des threads se fasse comme attendu, nous avons dû apporter quelques modifications aux méthodes préalablement définies. Ainsi, dans toutes les fonctions qui contiennent une attente sur une condition, nous avons ajouté une vérification de l'attribut `stopped` avant et après l'attente sur la condition. Si on se trouve avant l'attente et que l'attribut `stopped` est à `true`, on sort du moniteur et on lance une exception. Si on se trouve après l'attente et que l'attribut `stopped` est à `true`, on signale sur la condition pour débloquer les threads qui attendent dessus, on sort du moniteur et on lance une exception. Cela permet en quelque sorte de réveiller les threads en cascade pour être sûr qu'ils s'arrêtent tous.

La méthode `continueWork` a aussi été modifiée afin de retourner `false` si l'attribut `stopped` est à `true`, afin de signifier que le travail en cours doit s'arrêter.

4.2 Tests

L'environnement de test comporte 3 types de calculs différents, 2 calculateurs de type A, un calculateur de type B et un calculateur de type C. Les calculs de type A sont plus longs que les calculs de type B qui sont eux même plus longs que les calculs de type C.

- Les Google tests concernant l'étape 4 passent tous.
- Les méthodes `requestComputation`, `getWork` et `getNextResult` lancent une exception si le buffer est arrêté si elles allaient entrer ou sortent d'une phase d'attente. C'est le comportement attendu.
- Lorsqu'on arrête le buffer, les threads qui attendent sur une condition sont débloqués et s'arrêtent. C'est le comportement attendu.
- Lorsqu'on arrête le buffer, tous les threads s'arrêtent (i.e on arrive à `join` sur tous les threads). C'est le comportement attendu.
- Lorsqu'on arrête le buffer, on ne peut plus ajouter/ récupérer de calculs. C'est le comportement attendu.
- Lorsqu'on arrête le buffer, les derniers calculs sont récupérés dans l'ordre de leur demande et ne sont pas mélangés. C'est le comportement attendu.