

sujet de stage M2

Ordonnement d'arbres de tâches malléables sous contrainte mémoire

Loris Marchal, Frédéric Vivien.

Mots-Clés : Algorithmique, ordonnancement, évaluation de performance.

Encadrants : Loris Marchal (chargé de recherche, CNRS), Frédéric Vivien (directeur de recherche, INRIA).

Emails : Loris.Marchal@ens-lyon.fr, Frederic.Vivien@inria.fr.

Sujet de stage

Le stage se déroulera dans l'équipe ROMA, du Laboratoire de l'Informatique du Parallélisme, sous la co-direction de Loris Marchal et Frédéric Vivien.

L'équipe ROMA s'intéresse à la conception d'algorithmes parallèles et d'ordonnement pour les plates-formes de calcul distribuées, en particulier pour les applications scientifiques. Ces applications sont souvent modélisées par des graphes de tâches complexes, et de plus en plus, elles impliquent l'utilisation de données de très grandes tailles. Cependant, les plates-formes parallèles sur lesquelles doivent être traitées ces applications disposent d'une mémoire limitée. Il est donc nécessaire de prendre en compte les contraintes mémoire lors de l'allocation des tâches sur les machines de calcul. En particulier, nous nous intéressons dans le cadre du projet ANR SOLHAR¹ aux arbres de tâches qui apparaissent lors de la factorisation de matrices creuses par les méthodes directes multi-frontales.

Dans une étude préliminaire, nous avons considéré le problème de l'ordonnement sous contraintes mémoires d'arbres de tâches [2]. La principale limitation de ce travail concerne le parallélisme : bien que les heuristiques proposées soient parallèles, elles ne considèrent que le « parallélisme d'arbre » : chaque tâche est séquentielle et ne peut donc être traitée que par un processeur. Au contraire, dans la plupart des applications, le traitement de chaque tâche peut être distribué sur plusieurs processeurs. Le choix du meilleur nombre de processeurs par tâche pour réduire le temps de traitement total est d'ailleurs un problème complexe. Dans le cadre des arbres de tâches provenant des méthodes de factorisation multi-frontales, Beaumont et Guermouche [1] ont récemment proposé d'utiliser le modèle des tâches malléables pour appliquer les résultats de Prasanna et Musicus [3] : à l'aide de la théorie du contrôle optimal, ceux-ci parviennent à calculer le nombre optimal de processeurs à assigner à chaque tâche de l'arbre. Lorsque toutes les tâches ont le même modèle de performance (temps de traitement en $\frac{L}{p^\alpha}$ pour un nombre de processeur p et un même paramètre α), la solution est particulièrement régulière et facile à calculer, et ce modèle semble bien convenir aux arbres que nous voulons étudier. Cependant, dans la solution proposée toutes les feuilles de l'arbre sont traitées simultanément, ce qui est connu pour consommer beaucoup de mémoire. Il est donc nécessaire de trouver un compromis dans l'allocation des processeurs entre mémoire et temps d'exécution.

L'objectif principal de ce stage est d'étudier le problème de l'ordonnement parallèle d'arbres de tâches modélisées par des tâches malléables de même paramètre α sous contrainte mémoire. Il s'agira tout d'abord de s'appropriier les résultats de [3] pertinents dans ce contexte en les justifiant par des arguments d'ordonnement, puis d'étudier la complexité bi-critère temps/mémoire sous ce modèle et enfin de proposer des algorithmes (exacts ou heuristiques si le problème se révèle NP-complet) et d'étudier la qualité des solutions produites (en prouvant des rapports d'approximations ou à l'aide de simulations).

1. <http://solhar.gforge.inria.fr/doku.php>

Dans un deuxième temps, et en fonction des résultats obtenus par le stagiaire pour le premier objectif, nous nous intéresserons à l'ordonnancement d'arbre de tâches malléables sur plate-forme hybride : les machines de calcul actuelles comportent souvent des accélérateurs comme les GPU. Sur une telle plate-forme, chaque tâche peut-être effectuée soit sur un CPU soit sur un GPU (ou un autre accélérateur). Le modèle de performance dépend de la ressource, on peut donc imaginer différencier le paramètre α selon la ressource utilisée : α_{CPU} et α_{GPU} . L'objectif sera alors d'adapter les résultats de [3] pour une telle plate-forme hybride.

En plus des interactions avec les encadrants de stage, le stagiaire pourra être amené à collaborer avec les partenaires de l'équipe ROMA dans l'ANR SOLHAR, et en particulier avec Alfredo Buttari (équipe APO, laboratoire IRIT, Toulouse) et Abdou Guermouche (équipe HiePACS, INRIA Bordeaux Sud-Ouest et LaBRI, Bordeaux) qui étudient des problèmes similaires d'un point de vue plus applicatif.

Bibliographie

- [1] Olivier Beaumont and Abdou Guermouche. Task scheduling for parallel multifrontal methods. In *Proceedings of the 13th International Conference on Parallel Processing (Euro-Par)*, pages 758–766, 2007.
- [2] Loris Marchal, Oliver Sinnen, and Frédéric Vivien. Scheduling tree-shaped task graphs to minimize memory and makespan. In *Proceedings of the 27th IEEE International Symposium on Parallel and Distributed Processing (IPDPS)*, pages 839–850, 2013.
- [3] G. N. Srinivasa Prasanna and Bruce R. Musicus. The optimal control approach to generalized multiprocessor scheduling. *Algorithmica*, 15(1) :17–49, 1996.