

Optimisation du modèle de programmation parallèle PGAS pour architectures manycore cc-NUMA.

Mots clés :

- **Directeur de thèse** : Alain Greiner
- **Co-encadrant(s)** :
- **Unité de recherche** : Laboratoire d'informatique de Paris 6
- **Ecole doctorale** : École Doctorale Informatique, Télécommunications, Électronique de Paris
- **Domaine scientifique principal**: Divers

Résumé du projet de recherche (Langue 1)

L'évolution de la performance des processeurs jusqu'au milieu des années 2000 profitait de l'augmentation de la fréquence de fonctionnement. Ce moyen a malheureusement atteint ses limites à cause de la consommation énergétique des circuits intégrés. La performance passe désormais par l'augmentation du nombre de coeurs de processeur qui double tous les 18 mois. C'est sur constat que le projet TSAR a débuté. Ce projet propose une architecture clusterisée manycore NUMA à mémoire partagée où la cohérence mémoire est assurée par le matériel. Chaque cluster (jusqu'à 1000) gère un segment de l'espace physique de la mémoire. La cohérence étant assurée par le matériel, les applications parallèles développées pour les architectures multicore actuelles restent exécutables sur cette nouvelle architecture. Toutefois, on ne peut obtenir une bonne performance que si le système d'exploitation renforce la localité des accès mémoire en réalisant un placement conjoint des tâches sur les coeurs de processeurs et des données sur les bancs de mémoire physique proches des coeurs qui les utilisent. Le projet ALMOS (thèse de Ghassan Almaless) a permis d'expérimenter et de valider différentes techniques permettant d'atteindre cet objectif. ALMOS est construit autour d'un noyau monolithique dans lesquels toutes les structures internes sont distribuées dans les différents clusters pour que l'accès à ces structures ne soit pas le goulot d'étranglement. Cependant les performances dépendent de la façon dont les applications ont été parallélisées. À l'heure actuelle, il existe un certain nombre de paradigmes (langage OpenMP, modèle de programmation PGAS, pattern Map&Reduce) permettant d'exprimer des programmes parallèles tout en masquant la topologie exacte de l'architecture matérielle cible. Ces modèles masquent volontairement l'architecture afin de garder un code générique. Cependant, cette généricité (cette ignorance) engendre dans beaucoup de cas une perte de performance par rapport à un placement optimisé. Le but général de la thèse est d'analyser comment les environnements de programmation parallèles basés sur le modèle PGAS (Partitionned Global Address Space) peuvent être optimisés pour exploiter efficacement les mécanismes de placement conjoint proposé par l'OS.

Résumé du projet de recherche (Langue 2)

Les expérimentations réalisées dans cette thèse s'appuieront sur le prototype existant de l'architecture manycore TSAR car celle-ci permet d'instancier un grand nombre de coeurs et d'analyser les problèmes de passage à l'échelle ainsi que sur l'OS expérimental ALMOS. Toutefois, les mécanismes de coplacement fournis par ALMOS sont très généraux et peuvent s'appliquer à des architecture multicore plus classiques du type INTEL ou AMD.