

Parallélisme et calcul à haut rendement

Dates & Lieux :

- 14-18/12/2020 - classe à distance
- 22-26/03/2021 - classe à distance
- 31/05-04/06/2021 - classe à distance
- 08-12/11/2021 - classe à distance

Durée : 5 demi-journées (20 h)

Tarifs :

2000 €

-40% pour les académiques⁽¹⁾

-50% pour les doctorants⁽¹⁾

⁽¹⁾ dans la limite des places ouvertes à ces tarifs

Objectifs :

Acquérir une vision globale des grands enjeux du calcul parallèle, de la performance des codes et des infrastructures matérielles. À l'issue de la formation, un développeur pourra déterminer le potentiel et les limites des différents paradigmes et langages de programmation parallèles pour les architectures matérielles du marché (hors FPGA et DSP). Ces connaissances seront utiles au responsable d'infrastructure de calcul pour définir son parc sur la base de la puissance attendue, de l'efficacité énergétique des calculs et des efforts de développement logiciel induits.

Public cible :

- Développeur de codes de calcul scientifique intéressé par la performance
- Responsable d'infrastructure de calcul

Prérequis :

- Notions de base en programmation

Inscriptions : en ligne sur www.triscale-innov.com

Description

L'efficacité énergétique est le critère dominant qui guide la conception des ordinateurs depuis 20 ans. En stoppant brutalement l'augmentation de la fréquence horloge des ordinateurs, la contrainte énergétique a conduit les constructeurs à produire des machines dont la puissance résulte de l'assemblage d'un nombre toujours croissant d'unités de calculs. L'accélération des calculs informatiques repose alors principalement sur la division du travail en un nombre toujours plus élevé de tâches élémentaires s'exécutant en parallèle. En raison de cette contrainte énergétique, les architectures de calculs parallèles sont aussi de plus en plus spécialisées (CPU, GPU, TPU, DSP, FPGA) et leur programmation s'appuie sur des paradigmes et des langages adaptés.

Téléphone portable, laptop, desktop ou super-ordinateur : le parallélisme est omniprésent. La maîtrise de la programmation des machines parallèles n'est donc plus réservée aux laboratoires, mais fait partie du bagage nécessaire à tous les développeurs soucieux de la performance de leurs logiciels.

Cette formation vise à exposer les grands enjeux de la programmation et des machines parallèles et à guider les développeurs vers les outils les mieux adaptés à leurs besoins. En sus, la formation leur proposera une méthodologie pour concevoir des logiciels à haut rendement énergétique. Ces connaissances seront aussi utiles aux responsables de parcs informatiques pour établir leur stratégie d'équipement.

La formation inclut une large part de travaux pratiques. Dans le cadre de cette formation sous forme de cours interactifs à distance, un environnement numérique distant dédié est mis à disposition de chaque stagiaire.

Programme détaillé

1. Architectures des calculateurs HPC
 - Cœurs CPU et unités de calcul
 - Processeurs multi-cœurs
 - Hiérarchie de cache
 - Nœuds SMP
 - Clusters de SMP
2. Performance des calculs séquentiels
 - FLOPS : combien de cycles pour +, -, *
 - Mur de la mémoire (roofline model)
 - Combien de cycles pour /, sqrt, sin, exp...
 - RAM, bande passante, accès aléatoires
 - Tests et spéculation de branchement
3. Notions de bases du parallélisme
 - Loi d'Amdahl, scalabilité forte et faible
 - Dead locks, race conditions
 - Ordre des opérations et associativité des calculs flottants
4. Paradigmes de programmation parallèle
 - Vectorisation (SIMD)
 - Parallélisme en mémoire partagée
 - Threads, OpenMP, TBB, tâches
 - Parallélisme en mémoire distribuée
 - MPI, runtime parallèles
 - Calcul sur GPU
5. Profiling et debugging parallèle