

4 JOURS

2 900 EUROS

4-8 PERSONNES

GRENOBLE



OBJECTIFS

- Connaître les éléments clés des architectures de programmation avec GPU pour les utiliser au mieux
- Savoir porter et optimiser une application scientifique sur processeur graphique pour multiplier par 10 à 100 ses performances
- Savoir programmer en CUDA sur GPU
- Se familiariser avec le modèle de streaming CUDA et le multi GPU

LES + DE LA FORMATION

- Alternance entre apports théoriques et applications pratiques : 60 % de la formation se déroule sous forme de travaux pratiques avec des études de cas sur ordinateurs (1 poste par stagiaire)
- Codes sources des exemples traités sont fournis lors de la formation
- Un encadrement important lors des séances pratiques (au minimum 1 enseignant pour 5 stagiaires) permet une progression rapide adaptée à chacun

ATOUTS PÉDAGOGIQUES

Cette formation s'appuie sur les moyens techniques de Grenoble INP - Phelma, UGA école nationale supérieure de Physique, Electronique et Matériaux.

Les intervenants sont des enseignants-chercheurs et chercheurs du département Image et Parole au sein du laboratoire Grenoble Images Parole Signal Automatique (Gipsa Lab), spécialistes reconnus dans le domaine de l'adéquation algorithme architecture.



PROGRAMME

JOUR 1 - ARCHITECTURE DES GPU

Historique des GPU, pipeline graphique ; Architecture comparée des GPU (ATI, Nvidia, Intel) et Xeon Phi ; Présentation de l'API CUDA et de la chaîne de compilation ; Comparaison avec OpenCL ; Modèle de programmation et d'exécution (blocs, warps) ; Modèles de mémoires (textures, constantes, shared, globale, host) ; Méthodologie de portage sur GPU

APRÈS-MIDI : TRAVAUX PRATIQUES

- Prise en main du GPU et de son environnement de développement sur des exemples du SDK Nvidia (rotation d'image, filtre convolutif, tri...)
- Développement d'un filtre passe bas sur image en niveaux de gris avec textures

JOUR 2 - PROGRAMMATION PARALLÈLE DES GPU ET CPU

Points forts et faiblesses du modèle de programmation parallèle ; Architecture détaillée des familles Nvidia ; Multithreading, Hyperthreading ; Structuration des données (tiling, gestion des scratchpads) ; Précision des calculs ; Communication parallèle (débit, latence) ; Mémoires (types, accès, gestion) et cohérence des données (mémoires caches, partagées, globales, synchronisation) ; Outils d'analyse de performance (optimisation, profiler, debug...)

APRÈS-MIDI : TRAVAUX PRATIQUES

Développement d'un traitement de type Stencil avec mémoire shared, et réduction d'image avec histogramme

JOUR 3 - PROGRAMMATION AVANCÉE

Multi-GPU, fonctionnement asynchrone (streams) ; Lien avec MPI/OpenMP ; Schémas de parallélisation (Stencils, Reduction, ...) ; Optimisation parallèle (profiler) ; Lien CUDA / OpenGL ; Nouveautés CUDA 8 et 9

APRÈS-MIDI : TRAVAUX PRATIQUES

Développement d'un filtre complexe de détection de contour de Deriche avec mémoire shared et globale, avec lissage récursif et transposition optimisée

JOUR 4 - MISES EN SITUATIONS PRATIQUES

Développement d'une application multi-GPU

POUR QUI ?

Cette formation s'adresse à des ingénieurs ou techniciens impliqués dans la conception de logiciels de traitement de données à fortes contraintes de performance.

Pré-requis : Connaissance de la programmation en Langage C.

Méthode d'évaluation : Quizz d'acquisition des connaissances en début et en fin de formation.

Sanction de la formation : attestation de présence et de fin de formation.

CONTACT ET INSCRIPTION

Katia Plentay

04 76 57 45 03

formation-pro.stages@grenoble-inp.fr

Grenoble INP - UGA, Formation Pro
3 Parvis Louis Néel
38016 Grenoble Cedex 1
formation-pro.grenoble-inp.fr