

La programmation GPU en CUDA

Description de la formation

CUDA est une plate-forme informatique parallèle et un modèle de programmation développé par Nvidia pour l'informatique générale sur ses propres GPU. CUDA permet aux développeurs d'accélérer les applications gourmandes en ressources de calcul en exploitant la puissance des GPU pour la partie parallélisable des calculs.

Cette formation vous permettra d'acquérir les compétences et connaissances nécessaires pour mettre en pratique de manière efficace la programmation GPU en CUDA.

Objectifs pédagogiques

- › Maîtriser la programmation GPU en CUDA.

Prérequis

- › Bonne expérience de la programmation en langage C.

Modalités pédagogiques

- › **Modalité :** Formation réalisée en présentiel ou en distanciel selon la formule retenue.
- › **Méthode :** La formation se déroule entre 50% de théorie et 50% de pratique. Le formateur partage des points théoriques et des cas concrets, lance des discussions et échanges entre les stagiaires et propose des jeux / outils en relation avec le contenu et des mises en pratique.
- › **Support de formation :** Le support de formation utilisé par le formateur est remis au stagiaire à l'issue de la formation.

Modalités techniques

- › En format présentiel, le formateur dispose d'une présentation (support de formation), d'un vidéoprojecteur (ou TV), de tableaux blancs et de jeux / d'outils pédagogiques.
- › En format présentiel, le stagiaire a besoin d'un ordinateur permettant l'utilisation d'une machine virtuelle.
- › En format distanciel, le formateur dispose d'une présentation (support de formation), d'une plateforme de visioconférence et d'outils collaboratifs numériques.
- › En format distanciel, le stagiaire a besoin d'une connexion internet et d'un ordinateur équipé d'une webcam et d'un micro et permettant l'utilisation d'une machine virtuelle.
- › Dans tous les cas, des machines virtuelles Windows/Linux sur AWS sont mis à disposition des stagiaires pour apprendre sur un environnement représentatif.

Code

CUD100

Durée

3 jours (21 heures)

Nombre de participants

Entre 4 (minimum) et 10 (maximum) participants.

Profil des stagiaires

Personnes qui souhaitent acquérir une connaissance solide de la programmation GPU et de la technologie CUDA.

Sanction de la formation

Attestation de fin de formation.

Accessibilité

Accessible pour les personnes en situation de handicap et aménagement possible en fonction du type de handicap (prévenir avant le début de la formation).

Modalités et délais d'accès

10 jours minimum avant la formation pour une demande de prise en charge.

Modalités de suivi et d'évaluation

- › Evaluation préalable.
- › Autoévaluation des acquis au cours des exercices et mises en pratiques au cours de la formation.
- › Evaluation de fin de formation sous forme de test (QCM) afin de valider l'acquisition des compétences et des connaissances.
- › Questionnaire d'évaluation de la satisfaction en fin de formation.
- › Feuille d'émargement signée par le(s) stagiaire(s) et le formateur, par demi-journée de formation.
- › Attestation de fin de formation.

Intervenant

David est **Développeur Logiciel & Formateur** sur des technologies comme le Machine Learning, l'optimisation des performances, le Multi-core Programming, le **GPGPU (CUDA, OpenCL)** ou encore le Deep Learning. Ses compétences et connaissances couplées à sa pédagogie adaptée à vos besoins nous permet de vous proposer une formation de qualité.

Tarifs

- › Interentreprises : 1 600,00 € HT
- › Intra-entreprise : sur demande

Contenu de la formation

JOUR 01

INTRODUCTION A CUDA

- › Introduction générale à CUDA
- › Modèles physique et logique de CUDA
- › Evolution des matériels Nvidia
- › Le langage CUDA-C
- › Compiler un code CUDA
- › Optimiser un code CUDA
- › Mesurer la performance
- › Géométrie des noyaux CUDA
- › Problème d'alignement mémoire
- › Les warps CUDA
- › Problème de coalescence des accès
- › Branches et divergences
- › Déroulage de boucles
- › La mémoire constance
- › La mémoire texture
- › La mémoire surface
- › La mémoire partagée
- › Taux d'occupation
- › Introduction au Nvidia profiler

JOUR 02

TP D'OPTIMISATION

- › Optimisation d'un produit matrice constante
- › Optimisation d'un produit matrice vecteur
- › Optimisation d'une transposée de matrice
- › Gestion mémoire avancée
- › Alignement mémoire et padding
- › Sensibilisation au coût de la gestion mémoire
- › Allocation de mémoire verrouillée
- › Utilisation de mémoire « zéro-copie »
- › UVA

JOUR 03

NOTIONS D'ALGORITHMIQUE PARALLELE

- › Le pipe-line et les CUDA streams
- › La réduction
- › Opérations atomiques
- › Application du « parallélisme dynamique » au tri parallèle
- › Bibliothèques CUDA
- › cuFFT
- › cuBLAS

INTRODUCTION A LA PROGRAMMATION MULTI-GPU

- › Introduction générale à la programmation multi-GPU
- › Parallélisme à gros grain
- › Parallélisme à grain fin
- › L'accès « peer-to-peer » avec UVA

TP FINAL – ECRITURE D'UNE VERSION CUDA DU CODE MODIFIED GRAM SCHMIDT

IHMISEN

SAS au capital de 2.000 euros | Siège social : 5 impasse du Carlit 31490 Léguvin | N° SIRET : 91274766400015
N° TVA : FR31912747664 | Code APE : 7022Z | NDA : 76311232431 (auprès du préfet de région Occitanie)
+33 (0)6 88 28 29 62 | <https://www.ihmisen.com/>

Version du document : 02

Dernière mise à jour : lundi 1er septembre 2025