Optimisation et Parallélisation de Code Python sur CPU et GPU

Lien : https://innov-systems.com/formation/optimisation-et-parallelisation-de-code-python-sur-cpu-et-gpu

DURÉE	RÉFÉRENCE	CATÉGORIE
5 jours (35h)	**DEV336**	**Python, C# et .NET**

OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Comprendre et appliquer les concepts de parallélisation et d’optimisation en Python
- Identifier les limites du GIL et mettre en place des stratégies de contournement
- Exploiter efficacement les architectures CPU, GPU et clusters pour accélérer les traitements
- Utiliser les bibliothèques modernes (Numba, Dask, Ray, Polars, CuPy…) pour optimiser des cas concrets
- Développer une approche systématique de profiling et d’optimisation du code
- Acquérir une expertise pratique en traitement de données, calculs massifs et analyses distribuées

POUR QUI ?

- Développeurs Python souhaitant améliorer la performance de leurs applications
- Data scientists cherchant à traiter efficacement de grands volumes de données
- Ingénieurs logiciels et administrateurs système impliqués dans des projets haute performance

Innov Systems

Programme détaillé

1 / INTRODUCTION AUX ARCHITECTURES PARALLÈLES

- Différences entre CPU, GPU et clusters
- Comprendre le modèle mémoire et les goulots d'étranglement
- Limites théoriques de la parallélisation (loi d'Amdahl, loi de Gustafson)

2 / PROFILING ET ANALYSE DE PERFORMANCE

- Les différents types de profileurs (statistiques, événementiels)
- Mesurer le temps, la mémoire et l'énergie consommée
- Visualiser l'exécution avec SnakeViz, LineProfiler, Scalene

3 / MISES EN PRATIQUE

- Identification des ressources matérielles d'un poste
- Profiling d'un script Python simple (analyse mémoire, CPU, I/O)
- Visualisation des résultats pour cibler les optimisations

4 / LES BASES DE LA PARALLÉLISATION AVEC PYTHON

- Différence entre multithreading, multiprocessing et async
- Comprendre et contourner le GIL
- Synchronisation (verrous, files, sémaphores, évènements)

5 / OPTIMISATION DES PERFORMANCES CPU

- Pools de processus et parallélisation de tâches répétitives
- Accélération du code Python avec NumPy et vecteurisation
- JIT compilation avec Numba et Pythran

Innov Systems

6 / MISES EN PRATIQUE

- Parallélisation de tâches (analyse CSV, traitement images, web scraping)
- Optimisation CPU avec multiprocessing et Numba
- Libération du GIL avec Cython

7 / DISTRIBUTION DE TÂCHES À GRANDE ÉCHELLE

- Différence entre parallélisme local et distribué
- Présentation des frameworks : Ray, Dask, Celery
- Utiliser Dask pour paralléliser pandas et NumPy

8 / DÉPLOIEMENT DE CLUSTERS PYTHON

- Mise en place d’un cluster Dask local
- Distribution de tâches avec Ray
- Cas d’usage : calcul distribué pour analyse de logs massifs

9 / MISES EN PRATIQUE

- Parallélisation de workflows avec Dask et Ray
- Déploiement d’un cluster multi-machines simulé
- Monitoring des tâches distribuées

10 / PRINCIPES DU CALCUL GPU

- Différences entre CPU et GPU
- CUDA, OpenCL et interfaces Python
- Choisir les bonnes bibliothèques selon le cas d’usage

11 / PYTHON ET GPU DANS LA PRATIQUE

- Accélération avec Numba (CUDA JIT)
- Manipulation de dataframes avec cuDF et Polars

Inov Systems

- Calculs massifs sur matrices avec CuPy et Dask-CUDA

12 / MISES EN PRATIQUE

- Optimisation de pipelines data sur GPU
- Calcul scientifique (climat, finance, image) sur plusieurs GPU
- Comparaison CPU vs GPU en termes de performance

13 / STRATÉGIES D'OPTIMISATION COMPLÈTES

- Quand choisir le CPU, le GPU ou un cluster
- Combiner multiprocessing, GPU et calcul distribué
- Gestion des ressources et monitoring

14 / CAS PRATIQUES D'ENTREPRISE

- Traitement massif de données IoT
- Analyse de données financières en temps réel
- Simulation numérique à grande échelle

15 / ATELIER FINAL ET BONNES PRATIQUES

- Atelier fil rouge : optimisation et parallélisation d'un projet complet
- Check-list des bonnes pratiques de performance en Python
- Perspectives : Python et le futur du HPC (High Performance Computing)

Inov Systems

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Prochaines dates programmées

22 au 26 Fév. 2027	Présentiel - Casablanca
19 au 23 Avr. 2027	Présentiel - Casablanca
14 au 18 Juin 2027	Présentiel - Casablanca
09 au 13 Août 2027	Présentiel - Casablanca
04 au 08 Oct. 2027	Présentiel - Casablanca

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Réservation & Renseignements

Téléphone : +212 522 247 210
Email : contact@innov-systems.com
Web : https//www.innov-systems.com

Innov Systems — Tous droits réservés