Fiche du cours

70 h

Titre :

PYS400 - Optimisation & HPC Avancé (Vectorisation, Parallélisme & GPU)

Description :

Ce cours enseigne aux étudiants à transformer du code scientifique Python standard en code rapide, scalable et prêt pour le calcul intensif.
La première partie se concentre sur l'optimisation locale : profiling, vectorisation, parallélisation simple et exploitation efficace des ressources CPU.
La seconde partie introduit les architectures HPC modernes : parallélisation distribuée avec MPI4Py, accélération GPU avec CuPy ou Numba CUDA, et exécution sur clusters HPC.
Les étudiants conçoivent des simulations scientifiques qui passent du poste local à l'exécution massive.

Objectifs :

- Identifier et corriger les goulots de performance dans le code scientifique.*
- Maîtriser la vectorisation et l'optimisation CPU avec NumPy et Numba.*
- Paralléliser les calculs localement et distribuer des tâches sur plusieurs nœuds.*
- Exploiter les GPU pour accélérer des sections critiques.*
- Soumettre et gérer des jobs sur des clusters HPC.*
- Concevoir un pipeline complet, reproductible et performant.

Chapitres :

1. Profiling & optimisation : timeit, line_profiler, memory_profiler.*
2. Vectorisation avancée avec NumPy : slicing, broadcasting, élimination des boucles lourdes.*
3. Compilation JIT avec Numba pour accélérer les calculs CPU intensifs.*
4. Parallélisation locale : multiprocessing, concurrent.futures.*
5. Traitement distribué : introduction à Dask pour des volumes de données massifs.*
6. MPI avec MPI4Py : communications, broadcast, scatter/gather, réduction.*
7. Accélération GPU : CUDA avec Numba, CuPy et transferts mémoire optimisés.*
8. Soumission de jobs HPC : Slurm / PBS, gestion des ressources, scaling horizontal.*
9. Débogage et tuning en environnement HPC.*
10. Étude de cas pratique : optimisation et déploiement d'une simulation physique à grande échelle.

À la fin :
Vous serez capable de passer d’un code scientifique Python classique à une version hautement optimisée et massivement parallèle, capable de tourner sur plusieurs cœurs, plusieurs nœuds ou des GPU. Vous maîtriserez les techniques clés du calcul scientifique moderne, des optimisations locales au calcul distribué sur cluster HPC, avec une approche reproductible et structurée.