



ÉCOLE NATIONALE
SUPÉRIEURE
D'INFORMATIQUE
POUR L'INDUSTRIE
ET L'ENTREPRISE

FORMATION INITIALE D'INGÉNIEUR EN INFORMATIQUE

PARCOURS THÉMATIQUE

**CALCUL
INTENSIF ET
DONNÉES MASSIVES**

ARCHITECTURES HAUTE PERFORMANCE
PROGRAMMATION PARALLÈLE
APPRENTISSAGE STATISTIQUE
GESTION DES FLUX DE DONNÉES
SIMULATION

RESPONSABLE

PIERRE DOSSANTOS-UZARRALDE

pierre.dossantos-uzarralde@ensiie.fr

PARTENAIRES

EDF
CEA



NESTOR DEMEURE
PROMOTION 2017

En thèse au CEA et à l'ENS Paris-Saclay

« J'ai suivi le parcours Calcul Intensif de l'ENSIIE. Aujourd'hui, je fais une thèse sur l'étude de l'erreur numérique dans les calculs hautes-performances avec l'ENS Paris-Saclay et le Commissariat à l'Énergie Atomique. »



MÉTIERS

- Administrateur Système HPC
- Business Intelligence Manager
- Data Scientist
- Chief Data Officer
- Data Analyst
- Ingénieur Développement Logiciel Scientifique
- Master Data Manager
- Recherche et R&D
- Lead Data Miner
- Ingénieur Système HPC
- Ingénieur support applicatif HPC

EXEMPLES DE STAGES

- Développement des codes de simulation
ESI, CEA, ONERA, CNRS
- Développement de processus et d'outils de compilation
Intel, IBM, Nvidia
- Construction de systèmes HPC
Atos, IBM, CRAY
- Utilisation de systèmes HPC
EDF, Safran, Airbus, Total, L'Oreal, Michelin, Dassault, EADS, Air Liquide



ÉCOLE NATIONALE
SUPÉRIEURE
D'INFORMATIQUE
POUR L'INDUSTRIE
ET L'ENTREPRISE

FORMATION INITIALE D'INGÉNIEUR EN INFORMATIQUE

PARCOURS THÉMATIQUE

CALCUL INTENSIF ET DONNÉES MASSIVES

Ce parcours est mis en place lors de la deuxième année de formation à l'ENSIIE. Le semestre 3 fournit les compétences nécessaires pour entrer dans le monde des données, de la sciences des données lié à celui du calcul haute performance (HPC) et au développement de méthodes et techniques de programmation massivement parallèle (processeur multi-cœur, processeur graphique, supercalculateur, Cloud Computing). L'objectif du semestre 4 est d'exposer les étudiants aux techniques de programmation haute performance et de parallélisation massive en utilisant une gamme de bibliothèques de programmation multi-core, multi-thread ou GPU.

Le semestre 5 s'articule autour de sujets avancés spécialisés dans la gestion des données massives en particulier l'intelligence informatique pour l'analyse, l'exploration et la visualisation de données, le développement de systèmes de nuages, les méthodes d'apprentissage.

S3

**Initiation à la
Programmation scientifique**

Architecture d'un système d'exploitation

Programmation parallèle

**Projet informatique
et méthodes Agiles**

Système d'exploitation

Analyse des données (optionnel)

S4

Systèmes de Fichiers Parallèles

Parallélisme à base de thread

Réseaux pour Datacenter/HPC

Logiciel cluster

**Programmation
scientifique avancée**

Réseaux IP et administration (optionnel)

S5

**Simulation - Application
au calcul des incertitudes**

Compilation avancée

Virtualisation et Cloud

Machine Learning

Python for Data Science

ou Modèle de régression régularisée