HIGH PERFORMANCE COMPUTING / BIG DATA

HIGH PERFORMANCE ARCHITECTURE
PARALLEL PROGRAMMING
STATISTICS LEARNING
MANAGEMENT OF DATA FLOW
SIMULATION

JOBS OPPORTUNITIES
HPC System Administrator
Business Intelligence Manager
Data Scientist
Chief Data Analyst
Engineer in Scientific Software Development
Master Data Manager
Research and Development
Lead Data Miner
Engineer in HPC System
Engineer in HPC Applicative Support

EXAMPLES OF INTERNSHIPS
A performance study of parallel codes to GPU architectures
Atos France
Compilation optimisation for MPI calls
CEA
Specifications for Cloud Services
Ministère du numérique
Logo detection in images using Deep-Learning methods
Atos Senegal
Implementing sparse matrices on GPU in CUDA
Barcelone Supercomputing Center

CONTACT
PIERRE DOSSANTOS-UZARRALDE
pierre.dossantos-uzarralde@ensiie.fr

PARTNERS
EDF
CEA DAM

NESTOR DEMEURE
PROMOTION 2017
Doing a thesis at the CEA and at the ENS Paris-Saclay

This course is set up for 2nd Year students at the ENSIE. The 3rd semester provides the necessary skills to enter the world of Data, Data Science related to HPC Science and to the development of methods and technics of massive parallel programming (multi-core processor, graphic processor, supercomputer, Cloud Computing). The goal of the 4th semester is to expose students to the technics of high performance and massive parallel programming by using a range of multi-core programming, multi-thread or GPU libraries. The 5th semester is articulated around advanced topics, specialised in the management of Massive Data, particularly regarding Computer Science Intelligence for analysis, regarding the exploration and visualisation of Data, regarding the development of cloud systems but also regarding machine learning methods.

INITIAL ENGINEER TRAINING IN COMPUTER SCIENCE

THEMATRIC COURSE

HIGH PERFORMANCE COMPUTING / BIG DATA

Initiation to scientific programming
Exploitation System Architecture
Parallel Programming
Computer Science Projects and Agile Methods
Exploitation System
Data Analysis (optional)

Parallel Files Systems
Thread-based Parallelism
Data Centre/HPC Networks
Cluster Software
Advanced Scientific Programming
IP Networks and Administration (optional)

Simulation - Uncertainties
Calculation application
Advanced Compilation
Virtualisation and Cloud
Machine Learning
Python for Data Science
or Model of regulated regression