

# Inférence et modélisation du comportement des utilisateurs d'un grand centre de calcul

**Encadrant:** Frédéric Suter, Centre de Calcul de l'IN2P3. ([frederic.suter@cc.in2p3.fr](mailto:frederic.suter@cc.in2p3.fr))

**Durée:** 6 mois (possibilité de rémunération)

**Lieu de stage:**

Centre de Calcul de l'IN2P3  
21 avenue Pierre de Coubertin  
CS70202  
69627 Villeurbanne cedex

## Description

Le Centre de Calcul de l'Institut national de physique nucléaire et de physique des particules (CC-IN2P3) est l'un des quatre centres de calcul nationaux français. Le CC-IN2P3 dispose d'une infrastructure de calcul de près de 35 000 cœurs sur lesquels sont exécutés chaque mois près de trois millions de jobs. La majeure partie des expériences supportées par l'IN2P3, soit 2~500 utilisateurs participant à près de 80 collaborations internationales, utilisent massivement ces ressources pour traiter des données issues de grands instruments scientifiques dans le domaine de la physique, tels que des accélérateurs de particules, des satellites ou des télescopes et contribuer à des découvertes scientifiques majeures telles que le boson de Higgs par les détecteurs ATLAS et CMS du grand collisionneur d'hadrons (LHC) au CERN, ou la première observation des ondes gravitationnelles par les expériences LIGO et VIRGO. Il est cependant indispensable d'optimiser le fonctionnement du gestionnaire de jobs et de ressources du centre afin d'accélérer leur traitement et ainsi réduire le temps de production de ces résultats scientifiques.

L'objectif de ce stage est d'analyser l'ensemble des traces d'exécution du gestionnaire de jobs et de ressources du CC-IN2P3 afin de *caractériser* et *classifier* les calculs exécutés selon différents critères (utilisateur et groupe, taille et durée, distribution des soumissions, ...) et par rapport à différentes métriques sur la qualité de l'ordonnancement produit (temps d'attente, temps de réponse, ralentissement, ...). Cela permettra alors d'*inférer et modéliser le comportement* des utilisateurs, d'*optimiser l'ordonnancement des jobs* exécutés au CC-IN2P3 et, grâce à des *algorithmes d'apprentissage*, de mieux anticiper les variations de charge ou les évolutions nécessaires du système et de l'infrastructure.

Selon l'avancée du stage, il sera éventuellement possible de confronter les utilisateurs du centre de calcul aux résultats obtenus, en se rendant par exemple dans différents laboratoires de l'IN2P3, afin d'enrichir l'entraînement des algorithmes d'apprentissage d'un facteur humain d'une part et d'encourager de bonnes pratiques d'utilisation des ressources du Centre de Calcul de l'IN2P3 d'autre part.