

## Stage en informatique scientifique (H/F)

- Niveau d'études :** Master
- Durée du stage :** 5 à 6 mois, démarrage souhaité en janvier/février 2021
- Sujet :** Modélisation et simulation de systèmes de stockage
- Localisation :** CC-IN2P3 – Campus de la Doua, Villeurbanne
- Candidatures :** par email à [frederic.suter@cc.in2p3.fr](mailto:frederic.suter@cc.in2p3.fr) (joindre CV, relevé de notes et lettre de motivation).
- Gratification :** 591,51 euros/mois

### Contexte

Ce stage s'effectue dans le cadre d'une collaboration entre le Centre de Calcul de l'IN2P3 (Campus LyonTech la Doua), l'Université Grenoble Alpes, the University of Southern California et the University of Hawai'i, et est financé par la Région Rhône-Alpes Auvergne.

Le ou la candidate recherché(e) est en Master 2 en informatique ou en école d'ingénieur avec un profil "Système" et des compétences en Linux, C/C++, et administration système. La rigueur, l'autonomie et la curiosité sont également des compétences recherchées

### Sujet

L'accroissement des volumes de données manipulés par les sciences data-intensives s'accompagne d'une diversification des besoins en termes de types de données manipulées et d'accès à ces données.

On note également une diversification des systèmes de stockage aussi bien du point de vue matériel que logiciel. Les collaborations scientifiques et les centres de calcul sont alors confrontés à un problème complexe de dimensionnement : *quelle combinaison de matériel et de logiciel choisir pour obtenir le meilleur rapport performance/coût ?*

Ce type de choix repose toutefois généralement sur l'expertise, et parfois les préférences, des administrateurs de ces systèmes de stockage, les spécifications plus ou moins précises des utilisateurs et une étude expérimentale à petite échelle sur un ensemble de scénarios limités.

Pour s'affranchir de tels choix empiriques en offrant des indicateurs objectifs ainsi qu'une capacité à anticiper les goulots d'étranglement à large échelle ou à détecter des cas pathologiques, nous proposons de modéliser et simuler des infrastructures de stockage. Le but de ce stage sera de solidifier et valider les modèles de stockage et d'entrées sorties du logiciel SimGrid. Pour cela le ou la candidate mènera une campagne d'expériences sur la plate-forme Grid'5000 pour obtenir les mesures de performance nécessaires à la modélisation de différents types de disques





**Centre de Calcul**

de l'Institut National de Physique Nucléaire  
et de Physique des Particules



(i.e., mécaniques et SSD). Ces mesures de performance brute seront complétées par une vue plus applicative. Il/elle en dérivera un ou plusieurs modèles qui seront intégrés à SimGrid puis [in]validés lors d'une seconde campagne d'expériences.

## Environnement de travail

Le candidat effectuera son stage au Centre de Calcul de l'Institut national de physique nucléaire et de physique des particules (CC-IN2P3), situé sur le campus de la Doua à Villeurbanne.

Le CC-IN2P3 est l'un des quatre centres de calcul nationaux français. Il met à disposition de ses 2 500 utilisateurs regroupés dans 80 collaborations scientifiques près de 30 000 coeurs et 340 petaoctets de stockage. Cette infrastructure est principalement dédiée au stockage et au traitement de données issus de grands instruments scientifiques dans le domaine de la physique, tels que des accélérateurs de particules, des satellites ou des télescopes.

Le CC-IN2P3 a contribué à des découvertes scientifiques majeures lors de la dernière décennie 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.



21 avenue Pierre de Coubertin  
CS 70202 - F-69627 Villeurbanne CEDEX



Tél. : +33 (0)4 78 93 08 80  
Fax : +33 (0)4 72 69 41 70



Web : <https://cc.in2p3.fr>  
Twitter : @cc.in2p3