

Stage de Master 2-Informatique (english version below)

Optimisation de code HPC pour la simulation astrophysique

1. Contexte scientifique

Le projet CTA (Cherenkov Telescope Array) [1] a pour objectif la construction et l'exploitation d'un réseau d'une centaine de télescopes Cherenkov situés sur les sites de La Palma (Canaries, Espagne) et Paranal (ESO [2], Chili). CTA ouvre une nouvelle ère dans l'astronomie gamma de très haute énergie. Un premier grand télescope est en cours d'installation à La Palma et l'acquisition des premières données est prévue pour fin 2018. Le consortium de CTA regroupe environ 1200 membres, une centaine d'instituts de recherche de 25 pays dont la France qui en est un acteur majeur.

Ce stage s'effectue dans le cadre d'une collaboration entre informaticiens du laboratoire LIRMM (UMR 5506) et physiciens du laboratoire LUPM (UMR 5299). Le LUPM est en charge de la production et de l'analyse de simulations Monte Carlo sur la grille de calcul européenne EGI [3]. Le temps de calcul associé est très important, i.e. *environ 214 millions d'heures CPU HS06* [4] par an.

Ce stage participe aux travaux *d'optimisation des performances du logiciel corsika* [5] actuellement en cours au sein de l'équipe [6]. *Corsika* consomme environ 70% du CPU globale dans cette chaîne de simulation. Le moindre gain de performance permet à la fois une réduction des coûts d'utilisation des infrastructures de calcul et une amélioration scientifique grâce à une meilleure caractérisation de l'instrument.

2. Objectifs du stage

Le stage consiste d'abord à se familiariser aux outils d'optimisation automatique (état de l'art des compilateurs et environnements de compilation itérative) puis à les exploiter sur *corsika* pour obtenir une première version optimisée automatiquement. Dans un second temps il s'agit d'implémenter et évaluer un petit nombre d'optimisations manuelles (préalablement identifiées avec les encadrants) en exploitant entre autre des techniques de vectorisation (instructions SIMD) et en prenant en compte la qualité numérique des calculs. Une analyse comparative des optimisations manuelles et automatiques sera menée à l'aide d'outils spécialisés. Une synthèse mettra en avant des pistes d'amélioration du processus d'optimisation automatique par la prise en compte de la qualité numérique de l'application. Cette recherche se poursuivra dans le cadre d'une thèse.

3. Prérequis

Le/la candidat/e aura une très bonne aisance en programmation. Une connaissance de C/C++ et du Fortran sera un plus. Il/elle aura aussi des connaissances en architecture des machines et sur les processus de compilation.

4. Durée, lieu, encadrement et aspects logistiques

- Ce stage (4 à 6 mois) se déroulera au sein de l'équipe DALI à l'Université de Perpignan Via Domitia.
- Encadrants : Luisa Arrabito (LUPM) arrabito@in2p3.fr
David Parello (DALI/LIRMM/UPVD) david.parello@univ-perp.fr
- Logistique : Des facilités d'hébergement sur le campus (résidence universitaire) peuvent être organisées.
Nous contacter très vite à ce sujet.

5. Références

[1] <https://www.cta-observatory.org/>

[2] <https://www.eso.org/>

[3] <https://www.egi.eu/>

[4] <http://w3.hepik.org/benchmarking.html>

[5] <https://www.ikp.kit.edu/corsika>

[6] https://indico.cern.ch/event/587955/contributions/2938074/attachments/1679043/2707365/CHEP_2018_CTAOptSim_v2.pdf

Informatics master internship

HPC optimization for simulation code in astrophysics

1. Scientific context

The Cherenkov Telescope Array (CTA) [1], is the next-generation instrument in the field of very high-energy gamma-ray astronomy. It will be composed of an array of about one hundred of telescopes located at La Palma (Canary Island, Spain) and Paranal (ESO [2], Chile). The first telescope is currently under installation at La Palma, and the first data are expected by the end of 2018. CTA consortium gathers about 1200 members coming from about a hundred of research institutes in 25 countries, among which France plays a major role.

This internship will take place in the context of a collaboration between computer scientists of the LIRMM/CNRS laboratory and astrophysicists of the LUPM/CNRS laboratory. LUPM is in charge of the CTA Monte Carlo simulation production and analysis on the EGI grid infrastructure [3]. The associated computing time is very high, of the order of 200 millions HS06 [4] hours per year.

The student will participate to the on-going work [6] on the performance optimization of the corsika software [5] which consumes about 70% of the overall CPU time of the whole simulation chain. Reducing the CPU time devoted to simulations would allow to reduce e-infrastructure costs as well as to improve scientific performances thanks to a better instrument characterisation.

2. Objectives

In the first phase of the internship, the student will get familiar with automatic optimization tools (state of art of compilers and iterative compilation environments), then he/she will apply them on corsika to obtain a first automatically optimized version. In the second phase of the internship, he/she will implement a few manual optimization versions, with a particular focus on vectorization facilities (SIMD instructions) together with computing precision tuning. Finally, by means of specialised tools, he/she will conduct a comparative analysis of the automatic and manual optimizations. This study will allow to identify the potential improvements of the automatic optimization process, in which is incorporated the numerical precision tuning. These studies will be continued in the context of a future PhD thesis work.

3. Prerequisites

The retained candidate will have good programming skills. The knowledge of C/C++ and Fortran languages is a plus. He/she should have intermediate knowledge of processor architectures and compilation process.

4. Duration, location, supervision and logistics

- The internship will have a duration of 4 to 6 months. It will take place at the University of Perpignan Via Domitia, within the DALI/LIRMM team.

- Supervisors: Luisa Arrabito (LUPM) arrabito@in2p3.fr
David Parello (DALI/LIRMM/UPVD) david.parello@univ-perp.fr

- Logistics: Accommodation within the university campus can be arranged (résidence universitaire).
Contact us soon to benefit of accommodation facilities.

5. References

[1] <https://www.cta-observatory.org/>

[2] <https://www.eso.org/>

[3] <https://www.egi.eu/>

[4] <http://w3.hepik.org/benchmarking.html>

[5] <https://www.ikp.kit.edu/corsika>

[6] https://indico.cern.ch/event/587955/contributions/2938074/attachments/1679043/2707365/CHEP_2018_CTAOptSim_v2.pdf