

2 Offres d'emploi (CDD : 30 mois) : ingénieur (H/F) en développement d'applications scientifiques / calcul intensif, dans le milieu de la recherche.

Sujet : Développement de la bibliothèque XIOS
Domaine : Calcul intensif, modélisation en science du climat.
Date d'embauche : A partir de mars 2016

Contexte :

Les modèles de simulation du système climatique produisent un important volume de données au cours de simulations qui sont réalisées sur des calculateurs pouvant compter jusqu'à 100 000 cœurs de calcul. A cette échelle, les écritures sur disques et le post-traitement des données, deviennent alors un important goulet d'étranglement. Afin de générer efficacement et exploiter de manière performante l'énorme flux de données généré par ses simulations, l'IPSL (Institut Pierre Simon Laplace) et la Maison de la simulation collabore au développement de la bibliothèque XIOS.

Cette bibliothèque, dédiée au calcul intensif, permet de gérer efficacement et simplement les entrée/sortie parallèles des données sur les systèmes de stockage. Dans cette nouvelle approche, orientée client/serveur, des cœurs de calcul sont exclusivement dédiés aux I/O de façon à minimiser leur impact sur le temps de calcul des modèles. L'utilisation des communications asynchrones entre les modèles (clients) et les serveurs I/O permet de lisser les pics I/O en envoyant un flux de données constant au système de fichiers tout au long de la simulation, recouvrant ainsi totalement les écritures par du calcul.

Par ailleurs, XIOS implémente des fonctionnalités de workflow parallèle permettant d'effectuer le post-traitement des données « in situ », c'est-à-dire en parallèle et sur l'ensemble des cœurs de calcul d'une simulation. Ce système de workflow, décrit la chaîne de traitement allant de l'extraction des données en provenance des modèles (flux d'entrée), le chaînage des opérateurs transformant et combinant ces flux et leur sortie dans les différents fichiers associés. Cette description s'effectue de manière souple à partir d'un fichier d'entrée XML (analysé au moment de l'exécution), déchargeant les modèles de cette tâche. Ce fichier XML est structuré hiérarchiquement et implémente des concepts d'héritages parent-enfants, amenant à des définitions de la chaîne de traitement extrêmement souples et compactes.

La bibliothèque XIOS est écrite en C++ (~80 000 lignes de codes) et utilise la bibliothèque MPI pour ses communications asynchrones entre clients et serveurs. Afin d'agréger la bande passante du système de fichiers parallèle des calculateurs, XIOS fait appel aux I/O parallèles à travers les bibliothèques NETCDF4, HDF5 et MPI-IO. Aujourd'hui, XIOS est utilisé par différents instituts en France et en Europe, et sa base d'utilisateurs est en constante progression.

Description du poste :

Dans le cadre du projet Européen IS-ENES2 et du projet ANR CONVERGENCE, des centres d'excellence ESIWACE (HPC climat) et EoCoE (HPC énergie) le Laboratoire des Sciences du Climat et de l'Environnement (CEA/LSCE) et la Maison de la simulation (MdS) recrutent conjointement deux ingénieurs (CDD 30 mois) pour faire évoluer XIOS.

Les évolutions principales se feront sur 3 axes principaux:

- Implémentation d'une version massivement multithreadée : actuellement la parallélisme dans XIOS est géré exclusivement à travers la bibliothèque MPI et afin

de profiter des nouvelles architectures « many-core », nous souhaitons ajouter un niveau de parallélisme OpenMP.

- Développement de fonctionnalités de couplage entre modèles parallèles : les flux de données décrites à travers le workflow pourront avoir comme point de sortie un autre modèle.
- Ouverture de XIOS à d'autres communautés scientifiques nécessitant de relâcher certaines contraintes associées à la communauté d'origine (climat) et l'intégration de nouveaux formats de sortie de fichiers plus généraliste.

Au-delà de ces aspects, une attention particulière sera portée à la performance de la bibliothèque, aussi bien au niveau du parallélisme (MPI, OpenMP) que des I/O.

L'activité du poste implique une forte composante « développement » qui nécessite une solide expérience en développement applicatif C++ et un savoir-faire en programmation parallèle (MPI, OpenMP). Ce savoir-faire en programmation parallèle pourra éventuellement être renforcé au cours du contrat. D'autre part, la gestion et le maintien des sources sous le gestionnaire de sources SVN devront être assurés. Nous souhaitons également améliorer la documentation existante, ce qui nécessitera le goût et la capacité à rédiger en Anglais. Enfin, un « support » devra être assuré aux utilisateurs de XIOS, à la fois en les aidant à intégrer la bibliothèque dans leurs modèles mais aussi en suivant de près son utilisation (remontée de bugs, corrections,...).

Lieu de travail :

La mission se déroulera alternativement au LSCE (Laboratoire des sciences du climat et de l'environnement) et à la maison de la simulation sur le site du CEA/Saclay.

Salaire :

Le salaire est indexé suivant les grilles du CNRS et/ou du CEA, en fonction des diplômes et de l'expérience. A titre indicatif entre 25 K€ et 35 K€ brut.

Compétences indispensables :

- Bon niveau en C++.
- Expérience en développement d'applications scientifiques.

Compétences appréciées (par ordre d'importance) :

- Maîtrise du parallélisme en mémoire distribuée (MPI).
- Maîtrise du parallélisme en mémoire partagée, programmation multithread (OpenMP).
- Expérience en calcul intensif sur super-calculateur (connaissance du monde UNIX, gestionnaires de ressource, calculs parallèles).
- Connaissance du Fortran.
- Programmation sur accélérateur de calcul (GPU, Xeon Phi).
- Bon niveau d'Anglais.
- Goût pour la rédaction (documentation).
- Gestion de codes source sous Subversion (SVN).

Contact :

<u>CEA/LSCE</u> Yann Meurdesoif Tél : 01 69 08 17 56 Email : yann.meurdesoif@cea.fr	<u>CEA/Maison de la Simulation</u> Julien Dérouillat Tél : 01 69 08 60 70 Email : julien.derouillat@cea.fr
--	---