



Offre n°2024-07213

Doctorant F/H Analyse dynamique des données in situ et en transit pour l'Exascale Computing

Type de contrat : CDD

Niveau de diplôme exigé : Bac + 5 ou équivalent

Autre diplôme apprécié : Master's degree

Fonction : Doctorant

A propos du centre ou de la direction fonctionnelle

Le centre Inria Rennes - Bretagne Atlantique est un des huit centres d'Inria et compte plus d'une trentaine d'équipes de recherche. Le centre Inria est un acteur majeur et reconnu dans le domaine des sciences numériques. Il est au cœur d'un riche écosystème de R&D et d'innovation : PME fortement innovantes, grands groupes industriels, pôles de compétitivité, acteurs de la recherche et de l'enseignement supérieur, laboratoires d'excellence, institut de recherche technologique

Contexte et atouts du poste

Contexte

Le domaine du calcul hautes-performances vient d'atteindre une nouvelle étape, avec le dépassement du seuil exaflopique par les supercalculateurs les plus puissants du monde. Ces machines permettront de traiter de quantités de données à une échelle sans précédent, pouvant alimenter des simulations de phénomènes complexes avec une précision supérieure dans de nombreux domaines applicatifs : l'astrophysique, la physique des particules, la santé, la génomique, etc. A titre d'exemple, il est estimé que le projet SKA va traiter **un exaotet de données brutes par jour**. En France, l'installation du **premier supercalculateur exaflopique** est prévue en 2025. Des acteurs majeurs de la communauté scientifique française du domaine du calcul haute performance (*high-performance computing* - HPC) se sont regroupés au sein du programme PEPR NumPEX (<https://numpex.fr/>) pour des recherches visant à contribuer à la conception et à la mise en œuvre de l'infrastructure logicielle de cette machine. Cette thèse se déroulera dans le cadre du projet Exa-DoST de ce programme, qui s'intéresse aux **défis liés à la gestion des données à l'échelle exaflopique**.



Direction de thèse

- Gabriel Antoniu (Inria, directeur de thèse)
- Laurent Colombet (CEA, co-directeur)
- Julien Bigot (co-encadrant, Maison de la Simulation, CEA)

Localisation, mobilité

La thèse, co-encadrée par l'Inria et le CEA, sera hébergée par l'équipe KerData au **Centre de recherche Inria de l'Université de Rennes** et comprendra des visites régulières au CEA à Bruyères-le-Châtel et à la Maison de la Simulation. Rennes est la capitale de la Bretagne, dans l'ouest de la France. Elle est facilement accessible grâce à la ligne de train à grande vitesse vers Paris. Rennes est une ville dynamique, vivante et un centre important d'enseignement supérieur et de recherche : 25 % de sa population est étudiante.

Cette thèse comprendra également des collaborations avec des partenaires internationaux, notamment américains.

L'équipe KerData en quelques mots

- KerData est une équipe à taille humaine composée actuellement de 5 chercheurs permanents, 2 chercheurs contractuels, 1 ingénieur et 5 doctorants. Vous travaillerez dans un environnement chaleureux, offrant un **bon équilibre entre vie professionnelle et vie privée**.
- KerData mène de nombreux projets ambitieux dans des environnements collaboratifs nationaux et internationaux de haut niveau, notamment au sein du **Joint-Laboratory on Extreme-Scale Computing** : <https://jlesc.github.io>. Notre équipe collabore activement avec des institutions académiques de premier plan dans le monde entier (notamment aux États-Unis, en Espagne, en Allemagne ou au Japon) et avec l'industrie.
- Notre équipe privilégie fortement la **recherche expérimentale**, validée par la mise en œuvre et l'expérimentation de prototypes de logiciels avec des applications réelles sur des plates-formes réelles, y compris certains des superordinateurs les plus puissants au monde.
- L'équipe de KerData s'engage à fournir des conseils et un coaching personnalisés, afin d'aider les candidats au doctorat à se former et à se développer dans toutes les directions qui sont essentielles pour devenir des chercheurs performants.
- Pour en savoir plus sur l'équipe KerData, consultez notre site web <https://team.inria.fr/kerdata/>

Mission confiée

Introduction

Sans évolution majeure des pratiques, l'augmentation de la capacité de calcul des prochaines générations d'ordinateurs conduira à une explosion du volume des données produites par les simulations numériques. La gestion de ces données, de leur production à leur analyse, est un défi majeur.

L'exploitation des résultats des simulations repose sur un protocole calcul-stockage-calcul bien établi. Le différentiel d'évolution des capacités entre les calculateurs et les systèmes de fichiers rend inéluctable l'engorgement de ces derniers. S'il n'est pas envisageable de se passer d'un système de stockage, de nombreuses expérimentations visent à en réduire l'utilisation. Dans ce domaine, l'émergence d'un vocabulaire varié traduit la diversité des approches : *in situ processing*, *in transit processing*, staging node, helper cores. Toutes ces approches soulignent une volonté de remplacer le processus écriture-lecture, usuel pour l'enchaînement des applications, par une intervention en vol. Une analyse effectuée dans le même temps qu'une simulation est une capacité qui intéresse particulièrement les physiciens du CEA. Ce besoin a conduit à la réalisation des premières implantations de systèmes d'analyses *in situ* ou *in transit* dans les codes de simulation et à la création de middlewares spécifiques comme par exemple Damaris [1,2,3,4].

Contexte technologique et applicatif

Damaris (<https://project.inria.fr/damaris/>) est un middleware pour la gestion des E/S et le traitement en temps réel des données issues des simulations HPC basées sur MPI à grande échelle. Il proposait initialement de dédier des cœurs aux E/S asynchrones dans les nœuds multicœurs des plateformes HPC récentes, en mettant l'accent sur la facilité d'intégration dans les simulations existantes, l'utilisation efficace des ressources (grâce à l'utilisation de la mémoire partagée) et la simplicité d'extension par le biais de plug-ins. Au fil des ans, Damaris a évolué vers un système plus élaboré, offrant la possibilité d'utiliser des cœurs ou des nœuds dédiés pour effectuer le traitement et la visualisation des données *in situ*. Il propose une connexion transparente au logiciels de visualisation VisIt ou ParaView pour permettre la visualisation *in situ* avec un impact minimal sur le temps d'exécution de la simulation. Damaris fournit une interface utilisateur (API) extrêmement simple et peut être facilement intégré dans les simulations à grande échelle existantes. Il a été validé jusqu'à 14 000 cœurs sur les supercalculateurs tels que Titan (1er du Top500 au moment des expériences), Jaguar, Kraken, etc., avec de nombreux codes de simulations. Damaris est une des briques logicielles destinées à être utilisée sur le premier supercalculateur exaflopique qui sera installé en France.

PDI est une API qui permet un couplage faible entre codes de simulation et bibliothèques de gestion de données pour le calcul intensif. L'approche consiste à instrumenter les codes pour identifier où et quand des données deviennent présentes en mémoire et où et quand la mémoire va être ré-utilisée pour y stocker de nouvelles valeurs. Cette instrumentation est entièrement indépendante des bibliothèques de gestion de données utilisées. Un fichier séparé au format Yaml permet de spécifier que faire des données du code et avec quelle bibliothèque. Cette approche correspond à une démarche de programmation orientée aspects. Les différents aspects pouvant ainsi être pris en compte sont nombreux et incluent par exemple la lecture des paramètres du code, l'initialisation des données, le post-traitement *in situ*, la visualisation ou le stockage des résultats sur disque, la tolérance aux pannes, l'inclusion dans un couplage de code ou dans une simulation d'ensemble. Tous ces aspects peuvent être gérés, chacun grâce à un plugin donnant accès à une bibliothèque dédiée différente. PDI ne fait que arbitrer entre ces différents plugins, mais n'offre aucune fonctionnalité propre telle que la mise à disposition de cœurs dédiés dans Damaris. PDI est utilisé dans des codes de production comme Gysela qui exploite les machines Petaflop les plus puissantes du moment telles que Fugaku, Adastra, ou Exa1-HF. L'architecture intra-processus de PDI assure que la scalabilité offerte est exactement celle des bibliothèques utilisées en back-end.

Défis

Un premier retour d'expérience d'utilisateurs d'applications, utilisant un système in situ, affiche clairement la nécessité de développer un système permettant de gérer dynamiquement l'ajout ou la suppression d'analyses pendant l'exécution d'une simulation. Par exemple, il est fréquent, lors d'une étude par simulation du comportement d'un matériau sous contraintes, de demander des observations et donc des analyses différentes en fonction de la progression temporelle de la simulation. En effet, les propriétés élastoplastiques d'un matériau peuvent changer au cours du temps et donc entraîner le déclenchement de nouvelles analyses pour comprendre les phénomènes physiques sous-jacents comme la propagation de dislocations et les éventuels changements de phases (solide vs liquide ou solide vs solide). Dans ce cadre, pour économiser du temps et des ressources de calcul, il est donc important de déclencher l'activation des nouvelles analyses au bon moment lors du déroulement de la simulation. A noter que la détection de l'événement peut être faite soit par le code de simulation soit par une analyse. Bien évidemment, pour conserver de bonnes performances de restitution des résultats, il est impératif de bien gérer le placement des nouvelles analyses sur les nœuds du calculateur. Ces capacités permettant une gestion dynamique des analyses ne sont pas encore disponibles efficacement dans Damaris.

Coddex est un code de simulation qui résout les équations de la mécanique des milieux continus en hyperélasticité dynamique (chocs ou sollicitations rapides). Il incorpore également la description des discontinuités de comportement (cf. **Figure 1**) de changement ou maillage. Coddex est l'acronyme de Code de Dynamique des Discontinuités pour l'Étude des cristaux.

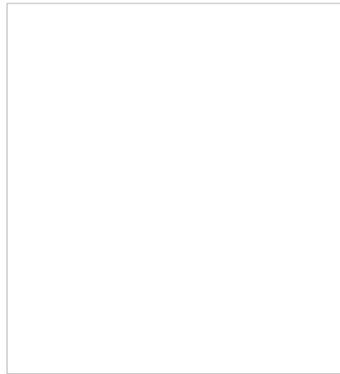


Figure 1 : Carte des déformations d'un polycristal de TATB, un matériau énergétique ultra anisotrope par le code Coddex

Exemple de scénarios pour la mise en œuvre du système in-situ dans Coddex via Damaris:

- Analyses « programmées » : l'utilisateur physicien définit une simulation physique complète et une liste d'analyses in-situ (liens Coddex-Damaris-Paraview) avec des fréquences d'exécution variables. Dans ce scénario les sorties sont réalisées sans l'aide d'opérateur (sorties programmées), sous forme de fichiers de statistiques ou de visuels 2D (images) ou 3D (par exemple sorties d'iso-surfaces 3D par Paraview sous forme de fichiers .obj).
- Analyses « à la volée » : l'utilisateur peut également (et indépendamment) lancer de nouvelles analyses via le système de pauses (attente de requêtes de l'orchestrateur). Une analyse déclenche la création et la mise en œuvre d'une nouvelle analyse avec une fréquence spécifique. De nouvelles sorties de données (initialement non-prévues) sont produites. La liste des analyses peut être modifiée à la volée par ce biais. Ce scénario créé « à la volée » pourra ensuite être incorporé dans une analyse « programmée ».
- Analyses « déclenchées » : le code détecte une discontinuité (création d'une nouvelle phase, par exemple) et déclenche une analyse adaptée, sélectionnée dans une banque d'analyses typiques.

Objectifs

Les travaux de recherches proposés dans le cadre de cette thèse, consistent à concevoir un modèle novateur de gestion dynamique des analyses in situ et in transit, d'en proposer une implémentation dans le middleware Damaris et de le valider grâce à des simulations réalisées via le code Coddex.

Principales activités

Après une étude de l'état de l'art et une prise en main de l'architecture de Damaris et du code Coddex, l'objectif est d'étudier, proposer et développer des solutions innovantes qui feront l'objet de publications dans les meilleures revues et conférences du domaine. La thèse se déroulera dans un environnement pluridisciplinaire (informatique et physique) grâce à la collaboration INRIA-CEA au sein du projet Exa-DoST du PEPR NumPEX. Ces activités permettront aussi un accès privilégié pour les expérimentations à de très grands calculateurs.

Références

- [1] M. Dreher, B. Raffin, "A Flexible Framework for Asynchronous In Situ and In Transit Analytics for

Scientific Simulations”, 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (2014).

[2] M. Dorier, G. Antoniu, F. Cappello, M. Snir, and L. Orf, “Damaris : How to Efficiently Leverage Multicore Parallelism to Achieve Scalable, Jitter-free I/O”, in CLUSTER – IEEE. International Conference on Cluster Computing. IEEE, Sep. 2012.

[3] M. Dorier, M. Dreher, T. Peterka, J. Wozniak, G. Antoniu and B. Raffin, “Lessons Learned from Building In Situ Coupling Frameworks”, in Proceedings of the First Workshop on In Situ Infrastructures for Enabling Extreme-Scale Analysis and Visualization, Austin, 2015.

[4] E. Dirand, L. Colombet, B. Raffin, “TINS: A Task-Based Dynamic Helper Core Strategy for In Situ Analytics”, in Proceedings of Asian Conference on Supercomputing Frontiers, Singapore 2018.

Compétences

- Un excellent master en informatique ou équivalent
- Solide connaissance des systèmes distribués
- Connaissance des systèmes de stockage et de fichiers (distribués)
- Capacité et motivation à mener des recherches de haute qualité, y compris à publier les résultats dans des revues spécialisées.
- Solides compétences en programmation (Python, C/C++)
- Une expérience professionnelle dans les domaines du calcul intensif et de la gestion des données massives (Big Data) est un avantage.
- Très bonnes capacités de communication en anglais, à l'oral et à l'écrit
- Ouverture d'esprit, fortes capacités d'intégration et esprit d'équipe.

Avantages

- Restauration subventionnée
- Transports publics remboursés partiellement
- Possibilité de télétravail à hauteur de 90 jours annuels
- Prise en charge partielle du coût de la mutuelle

Rémunération

rémunération mensuelle brute de 2100 euros les deux premières années et 2200 euros la troisième année

Informations générales

- **Thème/Domaine** : Calcul distribué et à haute performance
Calcul Scientifique (BAP E)
- **Ville** : Rennes
- **Centre Inria** : [Centre Inria de l'Université de Rennes](#)
- **Date de prise de fonction souhaitée** : 2024-09-01
- **Durée de contrat** : 3 ans
- **Date limite pour postuler** : 2024-07-20

Contacts

- **Équipe Inria** : [KERDATA](#)
- **Directeur de thèse** :
Antoniou Gabriel / gabriel.antoniu@inria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

L'essentiel pour réussir

Le candidat devra faire preuve de motivation, d'autonomie et d'une capacité à initier des liens entre les activités de recherche menées à l'INRIA et au centre CEA.

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des

Consignes pour postuler

Merci de déposer en ligne CV, lettre de motivation et éventuelles recommandations

Pour plus d'information, contactez gabriel.antoniou@inria.fr

Sécurité défense :

Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :

Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.