

PROPOSITION DE STAGE EN COURS D'ETUDES

Référence : DMPE-2026-34 (à rappeler dans toute correspondance)	Lieu : Châtillon
Département/Dir./Serv. : DMPE/PLM	Tél. : 01 04 73 43 16
Responsable(s) du stage : B. Andrieu, B. Maugars	Email. : bastien.andrieu@onera.fr

DESCRIPTION DU STAGE

Thématique(s) : Méthodes numériques et HPC

Type de stage : ☒ Fin d'études bac+5 ☐ Master 2 ☐ Bac+2 à bac+4 ☐ Autres

Intitulé : Industrialisation des tests de performance d'algorithmes massivement parallèles

Sujet : L'ONERA développe des codes de simulation numérique couvrant l'étude de nombreux phénomènes physiques tels que la mécanique des fluides, la mécanique des solides, l'acoustique ou encore l'électromagnétisme. Pour la plupart, ces logiciels sont conçus pour être utilisés sur des calculateurs massivement parallèles.

La bibliothèque ParaDiGM¹ est développée afin de répondre aux besoins HPC de ces codes en servant de réceptacle pour des algorithmes de manipulation de maillages efficaces dans un environnement massivement parallèle distribué. Cette bibliothèque repose sur une architecture incluant des protocoles d'échanges MPI optimisés, des structures d'arbres distribués (par exemple, pour l'indexation spatiale) et des algorithmes géométriques avancés (recherche de plus proches voisins, localisation de points, calcul de distance à une surface, intersection de maillages, etc). L'enjeu est de garantir la robustesse et l'efficacité de ces briques logicielles, notamment au passage à très grande échelle (dizaines de milliers de cœurs, maillages de plusieurs milliards d'éléments).

L'objectif de ce stage est de mettre en place une chaîne d'évaluation des performances systématique et pérenne pour l'ensemble des algorithmes de la bibliothèque. Ce travail est divisé en trois phases principales :

1. Délimitation du périmètre et instrumentation :

- Périmètre technique : Établir, en collaboration avec les développeurs, un périmètre d'étude précis des cas tests critiques qui poussent les algorithmes dans leurs retranchements (géométries complexes, fort déséquilibre de charge, communications creuses, large nombre de processus).
- Instrumentation : Isoler et instrumenter les blocs de code critiques des différentes familles d'algorithmes pour permettre une mesure précise de la charge CPU, des temps de latence et du volume de communication MPI ainsi que de la consommation mémoire.

2. Automatisation et campagne de tests :

- Automatisation : Concevoir et développer des scripts Python pour orchestrer le lancement des campagnes de tests sur notre cluster HPC interne (gestion des allocations de ressources, exécution, collecte des données brutes).
- Analyse de scalabilité : Mener des tests de scalabilité rigoureux afin d'identifier les goulots d'étranglement, l'efficacité des communications MPI et les lois d'échelle de chaque algorithme.
- Post-traitement : Développer des outils en Python pour synthétiser et visualiser les résultats de ces tests, permettant une analyse ciblée des régressions et des gains.

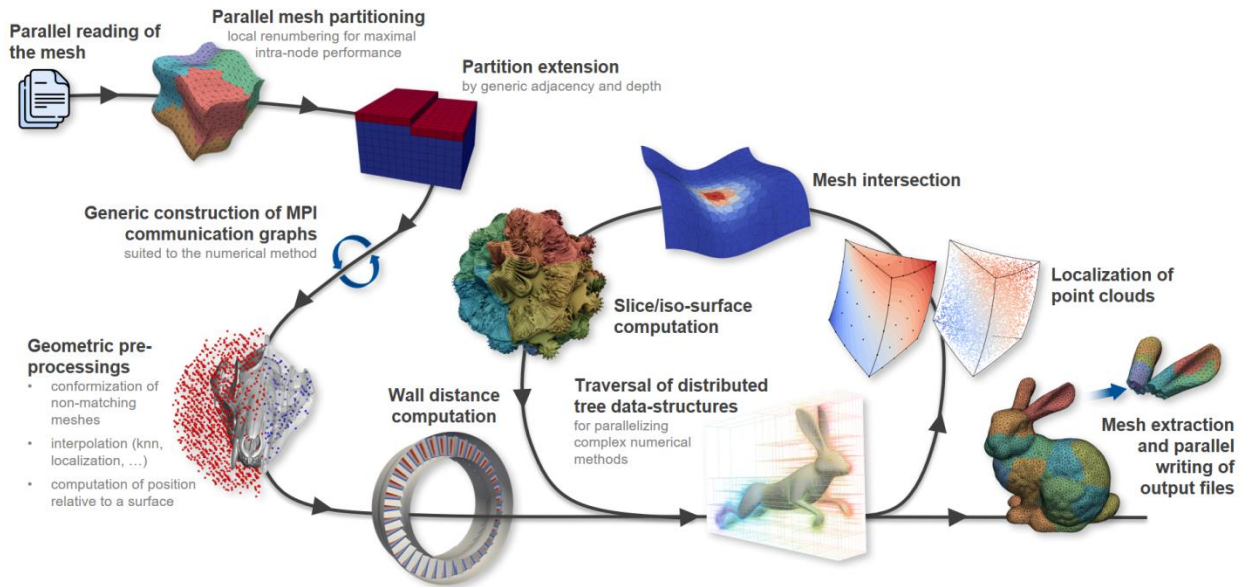
3. Pérennisation et intégration continue :

- Intégration : Intégrer la base de tests automatisée directement dans la *pipeline* GitLab CI/CD. Ce stage aboutira à la mise en place d'un job de performance capable de vérifier automatiquement les performances de la bibliothèque après chaque *Merge Request*.

¹ ParaDiGM (Parallel Distributed General Mesh) : <https://github.com/onera/paradigm>

- Rapports : Fournir une analyse finale détaillée et entrevoir des optimisations basées sur les données collectées.

Le stage se déroulera dans l'unité PLM du DMPE en collaboration les développeurs d'autres unités (DAAA/CLEF et DMPE/HEAT).



Exemple de chaîne de calcul utilisant les fonctionnalités proposées par ParaDiGM

Est-il possible d'envisager un travail en binôme ? Non

Méthodes à mettre en œuvre :

☐ Recherche théorique

☒ Recherche appliquée

☐ Recherche expérimentale

☒ Travail de synthèse

☒ Travail de documentation

☒ Participation à une réalisation

Possibilité de prolongation en thèse :

Non

Durée du stage :

Minimum : 5 mois

Maximum : 5 mois

Période souhaitée : Mars 2026

PROFIL DU STAGIAIRE

Connaissances et niveau requis :
Bonnes notions en HPC et calcul parallèle distribué (MPI).
Bonne maîtrise des langages C et Python.
Des connaissances des outils de gestion de version Git/GitLab et des plateformes de *batch scheduling* (SLURM) seraient un atout.

Ecoles ou établissements souhaités :
Ecole d'ingénieur ou master 2 recherche en informatique ou calcul scientifique