

Présentation

Code interne : EIN9-PROJ1

Description

La simulation numérique génère des volumes de données considérables, et les approches traditionnelles de post-traitement atteignent aujourd'hui leurs limites face aux contraintes de bande passante I/O et de capacité de stockage des supercalculateurs modernes. Le traitement in-situ propose une alternative en réalisant l'analyse directement pendant l'exécution de la simulation, réduisant ainsi les écritures disque et permettant un accès immédiat aux résultats.

Ce cours présente les enjeux, les principes et les solutions logicielles du traitement in-situ. Les étudiants découvrent ces concepts à travers un projet fil rouge basé sur une application de simulation de propagation d'ondes acoustiques utilisant la méthode des éléments spectraux. Ils mettent en place un workflow ad-hoc classique avec sauvegarde et post-traitement, puis implémentent progressivement des traitements in-situ incluant calculs statistiques et génération d'images. Le cours aborde également les techniques de compression de données et les formats de visualisation adaptés au contexte HPC.

L'accent est mis sur l'analyse comparative rigoureuse des deux approches selon des critères de performance, de flexibilité et de maintenabilité. Les étudiants apprennent à positionner leurs solutions selon la taxonomie de référence proposée par Childs et al. (2020), leur permettant de caractériser précisément les choix architecturaux de leurs workflows. À l'issue du cours, ils sont capables d'identifier les situations où chaque approche est pertinente et de concevoir des workflows adaptés aux contraintes spécifiques du calcul haute performance.

Heures d'enseignement

TD	Travaux Dirigés	14h
CI	Cours Intégrés	1,33h

Pré-requis obligatoires

C++, Python, OpenMP, MPI

Syllabus

Les technologies utilisées seront Python, C++, Paraview et Catalyst (un peu de parallélisme OpenMP+MPI). Les TP se baseront sur des exemples simples afin de bien appréhender la logique de structuration des données propre à Catalyst et les techniques de génération de code de visualisation Paraview-Python. Une deuxième partie des TP portera sur l'utilisation de la visualisation in-situ avec données distribuées. Le mini-projet permettra une mise en application des précédentes séances de TP dans un code de simulation existant.

Informations complémentaires

L'ère de l'ExaScale creusera encore plus l'écart entre la vitesse de génération des données de simulations et la vitesse d'écriture et de lecture pour analyser ces données en post-traitement. Le temps jusqu'à la mise à disposition des résultats sera donc grandement impacté et de nouvelles techniques de traitement des données doivent être mises en place. Les méthodes in-situ réduisent le besoin d'écrire des données en les analysant directement là où elles sont produites. Il existe plusieurs techniques, en exécutant les analyses sur les mêmes nœuds de calcul que la simulation (in-situ), en utilisant des nœuds dédiés (in-transit) ou en combinant les deux approches (hybride). La plupart des méthodes in-situ ciblent les simulations qui ne sont pas capables de tirer profit à 100% du nombre croissant de cœurs par processeur.

Le module se focalisera sur la visualisation de données in-situ dans le cadre d'une application scientifique.

Modalités de contrôle des connaissances

Évaluation initiale / Session principale

Type d'évaluation	Nature de l'évaluation	Durée (en minutes)	Nombre d'épreuves	Coefficient de l'évaluation	Note éliminatoire de l'évaluation	Remarques
Projet	Compte-Rendu			1		

Infos pratiques

Contacts

Intervenant

Alexis Bandet

✉ Alexis.Bandet@bordeaux-inp.fr

Responsable de la filière

Mathieu Faverge

✉ Mathieu.Faverge@bordeaux-inp.fr