



Hackathon HPC

Optimisation de codes industriels

Principe

Avec ce hackathon, nous visons à rassembler les étudiants HPC de niveau M2 dans une compétition virtuelle autour des codes de calcul fournis notamment par EDF (code Saturne et Telemac) et la CGG (noyau sismique). D'autres partenaires industriels pourraient venir compléter cette première liste.

Cette compétition s'appuiera sur les instances AWS basées sur les technologies Arm. En effet, les architectures cibles (processeur AWS Graviton 2 et 3) proposent certaines approches (écosystème logiciel, design) motivant un effort spécifique par rapport aux architectures classiques de type Intel ou AMD. Typiquement, sur le processeur Graviton3, on peut citer l'exploitation des unités SIMD (Scalable Vector Extension/SVE par rapport à l'AVX2 ou l'AVX-512 sur architectures x86) ou encore la disponibilité de 8 canaux mémoire de type DDR5.

Ce hackathon est structuré autour de codes de calcul, d'environnements logiciels et de solutions matérielles déjà éprouvés par les industriels. Les recettes de compilation, les phases d'optimisation ont été validées en amont de cet événement. Les étudiants seront donc dans un cadre proche d'une session de travaux pratiques guidés avec l'opportunité d'accroître leur compréhension des enjeux industriels autour de la simulation haute performance. Il n'est donc pas attendu de contributions spécifiques venant des étudiants en termes de portage ou d'optimisation (ces codes sont largement éprouvés dans un contexte de production et sur de multiples architectures matérielles). Les participants seront néanmoins crédités par les partenaires industriels (les codes ont un statut ouvert) en cas d'avancées significatives.

Modalités

Les étudiants sont encouragés à adopter une approche itérative pour le portage et l'optimisation de ces codes/noyaux industriels. Ces derniers sont de nature/complexité différentes et permettront de déployer les phases classiques de prise en main d'une application scientifique. L'évaluation des équipes (4 étudiants maximum par équipe) se fera notamment autour des points suivants

- **Portage** : il s'agira de valider l'application sur architectures Arm (Graviton2/Graviton3) en se focalisant sur le cas test fourni par le partenaire industriel. La validation s'effectuera par le biais d'une comparaison des fichiers résultats et/ou en comparant les résultats sur différentes plateformes (x86/Arm).
- **Profiling** : Il est demandé aux étudiants d'utiliser les outils classiques de profilage des applications permettant d'identifier les verrous en termes de performance... il s'agira notamment d'identifier les hotspots de ces applications (rapport du compilateur, analyse dynamique du code...)
- **Optimisation avancée** : Durant cette phase, les étudiants pourront apporter certaines modifications aux codes afin d'améliorer les performances. Pour les applications de taille modeste (e.g. code CGG, il pourra s'agir de rajouter des directives OpenMP ou de modifier l'organisation des boucles ...). Dans le cas des codes complexes, les participants pourront se concentrer sur l'impact des différentes chaînes de compilation et travailler à l'extraction de



certaines noyaux (mini-apps). Ces dernières pourraient par la suite faire l'objet d'efforts spécifiques.

- Les compilateurs Arm, Nvidia et GNU seront notamment disponibles durant ce hackathon. Les étudiants auront également accès à différents outils de profiling de codes (e.g. Arm Allinea Forge).
- La synthèse des résultats obtenus (recettes de compilation, procédure de validation, optimisation, extraction de boucles ...) sera mise à disposition sur un dépôt de type Git afin de faciliter l'évaluation et la reproductibilité.

Codes proposés / Noyaux identifiés (Juin 2022)

CGG

Noyau sismique : www.cgg.com

- Focus : validation numérique (comparaison x86 Intel/AMD), évaluation de différents compilateurs (Arm, Nvidia, GNU), Vectorisation .

EDF

Code_Saturne : <https://www.code-saturne.org/cms/web/>

- Focus : portage et validation numérique (comparaison intel/AMD), évaluation de différents compilateurs, éventuellement extraction de noyaux (hotspots) pour analyse avancée.

Code Telemac : <http://www.opentelemac.org/>

- Focus : portage et validation numérique (comparaison intel/AMD), évaluation de différents compilateurs, éventuellement extraction de noyaux (hotspots) pour analyse avancée.

Format

Planning :

- Annonce lors du forum 2022 : 14-15 juin 2022
- Inscription des équipes : septembre 2022
- Webinaire de présentation : **le 7 octobre de 16h00 à 17h30**
- Formation libre : octobre / novembre lancée par un webinar début octobre.
- Hackathon du **lundi 28 novembre 2022 à 9h00 au lundi 5 décembre 2022 à 9h00**
 - Travail estimé à 48h laissé au choix de l'institution dans le cadre de la semaine
 - Equipe de 4 étudiants
 - Support ARM/AWS

Prix

La compétition donnera lieu à un classement, et l'équipe « victorieuse » se verra attribuer un lot basé sur des processeurs ARM et qui sentent la pomme 🍏.

Ancrage industriel

Spécificités des codes :

Les codes HPC sont représentatifs de l'intérêt et de la préoccupation des partenaires industriels de TERATEC. La maîtrise du code source et des dépendances (possibilité de recompiler le code) constitue un point important, comme pour tout portage d'une architecture source vers une architecture nouvelle.

Chaque code sera présenté par le partenaire industriel pour expliquer son intérêt. Cette présentation aura lieu lors d'un webinar début octobre devant les étudiants inscrits lors duquel la plateforme sera également présentée.

ARM/AWS/UCIT

L'Europe a une stratégie pour construire un processeur européen (EPI) utilisant le jeu d'instructions basé sur Arm (AArch64) qui devrait arriver sur le marché en 2023 avec SiPearl. Ce processeur sera notamment utilisé par ATOS comme annoncé récemment avec leur dernière génération de supercalculateur).

Avec Amazon graviton2 en 2019, AWS a introduit sur sa plateforme Cloud la possibilité pour les communautés du monde entier de développer, construire et tester des applications à l'aide du jeu d'instructions arm64. L'adoption a été impressionnante avec de nombreuses applications open source ou commerciales exécutées sur AWS Cloud et des investissements importants pour améliorer la chaîne d'outils de développement logiciel.

Avec AWS graviton3 disponible depuis mi-mai, AWS ajoute désormais des mécanismes supplémentaires (SVE par exemple) pour prendre en charge une charge de travail technique avancée telle que le calcul haute performance. AWS veut fournir une alternative précieuse à la solution actuelle (principalement basée sur x86, avec une forte adoption des GPU ces derniers temps) et donnera accès aux chercheurs et ingénieurs pour construire, tester et prouver leur code de simulations numériques en utilisant une architecture qui sera très similaire à celui que fournira SiPearl Rhea.

UCit administrera et mettra à disposition les ressources de calcul nécessaires aux participants du Hackathon. Grâce à Cluster-in-a-Box (<https://ucit.fr/index.php/cluster-in-a-box/>), chaque équipe pourra disposer de son environnement HPC personnalisé sur le Cloud AWS et y accéder de façon transparente

Support avant Hackathon :

Formation à l'utilisation de la plateforme AWS + compilation et analyse du comportement par ARM et AWS.

Support pendant Hackathon :

L'utilisation de messagerie slack pour le support en mode asynchrone est envisagée afin de fournir une assistance technique aux différentes équipes. Chaque équipe aura un canal spécifique et des canaux techniques partagés seront disponibles (en fonction des applications par exemple), afin de permettre aux équipes techniques Arm et AWS de lever d'éventuels verrous.

Contacts :

TERATEC : Emmanuelle Vergnaud – emmanuelle.vergnaud@teratec.fr

ARM : Conrad Hillairet - Conrad.Hillairet@arm.com

AWS : Gilles Tourpe - gtourpe@amazon.com