# Réunion calcul CPHT du 29/06/2021

Présents :

ITA : équipe informatique, A. Canou

Chercheurs : A. Georges, S. Backes, H. Lutjens, A. Subedi, B. Van-Rees, L. Pourovskii, A. Couairon, S. Huller

Ordre du jour :

1. Retour questionnaire sur l’état des lieux du calcul au CPHT
2. Mise en production / lancement officiel du cluster IDCS Cholesky

# Questionnaire sur l’état des lieux du calcul

Impressions générales :

*9 personnes ont répondu. C’est à la fois peu par rapport à la totalité des gens qui calculent dans notre unité mais suffisamment représentatif des différentes équipes de recherche (et leurs diversités de besoin).*

## Les besoins

### Quels types de calcul ?

Les deux types : développement et production

### Quelles machines de calcul ?

Machine locale + machine GENCI

### Type de jobs

Majoritairement des calculs parallèles

### Nombre de cœurs CPU par job

2 ont répondu > à 512 coeurs

### Nombre de cœurs GPU par job

*Ambiguïté de la question : 0 ou Moins de 8 ? Y’a-t-il des besoins ?*

*Hinrich : pour le portage de son code dans le cadre du projet EUROfusion, une équipe européenne y travaille. En attente de retour (2022 ?).*

*Pour les autres utilisateurs, il n’y pas eu d’expression de besoin en ressources GPU à court et moyen terme. Compte tenu du retour d’expérience d’Aurélien, cela demande un effort et un temps de portage du code (Cuda, OpenCL) qui n’apporterait pas forcement d’amélioration significative pour le moment. Cela dépend du type code utilisée. Les calculs GPU ont la côte dans les domaines du Big Data, du Machine Learning et de l’IA.*

### Taille des données que vous générez ?

Hinrich : > 512 G *Pourquoi autant ? nature des données ? besoin d’archivage ?*

*Archivage des données sur 3 / 5 ans (durée d’une thèse).*

*Hinrich constate des problèmes de débit réseau dans son bureau vers/depuis l’extérieur (4 à 6 Mo/s). Cela nécessite de qualifier le problème.*

*Pour le stockage, il faut prévoir une réunion pour en discuter.*

### Pour vos calculs, avez-vous besoin d’un réseau ?

Réponses partagées entre un réseau à très faible latence comme l’IB et Ethernet.

*Il serait intéressant de tester via un POC (que l’équipe se propose de monter) s’il y’a des différences notables de performances entre les 2 technologies via plusieurs types de codes parallélisés, sachant que la technologie Ethernet évolue rapidement à un cout plus « raisonnable ». Et de voir ensuite s’il est intéressant d’investir encore dans la technologie IB compte tenu d’une part du prix prohibitif des équipements (switchs) et de la topologie réseau (FAT TREE) qui peut devenir complexe et onéreuse modulo le nb de nœuds de calcul.*

## Les clusters du CPHT et Hopper

Proposition d’une stratégie :

* **Pas d’investissement dans les clusters locaux**. On maintient l’existant à minima (ex. changement de disque si possible en cas de panne).
  + Pour les nœuds affichés « down » sur Montblanc, il est possible que cela provient d’alimentations défectueuses. Il serait envisageable de tester le remplacement et d’acheter le cas échéant de nouvelles alimentations.
  + Quid du matériel acheté via un financement d’ERC après la fin du projet ?
* **Utilisation du cluster IDCS Cholesky** : il faut inciter les utilisateurs à utiliser la machine. D’autre part un nouvel achat est prévu en juillet de 5 à 7 châssis de nœuds « compute » supplémentaires (5x4x40 cœurs CPU=> 800 cœurs) et 2 nœuds GPU (8 cartes) => au total : **2400 cœurs CPU et 16 GPU).** Certains utilisateurs au CPHT (de clusters locaux) ont malgré tous des réticences sur la puissance au regard du nombre de laboratoires qui vont utiliser cette ressource. Une autre remarque a été soulevé sur le manque de « transparence » et d’information de la part du comité de pilotage du mésocentre IDCS concernant la politique des queues, du « fair-share », des priorités (poids) en fonction des différents investissements des laboratoires et de l’École. Ces informations devraient être communiquées au prochain comité du pilotage (à la rentrée ?). L’équipe informatique va remonter cette remarque à la direction d’IDCS. Néanmoins pour rappel les informations concernant les queues/partitions ont déjà été communiqués il y’a 2-3 mois (au lancement de la phase de test) sur le site de documentation de Cholesky : <http://meso-ipp.gitlab.labos.polytechnique.fr/user_doc/>
* **Cluster local dédiée au développement** : montage d’un POC, création d’une stack/solution logicielle basé sur le travail de Jean-Luc Bellon sur « calculator ».
  + Objectif : cette solution logicielle doit pouvoir être déployée sur une architecture matérielle la plus générique (réseau haut débit type Ethernet) et notamment chez un prestataire dans un cloud public (voir offre « OVHCloud » de l’accord entre l’UGAP et OVH). Pour ce dernier prévoir un budget (20 KE ?) pour tester la solution.
  + Pas d’investissement matériel : utilisation de matériel existant (ex. nœuds de clusters, reprendre la machine « calculator », voir les nouveaux hyperviseurs équipés de processeurs AMD bi-proc 32 cœurs)

## Les demandes d’ANR/ERC

* Demande de ligne « Matériel »

Réponses : ligne « matérielle » plutôt que des heures de calcul.

*Compte tenu de la politique du mesocentre IDCS, il n’est pas possible sur Cholesky de dédier des queues/partitions à des projets de type ANR/ERC. Le matériel investi est mutualisé.*

*Cela pose 2 problèmes :*

* *le problème d’hétérogénéité du matériel, c.à.d. de génération de CPU/GPU.*
* *le financement d’un cluster via ANRs n’est pas réaliste*
* Demande de ligne « heures de calcul »

*Pour une demande « d’heures de calcul », le prix de l’heure n’est pas encore établi au niveau de la DAER (et soumis au CA de l’Ecole ?). Dans ce cadre, le mesocentre IDCS souhaiterait « thésauriser » l’argent récolté en complément de financements provenant de l’école, pour pouvoir ensuite investir dans une nouvelle machine (ou éventuellement une nouvelle queue).*

* Demande de moyens humains

Réponses : 1 seul a répondu « Oui ».

Les autres « Non ».

*Dans le cas des ANR, les budgets alloués ne sont pas suffisants pour l’obtention de moyens humains en soutien. Dans le cas des ERC, il n’est pas exclu que ce soit impossible mais il faudrait se rapprocher des services concernés de l’École.*

## L’externalisation

Des doutes et réticences sur l’externalisation ont été exprimés sur les aspects financiers, humains et techniques.

*Nous n’avons pas eu le temps d’approfondir ce sujet pendant la réunion.*

*Néanmoins pour les aspects financiers, un accord-cadre entre l’UGAP (centrale d’achat public) et OVH a été signé avec des tarifs compétitifs.*

*L’équipe informatique se donne comme objectif à moyen terme de développer une pile logicielle de calcul « agnostique » (sans dépendance vis-à-vis du matériel et du système d’exploitation) à iso fonctionnalités en termes d’outils (slurm, modules, etc) de ceux présents sur les clusters. Celle-ci sera dans un premier temps déployé et testé sur une infrastructure matérielle locale (type « calculator »). Dans un second temps déployé et testé sur un cloud public.*

*Objectif : fournir une alternative aux clusters locaux et mutualisée comme Cholesky.*

## Formation et compétences

* Les besoins en formation ?

*L’équipe informatique propose d’organiser 1 à 2 session(s) de formation / an en priorité pour les doctorants/post-doctorants/stagiaires. Ces formations de 3h environ, placées en fin d’année (courant novembre) ont pour objectif de fournir des connaissances seront dispensés par l’ingénieur calcul Aurélien Canou et accompagné éventuellement d’un permanent du CPHT. A la rentrée de septembre/octobre (ex. arrivée des doctorants), un sondage pourra être proposé à l’ensemble des managers d’équipe pour estimer le nb de personnes intéressés. Une attention devra être portée aux prérequis (ex. commandes de base unix) des participants, pour si besoin orienter en amont de la formation les participants vers des pointeurs de documentations (auto-formation).*

*2 formations complémentaires pourraient être proposés :*

* + *Une formation sur les « Aspects fondamentaux de l’utilisation des moyens de calculs » comme la soumission de jobs par script batch, choix des queues en fct des besoins, etc.*
  + *Une formation sur les « Bonnes pratiques de développement » comme l’utilisation de la plateforme Gitlab et des dépôts GIT*

*Pour des formations un peu plus « avancées » et techniques l’équipe informatique propose d’orienter le personnel vers des formations dispensés par :*

* *l’IDRIS (*[*http://www.idris.fr/formations/formations.html*](http://www.idris.fr/formations/formations.html)*),*
* *la maison de la simulation (*[*https://www.maisondelasimulation.fr/*](https://www.maisondelasimulation.fr/)*)*
* *PRACE*
* Tenez-vous à la proximité des ingénieurs pour vous aider dans vos calculs ?

Les réponses sont partagés : « Oui pour des questions de portage de code / développement-optimisation de mes codes» et « Oui pour des questions de gestionnaire de queue/rapatriement de données ».

*Il y’a une demande de soutien sur du développement, du portage de code, de l’optimisation au CPHT mais aussi dans les autres laboratoires partenaires d’IDCS. Une demande de poste à l’IDCS (mutualisé) a été exprimé pour l’obtention d’un ingénieur calcul l’année prochaine. Un candidat à ce poste est d’ailleurs pressenti au LPP.*

* Utilisez-vous des outils d'optimisation et d'analyse de performance (gprof, cachegrind, Vampir, Scalasca…)?

Réponses : 2 ont répondu « Faute d’avoir du temps ou de soutien pour cela ».

*Même remarque que la question précédente.*

# Lancement officiel du cluster IDCS Cholesky

Pour soumettre les demandes d’accès au cluster :

* Remplir / compléter une fiche descriptive de projet (fichier README.md)
* Envoyer la fiche par email : [idcs.meso.support@polytechnique.fr](mailto:idcs.meso.support@polytechnique.fr)

La documentation : <http://meso-ipp.gitlab.labos.polytechnique.fr/user_doc/>