

## Compte-rendu de la rencontre C4-CINES-GENCI Vendredi 6 Octobre 2017 au CINES à Montpellier

<p><u>Présents C4 :</u></p> <ul style="list-style-type: none"><li>- Sébastien Theetten (CT1)</li><li>- Julien Bodart (CT2A)</li><li>- Florent Duchaine (CT2B) : visio-conférence</li><li>- Alain Minuissi (CT4)</li><li>- Virginie Grandgirard (CT5)</li><li>- Rachel Nuter (CT5) : visio-conférence</li><li>- Michel Kern (CT6) : visio-conférence</li><li>- Nicolas Floquet (CT7)</li><li>- Sébastien Le Roux (CT8)</li><li>- Evelyne Lampin (CT9) : visio-conférence</li></ul>	<p><u>Présents CINES :</u></p> <ul style="list-style-type: none"><li>- Boris Dinstrans (directeur du CINES)</li><li>- Olivier Rouchon</li><li>- Gérard Gil (DCI admin)</li><li>- Romain Couturat (DCI admin)</li><li>- Jean-Christophe Penalva (DCI support)</li><li>- Mathieu Cloirec (DCI support)</li><li>- Philippe Prat (DCI eDARI)</li><li>- Marion Massol (DAD)</li><li>- Jérôme Chapelle (DS2I)</li></ul> <p><u>Présents GENCI :</u></p> <ul style="list-style-type: none"><li>- Arnaud Valois</li><li>- Jean-Philippe Proux</li></ul>
---	--

A la demande du nouveau directeur du CINES, Boris Dinstrans, la réunion s'est déroulée en trois parties : une première interne aux membres du C4, une deuxième entre les membres du C4 et le CINES et une dernière incluant GENCI. Il est aussi demandé d'accentuer la fréquence des réunions C4 pour en avoir 4 par an : deux au CINES et deux en visioconférence.

### 1. Réunion interne C4

Les questions à aborder avec le CINES et GENCI sont discutées. Elles concernent majoritairement les difficultés que rencontrent encore certains utilisateurs à pouvoir tourner correctement sur la machine OCCIGEN. La liste des codes impactés a été transmise par avance au CINES pour qu'ils puissent nous faire un point sur chacun d'entre eux. La priorité affichée de cette réunion C4 était clairement de réussir à faire avancer ces dossiers.

### 2. Réunion C4-CINES : Questions-réponses C4 à CINES

#### 2.1) Point sur les codes rencontrant encore des difficultés à tourner

*C4 : Nous avons de nombreux retours d'utilisateurs se plaignant de ne pas avoir pu calculer du tout jusqu'à Juin principalement dû à des programmes qui tournaient avant ou sur d'autres machines et qui n'arrivaient pas à fonctionner sur Occigen2. Au moins 4 codes, qui concernent une même communauté (LAMMPS, CPMD, CP2K et ADF), rencontrent encore des problèmes de stabilité. Les calculs s'arrêtent, aléatoirement d'un job à l'autre, et à un moment aléatoire de la simulation, et sans qu'il n'y ait d'allocation dynamique de la mémoire.*

CINES : Le CINES n'a visiblement pas mesuré le fait qu'il y avait un problème pour une communauté entière. Ceci doit être dû au fait que les incidents sont traités par tickets individuels, et que pour certains des codes en questions des utilisateurs semblent ne pas rencontrer de problème.

*C4 : Le C4 propose qu'une concertation soit mise en place auprès des utilisateurs concernés pour aider le CINES à identifier la présence de facteurs communs qui pourraient aider à identifier des pistes de recherche.*

CINES : Le CINES est tout à fait pour et propose que le C4 se charge au plus vite d'un appel dans les Comités Thématiques concernés (7, 8 et 9) afin de répertorier les utilisateurs qui rencontrent ou non des problèmes sur ces codes, et identifier ces problèmes.

*C4 : Certains utilisateurs ont l'impression d'avoir à réexpliquer systématiquement leurs problèmes en reprenant tout depuis le début. Comment se passe le suivi des tickets ?*

CINES : Le même interlocuteur CINES suit un ticket ouvert mais à partir du moment où celui-ci est fermé et un autre ouvert l'interlocuteur peut changer. Il est donc conseillé pour un problème déjà abordé et non résolu, de rouvrir le même ticket.

Concernant la gestion des tickets, ils sont fermés par le CINES une semaine après la relance finale à l'utilisateur lui demandant de valider la solution apportée.

*2.1.2) C4 : Installation d'OpenFoam ? a) Pourquoi a-t-elle pris autant de temps (opérationnel uniquement depuis le 28 Juin 2017) ? b) OpenFoam 2.4 a-t-il vocation à rester dans les logiciels proposés ?*

CINES : La version 4.1 est effectivement disponible depuis le 28 juin 2017. L'installation a été longue et complexe pour obtenir une version optimisée qui semble maintenant satisfaire la plupart des utilisateurs. Une version antérieure 2.4 a été installée à la demande d'un utilisateur mais elle n'est pas publiée dans `module`. Cette version est figée et aucun support ne pourra être apporté.

Il est rappelé que, quel que soit le logiciel, le maintien de versions antérieures a un coût non négligeable et que le CINES privilégie la limitation du nombre de versions d'un même code et l'aide à la migration vers les nouvelles versions.

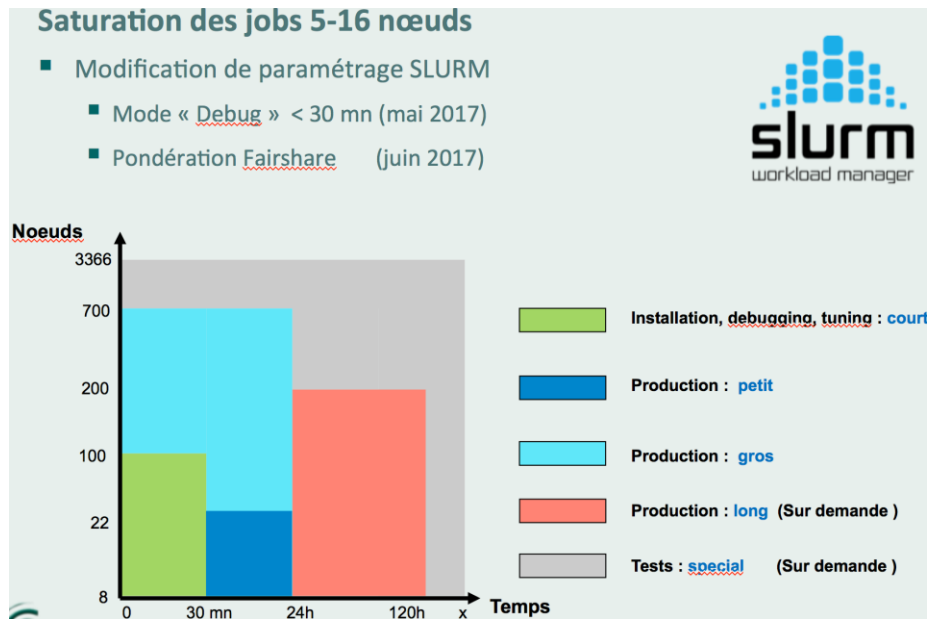
### **2.3) Politique des queues batchs**

*C4 : Un utilisateur (CT5) s'étonne de la différence brutale entre 2017 et les années antérieures concernant la saturation de la machine pour les jobs de 5 à 16 nœuds, alors que la machine a été étendue. Cette année ses jobs (15h à 20h sur 5 à 16 nœuds donc) sont restés généralement 3 à 5 jours en queue avant de démarrer, alors qu'ils partaient dans l'heure les années précédentes.*

CINES : Attention il est rappelé que les machines Haswell et Broadwell sont deux machines distinctes. [La commande `etat\_machine` permet de vérifier la charge courante de chacune d'elles.](#)

Les temps d'attente sont dûs au fait que la machine est surchargée. Le fait que les petits jobs (5-16 nœuds) passent plus difficilement que précédemment est dû à deux changements opérés récemment sur le paramétrage de SLURM : (i) le mode « debug » a été limité à des cas inférieurs à 30 minutes (mai 2017) pour éviter que des jobs en mode « pseudo-production » l'utilisent et (ii) le fairshare a été pondéré pour donner une priorité plus haute aux gros jobs. Les plus petits jobs ou les jobs longs ont une priorité plus basse.

Le CINES en profite pour rappeler qu'il est possible de passer des jobs dépassant la limite fixée à 700 nœuds en faisant une demande spécifique.



## 2.4) Diffusion de l'information. Documentation

*C4 : Plusieurs utilisateurs se plaignent du manque de documentation pour les nouveaux arrivants sur Occigen (documentation éparse, incomplète avec parfois des contradictions)*

CINES : Le CINES reconnaît que certaines rubriques du site web ne sont pas encore actualisées, ainsi que les documentations spécialement pour les nouveaux utilisateurs. Il est conscient du problème, en partie dû au fait qu'il n'y a actuellement plus personne en charge de la communication de leur côté. Ce point devrait être amélioré dans le futur puisqu'un poste dédié devrait être prochainement créé.

Le CINES rappelle toutefois qu'il ne faut jamais hésiter à contacter le support SVP pour toute question.

*C4 : En termes de communication, les utilisateurs ne constatent pas d'amélioration depuis la dernière réunion du CU. A savoir toujours aussi peu d'informations remontées aux utilisateurs lorsqu'il y a des problèmes ou instabilités sur la machine et communication minimale sur les opérations de maintenance.*

CINES : En ce qui concerne les informations délivrées par le CINES, il y a 2 canaux à l'heure actuelle : le site web et les mails envoyés à l'ensemble des utilisateurs. Il est proposé **d'ajouter une commande en ligne journal** qui permettrait aux utilisateurs de savoir s'il y a un problème au moment où ils sont sur la machine mais aussi d'avoir accès à des news que le CINES jugerait importantes de diffuser. La proposition d'un **forum utilisateurs**, déjà abordée lors de la précédente réunion du C4, est rediscutée. Le CINES y est tout à fait favorable, mais il propose une implication des membres du C4 pour gérer le forum par Comité Thématique.

Le CINES propose également d'utiliser l'application <https://reser.cines.fr/> pour faciliter la communication entre les utilisateurs. Parmi les idées, il est envisagé d'ajouter les coordonnées des représentants au C4 de la thématique du projet dans l'espace utilisateur.

## 2.5) Quota disques

*C4 : Un utilisateur signale qu'il serait bien d'avoir des notifications pour la suppression des fichiers sur le scratch sachant qu'un délai d'un mois est très court.*

CINES : Le CINES rappelle que le SCRATCH n'est pas un espace sauvegardé et qu'il se peut toujours que lors d'un reboot les fichiers disparaissent.

A l'heure actuelle, seuls les quotas physiques sont actifs. Mais **à partir de la prochaine opération de maintenance (initialement ciblée autour du 6 Novembre 2017, mais planifiée depuis la réunion au 13) des quotas « virtuels » seront activés ainsi que de nouvelles limites sur le nombre de fichiers.**

Nom	Quota (évolutif)	Cible	Accessible depuis	Type d'utilisation	Performance	Sauvegarde
/home	40 Go 20 000 fichiers	Utilisateur	nœuds de login + nœuds de calcul	compilation, stockage exécutable + bibliothèques nécessaires à l'exécution de vos jobs	moyenne	OUI
/scratch	4 To 200 000 fichiers	Groupe / projet	nœuds de login + nœuds de calcul	espace de travail : lancement des jobs à partir de cet espace, écriture des fichiers temporaires de sortie pendant l'exécution etc...	élevée	non
/store	1 To 10 000 fichiers	Groupe / projet	nœuds de login + nœuds de calcul	espace de stockage <b>sécurisé</b> et partagé par les calculateurs	moyenne	OUI

**Rappel sur les « Quotas Virtuels »** (Outil de gestion de l'occupation des espaces de stockage) :  
/home et /scratch (aujourd'hui) :

- Dépassement d'une limite
  - o Les travaux en cours peuvent se terminer (dans la limite d'un quota « physique » : garde fou)
  - o Les travaux en file d'attente sont basculés dans une partition Blocked\_xxx
  - o La soumission de nouveaux travaux refusée : message d'erreur clair en expliquant la raison.
  - o Transmission d'un message à l'utilisateur ou à tous les utilisateurs du groupe : information sur dépassement et blocage en cours avec le détail des utilisations.
- Retour sous une limite
  - o Rétablissement de l'autorisation de soumission et déblocage des travaux en file d'attente

#### **Evolution prévue : couplage RobinHood**

- Récupération de critères qualitatifs: âge des fichiers, ...
- Suppression des vieux fichiers (+30 jours) après avertissement et relances de l'utilisateur.

Rappel : la commande `etat_projet` permet de connaître les quotas et les niveaux d'utilisation pour son projet.

Le nettoyage des espaces est à la charge de l'utilisateur. Il est responsable de la gestion de ses espaces de stockage : suppression de fichiers, archivage et transfert vers /store (prévu pour conserver les fichiers volumineux et/ou qui ne sont plus nécessaires aux calculs). Sur l'espace /store, les utilisateurs peuvent/doivent diminuer le nombre de fichier en créant des archives via la commande `tar`. Le CINES (Jérôme Chapelle) conseille des fichiers archives de taille supérieure à 4 Go

mais en en-dessous de 1 To. Attention lors des archivages à bien contrôler que tout a été transféré correctement et qu'il n'y a pas eu de plantage en cours de procédure. [svp@cines.fr](mailto:svp@cines.fr) est là pour aider dans des cas particuliers, au cas par cas. Il existe un outil au TGCC qui crée automatiquement des archives dans un répertoire.

## **2.4) Questions diverses**

*C4 : Où en est la commande qui est censée donner les consommations mémoire d'un job (sacct). Est-elle exploitable ?*

CINES : La commande sur la consommation mémoire sacct est-elle opérationnelle ? SLURM ne collecte pas l'info correctement sur la consommation mémoire des travaux (les valeurs retournées par la variable MaxRSS sont fausses), le CINES travaille sur une autre solution (commande cgroup) mais qui ne marche pas pour l'instant. Un travail est en cours avec ATOS.

*C4 : Où en sont les problèmes liés aux i/o sur le scratch ? (escaladé chez RedHat ? NdE)*

CINES : Il n'y a plus de problèmes purement i/o identifiés actuellement mais un problème LUSTRE vient d'être détecté provoquant l'interruption de certains travaux. Ce problème est en cours d'investigation.

*C4 : Un utilisateur signale que depuis jeudi 28 septembre, il constate que certains de ses jobs apparaissent en running alors qu'en fait l'exécutable ne paraît pas être lancé ou du moins n'exécute rien.*

CINES : Pas de réponse claire apportée en séance. Cas à étudier avec un exemple concret via le support.

## **3. Réunion C4-CINES-GENCI :**

### **3.1) Questions-réponses C4 à CINES (suite)**

La séance de questions-réponses entre le CINES et le C4 n'étant pas terminée à l'heure du repas, elle s'est poursuivie en présence des membres de GENCI.

#### **3.1.1) Travaux bonus**

*C4 : Concernant les travaux bonus, certains utilisateurs se plaignent de ne pas avoir pu s'en servir correctement (fermée jusqu'à fin juillet, ouverte ensuite). Quelle est la politique associée ? Pourquoi ne pas laisser la queue ouverte en permanence ?*

CINES : Après quelques ratés au démarrage, le système est pleinement opérationnel depuis fin juillet. On peut utiliser au maximum 100 nœuds pendant 24h, à hauteur de 20% du nombre d'heures allouées au projet. Ce sont des jobs à priorité nulle, et qui ne passent que si la machine est libre. Les jobs 24h ne passent que le weekend. Il faut garder à l'esprit que la machine est très chargée et que du coup les jobs ne passent pas forcément.

Pour plus de détails sur le fonctionnement de cette queue, on peut se référer aux pages web dans la rubrique « trucs et astuces » (<https://www.cines.fr/tt8-soumission-de-jobs-en-classe-bonus/>).

- Bilan d'utilisation : août : ~6000 jobs / 5 Mh, septembre ~ 2100 jobs / 2 Mh.

#### **3.1.2) Visualisation distante au CINES ?**

*C4 : Dans la dernière réunion du C4, F. Daumas avait demandé un rapport interne sur la situation concernant la visualisation distante pour mi-mai 2017. Où en est-on ?*

CINES : Pour le problème de la visualisation, il n'y a pas eu d'avancée car un problème matériel vient se greffer au choix d'un logiciel compatible avec les règles de sécurité du CINES : la machine CRISTAL (dédiée en partie à la visualisation) arrive en fin de vie (fin 2017). Il a été décidé qu'il serait trop coûteux de mettre en place une solution temporaire pour pallier au manque de visualisation distante d'ici la fin de l'année.

*C4 : Le C4 déplore ce choix et se demande pourquoi la solution HPCdrive mis en place et opérationnelle au TGCC n'a pas été retenue.*

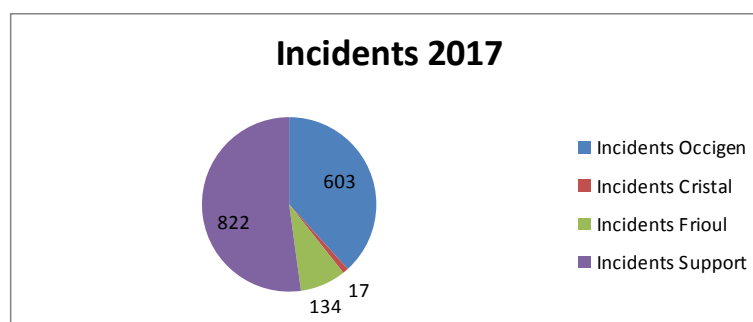
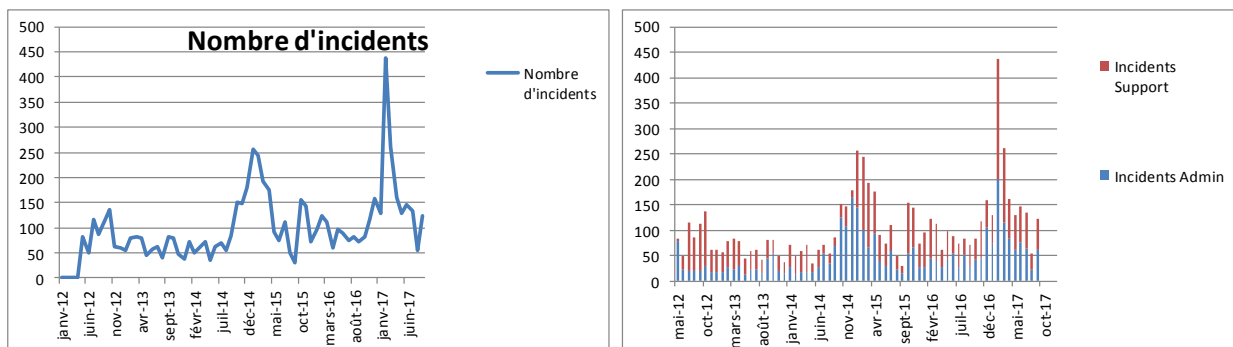
CINES : Les équipes CINES se sont rapprochées de celles du TGCC sur ce point mais HPCdrive ne sera pas maintenu au-delà de 2017. Le TGCC réfléchit à poursuivre cette maintenance ou à se tourner vers un autre outil.

CINES : L'achat d'une nouvelle machine ainsi que le choix d'une nouvelle solution « portail » sont à l'étude. Le CINES propose de faire une enquête rapide auprès des utilisateurs via le C4 pour qu'ils définissent leur besoin afin de les aider à dimensionner cette future machine. Il est clair que le budget alloué ne pourra être que faible celui étant déphasé de l'arrivée d'Occigen2. A l'avenir tout achat de machine GENCI devrait intégrer des nœuds de visualisation comme ce sera le cas pour la future machine TGCC.

*C4 : Le C4 émet le souhait qu'une concertation au niveau des 3 centres ait lieu pour offrir une solution homogène de visualisation distante sur les 3 centres.*

### 3.1.3) Statistiques de gestion des tickets utilisateurs

CINES : La hausse du nombre de tickets a été très nette autour de la période de déploiement de l'extension d'Occigen (+300% en Mars, Avril, Mai), entraînant des temps de réponse ou de traitement plus important. La tendance est à un retour à la normale depuis la fin de l'été.



### 3.2) Présentation actualités GENCI

Dans ce qui suit, nous résumons brièvement les informations transmises par GENCI au cours de la réunion. Pour plus de détails n'hésitez pas à consulter les transparents accessibles sur le site du CINES.

#### 3.2.1) Actualité sur les centres nationaux :

##### IDRIS :

Renouvellement Ada / Turing

- Calculateurs + stockage: publication appels d'offres fin 2017
- Maintien Ada / Turing en 2018
- Cible de mise en production des nouvelles ressources T2 2019 pour allocation A6 du DARI

Evolution stockage CMIP6

- AO publié / production fin 2017

##### CINES :

- La mise en place des travaux bonus sur Occigen a permis 10 % de gain en août.
- La **journée de restitution des grands défis est prévue le 7 décembre 2017.**

##### TGCC :

- **L'installation de la machine Joliot-Curie** (nom de code : Irene), successeur de Curie, se fera progressivement au 2<sup>e</sup> trimestre 2018 pour une mise en production en **juillet 2018**. Il s'agira d'une configuration équilibrée entre puissance de calcul, capacité mémoire et capacité I/O :
  - o une partition avec des processeurs Intel Skylake : 79 488 cœurs où 1 nœud est composé de 48 cœurs et de 192 Go de mémoire DDR4.
  - o une partition avec des processeurs Intel KNL : 45288 cœurs où 1 nœud est composé de 68 cœurs avec 96 Go de mémoire DDR4.
  - o 20 nœuds de visualisation distante
  - o 5 nœuds à grande mémoire pour le pré et post-traitement.
- Cette machine contiendra des modules refroidis directement par eau.

#### 3.2) Actualité PRACE :

- CALL 15 : Le comité d'attribution des heures s'est réuni le 14 Septembre 2017. Les allocations ont démarré le 1<sup>er</sup> octobre 2017.
- **CALL 16 : L'appel à projet est en cours** (du **26 septembre au 21 novembre 2017**). Le démarrage des allocations est prévu le 3 avril 2018. Concernant les heures allouées sur Curie, elles débuteront sur Curie et seront ensuite transférées sur Irene à partir de juillet 2018.

#### 3.3) Utilisation des moyens de calculs nationaux (projets DARI)

- L'appel à projet pour l'allocation A3 a été clos le 5 septembre 2017. Les allocations d'heures commenceront au 1<sup>er</sup> Novembre 2017 pour une durée d'un an.
- Le **prochain appel à projet pour l'allocation A4 et compléments A3** aura lieu du **8 janvier au 8 février 2018**, avec un comité d'évaluation et d'attribution début Avril pour un démarrage des allocations au 1<sup>er</sup> mai 2018.
- **L'homogénéisation du processus d'ouverture de compte** demandé de longue date par les utilisateurs sera effective à partir de l'allocation A3. Pour ce faire le CEA a développé un portail de

saisie en ligne utilisable par les 3 centres de calcul. Celui-ci permettra à l'utilisateur de réaliser une seule demande pour les 3 centres.

### **3.4) Forum des utilisateurs**

Il est rappelé que le forum des utilisateurs GENCI se tiendra le mercredi 11 octobre 2017 au siège du CNRS à Paris. Ce forum a eu lieu depuis. L'ensemble de la journée a été filmé et peut être consulté sur le web <https://webcast.in2p3.fr/live/forum-utilisateurs-genci>. Un retour sera prochainement fait aux utilisateurs.

### **Actions à prévoir d'ici la prochaine réunion :**

1. Le C4 se charge (via les représentants de CT concernés) de contacter les utilisateurs des CT7, CT8 et CT9 pour répertorier l'ensemble des personnes ayant des problèmes de stabilité sur la machine en leur demandant d'identifier au mieux les problèmes rencontrés.
2. Le CINES se charge d'analyser les résultats de cette enquête et d'organiser des séances de discussion en visioconférence ou en présentiel pour que les spécialistes de chaque code puissent confrontés leurs expériences.
3. Le CINES prépare un questionnaire qui lui permettrait d'identifier au mieux les besoins des utilisateurs en termes de visualisation distante et de post-traitement.
4. Le C4 se chargera de la diffusion du questionnaire et de la collecte des réponses.