

Compte-rendu de la réunion du C4

« Comité des Chercheurs Calculant au CINES »

23 mai 2014, CINES, Montpellier

Membres du C4 présents :

CT1 : Sébastien Theetten
CT4 : Benjamin Pajot
CT5 : Virginie Grandgirard (présidente)
CT7 : Jérôme Hénin
CT8 : Sébastien Le Roux
CT9 : Évelyne Lampin

Membres du C4 excusés :

CT2 (a&b) : Florent Duchaine
CT8 : Marie-Pierre Gageot
CT10 : Razvan Caracas & Kristelle Roidot

La journée s'est déroulée en 2 parties : la 1^{ère} moitié de la journée était dédiée à une réunion interne aux représentants du C4, la 2^{nde} était un échange avec les ingénieurs du CINES.

0. Remarques diverses

Il n'y avait pas de représentant du GENCI, le remplaçant de Catherine Le Louarn ayant pris ses fonctions depuis peu. La présence du GENCI sera en principe assurée aux prochaines réunions avec le C4.

Remarque générale : Jade est peu chargée ces temps-ci (80 %). La file d'attente étant courte, c'est le moment de faire tourner ses calculs.

Un arrêt électrique prolongé du CINES est programmé entre le mardi 15 juillet 20 heures et le mardi 22 juillet dans la soirée. Il s'agit d'une maintenance annuelle des installations électriques et de l'amélioration des infrastructures de climatisation et d'électricité en particulier en vue de préparer l'installation de la nouvelle machine. En conséquence, les calculateurs Jade, Ambre, Cristal et Yoda seront indisponibles, ainsi que l'accès aux données

1. Point sur le changement de la politique d'accès ZRR

L'accès physique et la connexion par le réseau sont soumis à la ZRR, mais la plupart des utilisateurs seront concernés par le second point, l'accès à distance. Ce changement de politique d'accès a déjà commencé en 2014, puisque les non-ressortissants de l'UE ont dû demander une autorisation au haut fonctionnaire de défense (HFD). Le délai de réponse peut atteindre 2 mois. La procédure sera **généralisée à tous les utilisateurs** à partir de 2015. Afin de prendre en compte le délai de réponse du HFD, la demande de login devra être faite en même temps que la demande DARI, en

octobre/novembre. Cette demande sera valable pour 5 ans et sur tous les sites gérés par GENCI (CINES, IDRIS, TGCC)

Il existe une application du CNRS pour demander une autorisation personnelle pour l'accès au ZRR (procédure lourde).

D'autres modalités d'accès sont évoquées : certificats personnels, filtrage IP comme à l'IDRIS.

Les centres sont en communication entre eux à ce sujet. Un travail d'harmonisation est envisagé.

Cette nouvelle politique d'accès devra trouver un compromis entre recherche efficace et sécurité fiable...

Le CINES rappelle qu'il ne faut **en aucun cas communiquer son login/mot de passe**. Il y a possibilité de créer plusieurs logins par projet en cas de multi-utilisateurs.

2. Arrêt de Jade et période transitoire

Jade a été installée en 2 phases. La première phase (processeurs Harpertown) a été réalisée en juillet 2008. La fin de la période de maintenance pour ces machines est prévue en juillet 2014. La deuxième phase (processeurs Nehalem) a été réalisée en novembre 2009. La fin de la maintenance de ces machines est prévue en novembre 2014. Le principe du meilleur effort sur les parties non-maintenues sera appliqué (*e.g.* si panne, retrait des nœuds incriminés, donc moins d'heures disponibles théoriquement).

Il n'y aura donc pas dans un premier temps d'arrêt brutal des machines, mais à cause de cet arrêt de la maintenance, la disponibilité des nœuds Harpertown n'est plus garantie après le 1^{er} août, de même que la disponibilité des nœuds Nehalem ne l'est plus après le 10 novembre. Cette réduction de disponibilité potentielle a été prise en compte dans le nombre d'heures proposées aux utilisateurs. Toutefois, il est prudent de ne pas reporter ses calculs en toute fin d'année, le risque augmentant que des nœuds ne soient plus disponibles.

L'arrêt total de jade est prévu le 31 décembre 2014, elle n'apparaîtra donc plus dans la demande DARI 2015.

Un appel d'offre a été lancé pour une nouvelle machine, l'annonce sera faite fin juin/début juillet. L'achat est financé par GENCI (dont les partenaires sont : CEA, CNRS, CPU Enseignement Supérieur et Recherche, INRIA). La nouvelle machine sera installée pendant l'été, et la vérification de son fonctionnement sera faite avant fin 2014 pour une mise en production en janvier 2015. Cette phase de mise en production sera précédée d'une période de « Grands Défis » en novembre/décembre. Il s'agit de charger la machine afin de tester son fonctionnement. Dans ce but, les utilisateurs peuvent proposer un Grand Défi demandant des moyens de calculs très importants (plusieurs milliers de cœurs) en argumentant de son intérêt scientifique et de ses retombées « sociétales » potentielles. C'est donc également pour les utilisateurs choisis une opportunité rare de pouvoir faire tourner des jobs lourds sur une partie importante de la machine. L'appel à candidature aura lieu fin juin/début juillet, les jobs tourneront fin 2014, et les résultats de ces Grands Défis devront faire l'objet d'un rapport puis d'une présentation lors d'une journée organisée début 2015.

Les membres du C4 demandent quels sont les benchmarks qui ont été utilisés pour l'appel d'offre de la nouvelle machine, et quels sont leurs résultats ? Cette information pourrait aider les utilisateurs à choisir la machine cible.

Problème : l'information est difficile à diffuser pour cause de NDA (accord de non-divulgaration). Le CINES va demander au GENCI s'il peut divulguer certaines de ces informations.

Rappel des quelques dates clés :

Fin juin / début juillet 2014 :

- Annonce de l'offre retenue pour le remplacement de JADE.
- Ouverture de l'appel à candidature "Grands Défis" pour un accès au calculateur en novembre-décembre 2014.

15 juillet 2014 au soir au 22 juillet 2014 inclus :

- Arrêt électrique prolongé (aspect électrique de la nouvelle salle pour la nouvelle machine et gros travaux sur la redondance électrique) .

1er août 2014 :

Fin de maintenance de Jade 1 (nœuds Harpertown)

Été / automne 2014 :

Installation de la nouvelle machine.

10 novembre 2014 :

Fin de maintenance de Jade 2 (nœuds Nehalem)

Novembre-décembre 2014 :

Utilisation de la nouvelle machine pour les "grands défis".

Fin 2014

Période de vérification de service régulier pour la nouvelle machine. Elle n'est pas disponible pour l'ensemble des utilisateurs, uniquement pour les grands défis.

31 décembre 2014 :

Arrêt total de Jade.

Janvier 2015 :

Exploitation normale de la nouvelle machine : ouverture à l'ensemble de la communauté scientifique.

3. Gestion des espaces disques

Rappel des bonnes pratiques de l'utilisation des disques :

- ⤴ l'espace disque /scratch est prévu pour des données temporaires, c'est là que se trouvent les répertoires depuis lesquels on lance les jobs
- ⤴ l'espace disque /store est un espace sauvegardé, y compris 6 mois après la fin de login
- ⤴ l'espace disque /work/home est prévu pour la compilation et le stockage des exécutables

- Espace /scratch :

Trop d'utilisateurs stockent leurs données sur le /scratch. Cela risque de dégrader les performances de la machine, et l'utilisateur s'expose à un risque de perte de ses données puisque le scratch n'est pas sauvegardé. Une nouvelle politique de gestion du /scratch va être mise en place pour contraindre les utilisateurs à ne pas y stocker durablement leurs données. A j+30, sans déplacement

de données, un mail d'alerte sera envoyé à l'utilisateur. Sans réaction aux diverses mises en garde, les données seront supprimées jusqu'à j+60. Une évolution envisagée est de passer le quota sur le scratch de 4 à 20 To mais sur une durée beaucoup plus courte (durée de 48h avec fenêtre glissante, notion d'espace temporaire : pas de problème de quota pour la durée des calculs), pour permettre de répondre à des besoins de stockage plus importants, comme des fichiers de restart intermédiaires qu'on efface en fin de calcul.

La bonne pratique est d'écrire les fichiers que l'on veut garder sur l'espace /store après exécution du job, en utilisant notamment des tar pour créer moins de fichiers.

La bande passante vers le scratch est de 20 Go/s, elle passera à 100 Go/s dans la nouvelle machine.

- Espace /store :

La capacité de stockage actuelle de /store (espace sauvegardé) est actuellement de 350 To, elle passera à 2Po. Les comptes utilisateurs font 200Go par défaut. La bande passante de la communication store-work/home est actuellement de 5Go/s, elle passera à 50 Go/s dans la nouvelle machine. Avec l'amélioration du /store , le traitement des fichiers devrait y être plus rapide (pour les utilisateurs devant créer des archives de très nombreux fichiers).

Un projet est en cours pour proposer un document de « bonnes pratiques » dont le but sera de donner une marche à suivre aux utilisateurs.

Le C4 évoque le problème du stockage à moyen terme (5 à 10 ans) qui se trouve à la limite entre le stockage sur /store qui dure 1 an et demi et l'archivage sur plusieurs décennies.

4. Interconnexion avec les autres centres nationaux

Actuellement, chaque site a sa politique de sécurité propre empêchant une machine corrompue d'accéder aux autres. Les connexions par RENATER, sont ouvertes à tous les utilisateurs, avec une caractéristique de 10 Gbe -Gigabit Ethernet- pour le CINES. Il suffit d'utiliser les outils de connexion ou transfert comme ssh, scp, sftp, rsync. A noter que pour PRACE, il y a un réseau de 10 Gbe dédié, réservé aux seuls utilisateurs PRACE.

Pour l'avenir, il y a un projet de monter un réseau "CCFR" (Centre de Calcul FRançais) d'interconnexion des 3 centres nationaux (TGCC, IDRIS, CINES) par un réseau dédié 10 Gbe pour mettre à la disposition des utilisateurs des services simples et utiles (accès interactif et transfert de données *via* ssh, gsissh, scp, sftp...). Ce projet d'interconnexion est à l'état de réflexion pour le moment.

5. Retour sur la campagne DARI 2014 et nouveautés sur la campagne DARI 2015

En 2014, 288 dossiers ont été déposés pour 90 millions d'heures de calcul. Au vu des chiffres présentés par le CINES, il y a le constat d'une inhomogénéité entre la demande d'heure et l'attribution de ces heures, ainsi qu'une progression importante du CT1. Il est rappelé qu'une régulation mensuelle est mise en place pour les projets de plus d'1 millions d'heures, et au bout de 6 mois pour les projets plus petits.

A partir de DARI-2 2014, les porteurs de projets ANR devront déposer un dossier DARI-ANR pour les 3 années à venir. L'objectif est de faciliter la demande d'heure lorsque le projet est identifié DARI-ANR. L'expertise DARI est valable pour la durée du contrat ANR. L'objectif est de promouvoir une expertise collaborative entre les experts ANR et DARI.

Les membres du C4 font remarquer les contraintes que cela impose, notamment vu le faible taux de succès auprès de l'ANR.

Par ailleurs un utilisateur a demandé s'il serait possible de faire des demandes d'heures pour une durée de 2 ans : la réponse est non, car a priori un projet accepté une année devrait être renouvelé la suivante.

6. Évolution du site web

Le nouveau site web du CINES est désormais accessible !

Ce dernier contient 4 axes principaux :

- calcul,
- archivage pérenne,
- MOOC-FUN (France Université Numérique),
- visualisation.

A noter que les CT 3, 5, 6, 7 et 8 n'ont pas encore fourni de vidéo pour le site du CINES. Il n'est pas trop tard !

On peut faire remonter des remarques éventuelles, une mise à jour étant prévue tous les 6 mois.

7. Projet de FAQ calcul

Cette FAQ sera ajoutée au nouveau site web, on peut proposer des questions. Un « best practice guide » est en cours d'écriture, ou comment utiliser jade en 15 minutes !

8. Projet OTRS (suivi des tickets d'incidents sur le Web)

Il y a un projet de mise en place d'un outil de suivi avec interface web en plus de svp (@cines.fr).

La mise en service est prévue pour le 2^{ème} semestre 2014.

9. Formation calcul du CINES

La formation a été réactivée en 2012. Elle est composée :

▲ d'un cycle standard en 3 niveaux :

- introduction aux architectures parallèles et au calculateur jade
2/an, durée 1 jour, prochaine en janvier 2015
- calcul parallèle
1/an, durée 4 jours, prochaine au printemps 2015
- debugging & optimisation (agrément PATC, coordonnée par la maison de la simulation, formation en anglais destinée à un public européen-)

1/an, durée 4 jours, prochaine en octobre 2014

Ces formations ont un coût forfaitaire de 100 euros.

✧ des formations spécifiques :

- les accélérateurs de calcul (en discussion) pour l'automne 2014
- coupleur OpenPALM (PATC, celle de 2014 annulée par manque de participants)

Ces formations sont gratuites.

Consulter le site web pour plus d'informations.

10. Journée jeunes chercheurs

La précédente avait eu lieu en 2005. Cette journée, destinée aux chercheurs de moins de 35 ans, sera organisée sous forme de colloque, avec une introduction généraliste puis des présentations des jeunes chercheurs.

Le C4 soulève les questions de l'intérêt pour les jeunes chercheurs d'y participer, de l'invitation de conférenciers senior et du couplage de cette journée à celle des résultats des grands défis, pour davantage de brassage, motiver plus de participants, et pour des problèmes de plus en plus lourds de financement des missions. Cette dernière proposition est notée par le CINES, qui va l'étudier.

Initialement planifiée pour le 28 novembre, elle va peut-être être reportée pour être couplée à une demi-journée retours sur les grands défis, de façon à attirer plus de public, motiver les « jeunes chercheurs » en leur montrant des résultats de grande ampleur, donner plus de visibilité à leurs présentations, et justifier le financement de leur déplacement.

A noter que probablement un prix, du type attribution d'heures de calcul, sera donné à la meilleure présentation.