

# Réunion des Expériences du 22 janvier 2015

<https://indico.in2p3.fr/event/10875>

mot de passe : re2015

## I - Compte-rendu

- |   |                              |
|---|------------------------------|
| a. Le CC-IN2P3 : Status et perspectives               | MACCHI Pierre-Etienne        |
| b. Bilan de l'utilisation du Centre de Calcul en 2014 | BOUVET David                 |
| c. Cycle de vie des données                           | NIEF Jean-Yves               |
| d. Service cloud au CC-IN2P3                          | BOUVET David et PUEL Mattieu |
| e. Expérience AMS au CC-IN2P3                         | POIREAU Vincent              |
| f. Expérience AGATA au CC-IN2P3                       | STEZOWSKI Olivier            |

## I - Compte-rendu

### a. Le Centre de Calcul : Statut et perspectives

MACCHI Pierre-Etienne

Pierre-Etienne MACCHI présente le budget et la répartition des dépenses au cours de l'année 2014. Une grosse partie du budget informatique a été consacrée au maintien des capacités. Notons également qu'une plate-forme QServ pour l'expérience LSST a été mise à disposition par Dell. En ce qui concerne l'effectif global, il est en baisse, par contre celui du personnel permanent est en progression.

Pour 2015, le budget restera constant. Toutes les demandes de ressources devraient être couvertes cette année. Nous envisageons une augmentation de 3 Po de disques, 389 To pour GPFS et 10,6 Po de bandes. Nous avons reçu des requêtes de GPU de la part de plusieurs groupes. La satisfaction de ces requêtes GPU pourrait rentrer dans le cadre du renouvellement envisagé de la ferme de calcul MPI.

Des investissements lourds sont à prévoir en 2017-2018 en particulier pour les expériences du LHC et d'Astroparticules : LSST, EUCLID, CTA, etc. En amont, il faut participer de façon plus précoce aux futures expériences d'Astroparticules, développer des synergies plus concrètes entre la recherche en informatique, les utilisateurs et les centres de production, et approfondir nos relations avec les autres centres européens (initiative EU-T0). Le COCIN (Comité d'Orientation du Calcul Intensif du CNRS) se propose de mieux coordonner les ressources de calcul dans un contexte de forte croissance de la demande, du foisonnement de petits centres et des demandes non coordonnées.

#### **Discussion**

*Il y a 20 ans, avec le même budget informatique nous avons moins de puissance de calcul et de stockage mais depuis, les capacités informatiques ont beaucoup évoluées.*

### b. Bilan de l'utilisation du Centre de Calcul en 2014

BOUVET David

David BOUVET présente l'utilisation des ressources informatiques en 2014 par domaines scientifiques. La capacité de calcul disponible sur l'année était de 95% contre 93% en 2013. Le pourcentage consommé a été de 75% contre 80% en 2013. Il y a eu moins de jobs des groupes LHC à cause de l'arrêt du LHC. Malgré l'augmentation des limites des autres groupes, cette baisse du LHC n'a pas pu être complètement compensée.

Une nouvelle version du système de Batch UGE sera mise en place prochainement. Rappelons quelques bonnes pratiques concernant l'utilisation du système de Batch : avoir un nombre de jobs en queue inférieur à 1000, veiller à bien spécifier les ressources, s'assurer que les espaces utilisés ne sont pas

pleins, éviter les jobs courts, limiter la fréquence d'utilisation des commandes « qstat/qacct » et si possible informer le centre de calcul lors d'une planification de grosses productions.

En ce qui concerne la gestion des comptes, les utilisateurs seront désormais notifiés avant l'expiration de leur compte ce qui permettra de préparer la gestion de leurs données en cas de départ ou de prolonger le compte si besoin. Pour les données orphelines, il serait possible de fournir aux czars la liste des comptes désactivés ainsi que la liste des données associées.

Pour l'instant, aucune décision n'a encore été prise sur le remplacement de AFS. Il n'y a pas d'urgence tant que IPv6 n'est pas déployé. L'espace GROUP\_DIR pour les nouveaux groupes sera néanmoins remplacé par un espace SPS.

Une réflexion est cours sur l'amélioration de la procédure d'escalade des tickets ainsi que sur la mise en place d'un catalogue de services.

### **Discussion**

- Est-ce qu'une authentification par clé publique/privée sera possible avec la nouvelle version de Kerberos ?  
Non, pour des raisons de sécurité : cette fonctionnalité ne sera pas disponible car nous ne pouvons contrôler la façon dont les utilisateurs gèrent leurs clés.

- Pourra-t-on utiliser les « cgroups » pour la « resident memory » ou la « mémoire virtuelle » ?  
Normalement oui, il faudra tester.

- Sera-t-il possible de choisir l'une ou l'autre des deux mémoires ?  
En principe oui, c'est déjà le cas.

- La prochaine migration d'OS sera vers Centos ou SL ?

Cela dépendra des prochaines tendances mais également des requêtes des utilisateurs.

## **c. Cycle de vie des données**

NIEF Jean-Yves

Jean-Yves NIEF expose la problématique du cycle de vie des données au CC-IN2P3. Beaucoup de données sont stockées par des utilisateurs et des groupes aujourd'hui inactifs (environ 40 groupes). Le CC-IN2P3 se pose la question de la bonne gestion de ces données. Il en va de même pour les groupes ayant accumulé des données devenues obsolètes ou d'utilisateurs toujours actifs mais ayant changé de groupe.

Voici quelques éléments de solution :

- Améliorer le cycle de vie des données en identifiant rapidement les données en fin de vie, en faisant un audit quotidien des données, en facilitant et en accélérant la prise de décision sur le devenir de ces données.

- Établir un plan de gestion des données en définissant une politique claire et pré-établie en accord entre le CC-IN2P3 et les expériences ou projets.

- Optimiser l'usage des espaces de stockage ainsi que la migration des données vers des médias plus adaptés.

- Réfléchir sur l'archivage à moyen et long terme.

### **Discussion**

- Que deviennent les données personnelles des utilisateurs qui partent ?

Légalement, les données appartiennent à l'employeur sauf si elles sont dans un répertoire explicitement nommé « privé ». Dans certains laboratoires par exemple, les données sont transmises au responsable du groupe de physique.

- Peut-on appliquer ce procédé également au CC-IN2P3 systématiquement ?

Pour les grosses expériences, cela pourrait accroître les responsabilités des czars. Il serait également intéressant de sensibiliser les nouveaux utilisateurs afin d'anticiper la gestion de leurs données avant leur départ.

- Au bout de combien de temps peut-on considérer que les données sont caduques ?

Cette durée dépendra du type de données et du modèle de calcul de l'expérience. Par exemple pour les expériences dont le CC-IN2P3 est T0, les données pourraient être amenées à être conservées sur une plus longue durée au CC-IN2P3. Une solution envisageable serait de fixer une durée de rétention des données à travers un DMP (Data Management Plan). Cela serait trop fastidieux de l'établir pour chaque groupe spécifi-

quement mais quelques DMP génériques pourraient être proposés aux expériences suivant les besoins typiques les plus courants.

- Serait-il plus pertinent de définir des zones de stockage avec des politiques de rétention par zone ?

Il est en effet difficile de savoir dans une arborescence ce qui doit être conservé ou pas lorsque les différents types de données sont hébergés indifféremment dans une même zone de stockage.

- Comment bien archiver les données et permettre de les relire ?

Au CNES, il existe un service spécial qui s'occupe de l'archivage des données. Un gros travail est fait en amont avec les équipes sur le niveau d'archivage que requiert chaque projet.

#### d. Service cloud au CC-IN2P3

BOUVET David et PUEL Mattieu

Mattieu Puel et David Bouvet présentent le service Cloud au CC-IN2P3. Actuellement, le service Cloud du CC-IN2P3 est limité à de l'hébergement de services. Les serveurs sont fournis à la demande. Le choix des serveurs peut se faire selon plusieurs critères (CPU, RAM, disque, ...) . De nombreux OS sont également disponibles : Ubuntu, SL, CentOS, Fedora, etc.

Le Cloud au CC-IN2P3 est basé sur la technologie Openstack. La gestion des instances peut se faire via une interface web, par lignes de commandes, ou par APIs compatibles avec NOVA et EC2.

Des quotas sur les ressources Cloud sont attribués à chaque groupe (laboratoire, expériences, projets). Certaines expériences (BioAster, eTricks, AML, etc) utilisent déjà le service Cloud du CC-IN2P3 pour y déployer leurs services.

Grâce au cloud, l'hébergement de services est plus simple à fournir et devient facilement ajustable au besoin. Les machines peuvent être administrées par les utilisateurs eux-mêmes. L'entretien des machines hôtes est assuré par l'infrastructure du CC-IN2P3.

A l'avenir, le service Cloud du CC-IN2P3 pourrait s'étendre au stockage objet (Swift).

#### Discussion

- Le choix de l'adressage Ipv6 est-il disponible ?

C'est possible mais pas dans l'immédiat. Cela permettra également de lever les limites sur le nombre de subnets et VLAN.

- Quels sont les ports qui sont ouverts par défaut ?

Seul le port 22 est ouvert. D'autres ports peuvent être ouverts sur demande.

- Est-il envisageable d'utiliser ce service pour héberger des serveurs de licences ?

La génération des adresses MAC à la volée peut être un obstacle à l'hébergement des serveurs de licences.

- Comment sont évalués les quotas ?

Les quotas sont évalués de façon empirique, selon le parc et la taille estimée du service. Le quota RAM reste le plus critique à évaluer.

#### e. Expérience AMS au CC-IN2P3

POIREAU Vincent

Vincent POIREAU présente l'expérience AMS au CC-IN2P3. AMS est un détecteur de particules dans l'espace. Le détecteur a été lancé depuis la base de Cap Canaveral en 2011. Le détecteur récolte chaque jour environ 40 millions d'événements et transfère environ 100 Go. Ainsi par an, 39 To de données sont transférées et 200 To sont reconstruites.

L'expérience AMS est une collaboration de 60 instituts dans 16 pays, dont la France où elle compte deux instituts : le LAPP et le LPSC de Grenoble (26 utilisateurs français dont environ 13 actifs sur une année). Elle a pour objectif de mesurer le flux des rayons cosmiques, de rechercher l'existence de matière noire et d'antimatière.

Le Framework AMS est essentiellement basé sur des fichiers « Root » et chaque fichier fait entre 4 et 10 Go. Au CC-IN2P3, les données de AMS sont stockées sur HPSS, 1228 To en 2014 et 600 To en plus ont été

demandés pour 2015. L'accès à ces données se fait via XROOTD avec environ 500 To sur disque. Les résultats issus des analyses sont stockés sur SPS, environ 15 To et 10 To en plus pour 2015.

L'expérience AMS utilise les ressources du CC-IN2P3 pour faire de l'analyse de données et de la production Monte Carlo. Elle a consommé en terme de temps CPU 39,5 millions d'heures HS06 en 2014 pour la production Monte Carlo (développement et tests) et 112 millions d'heures HS06 requises pour 2015 à cause de l'augmentation de la production Monte Carlo.

Dans le but d'optimiser la lecture/écriture des fichiers de données AMS, un nouveau format de fichier nommé ACQt est actuellement en cours d'expérimentation. Ce format viendra remplacer le format Root et a pour avantages d'être moins lourd, de permettre l'analyse de plusieurs chunks en parallèle et d'être fusionnable en gros fichiers Multi-ACQt. Cela induit un gain considérable en temps d'exécution des jobs de l'expérience AMS (quelques heures au lieu de quelques jours).

L'utilisation du format ACQt au CC-IN2P3 requiert quelques conditions : l'installation du logiciel ACsoft, la mise à disposition d'environ 50 To d'espace disque et d'une ferme pour les calculs parallèles.

L'expérience AMS prévoit en 2015 d'améliorer son code de reconstruction et de simulation et de faire tourner de façon constante 10 à 20 % de toute sa production Monte Carlo au CC-IN2P3.

## **f. Expérience AGATA au CC-IN2P3**

STEZOWSKI Olivier

Olivier STEZOWSKI présente l'expérience AGATA au CC-IN2P3. Le détecteur mis en oeuvre est le fruit d'une collaboration Européenne qui regroupe une douzaine de pays et plus de 40 instituts impliqués dans la recherche en Physique Nucléaire. L'objectif final est de reconstruire la trace de rayonnements gamma dans un multi-détecteur en Germanium (couverture proche de 4PI) afin d'étudier la structure des noyaux atomiques avec une sonde de sensibilité inégalée jusqu'ici. Ce détecteur, nomade, fonctionne par campagne dans différents laboratoires.

L'expérience peut générer jusque 10 To de données brutes par semaine dans les cas les plus extrêmes: le taux de production étant en effet fortement conditionné par la physique étudiée et par les détecteurs additionnels. Les données sont transférées par périodes sur le système de stockage de grille du CC-IN2P3.

Pour l'instant, il n'y a pas d'utilisation de ressources de calcul au CC-IN2P3.

La collaboration AGATA utilise aussi un large panel de services mise en oeuvre au CC-IN2P3 et ceci depuis 2005 : hébergement de sites web, de serveurs svn et de bases de données ORACLE. En 2014, la composante AGATA France a été pilote pour le projet ATRIUM dont l'utilisation se démocratise peu à peu dans la collaboration. Il est aussi envisagé d'utiliser en 2015 la plateforme d'intégration continue du CC-IN2P3 pour les différents software développés dans la collaboration.