



À LA  
UNE

## L'infrastructure TIDRA tente de séduire les industriels

Interview  
CCRI



Depuis plusieurs années, les laboratoires de l'IN2P3 et de l'INSB du CNRS ont développé une expertise dans le domaine de l'informatique distribuée à grande

échelle. Cette expertise est aujourd'hui mise à disposition des chercheurs et des industriels de la région Rhône-Alpes, à travers le projet TIDRA (Traitement de données et Informatique Distribuée en Rhône-Alpes).

Le projet TIDRA met potentiellement à disposition des chercheurs et des industriels rhônalpins une infrastructure informatique de plus de 10 000 cœurs (CPUs). Cette infrastructure est actuellement constituée des nœuds suivants

: le LAPP (Annecy), le LPSC (Grenoble), l'IPNL (Lyon), l'IBCP (Lyon) et le CC-IN2P3 (Lyon). A terme, tous les organismes de recherche de la région Rhône-Alpes qui ont une infrastructure informatique à partager sont susceptibles de participer à ce projet.

L'objectif du projet TIDRA est de rassembler les moyens de calcul des laboratoires de la région ayant des ressources à mettre à disposition d'acteurs de la région afin d'accélérer la recherche et l'innovation en Rhône-Alpes.

TIDRA permet déjà à plusieurs laboratoires de traiter leurs données et s'ouvre aujourd'hui au monde industriel.

L'une des applications pilote concerne les sciences du vivant et le laboratoire LBBE (Laboratoire de Biométrie et Biologie (...))



Comité de Coordination du Réseau des Informaticiens  
[lire l'interview](#)

[lire la suite](#)

**Recherche matériel informatique ancien**  
Le Centre de Calcul de l'IN2P3 recherche du matériel informatique ancien afin d'enrichir son musée (...)  
[en savoir plus](#)

### ■ Conférence



#### Super Computing 2010

L'édition 2010 de la conférence Super Computing (SC10) s'est tenue à la mi-novembre à la Nouvelle Orléans. Ce rassemblement de plus de 10 000 personnes, et très fortement orienté sur le calcul haute performance, est l'occasion de réunir des constructeurs et des utilisateurs d'horizons très divers.

[lire la suite](#)

### ■ Standardisation



#### Le standard SAGA à l'OGF

Les utilisateurs de grilles ayant besoin d'accéder à une grande variété d'infrastructures utilisant des intergiciels hétérogènes, il est nécessaire de leur fournir des interfaces uniformes et standardisées. L'OGF (Open Grid Forum) est un consortium dont la mission est de créer et de promouvoir des standards ouverts pour les grilles et autres systèmes distribués (cloud computing...). Les groupes de travail de cette organisation, dédiés à des domaines particuliers, se rencontrent régulièrement lors de colloques.

[lire la suite](#)

### ■ Conférence



#### CHEP 2010

CHEP (Computing in High Energy Physics) est une série de grandes conférences se déroulant tous les 18 mois sur un continent différent. Cette conférence, regroupant des physiciens et des informaticiens, permet de cibler les besoins de la recherche et de dessiner les évolutions futures. Cette année, la conférence qui a eu lieu à Taïpei du 18 au 22 octobre, était organisée par le Centre de Calcul de l'Academia Sinica.

[lire la suite](#)

### ■ Aspera Workshop



#### Computing and Astroparticule Physics

Les 7 et 8 octobre, un atelier organisé par ASPERA a rassemblé au Centre de Calcul de l'IN2P3 (CC-IN2P3) une cinquantaine de participants autour du calcul pour l'Astroparticule.

[lire la suite](#)

### ■ Agenda

**Arrêt du Centre de Calcul de l'IN2P3**  
Un arrêt des systèmes informatiques du CC-IN2P3 est programmé pour le 14 Décembre 2010. Le blocage (...)  
[en savoir plus](#)

### ■ Archives

#### Abonnement

Pour vous abonner/désabonner, suivez ce [lien](#).

#### Proposer un article

Vous souhaitez proposer un article ? Envoyez un mail à [LettreInformatique@in2p3.fr](mailto:LettreInformatique@in2p3.fr).



© 2010 CCIN2P3



Equipe

**Responsables éditoriaux :** Dominique Boutigny et Cristinel Diaconu

**Comité de rédaction :** Virginie Dutruel, Sébastien Grégoire, Eric Legay, Gaëlle Shifrin et Tiffany Thome



**n°14**  
 Décembre  
 2010

# La lettre IN2P3 Informatique

Réseau des Informaticiens de l'IN2P3 et de l'IRFU

**A LA  
 UNE**

## L'infrastructure TIDRA tente de séduire les industriels

Mardi 14 décembre 2010 à Villeurbanne



Depuis plusieurs années, les laboratoires de l'IN2P3 et de l'INSB du CNRS ont développé une expertise dans le domaine de l'informatique distribuée à grande échelle. Cette expertise est aujourd'hui mise à disposition des chercheurs et des industriels de la région Rhône-Alpes, à travers le projet TIDRA (Traitement de données et Informatique Distribuée en Rhône-Alpes).

Le projet TIDRA met potentiellement à disposition des chercheurs et des industriels rhônalpins une infrastructure informatique de plus de 10 000 cœurs (CPUs). Cette infrastructure est actuellement constituée des nœuds suivants : le LAPP (Annecy), le LPSC (Grenoble), l'IPNL (Lyon), l'IBCP (Lyon) et le CC-IN2P3 (Lyon). A terme, tous les organismes de recherche de la région Rhône-Alpes qui ont une infrastructure informatique à partager sont susceptibles de participer à ce projet.

L'objectif du projet TIDRA est de rassembler les moyens de calcul des laboratoires de la région ayant des ressources à mettre à disposition d'acteurs de la région afin d'accélérer la recherche et l'innovation en Rhône-Alpes. TIDRA permet déjà à plusieurs laboratoires de traiter leurs données et s'ouvre aujourd'hui au monde industriel.

L'une des applications pilote concerne les sciences du vivant et le laboratoire LBBE (Laboratoire de Biométrie et Biologie

Evolution de l'Université Claude Bernard Lyon 1). Les recherches du LBBE conduisent les scientifiques à maintenir de nombreuses bases de données dédiées à la phylogénie et à l'évolution moléculaire. La construction et la mise à jour de ces bases reposent sur la comparaison entre elles de centaines de milliers de séquences biologiques (voire des millions pour certaines bases) ainsi que sur le calcul de centaines de milliers d'alignements de séquences et d'arbres phylogénétiques. Le LBBE utilise l'infrastructure TIDRA pour ses recherches, et accélère ainsi considérablement le temps de traitement des données.

Le projet TIDRA s'inscrit également comme maillon régional de la grille de calcul française France-Grilles, celle-ci étant elle-même un élément de la grille Européenne EGI (European Grid Initiative).

Un Colloque de Lancement de TIDRA est organisée le mardi 14 décembre 2010 à l'Espace Tête d'Or à Villeurbanne pour présenter l'infrastructure TIDRA et les services associés, et répondre aux questions des utilisateurs actuels et potentiels.

Pour plus d'informations et s'inscrire : [www.lancement.tidra.org](http://www.lancement.tidra.org)





n°14  
Décembre  
2010

# La lettre IN2P3 Informatique

Réseau des Informaticiens de l'IN2P3 et de l'IRFU



CCRI

Comité de Coordination du Réseau des Informaticiens



**Tout d'abord, pouvez-vous nous dire qui composait le CCRI ?**

*Réponse de Cristinel Diaconu*

Le CCRI est composé de 9 membres :

- Le directeur adjoint scientifique de l'IN2P3
- Le chargé de mission pour l'informatique
- Deux animateurs du réseau des informaticiens avec des compétences complémentaires
- Trois membres chargés du contact avec les laboratoires
- Un responsable communication
- Un représentant du centre de calcul

Ce groupe est complété par un contact IRFU et les contacts RESINFO de l'IN2P3.

**Quel était le rôle exact des membres chargés de la communication vis-à-vis des labos ? Comment étiez-vous organisés ?**

*Réponse de Solveig Albrand*

Un élément très important du mandat du CCRI est d'assurer la circulation de l'information au sein de la communauté des informaticiens de l'IN2P3. Pour cela nous avons tenté de mettre en place des canaux de communication entre le CCRI et les services informatiques. Chaque service avait un "correspondant CCRI", et les 3 contacts (Catherine, Sabine et Solveig) ont chacune pris une sous-partie de la liste de laboratoires et se sont mises en relation avec eux. Notre première idée était d'essayer de mettre en place des discussions régulières par téléphone avec chacun de nos contacts, afin de ressentir le pouls de l'informatique dans l'IN2P3. Rapidement, il est apparu que cet investissement était trop lourd pour nous, et que nos interlocuteurs n'avaient rien à signaler de particulier. Les échanges ont été plus riches lorsque nous avons nous mêmes posé des questions spécifiques telle que "Qu'est-ce que vous voudriez voir comme thème pour l'école informatique cette année ?".

**Quel est votre ressenti par rapport à la mission qui vous a été confiée ? Que vous a-t-il manqué ou que reste-t-il à améliorer ?**

*Réponse de Solveig Albrand*

J'ai toujours été partisane des initiatives qui renforcent la communauté, ayant ressentie un grand isolement au début de ma carrière à l'IN2P3. J'ai vu les choses changer pour le mieux au fil des années, et pour moi l'établissement du CCRI a été très positif ; j'espère qu'il a été un catalyseur. Sans doute reste-t-il beaucoup à faire. Nous pensons par exemple qu'il faudrait étendre le rôle des contacts dans les labos, peut-être en faisant plus de réunions CCRI "ouvertes". Mais comme toujours on manque de temps pour ces activités communautaires, qui sont toujours en supplément à notre travail de tous les jours. En tout

*Réponse de David Chamont*

Issus des domaines complémentaires de l'opération des infrastructures et de la production du logiciel, les deux animateurs devaient suivre l'activité des groupes de travail existants et, le cas échéant, stimuler la création de nouveaux groupes. Au delà de cet objectif principal, les interventions des animateurs dans les réunions du CCRI, dans les comités des Journées Informatique et dans les forums du site web, ont permis de faire valoir les points de vue parfois différents des administrateurs et des développeurs.

*Réponse de Jean-Michel Barbet*

En résumé, je crois que le principal travail de l'animateur est de tisser des relations avec tous les ASRs / développeurs et d'aller à la pêche aux informations pour susciter des présentations, des partages d'expérience. Le support (listes, forums, webinaires) n'est pas le plus important.

**Quelles actions avez-vous menées ?**

*Réponse de David Chamont*

Nous nous sommes efforcés, dans les actualités du site web, d'annoncer tous les événements susceptibles d'intéresser notre communauté. Nous avons également amorcé quelques discussions dans les forums. Bien qu'ayant rappelé régulièrement que le site était ouvert à tous, nous n'avons pas vu fleurir beaucoup de contributions spontanées. J'espère qu'elles seront plus nombreuses quand notre communauté et notre site web auront gagné en maturité. En ce qui concerne le développement, je me suis focalisé sur la mise en place d'un groupe de travail sur les forges, qui me semblait urgent et prioritaire. Suite aux échanges sur les besoins et les forges existantes, le groupe a encouragé le centre de calcul à ouvrir plus largement le serveur Redmine qu'il avait installé en interne. Il reste à transformer l'essai en renforçant les moyens attribués au projet. Parmi les chantiers à privilégier par les futurs animateurs, je retiendrais le référencement dans Plume des développements de notre communauté, et l'initiation de groupes sur les bonnes pratiques de développement et/ou les environnements de développement (tel Eclipse).

*Réponse de Jean-Michel Barbet*

D'un point de vue ASR, la principale réussite jusqu'à présent est due au travail de Serge Borderes et de son groupe avec la production du rapport Nomadisme. Encore faut-il que ce travail soit valorisé par la suite. Quelle est la situation ? Un travail au moins similaire est à faire pour les outils collaboratifs suite aux discussions intéressantes aux JI, travail qui nécessitera d'être complétés avec des tests de solution, mais rien ne s'est passé depuis 6 mois. Il est à noter qu'une personne a signalé son intérêt pour des enquêtes régulières afin de suivre l'évolution des tendances du métier d'ASR dans les autres laboratoires. La question est aussi de savoir si cette personne est la seule à ressentir ce besoin. Il y a très peu de retour sur les actions menées ou non-menées, ce problème doit être pris en compte par les futurs membres du CCRI.

**Quel était le rôle du responsable communication ? Quelles actions ont été menées ?**

*Réponse de Eric Legay*

L'objectif du responsable de la communication était de mettre en place les outils nécessaires à la circulation des informations. Très rapidement, le besoin d'un site communautaire c'est est fait sentir, le besoin d'un endroit où l'on pouvait regrouper les informations intéressantes pour tous les informaticiens de l'IN2P3. Ce site existe depuis deux ans, et je tiens à remercier

cas j'ai eu personnellement beaucoup de plaisir à travailler dans le CCRI, et je souhaite bon vent aux nouveaux membres.

*Réponse de Catherine Seznec*

Pour compléter la réponse de Solveig, je souhaiterais ajouter que le rôle des contacts labos ne devrait pas se limiter à un échange de mails, que le téléphone n'est pas la bonne solution, et qu'il faudrait pouvoir trouver du temps pour que les contacts et correspondants dans les labos puissent se rencontrer pour échanger et communiquer plus régulièrement.

*Réponse de Sabine Elles*

En tant que correspondante entre autres du laboratoire du CPPM dont est issu l'idée de création des e-séminaires RI3, j'ai consacré mon temps CCRI à la mise en place de ces derniers : tests et réunions techniques, rédaction de la documentation, pages e-séminaires du site du RI3, recherche des premiers orateurs,... et participé ainsi à la concrétisation de cette idée. Par conséquent, j'ai été complètement absente, faute de temps, concernant le rôle premier de contact laboratoire qui m'avait été confié. J'ai certes communiqué par mail avec mes contacts CCRI, mais nos échanges sont restés trop informels. Je pense que la définition du rôle de contact CCRI est resté trop imprécise lors de ce premier mandat et leur isolement a compliqué leur action au sein de leurs laboratoires respectifs. L'idée d'organiser des rencontres entre contacts CCRI et contacts laboratoires me semble donc prometteuse pour dynamiser les échanges.

**Quel était le rôle des animateurs ? Comment étiez-vous organisés ?**

David Chamont, Christian Helft et Sébastien Grégoire pour leur investissement dans ce site. Cependant de mon point de vue, ce site n'est pas satisfaisant. La partie communautaire n'est pas assez visible et ne joue pas son rôle de fédérateur. L'autre outil mis en place au cours de ce mandat est la lettre informatique, lettre électronique paraissant tous les deux mois principalement grâce au travail de Gaelle Shiffrin.

**Globalement, après 3 ans de fonctionnement, quel bilan retenez-vous des actions du CCRI ?**

*Réponse de Cristinel Diaconu*

Le CCRI a mis ensemble les principales activités à l'intérieur du réseau et a réussi le démarrage de quelques projets nouveaux. Le CCRI a assuré assuré en effet une « permanence » au sein du réseau, pour faire circuler l'information et pour ressortir et croiser les compétences au sein de l'institut. Sans faire de l'autosatisfaction, le premier CCRI a rempli son mandat, malgré un certain nombre d'aspects qui restent à améliorer. Bien sur, le travail en soi n'est pas terminé, puisque cette action d'animation du réseau doit s'inscrire dans la durée et devrait être renforcé. Si la première équipe a démontré la validité du concept, ce sera la mission de l'équipe suivante de mettre en place des projets et des moyens de fédération plus forts à l'intérieur du réseau, pour aboutir à des actions plus implantées dans les laboratoires et avec un impact visible dans les projets de recherches de l'IN2P3.

**Le CCRI**



n°14  
Décembre  
2010

# La lettre IN2P3 Informatique

Réseau des Informaticiens de l'IN2P3 et de l'IRFU



## Super Computing 2010



Le stand de la NASA à SC10

L'édition 2010 de la conférence Super Computing (SC10) s'est tenue à la mi-novembre à la Nouvelle Orléans. Ce rassemblement de plus de 10 000 personnes, et très fortement orienté sur le calcul haute performance, est l'occasion de réunir des constructeurs et des utilisateurs d'horizons très divers.

Le point focal de la conférence est un hall de plus d'un hectare dans lequel cohabitent l'exposition des matériels les plus récents et les démonstrations préparées par les communautés utilisatrices. Pour l'occasion, constructeurs et experts bénévoles assemblent le réseau SCinet qui cette année offrait une bande passante agrégée de 260 Go/s et mettait en œuvre une liaison longue distance à 100 Gb/s.

Traditionnellement Super Computing est l'occasion d'annoncer l'un des deux classements annuels du Top 500 des machines les plus puissantes de la planète. Cette année, la première machine mondiale est chinoise et se nomme Tianhe-1A (la voie lactée). C'est un ordinateur hybride basé sur un assemblage d'INTEL Xeon et de processeurs graphiques (GPGPU) NVIDIA. Tianhe-1A affiche une puissance de 2.6 pétaFlops sur l'étalon LINPACK caractéristique du Top 500, surpassant ainsi le CRAY XT5 JAGUAR de 1.7 pétaFlops installé à Oak Ridge. La première machine française est classée 6ème, il s'agit du ordinateur BULL / TERA100 du CEA militaire.

Lors de l'édition 2009 à Portland, les thèmes abordés lors des conférences ou sur les stands montraient une forte activité au niveau de la problématique du refroidissement des machines. Cette année, il semble que la solution des armoires réfrigérées et du confinement des allées chaudes se soit imposée et que le refroidissement ne soit plus un problème immédiat. Par contre, il est évident que les processeurs des machines de demain (K-Computer de Fujitsu, IBM Blue Water,...) disposeront d'un refroidissement liquide au niveau de la carte mère elle-même.

SC10 confirme la percée de l'utilisation des cartes graphiques GPGPU en tant que processeurs afin de traiter des problèmes nécessitant un haut degré de parallélisme. Les difficultés de programmations inhérentes à ces cartes détournées de leur usage initial, sont de mieux en mieux surmontées et on observe que les « road maps » de plusieurs grands fabricants sont basées sur des processeurs dont l'architecture est directement inspirée des GPGPU.

Au niveau des perspectives, on constate que la puissance de calcul des premières machines du Top 500 augmente d'un facteur 1000 tous les 10 ans. Les experts s'accordent à dire que ce sera encore le cas dans l'avenir et que la barre de l'exaFlop sera franchie dès 2018.

La physique des particules a été longtemps en pointe et a tiré la communauté du calcul scientifique, notamment sur tous les aspects touchant aux transferts et aux traitements des données. Force est de constater que c'est de moins en moins le cas, la problématique des données étant omniprésente dans toutes les disciplines et les superordinateurs étant eux-mêmes devenus des sources de masses de données. Si notre discipline veut continuer de jouer un rôle de leader comme elle en a été capable avec la mise en œuvre des grilles à l'échelle mondiale, il lui faudra être inventive et imaginer les modèles de traitement des données de demain. À cet égard, l'expérimentation que va prochainement mener RENATER sur la liaison CC-IN2P3 – CERN à 100 Gb/s est très importante.

SC11 se tiendra en novembre 2011 à Seattle, le thème principal sera justement focalisé sur les données. C'est peut-être l'occasion pour l'IN2P3 et l'Irfu de renforcer leur présence à cette importante conférence.

**Dominique Boutigny**



## Le standard SAGA à l'OGF



Les utilisateurs de grilles ayant besoin d'accéder à une grande variété d'infrastructures utilisant des intergiciels hétérogènes, il est nécessaire de leur fournir des interfaces uniformes et standardisées. L'OGF (Open Grid Forum) est un consortium dont la mission est de créer et de promouvoir des standards ouverts pour les grilles et autres systèmes distribués (cloud computing...). Les groupes de travail de cette organisation, dédiés à des domaines particuliers, se rencontrent régulièrement lors de colloques.

L'un des standards de l'OGF, SAGA (Simple API for Grid Application) définit une spécification permettant aux applications d'accéder de façon simple et transparente aux ressources (calcul, stockage) de différentes grilles. Le Centre de Calcul de l'IN2P3 / CNRS (CC-IN2P3) propose une implémentation de ce standard utilisable en Java (JSAGA), et une couche d'adaptation en Python (JPySAGA).

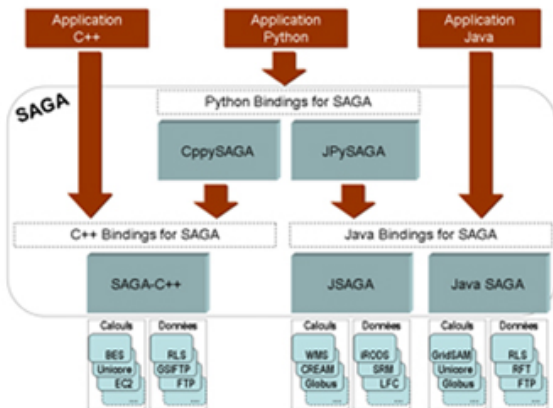
Le groupe de travail chargé de définir ce standard s'est réuni à l'occasion de trois sessions à OGF30, qui s'est déroulé à Bruxelles du 25 au 28 octobre 2010, conjointement à la conférence Grid2010. Les membres du groupe SAGA ont montré, au travers de démos, la faisabilité et la facilité pour une application cliente d'utiliser plusieurs infrastructures de grilles simultanément. A cette occasion, le CC-IN2P3 a présenté une démo dont l'objectif était de montrer qu'il est possible de faire cohabiter plusieurs implémentations de SAGA, de manière à les utiliser conjointement au sein d'une même application et avec un même code source. Cette démo montrait ainsi une application Java utilisant les implémentations JSAGA et JavaSAGA [1], puis des applications Python utilisant les implémentations JSAGA et SAGA-C++ [2]. L'intérêt de telles cohabitations est de permettre à une application d'accéder à un nombre encore plus important d'infrastructures de grille et de fonctionnalités (voir schéma ci-dessous). Les modifications de la spécification proposées par le CC-IN2P3 afin de permettre ces cohabitations ont été accueillies favorablement par le groupe de travail.

Cette démo a par ailleurs été rejouée lors d'une session du groupe de travail GIN (Grid Interoperability Now). L'objectif de ce groupe est de sélectionner les standards à implémenter pour obtenir des composants interopérables. Suite aux discussions au sein de cette session, il a été décidé d'inclure SAGA comme un des éléments essentiels de cette interopérabilité. SAGA permet en effet de proposer une interface uniforme aux différents composants de grille, que ceux-ci implémentent une interface standard ou non. De plus, il n'est pas envisageable, ni à court ni à long terme, que tous les intergiciels de toutes les grilles (académiques/industrielles, internationales/régionales...) et autres systèmes distribués (clusters, grilles de type RPC, grilles de bureau, clouds et les technologies à venir...) implémentent les mêmes interfaces. Les groupes GIN et SAGA ont par ailleurs décidé de collaborer sur la mise en place de démonstrations d'interopérabilité, dont la première aura lieu lors de la prochaine conférence Super Computing en Louisiane.

D'autres groupes de travail, tels que GRAAP (Grid Resource Allocation Agreement Protocol) et GLUE (Grid Laboratory for a Uniform Environment), ont discuté des rapprochements possibles avec le standard SAGA, comme cela avait déjà été le cas dans le passé, notamment avec les standards DRMAA (Distributed Resource Management Application API) et GridRPC (Grid Remote Procedure Call). SAGA a également été mentionné en session plénière [3] et dans la conférence Grid 2010.

Enfin, l'intégration de SAGA dans les intergiciels de la grille américaine de supercalculateurs (Teragrid) et de la grille japonaise (NAREGI) a été évoquée, et le centre de calcul japonais KEK a manifesté son intérêt pour utiliser JPySAGA et JSAGA conjointement à l'implémentation C++ de SAGA.

Julien DEVÉMY, Sylvain REYNAUD, Lionel SCHWARZ



[1] <http://sagacct.lsu.edu/implementation...>

[2] <http://saga.cct.lsu.edu/implementation...>

[3] Voir la présentation de Shantenu Jha : [http://www.ogf.org/gf/event\\_schedul...](http://www.ogf.org/gf/event_schedul...)



n°14  
Décembre  
2010

# La lettre IN2P3 Informatique

Réseau des Informaticiens de l'IN2P3 et de l'IRFU



## CHEP 2010



**CHEP (Computing in High Energy Physics) est une série de grandes conférences se déroulant tous les 18 mois sur un continent différent. Cette conférence, regroupant des physiciens et des informaticiens, permet de cibler les besoins de la recherche et de dessiner les évolutions futures. Cette année, la conférence qui a eu lieu à Taïpei du 18 au 22 octobre, était organisée par le Centre de Calcul de l'Academia Sinica.**



Les sessions plénières et parallèles ont couvert un éventail de sujet allant des outils et du fonctionnement actuels du calcul à toutes les améliorations futures en logiciels et infrastructures. En voici un aperçu.

### La montée en puissance du LHC

Le point principal de cette « session » a été l'arrivée des nouvelles données prises auprès du collisionneur LHC par les 4 expériences Alice, Atlas, CMS et LHCb, ainsi que le fonctionnement de toute la chaîne de calcul pour parvenir aux résultats de physique à publier. Les systèmes « online » de prise de données ont fonctionné et l'efficacité d'utilisation des données prises par les détecteurs a dépassé les 90%. Le traitement de celles-ci a permis d'obtenir des résultats de physique très rapidement, parfois en moins d'une semaine. Les infrastructures de calcul ont été intensivement sollicitées et ont fonctionné bien au-delà du design initial. De manière générale, il a été reconnu que le réseau est la ressource disponible la plus fiable, même si, étant donné sa taille, l'infrastructure est difficile

à maintenir disponible à 100%.

### CPU « manycores »

Comment exploiter le potentiel des nouveaux CPU (Sverre Jarp, Roger Goff...)? Les CPU sont composés d'un nombre croissant de cœurs qui comportent des unités de calculs complexes et fonctionnant de façon de plus en plus vectorielle. Un changement de traitement est nécessaire pour permettre d'en exploiter tout le potentiel : il faut « penser parallèle » aussi bien pour l'organisation des données que pour celles des tâches à accomplir. Il faut en effet penser les futures applications de façon à ce qu'elles puissent fonctionner le plus possible en parallèle et qu'elles s'adaptent naturellement aux améliorations des futurs systèmes. La physique des hautes énergies présente un avantage : les données exhibent déjà un « pattern parallèle ». Elles sont en effet classées par événements indépendants, traces, vertex, etc. Des essais de parallélisation des programmes ont été présentés dans les sessions parallèles. Ils montrent des résultats prometteurs lorsqu'ils sont appliqués à GEANT4 ou à des logiciels d'expérience.

### Calcul sur les nuages

Beaucoup d'activités s'orientent vers le « cloud computing », c'est à dire vers une infrastructure abstraite du point de vue de l'utilisateur mais sur laquelle il trouve les ressources de calcul, de stockage ou de bande passante sur une plateforme qui répond à sa requête particulière. Il existe des nuages commerciaux et des nuages publics et leur utilisation est déjà rentrée dans sa phase de production bien qu'ils ne remplacent pas le calcul sur grille.

### L'électronique «green»

Les supports de stockage sont actuellement les plus gros consommateurs d'énergie parmi les composants électroniques d'un centre de calcul. Ils génèrent plus du quart des émissions de CO<sub>2</sub> attribuées à la technologie de l'information. Des présentations d'entreprises commerciales ont montré les bénéfices des SSD (Solid State Drives). Ceux-ci ont en effet l'avantage d'un petit volume pour une efficacité accrue. Des tests ont montré leur haute performance en terme de lecture/écriture, et leur grande fiabilité. L'accent a été en outre mis sur leur composante « verte ». Les SSD n'ayant pas de composants mécaniques, ils consomment moins d'énergie et émettent moins de chaleur. Bien que leur prix est à ce jour relativement élevé, il devrait progressivement diminuer et les SSD pourraient bien à terme remplacer les types de stockage d'informations utilisés jusqu'à présent (HDD par exemple).

D'autres sujets ont été traités, tels les nouveaux challenges de communication et d'accès à la science par le public, la préservation et la possibilité d'exploitation future des données des expériences déjà arrêtées. Pour conclure, les participants se sont donnés rendez-vous dans dix-huit mois sur le continent américain, à New-York.

Ghita Rahal



n°14  
Décembre  
2010

# La lettre IN2P3 Informatique

Réseau des Informaticiens de l'IN2P3 et de l'IRFU



## Computing and Astroparticule Physics



Les 7 et 8 octobre, un atelier organisé par ASPERA a rassemblé au Centre de Calcul de l'IN2P3 (CC-IN2P3) une cinquantaine de participants autour du calcul pour l'Astroparticule.

ASPERA est un réseau d'agences gouvernementales, responsable de la coordination et du financement des projets européens en astroparticule. Certaines grandes infrastructures proposées par la feuille de route d'ASPERA vont être confrontées à des besoins ambitieux en terme de calcul. Une coordination au niveau européen sur cette thématique devient nécessaire. Cet atelier proposé par Stavros KATSANEVAS a permis de cerner certaines opportunités de collaboration entre les projets et les organismes invités.

Certaines présentations décrivent en détail des modèles de calcul pour les rayons cosmiques, les ondes gravitationnelles et l'énergie noire. Sans représenter une vue exhaustive de l'ensemble des activités en astroparticule, ces présentations mettent en évidence la variété des besoins spécifiques en question.

La grille utilisée pour le LHC permet de satisfaire d'énormes besoins en calcul et stockage. Les rayons cosmiques utilisent des méthodes de calcul similaires aux expériences de physique des particules. Cependant, si le réseau est extrêmement fiable, la gestion de la grille reste encore lourde. Le traitement des données est rendu difficile par les incidents fréquents affectant le hardware. Une amélioration de la tolérance aux incidents s'appuyant sur la fiabilité du réseau pourrait être envisagée. Giovanni LAMANNA, pour le CTA, et Christian NEISSNER, pour Magic, ont partagé leur expérience dans la mise en place du calcul sur la grille et mis en évidence que l'utilisation de la grille par un projet offre des avantages dans un contexte très international. Jean-Noël ALBERT a implémenté pour AUGER une solution mixte prenant en compte non seulement le calcul intensif nécessaire pour les simulations, mais aussi le volume relativement limité de données. Ainsi les simulations pour AUGER se font sur la grille mais le traitement des données se fait en local au CC-IN2P3. Dans le domaine des ondes gravitationnelles, la grille est utilisée pour LIGO et VIRGO.

A la suite de la grille européenne EGEE, EGI s'attache à rendre cette infrastructure pérenne. Les interactions entre EGI et les communautés de recherche virtuelles sont d'avantage formalisées. La grille fournit aussi un cadre adapté pour les outils interactifs de contrôle et de surveillance de télescopes distants. Dans un domaine connexe à la Grille, le Cloud pourrait apporter un degré supplémentaire de flexibilité et de virtualisation des inhomogénéités entre sites, comme par exemple le projet StratusLab.

La participation de centres de calcul européens à cet atelier ouvre la perspective d'une meilleure coordination pour le calcul en astroparticule. Des descriptions détaillées de leurs activités ont été présentées et, lors de la table ronde, les Tiers-1 ont confronté leurs activités présentes et leurs perspectives futures.

Le partage, la fusion et l'accès à des données très hétérogènes, mais complémentaires en terme d'objectifs scientifiques, représentent un enjeu majeur pour l'astroparticule. Le fonctionnement de l'observatoire virtuel IVOA, déjà largement utilisé en astronomie pour la publication de quantités considérables de données au travers d'une interface web, a été détaillé durant ce colloque. Ce partage des données est aussi intéressant pour l'astroparticule car il permet par exemple de fusionner des observations indépendantes utilisant des particules messagères différentes et provenant d'une même source astrophysique. Ces données peuvent aussi être réutilisées par des physiciens ayant des objectifs scientifiques variés pour en extraire des informations pertinentes à leur travail. Comme l'a souligné Roberto COSSU à la suite de sa présentation, un partage de données avec les sciences environnementales pourrait se révéler mutuellement fructueux. En effet certains instruments utilisés en astroparticule effectuent des mesures couvrant des espaces considérables sous les mers ou sur les terres correspondant aux échelles caractéristiques de géophysique.

En ce qui concerne le hardware, certains besoins en astroparticule bénéficient de solutions spécifiques. Le gain énorme obtenu avec les processeurs de type GPU, déjà très largement utilisés pour les ondes gravitationnelles, a été mis en avant. Les GPUs sont parfaitement adaptés à l'algèbre linéaire notamment pour le traitement du signal. Ce type de processeur pourrait aussi être performant dans certaines analyses utilisant les données du type LSST. Les expériences d'énergie noire nécessitent des accès très intensifs aux données et donc des disques particulièrement performants. Aussi un modèle de calcul totalement distribué entre différents sites n'apparaît pas envisageable. Le LSST par exemple nécessite une infrastructure de calcul très ambitieuse. Les objectifs scientifiques de l'expérience spatiale EUCLID quant à eux demandent une analyse de données provenant d'autres expériences notamment du LSST. Un traitement spécifique de ces données externes est nécessaire pour les rendre utilisables par EUCLID. Cela représenterait un supplément énorme en calcul.

Pour plus d'informations sur ce colloque : <http://indico.in2p3.fr/conferenceDisplay.py?confId=3845>

Le prochain atelier, prévu à Barcelone fin Mars, permettra d'approfondir la coordination des activités et d'étudier certains sujets importants non abordés lors de ce premier atelier. Ainsi l'optimisation et la structuration du software pourraient être discutés ainsi que les bases de données en particulier pour le LSST. Parallèlement à ces ateliers, des interactions plus bilatérales et ciblées entre projets et infrastructures seront concrétisées.

Rachid Lemrani





**n°14**  
Décembre  
2010

# La lettre IN2P3 Informatique

Réseau des Informaticiens de l'IN2P3 et de l'IRFU



## Arrêt du Centre de Calcul de l'IN2P3

14 Décembre 2010

Un arrêt des systèmes informatiques du CC-IN2P3 est programmé pour le 14 Décembre 2010.

Le blocage de l'entrée des jobs utilisant HPSS, XROOTD, ORACLE, SRB et IRODS se fera la veille, le 13 Décembre à partir de 19:00. Ces ressources seront débloquées l'après-midi du 14 Décembre.

Les opérations de maintenance concernent :

- HPSS : indisponible toute la matinée jusqu'en début d'après-midi. Cette intervention impactera aussi XROOTD, DCACHE, SRB et IRODS car il n'y aura pas d'accès aux bandes.
- ORACLE - mises à jour : arrêt de toutes les bases toute la matinée jusqu'en début d'après-midi.
- ORACLE - migration : dans l'après-midi impactant fts, lfc (sauf lfc-lhcb-ro), voms (accessible en Read-only).



**n°14**  
Décembre  
2010

# La lettre IN2P3 Informatique

Réseau des Informaticiens de l'IN2P3 et de l'IRFU



## Recherche matériel informatique ancien

Le Centre de Calcul de l'IN2P3 recherche du matériel informatique ancien afin d'enrichir son musée de l'informatique. Notamment des mémoires à tores de ferrite, un modem à couplage acoustique, des plateaux de disque 14" ou plus, un disque dur Winchester 8", un disque dur 8", etc.

Si vous disposez de ce type de matériel, ou de tout autre qui puisse venir enrichir la collection du CC-IN2P3, merci d'envoyer un courriel à Fabien Wernli ([wernli@in2p3.fr](mailto:wernli@in2p3.fr)).