



A LA  
UNE

## CHEP, le rendez-vous international incontournable



L'édition 2012 organisée par le laboratoire de Brookhaven (BNL), s'est tenue du 21 au 25 mai dernier, à l'Université de New-York (NYU) en plein cœur de Greenwich Village. Plus de cinq cents personnes étaient réunies pour évoquer le futur des outils informatiques et les évolutions du calcul en Physique Nucléaire et Physique des

Particules. La « délégation française » était, comme à chaque fois, constituée d'une bonne vingtaine de collègues de l'IN2P3 et de l'IRFU qui ont suivi assidûment les différentes sessions parallèles.

Le programme de CHEP aborde le traitement des données dans tous ses aspects, depuis le traitement en ligne au plus près des détecteurs jusqu'à l'analyse finale) sans oublier la préservation des données à long terme et les outils collaboratifs. Cette année a été, très logiquement, marquée par une présence forte du calcul LHC, les perspectives et les besoins étant essentiellement portés par le projet LSST (Large Synoptic Survey Telescope). Cet article met en lumière de façon subjective quelques idées marquantes issues de la pléthore de sujets et de présentations.

GEANT et ROOT restent les deux piliers du Computing HEP : le premier pour la simulation Monte Carlo, le second pour l'analyse de données à l'aide d'histogrammes. Depuis le milieu (...)

### Interview

**Michel Jouvin : "Le GDB a pour mission de fournir aux expériences LHC une infrastructure informatique stable et fiable."**



Nouveau chairman du Grid Deployment Board de LCG

[lire l'interview](#)

[lire la suite](#)

### Agenda

#### Appel à proposition pour les Journées Informatique 2012

Le réseau des Informaticiens IN2P3 IRFU organise tous les deux ans des Journées Informatique (...)

[en savoir plus](#)

#### 1-3 oct. 2012 - Journées scientifiques mésocentres et France Grilles

Le GIS France Grilles et le Groupe Calcul organisent pour la première fois en coordination les (...)

[en savoir plus](#)

#### EGI Technical Forum - sept. 2012

L'inscription pour le Technical Forum, qui aura lieu à Prague du 17 à 21 septembre 2012, vient (...)

[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).

#### Visioconférence



#### Le CC IN2P3 remet la responsabilité du service RMS à Renater

Dès 1999, sur les traces des premières expériences de visio-conférences multicast sur BETEL, une politique volontariste de l'IN2P3 a décidé de déployer une infrastructure de visioconférence dans ses laboratoires. Rejointe par la DSI du CNRS qui a déployé ses propres matériels dans ses DR dès 2000. La nécessité est alors vite apparue de donner aux composantes de ces deux infrastructures, voisines et parallèles, des moyens d'interconnexion entre elles, et avec le reste du monde, mais surtout à plusieurs points simultanément. La conférence multisites devenait ainsi possible avec l'achat en commun, et avec l'INSERM, puis l'INRA, du premier MCU, géré pour la communauté de ces EPST, par le CC-IN2P3.

[lire la suite](#)

#### Collaboration



#### L'IN2P3 officialise ses relations avec Israël

Dans le cadre du Projet International de Coopération Scientifique mis en place cette année entre l'IN2P3 et l'Université de Tel-Aviv, une délégation s'est rendue fin juin en Israël.

La délégation était composée de Dominique Boutigny, directeur du CC-IN2P3, Mattieu Puel, administrateur système en charge du déploiement de l'infrastructure cloud au CC-IN2P3, Loïc Tortay, expert stockage, Gaëlle Shifrin, responsable des partenariats et de la communication, et Frédéric Suter, chercheur en informatique, chargé de la liaison entre les communautés de recherche et de production.

[lire la suite](#)

#### Partenariat

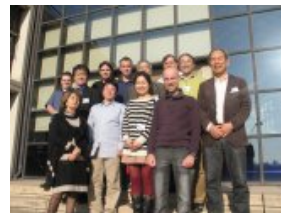


#### Le CC-IN2P3 s'associe à DELL pour le développement de solutions

Le CNRS, via le Centre de Calcul de l'IN2P3, vient de signer une convention de partenariat d'une durée de trois ans avec la société DELL, visant au renforcement de l'expertise et au développement de solutions autour du traitement massif de données et des moyens de calcul pour la recherche. Ce partenariat public - privé prend la forme d'une mise à disposition d'une plateforme DELL à la pointe de la technologie. Cette plateforme sera mise à rude épreuve par les ingénieurs du CC-IN2P3 puisqu'elle leur permettra de mener des tests sur des cas d'utilisation concrets et d'évaluer la pertinence de solutions informatiques innovantes non encore mises sur le marché.

[lire la suite](#)

#### International



#### Les LIA avec l'Asie se portent bien

Depuis 5 ans, l'IN2P3 a entrepris sous l'impulsion de François Le Diberder de développer la collaboration avec l'Asie. Plusieurs Laboratoires Internationaux Associés sont nés de cette volonté de tisser des liens plus étroits avec les établissements de recherche asiatiques. Ces laboratoires sans murs sont co-dirigés et co-financés respectivement par les pays partenaires et proposent des cadres propices à la croissance de collaborations. Quatre Laboratoires Internationaux Associés ont ainsi vu le jour depuis 2006 : le FCPPL avec la Chine, le FJPPL avec le Japon, le FKPPPL avec la Corée et le FVPPL avec le Vietnam. Aujourd'hui, près d'une centaine de projets scientifiques impliquent plusieurs centaines de chercheurs des établissements de recherche français et asiatiques

[lire la suite](#)

impliqués dans ces LIAs dans les domaines de la physique des 2 infinis et des disciplines connexes, dont la grille informatique.

[lire la suite](#)



**Equipe**

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

**Comité de rédaction** : Dominique Cathala, Virginie Dutruel, Sébastien Grégoire, Eric Legay, et Gaëlle Shifrin

---



n°21  
Juillet  
2012

# La lettre IN2P3 Informatique

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

A LA  
UNE

## CHEP, le rendez-vous international incontournable

Conférence



L'édition 2012 organisée par le laboratoire de Brookhaven (BNL), s'est tenue du 21 au 25 mai dernier, à l'Université de New-York (NYU) en plein cœur de Greenwich Village. Plus de cinq cents personnes étaient réunies pour évoquer le futur des outils informatiques et les évolutions du calcul en Physique Nucléaire et Physique des Particules. La « délégation française » était, comme à chaque fois, constituée d'une bonne vingtaine de collègues de l'IN2P3 et de l'IRFU qui ont suivi assidûment les différentes sessions parallèles.



développe un nouveau système de gestion de disques, EOS, qui après un an d'exploitation se révèle beaucoup plus adapté pour l'accès aléatoire aux données. Pour preuve, près de 9 Pétaoctets de données essentiellement ATLAS et CMS, sont aujourd'hui stockés dans EOS. Décoder le sigle ne vous sera d'aucune utilité. Comprenez qu'il s'agit d'une nouvelle approche dite « multi-tiers » voulue, non seulement pour découpler le stockage disque du stockage de masse toujours géré par CASTOR, mais aussi pour hiérarchiser les niveaux disque. Ainsi, la couche supérieure autorisant les accès à faible latence est séparée de la couche assurant la volumétrie disque. EOS se charge de répartir les fichiers par bloc entre les différents serveurs de disques et de gérer l'espace de nommage en mémoire. On parle alors d'architecture RAIN (Redundant array of Inexpensive/Independent Nodes) qui est aux serveurs ce que le RAID est aux disques.

Fédérer le stockage à l'échelle de la grille est l'autre idée qui anime la communauté des sites LCG (LHC Computing Grid). Les possibilités offertes par le réseau permettent de l'envisager aujourd'hui. Ce serait enfin mettre le stockage au niveau du CPU et introduire de la souplesse dans les modèles de gestion de données ! Imaginer un accès uniforme aux données, décloisonner le stockage des sites, éviter la distribution a priori de données, ignorer la localisation des données, optimiser l'accès aux données en fonction de leur popularité, disposer d'un espace de nommage unique et cohérent, s'affranchir des défaillances du stockage local, exploiter des sites ne disposant pas de stockage, la liste des avantages est longue et justifie parfaitement l'intérêt actuel. ALICE exploite depuis assez longtemps cette idée et a promu une approche basée sur xrootd. Plusieurs fédérations basées sur xrootd sont également déployées aux Etats-Unis pour ATLAS et CMS. Mais la réflexion actuelle doit en partie s'affranchir des implémentations (xrootd, http/webDAV...) pour aborder des points importants liés aux modèles et algorithmes de cache de données, à la connaissance et la surveillance de la topologie réseau, à la gestion des autorisations et plus généralement des métadonnées : ACLs, checksums etc.

Côté infrastructures, à l'évidence, les choix se diversifient. Après étude des différents critères (coût, performance énergétique, localité, pérennité), l'expérience LHCb a choisi une solution de « data center en conteneur » pour héberger la nouvelle ferme online qui sera nécessaire après le second arrêt du LHC (LS2) prévu en 2017. Autre besoin, autre choix : l'extension du Tier-0 pour les besoins du calcul LHC au CERN a fait l'objet d'un contrat d'hébergement avec un centre situé à Budapest suite à un appel d'offre. Ces choix sont autant d'alternatives à la construction de nouvelles infrastructures de calcul au CERN et marque un tournant. Le temps où la physique des particules, le CERN en particulier, devait satisfaire les besoins des utilisateurs « les plus gourmands en ressources qui soient » en déployant des solutions totalement spécifiques, est en passe d'être définitivement révolu. A tel point que le CERN prend aujourd'hui exemple sur les datacenters extérieurs pour promouvoir une approche basée sur l'industrialisation, la sous-traitance et l'agilité (osons le terme !) de ses services et activités IT. A ceci, s'ajoute le souci de l'efficacité énergétique pour un « calcul écologique ou vert ». « L'informatique au plus juste » pourrait être le maître mot aujourd'hui.

A l'ère du « Big data », la communauté HEP est un acteur parmi d'autres en matière de déploiement d'infrastructures IT à grande échelle. Les solutions à grande échelle ne peuvent plus et ne doivent plus être spécifiques. Chacun pensera aux solutions Cloud que nous n'avons malheureusement pas abordées ici mais qui ont été largement évoquées à CHEP dans le cadre de la session « Traitement distribué et Analyse sur Grille & Clouds ».

Le programme de CHEP aborde le traitement des données dans tous ses aspects, depuis le traitement en ligne au plus près des détecteurs jusqu'à l'analyse finale) sans oublier la préservation des données à long terme et les outils collaboratifs. Cette année



a été, très logiquement, marquée par une présence forte du calcul LHC, les perspectives et les besoins étant essentiellement portés par le projet LSST (Large Synoptic Survey Telescope). Cet article met en lumière de façon subjective quelques idées marquantes issues de la pléthore de sujets et de présentations.

**GEANT et ROOT** restent les deux piliers du Computing HEP : le premier pour la simulation Monte Carlo, le second pour l'analyse de données à l'aide d'histogrammes. Depuis le milieu des années 90, ces deux logiciels que l'on a plutôt l'habitude de qualifier de « frameworks » (ou cadriciels en bon français), restent au cœur du programme de CHEP. Leurs évolutions sont intimement liées à l'histoire de la conférence. GEANT4 a été initié lors de la tenue de CHEP 1994 à San Francisco. Le cadriciel GEANT est aujourd'hui de loin le plus cité dans des publications scientifiques avec près de 3000 articles, laissant le deuxième (ROOT) loin derrière avec environ 500 citations jusqu'au début 2012.

Cette année, après plus 30 ans d'activité au CERN, **René Brun** a été salué par la communauté en ouvrant la première séance plénière et en livrant sa vision de l'évolution du software HEP des années 1970 jusqu'à cette année qui sera marquée par son départ à la retraite. Évoquant dix-sept années de développement de ROOT, il n'a pas éludé les difficultés actuelles de nos logiciels : utilisation de la mémoire RAM, structuration et format de données (traces, vertex, événement) incroyablement complexes, absence, dans la majorité des cas, de tests systématiques et automatisés. La structuration, l'agrégation des données et les performances des entrées/sorties sont au cœur des enjeux depuis près de 25 ans. ROOT, à partir du concept de TTree et avec le support de la persistance, a certes influencé le format de stockage des données à l'interface entre le traitement en batch et l'interactif. Mais, face aux volumes de données croissants, les questions restent les mêmes : « Comment naviguer dans un grand nombre de fichiers en n'utilisant qu'une petite fraction de chaque fichier ? Comment limiter l'impact des étapes séquentielles comme la fusion (ou merging) de fichiers ? ». En outre, les processeurs multi-cœurs ont déjà un gros impact sur l'architecture des logiciels. Demain, la communauté HEP devra être capable d'exploiter le parallélisme des nouvelles architectures. Pour cela, il faudra être en mesure d'ajuster, de façon dynamique, les ressources critiques à tous les niveaux possibles : mémoire, cache, bande passante.

DELL, principal sponsor de cette édition, a évoqué le **passage à l'Exascale Computing** repoussant cette échéance vers 2018-2019 dans le meilleur des cas. Chacun sait que la fréquence des processeurs stagne depuis dix ans autour de 3 GHz et que les facteurs d'amélioration se trouvent ailleurs, d'autant que la principale préoccupation reste la consommation électrique. Autre difficulté, la puissance de calcul ne suit pas nécessairement l'évolution de performances des processeurs à cause des débits mémoire. Dans ce contexte, les acteurs envisagent des approches hybrides combinant GPU et CPU avec des processeurs ARM issus du marché des Smartphones et tablettes qui seraient en mesure de venir concurrencer les processeurs à architecture x86.

Une autre facette de la difficulté d'exploiter les avancées techniques selon « la loi de Moore » était représentée par le débit de transfert des réseaux et le stockage. La loi de Moore ayant été formulée initialement concernant le degré d'intégration (de densité) des transistors sur les puces des circuits intégrés, elle était vite généralisée sur les autres paramètres de performance comme la vitesse des CPU, taille mémoire, les débits de réseau, vitesse et volume de stockage. Or, la fréquence d'ordonnement des CPU stagnante depuis quelques années (« Free lunch is over »), alors que les autres caractéristiques continuent à s'améliorer, la cohabitation entre parallélisme et accès sériel aux canaux d'une interface unique pose et continuera à poser quelques défis aux développeurs.

**Infrastructures de calcul, grilles de production et réseaux**, cette session, parmi les plus suivies, a réuni en permanence une assistance de 70 à 80 personnes. Près d'un quart des présentations traitait des systèmes de stockage et des performances des entrées/sorties toujours en relation avec les besoins d'analyse. Outre l'analyse des goulots d'étranglements, la comparaison entre les performances des différents protocoles, les promesses de NFS version 4.1, de nouvelles

l'avenir passe aussi par **les FPGAs et les GPUs** qui vont permettre d'accroître la puissance de calcul disponible en ligne au plus près des détecteurs. Tous les efforts de portage et d'optimisation des algorithmes de reconstruction sont requis et seront utiles pour le online comme pour le offline. Avec l'amélioration des performances du LHC, il faudra à la fois gérer l'empilement des événements et améliorer la pertinence et les performances de la sélection en ligne. En résumé, il faut traiter plus de données, plus rapidement et être capable d'exécuter en ligne de nouveaux algorithmes, de reconstruction. Aujourd'hui, la sélection en ligne se base sur les données des détecteurs à muons et des calorimètres. Demain, il est envisagé d'exploiter en ligne les informations fournies par les détecteurs de traces, jusqu'alors considérés comme nécessitant des algorithmes de reconnaissance de particules trop complexes pour fournir des résultats utilisables et rapides. Tout l'intérêt des GPU pour le tracking et l'amélioration des triggers est là, se reflétant dans une session plénière complète (Event processing) traitant exclusivement des GPU et de leurs applications.

Outil devenu indispensable au point d'oublier parfois les exploits technologiques qui étaient nécessaires pour leur mise en œuvre, **les outils collaboratifs** font l'objet de deux séances parallèles, récentes dans l'histoire de CHEP, mais en constante augmentation de fréquentation et de présentation. Le choix crucial d'un système de visioconférences universel qui semble décidé du fait que le CERN a initié la migration définitive vers le produit Vidyo a été remis en discussion. Philippe Galvez du Caltech a présenté en séance plénière une étude de marché comprenant des offres commerciales, mais surtout les grands joueurs ECS/ESnet et EVO (ainsi que Vidyo). Encore une fois « HEP » mène le palmarès des utilisateurs les plus demandeurs. Rien que sur EVO, 60 millions de minutes annuelles de conférences en vidéo, audio et de partage d'écran mélangés sont gérées par le système. Rappelons qu'une année ne compte 526 000 minutes ! Le niveau d'utilisation est donc équivalent à 120 utilisateurs en moyenne, connectés en permanence. Les documents associés à une expérience sont gérés sans les contraintes de fonctionnement et de disponibilité 24/7, mais la volumétrie est tout aussi impressionnante, comme nous l'a montré Lucas Taylor, imaginant tous les documents de CMS imprimés et empilés (voir ci-dessous).



Lucas Taylor, imaginant tous les documents de CMS imprimés et empilés.

Pour la première fois, le **groupe international DPHEP** qui développe des idées et solutions concernant la préservation des données et résultats de physique des expériences, a présenté le résultat de ses travaux en séance plénière, ayant également organisé un atelier d'une demi-journée en parallèle avec le programme CHEP.

**Est-ce que les paradigmes modernes du développement logiciel s'appliquent à HEP ?** Telle était la question soulevée par Federico Carminati en séance plénière. Après avoir fait une analyse comparative entre l'« agile » (agile), l'« extrême » (extreme programming, XP) et les habitudes dans notre domaine, il a conclu que la communauté HEP avait en réalité développé, sans s'en rendre compte, sa propre méthode de développement : une méthode itérative et incrémentale, menée dans un esprit collaboratif avec juste ce qu'il faut de formalisme. Resterait juste à la communauté HEP à expliciter, à adapter et à optimiser ce qu'elle fait naturellement et spontanément ! Chacun appréciera.

NIKHEF vous donne rendez-vous à Amsterdam du 14 au 18 octobre 2013 pour la 20<sup>e</sup> édition de CHEP, 30 ans après ses débuts au même endroit !

approches (hiérarchisation et fédération notamment) viennent enrichir la réflexion du côté du stockage des données.

*Pour en savoir plus : <http://www.chep2012.org/>*

**Frédérique CHOLLET, Dirk HOFFMANN**

L'approche HSM classique (disques et bandes), typiquement incarnée par CASTOR et mise en place pour les besoins du stockage des données LHC au CERN, n'est clairement pas adaptée aux besoins de l'analyse. Depuis mi 2010, le CERN



n°21  
Juillet  
2012

# La lettre IN2P3 Informatique

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



## Michel Jouvin : "Le GDB a pour mission de fournir aux expériences LHC une infrastructure informatique stable et fiable. "

Nouveau chairman du Grid Deployment Board de LCG



- Michel, tu viens d'être nommé chairman du GDB, peux-tu nous rappeler en quelques mots ce qu'est ce GDB ?

GDB veut dire Grid Deployment Board. Réuni une fois par mois, c'est un des comités du projet WLCG qui coordonne l'informatique des expériences LHC. Cette informatique repose uniquement sur la grille, en fait trois infrastructures de grilles : EGI en Europe, OSG aux Etats-Unis, NorduGRID dans les pays nordiques (+ la Slovénie !). A l'origine le GDB a été créé pour coordonner le déploiement des services grilles en vue de leur utilisation par les expériences LHC. A ce titre, il regroupe les principaux sites offrant des ressources à WLCG (en particulier le CERN et les T1s) et les quatre expériences LHC ainsi que la coordination technique de WLCG localisée au CERN. Depuis le début de la prise de données en 2010, le GDB a évolué en une coordination technique entre les sites et les expériences où sont discutés les principaux problèmes et où se discute les évolutions techniques. Les sujets discutés concernent les services utilisés par les expériences et non pas l'administration du site proprement dit : en ce sens, le GDB est très complémentaire de HEPiX avec lequel il entretient des liens explicites. Ce n'est pas une structure décisionnelle avec une représentation formelle des expériences mais plutôt un forum où s'élaborent des consensus sur lesquels s'appuieront d'éventuelles décisions du Management Board. En France les principaux participants sont un représentant du CC et la direction technique de LCG France, la « branche française » de WLCG.

- Quel impact peut avoir le GDB sur les structures de calcul actuel ?

L'influence directe du GDB se limite aux ressources WLCG. Pour WLCG, c'est le lieu où s'élabore les solutions techniques mises en œuvre par les sites. Comme pour beaucoup de sites, WLCG est un client majeur, cela façonne certains choix des sites. Pour optimiser la gestion, ceux-ci essaient généralement de mutualiser les solutions. Indirectement le GDB peut donc avoir aussi une influence sur certains services offerts à d'autres communautés mais cela ne fait pas parti de ses missions. Ceci dit, les sites n'ont pas à redouter de comportements arbitraires du GDB. Le GDB a pour mission de fournir aux expériences LHC une infrastructure informatique stable et fiable. De ce point de vue, le GDB est plutôt conservateur ! Toute évolution est décidée de façon à permettre une continuité de service. C'est assez important dans une infrastructure distribuée à l'échelle de la grille.

Pour illustrer ces propos, un exemple : l'authentification dans la grille repose sur les certificats. Les certificats actuels utilisent un algorithme d'encryptage de la clé privé qui s'appelle SHA1. Il se trouve qu'avec l'augmentation de la puissance de calcul disponible aujourd'hui, SHA1 commence à être considéré comme vulnérable et il va falloir utiliser SHA2 qui est plus complexe et donc plus difficile à casser. Pour différentes raisons, cette transition est complexe et nécessite la mise à jour d'un certain nombre de services avant la première utilisation d'un encryptage SHA2. Le GDB joue un rôle moteur sur ce problème

mais bien évidemment il ne peut pas le faire sans concertation avec les différentes grilles, EGI, ... car cela aura un impact pour tout le monde à relativement court terme.

- Comment vois-tu l'évolution du calcul pour les années à venir, entre autre avec l'upgrade du LHC ?

Je ne suis pas sûr que l'upgrade du LHC en 2018 est un facteur déterminant de l'évolution. C'est encore trop loin. La principale motivation pour les évolutions prévues à court et moyen terme est plutôt l'expérience acquise depuis le démarrage de la prise de données. Contrairement à ce que certains redoutaient, l'infrastructure informatique des expériences LHC reposant sur la grille a relevé le défi et a permis une analyse rapide de toutes les données acquises. Mais derrière ce succès se cache une complexité importante qui nécessite un effort humain non négligeable pour obtenir le niveau de service requis. Avec la baisse des budgets et des moyens humains partout, il est indispensable de simplifier l'opération des services. Dans certains cas, il peut s'agir d'améliorer le service lui-même, dans d'autres cas cela passe par une refonte plus importante d'un service. Avec deux ans d'expérience, nous sommes maintenant dans une bonne position pour analyser ce qui est inutilement complexe ou ce qui n'est pas assez fonctionnel ou flexible.

Un autre axe de l'évolution consiste à prendre en compte l'explosion de la problématique « Big Data » dont le LHC a été précurseur. Il y a 10 ans, il a fallu tout inventer. Aujourd'hui un certain nombre de services et protocoles standards sont disponibles pour traiter les besoins des expériences LHC. Notre volonté est donc de favoriser de plus en plus l'utilisation de ces solutions en remplacement des solutions spécifiques WLCG.

Enfin, nous devons prendre en compte que la technologie qui concentre les efforts aujourd'hui est le cloud. Le cloud, c'est à la fois des offres commerciales et une technologie qui s'appuie sur la virtualisation. Il est probable qu'à l'avenir de plus en plus de ressources seront mises dans le cloud plutôt que dans la grille. Malheureusement le cloud n'est pas une alternative à la grille et la réflexion est en cours pour déterminer comment la grille et/ou les expériences peuvent tirer parti efficacement de ressources de type cloud.

- D'après toi, quel est l'impact de ces évolutions pour le calcul au sein de l'IN2P3 et de l'IRFU ?

Pour le court terme, je pense que l'impact restera assez faible. Il existe un réseau de compétences solide, habitué à gérer les ressources grilles et les évolutions régulières des services. La simplification des services en cours a vraiment pour but de simplifier la vie des sites !

A un peu plus longue échéance, si la grille évolue pour permettre l'utilisation de ressources de type cloud, cela peut impacter les services mis en œuvre par les sites. De plus, il n'est pas complètement évident que dans ces évolutions, la forte distribution des ressources, avec des sites de taille très différente, persiste. On verra peut être une évolution de la « géographie » des ressources et cela pourrait impacter le réseau de compétences très diffus qui fait une des forces de la grille en France et que je tends à considérer comme une des réussites de la grille. Cette grande capillarité de l'expertise grille bénéficie aux utilisateurs qui ont des experts près d'eux. Mais sur ce point, je pense que l'évolution de la politique nationale a au moins autant de poids, sinon plus, que le GDB...

Lien utile : <http://lcg.web.cern.ch/lcg/gdb.htm>.

PROPOS RECUEILLIS PAR E. LEGAY



n°21  
Juillet  
2012

# La lettre IN2P3 Informatique

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



## Le CC IN2P3 remet la responsabilité du service RMS à Renater



interrogations parmi nos collègues IN2P3 et alliés : "Qu'en serait-il de notre liberté d'usage et de nos facilités ?" Aucun changement n'est à prévoir quant aux possibilités d'usage par nous, IN2P3, bien entendu.

Le principe d'accueil est le même que celui que nous appliquons : faire partie d'une EPST "éduc-recherche" (hors enseignement pur). La hot-line sera désormais assurée par un service dédié de Renater qui, fort de sa meilleure influence et pénétration dans le tissu éduc-recherche devrait mettre en place une hot-line à deux niveaux.

Chaque organisme ou EPST devra, pour être accepté, indiquer un contact référent (qui pourra être démultiplié ensuite localement suivant l'organisation, université par exemple), lequel pourra s'adresser à Renater directement en cas de problème non résolu à son niveau. Logique, tout ça !

Il est donc prévu que la cellule visio de l'IN2P3 continue d'apporter son support et son expertise à la demande, en cas de problème technique (de toutes façons déjà peu nombreux !). Ce, pendant une période de transition d'environ un an, depuis le 1er janvier 2012. Elle continuera ensuite si nécessaire à apporter son aide, mais à nos seuls collègues IN2P3. Renater a toujours démontré son assurance et sa volonté de maintenir tel quel l'usage déjà existant pour les premiers utilisateurs "historiques". Nous ne sommes d'ailleurs plus les plus importants en nombre, l'INRA ayant bien mieux déployé et utilisé cette infrastructure que nous, qui avons aussi d'autres choix possibles avec notamment VRVS, puis EVO, et maintenant Vidyo.

Fort de ce point rassurant, l'IN2P3 a-t-elle une stratégie à 5 ans ? Disons que bien avant la reprise du service par Renater, nous, "gérants" au CC-IN2P3, étions d'avis de laisser le système RMS tel qu'il était, sans gros développement matériel ou logiciel. Il remplit toujours bien son rôle. Mais l'avenir des systèmes de visio semble désormais étroitement lié (à l'IN2P3 comme ailleurs) à leur intégration dans les systèmes existants et futurs, de travail collaboratif, incluant mail, messagerie instantanée, agenda partagé, planning, mots-clés : Indico, Zimbra, etc.

A la marge de la visio-conférence "réunion de groupe", il y a par ailleurs une ancienne et forte demande, pour des capacités de diffusion/enregistrement automatique de conférences ou de "cours", et autres formations. Elle ne peut être satisfaite actuellement dans de bonnes conditions, ni par ce qui existe dans RMS, ni par les capacités, limitées en "manpower", de la cellule webcast in2p3, même récemment renforcée. Un autre sujet porteur sera certainement le développement de salle de téléprésence HD, voire holographiques, hors de portée des budgets actuels

A la veille de la transmission de responsabilité à Renater, qu'il me soit permis ici d'évoquer ici le rôle important de "gourou" et prosélyte de la visioconférence tenu à l'IN2P3 et dans la communauté par Christian Helft (LAL), tout au long de ces années, depuis les premières expériences des années 90. Associé à notre action, ici au CC-IN2P3, il a permis à l'IN2P3 d'acquiescer un statut de référence dans la visioconférence académique.

Gerard DREVON (CC-IN2P3)

Dès 1999, sur les traces des premières expériences de visio-conférences multicast sur BETEL, une politique volontariste de l'IN2P3 a décidé de déployer une infrastructure de visioconférence dans ses laboratoires. Rejointe par la DSI du CNRS qui a déployé ses propres matériels dans ses DR dès 2000. La nécessité est alors vite apparue de donner aux composantes de ces deux infrastructures, voisines et parallèles, des moyens d'interconnexion entre elles, et avec le reste du monde, mais surtout à plusieurs points simultanément. La conférence multisites devenait ainsi possible avec l'achat en commun, et avec l'INSERM, puis l'INRA, du premier MCU, géré pour la communauté de ces EPST, par le CC-IN2P3.

Un logiciel de gestion RMS, entièrement développé au CC, par Daniel Charnay, a permis ensuite de le gérer au mieux de ses possibilités. Lequel s'est adapté si bien au second MCU acheté (le CERN ayant rejoint le "consortium"), que RMS est devenu l'acronyme du service de visioconférence. Ses qualités ajoutées aux performances remarquables de stabilité opérationnelle du MCU Codian ont permis l'extension et l'ouverture de ce service, bien au delà des EPST de départ, devenant de fait, au cours des six dernières années, le premier service pour l'ensemble de la communauté recherche-éducation publique française.

Cet objectif non déclaré, une fois atteint, a permis de cerner les limites, en moyens humains et en responsabilité, de l'infrastructure actuelle, et a amené ses gestionnaires, personnels du CC-IN2P3, à engager avec Renater des négociations dès 2010. Renater, opérateur d'infrastructure que nous connaissions bien par ailleurs au Centre de Calcul, nous semblait en effet le partenaire idéal pour prendre en charge, le devenir de celle-ci.

Pour ce transfert de responsabilité, opération de longue haleine, une convention en réglant les conditions a été finalisée entre le CC-IN2P3, le CNRS et Renater fin mai 2012.

L'annonce, un peu prématurée sans doute, de Renater aux dernières JRES fin 2011 avait d'ailleurs suscité quelques





n°21  
Juillet  
2012

# La lettre IN2P3 Informatique

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



## L'IN2P3 officialise ses relations avec Israël



Les membres de la délégation et leurs partenaires de l'Université de Tel-Aviv. © CC-IN2P3 / CNRS

Dans le cadre du **Projet International de Coopération Scientifique** mis en place cette année entre l'IN2P3 et l'Université de Tel-Aviv, une délégation s'est rendue fin juin en Israël.

La délégation était composée de **Dominique Boutigny**, directeur du CC-IN2P3, **Mattieu Puel**, administrateur système en charge du déploiement de l'infrastructure cloud au CC-IN2P3, **Loïc Tortay**, expert stockage, **Gaëlle Shifrin**,

**responsable des partenariats et de la communication**, et **Frédéric Suter**, chercheur en informatique, chargé de la liaison entre les communautés de recherche et de production.

Cette mission avait pour objectif la poursuite des relations avec Israël, amorcées l'année dernière de façon informelle avec la visite des membres du département de physique de l'Université de Tel-Aviv.

Cette année, cette collaboration a été formalisée par un PICS dont les crédits alloués ont permis de financer une nouvelle visite sur place. Au programme : une réunion avec les partenaires de l'Université de Tel-Aviv avec qui des discussions se sont engagées sur le cloud computing et une visite de l'Israel Institute of Technology (Technion), situé à Haïfa (nord du pays). L'objectif de cette visite était de découvrir ce prestigieux institut et d'identifier des pistes de collaborations de recherche dans les domaines du cloud computing, storage cloud, etc.

Enfin, la délégation s'est également rendue chez Exanet, société récemment acquise par DELL, et au laboratoire de recherche IBM où les discussions ont particulièrement portées sur l'optimisation des solutions de virtualisation et sur le logiciel OpenStack, prochainement mis en production au CC-IN2P3.

**Gaëlle SHIFRIN**





n°21  
Juillet  
2012

# La lettre IN2P3 Informatique

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



## Le CC-IN2P3 s'associe à DELL pour le développement de solutions



(De g. à dr.) Dominique Boutigny, directeur du CC-IN2P3, Amandine Lhéritier, adjointe au délégué régional au CNRS Rhône-Auvergne et Thierry De Boishevallier, Directeur du Secteur Public de Dell France, lors de la signature de la convention. © CC-IN2P3 / CNRS

**Le CNRS, via le Centre de Calcul de l'IN2P3, vient de signer une convention de partenariat d'une durée de trois ans avec la société DELL, visant au renforcement de l'expertise et au développement de solutions autour du traitement massif de données et des moyens de calcul pour la recherche. Ce partenariat public – privé prend la forme d'une mise à disposition d'une plateforme DELL à la pointe de la technologie. Cette plateforme sera mise à rude épreuve par les ingénieurs du CC-IN2P3 puisqu'elle leur permettra de mener des tests sur des cas d'utilisation concrets et d'évaluer la pertinence de solutions informatiques innovantes non encore mises sur le marché.**

Ce partenariat et les infrastructures mises à la disposition du Centre de Calcul par Dell auront pour objectif d'évaluer et démontrer, au moyen de tests sur des cas d'utilisation concrets, la pertinence des technologies futures ou de concevoir des solutions informatiques HPC innovantes. Les résultats de ces travaux bénéficieront à la communauté des utilisateurs HPC au travers de livres blancs, bulletins technologiques et benchmarks, par l'adoption des nouvelles technologies de calcul et de stockage.

En raison de ses activités, le CC-IN2P3 utilise et déploie les technologies de traitement de données les plus en pointe et entretient une veille technologique permanente afin de pouvoir offrir à ses utilisateurs des architectures innovantes et performantes. C'est dans ce cadre, et afin de développer des échanges techniques de très haut niveau, que le CC-IN2P3 renforce aujourd'hui et au travers de ce partenariat, ses liens avec Dell, fournisseur global de solutions informatiques et partenaire historique du secteur de la recherche.

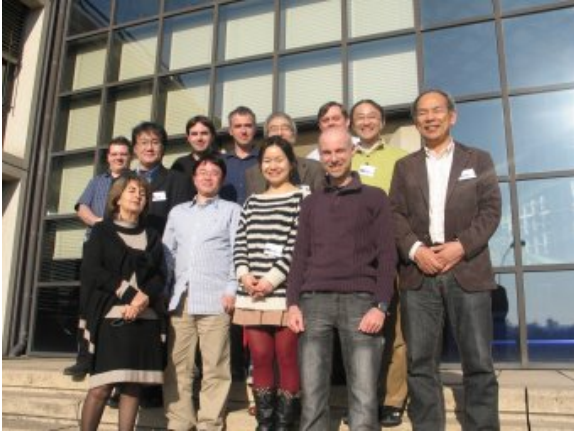
Le partenariat avec le CC-IN2P3 prolonge l'implication de Dell dans le secteur du calcul intensif. L'échange technique et le partage de connaissances entre les deux parties autour des solutions de traitement massif de données, permettra de mieux appréhender l'impact de l'évolution des technologies matérielles et logicielles et, ainsi, de construire les solutions optimisées en termes de performances, de rendement énergétique ou de simplicité de mise en œuvre.

Ce partenariat prometteur devrait ainsi aboutir prochainement à la mise en œuvre d'architectures novatrices en matière de calcul et de traitement massif de données, qui seront mises à la disposition des chercheurs pour accélérer leurs travaux.

**GAELE SHIFRIN**



## Les LIA avec l'Asie se portent bien



Les participants au colloque FJPPL de mars 2012 au CC-IN2P3.  
(© CC-IN2P3/CNRS)

Depuis 5 ans, l'IN2P3 a entrepris sous l'impulsion de François Le Diberder de développer la collaboration avec l'Asie. Plusieurs Laboratoires Internationaux Associés sont nés de cette volonté de tisser des liens plus étroits avec les établissements de recherche asiatiques. Ces laboratoires sans murs sont co-dirigés et co-financés respectivement par les pays partenaires et proposent des cadres propices à la croissance de collaborations.

Quatre Laboratoires Internationaux Associés ont ainsi vu le jour depuis 2006 : le FCPPL avec la Chine, le FJPPL avec le Japon, le FKPPL avec la Corée et le FVPPL avec le Vietnam. Aujourd'hui, près d'une centaine de projets scientifiques impliquent plusieurs centaines de chercheurs des établissements de recherche français et asiatiques impliqués dans ces LIAs dans les domaines de la physique des 2 infinis et des disciplines connexes, dont la grille informatique.

La grille a été tout de suite identifiée comme un thème important de tous les LIAs. En effet, le partage d'une infrastructure commune pour la modélisation, le stockage et l'analyse des données constitue un puissant vecteur de développement de collaborations dans tous les domaines scientifiques. Avec chacun des 4 pays, l'activité sur la grille s'est développée d'une façon différente :

■ avec la Chine, l'activité a démarré dans le cadre du projet européen EUChinaGrid et se poursuit dans le cadre d'une collaboration avec l'Institut de Physique des Hautes Energies de

Beijing autour de la mise en œuvre de l'infrastructure LCG pour l'expérience ATLAS et plus généralement de la gestion de données distribuées. Cette année a également vu l'émergence d'un nouveau projet de développement d'un outil de partage de données scientifiques dans le cloud.

■ avec le Japon, la collaboration avec le KEK s'est développée autour de la gestion de données distribuées et du développement de standards pour l'interopérabilité.

■ avec la Corée, la collaboration avec le KISTI a débouché sur la mise en place d'une première organisation virtuelle franco-asiatique utilisée par plusieurs communautés scientifiques (physique théorique, physique médicale, biochimie).

■ avec le Vietnam, l'activité s'est concentrée sur l'organisation de formations bi-annuelles à la grille et au cloud et le déploiement d'applications scientifiques par des chercheurs vietnamiens (sismologie, recherche de nouveaux médicaments).

Ces activités ont débouché sur un projet global cohérent structuré autour de l'Organisation Virtuelle France-Asie qui intègre aujourd'hui des ressources très importantes de calcul et de stockage en France (CC-IN2P3), en Chine (IHEP), en Corée (KISTI) et au Japon (KEK), offre des services avancés de gestion de données distribuées basés sur iRODS et bientôt un service DIRAC pour la soumission de tâches. Cette Organisation Virtuelle est opérationnelle et ouverte aux chercheurs des 3 LIAs mais aussi aux chercheurs vietnamiens. Elle constitue un environnement privilégié pour développer aussi des collaborations trilatérales impliquant les équipes françaises et des équipes de plusieurs pays asiatiques. Plusieurs projets communs à plusieurs LIAs ont ainsi été présentés au premier colloque commun des LIAs FJPPL et FKPPL avec respectivement le Japon et la Corée à Clermont-Ferrand au mois de Mai 2012.

Une retombée inattendue de l'excellente collaboration entre les LIA concerne le projet de création d'un système de sauvegarde et de reprise d'activité du centre informatique de KEK en cas de désastre majeur. Celui-ci consisterait à sauvegarder les données les plus critiques dans des centres en dehors du Japon de façon à pouvoir reconstituer le plus rapidement possible un système informatique fonctionnel après la survenue d'une catastrophe.

Point de contact pour la VO France-Asie : Yonny Cardenas

Vincent BRETON



n°21  
Juillet  
2012

# La lettre IN2P3 Informatique

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



## Appel à proposition pour les Journées Informatique 2012

Le **réseau des Informaticiens IN2P3 IRFU** organise tous les deux ans des **Journées Informatique** destinées à permettre aux informaticiens de l'IN2P3 et de l'IRFU de se rencontrer, échanger leurs expériences et réfléchir ensemble aux évolutions futures.

Les 8èmes Journées Informatique de l'IN2P3 et de l'IRFU se tiendront à **La Londe les Maures du 22 au 25 octobre 2012**.

Nous avons le plaisir de vous annoncer l'ouverture de l'appel à proposition sur le site des JI2012 : ***cliquer ici***.

La date limite de soumission est fixée au **1er septembre 2012**.

Vous pouvez aussi interagir avec le comité de programme par l'intermédiaire de l'adresse mail : [ji-program@in2p3.fr](mailto:ji-program@in2p3.fr).

Vous pouvez proposer des présentations et/ou des posters sur n'importe quel sujet même en dehors des thèmes des sessions déjà prévues (voir le site), thèmes qui ne sont cependant ni définitifs ni exhaustifs.

En espérant vous retrouver à Lalonde les Maures et en vous remerciant d'avance de votre collaboration active,

Le comité de programme des Journées Informatique 2012 :

Valérie Givaudan,  
Solveig Albrand,  
Laure Amélie Couturier,  
Eric Fede,  
Andrea Formica,  
Laurent Garnier,  
Pierre Girard,  
Pierre-François Honoré,  
Éric Legay,  
Maude Lejeune,  
Pierrick Micout,  
Jean-René Rouet,  
Françoise Virieux



## 1-3 oct. 2012 - Journées scientifiques mésocentres et France Grilles

Le **GIS France Grilles** et le **Groupe Calcul** organisent pour la première fois en coordination les rencontres scientifiques et les journées annuelles des mésocentres, à Paris, du 1er au 3 octobre 2012.

L'objectif de ces rencontres est double :

- réaliser un bilan d'étape des évolutions marquantes dans le paysage des mésocentres d'une part, des infrastructures de grilles et de cloud d'autre part.
- présenter des travaux scientifiques, dans toutes les disciplines, réalisés grâce au soutien des mésocentres, des infrastructures de grilles de calcul ou de cloud.

Vos contributions sont attendues sous forme de résumés étendus de trois à quatre pages en français, avec un résumé en anglais, la légende des figures en français et en anglais. Les dix meilleures contributions feront l'objet d'une présentation orale en session plénière. La contribution de type « poster » la plus marquante sera récompensée par un prix.

Pour déposer une contribution, veuillez vous inscrire sur ce site : <http://mesogrilles2012.sciencesconf.org>.

Les contributions doivent couvrir les points suivants : motivations, travail scientifique et technique, originalité, résultats, impact, choix des ressources de calcul utilisées pour ces travaux ainsi que bibliographie des articles liés à ces travaux.

Les thèmes privilégiés sont :

- Travaux scientifiques nécessitant l'usage de plusieurs infrastructures.
- Thématiques scientifiques émergentes : quels challenges pour demain ?
- Nouvelles applications, portails et architectures.
- Retours d'expérience d'utilisation de cloud.
- Quelles applications pour quelles infrastructures ? Adéquation aux besoins en calcul, stockage et visualisation.
- Impact de la mutualisation des ressources.
- Émergence de nouvelles communautés d'utilisateurs...

Les résumés publiés seront sélectionnés par un comité de lecture pluridisciplinaire, les comptes-rendus des journées feront l'objet d'une diffusion via la base de données HAL et seront répertoriés avec un numéro ISBN. Ils seront également disponibles ainsi que les présentations et les posters depuis le site web de France Grilles et le site du Groupe Calcul.

**La date limite ferme de soumission est le 12 juillet 2012 avant 13 heures.**

Les auteurs seront notifiés le 29 août 2012 de la sélection de leur contribution pour une présentation orale ou pour un poster ainsi que pour une publication dans les comptes-rendus.



## **EGI Technical Forum - sept. 2012**

L'inscription pour le Technical Forum, qui aura lieu à Prague du 17 à 21 septembre 2012, vient d'être annoncée : <http://tf2012.egi.eu/Registration>.

Le programme global des sessions est disponible à cette adresse : <https://indico.egi.eu/indico/conferenceTimeTable.py?confId=1019#20120917>.

Voici l'appel formulé par les organisateurs :

"The organisers of the EGI Technical Forum 2012 would like to invite the EGI community and everyone with an interest in distributed computing and its applications to research, to register online for the event.

The event will take place at the Clarion Congress Hotel in Prague, Czech Republic (17-21 September 2012), hosted by EGI.eu in partnership with CESNET, the consortium of Czech universities and the Czech Academy of Sciences that represents the country in the EGI Council.

The key goal of the Technical Forum is to reflect on the considerable progress made by the collaboration so far and looking forwards to the challenges of public availability, reusability and transparency of scientific methods and data in Open Science.

The aim is to continue the development of an open and sustainable EGI ecosystem that will support Open Science in the digital European Research Area."