



A LA  
UNE

## Top départ pour la salle informatique « vallée » de Virtual Data !



Virtual Data est le groupe de travail du labex P2IO (réseau de tous les laboratoires du Sud de l'île de France impliqués dans la physique de l'infiniment petit, de l'infiniment grand et de l'étude des conditions d'apparition de la vie) dédié à la réflexion sur l'évolution des besoins informatiques des laboratoires du labex.

Après des mois de réhabilitation d'un bâtiment de la vallée de

l'université Paris-Sud, une longue phase de planification et de très nombreux allers et retours entre les salles informatiques historiques des laboratoires P2IO et la nouvelle salle « vallée » de VirtualData, un grand nombre de serveurs ont rejoint leurs nouveaux emplacements, concluant ainsi brillamment le déménagement d'une grande partie de l'infrastructure informatique de nos laboratoires de la vallée ! Ces machines utilisent maintenant presque totalement 18 racks sur les 30 prévus (pour une puissance totale de 400 kW) installés dans les 100 m<sup>2</sup> de la première tranche de « Virtual Data », un projet porté par le Labex, mené à bien, dans les temps et dans le budget prévu, par un groupe de travail d'informaticiens des différents laboratoires de P2IO avec l'aide du service infrastructure du LAL.

Ce projet a pour but de regrouper les salles informatiques de huit laboratoires du campus au sein d'une plateforme commune, répartie sur une salle (...)

**Interview**  
"Assurer notre excellence dans le domaine du calcul pour servir nos communautés scientifiques reste une priorité"



*Ursula Bassler, directrice adjointe scientifique de l'IN2P3 en charge de la physique des particules et du calcul*

[lire la suite](#)

### ■ Stockage



#### Le service iRODS à l'IN2P3

Depuis plus d'une décennie, l'IN2P3 comme beaucoup d'autres acteurs du monde scientifique, est le témoin d'une explosion des quantités de données produites par les projets scientifiques qu'il soutient. La gestion de ces masses de données est un enjeu de première importance : il faut à la fois assurer la disponibilité et la pérennité de l'accès à celles-ci. Ces deux tâches sont rendues particulièrement délicates dans un environnement de travail qui est complexe à plus d'un titre : • Les collaborations scientifiques sont généralement dispersées géographiquement, ce qui nécessite de pouvoir partager, diffuser et gérer de façon efficace et sécurisée, les données réelles et les simulations numériques produites par les expériences, et ce à travers des sites de traitement eux aussi distants les uns des autres. • Les supports et systèmes de stockage employés peuvent différer grandement d'un centre de traitement à l'autre. A cette hétérogénéité, viennent s'ajouter l'évolution

### ■ Développement



#### Pyrame, un framework de prototypage rapide pour systèmes online

Pyrame est un framework léger développé au Laboratoire Leprince-Ringuet (IN2P3/École Polytechnique) qui permet de prototyper rapidement et facilement des systèmes online de complexité arbitraire. Pyrame est basé sur une série de modules, gérant chacun un matériel spécifique ou une fonctionnalité système. Ils communiquent entre eux au moyen de connections réseau standards (TCP) et d'un protocole très simple et ouvert, basé sur XML. Cette structuration permet un déploiement très facile des modules dans un environnement réseau, même sur des sites différents. Le module de commande est une machine virtuelle Python intégrant un serveur réseau et un parseur XML. C'est un composant passif qui exécute ses contrôles-commandes à chaque fois qu'il reçoit une requête. Toute la complexité réseau et protocolaire est masquée au développeur qui peut se concentrer sur le contrôle-commande proprement dit. Il dispose de deux primitives qui permettent de renvoyer un résultat ou de

### ■ Collaboratif



#### Accès nomade à Internet des personnels IN2P3 via « eduroam.in2p3.fr »

Le service eduroam vise à offrir un accès sans fil sécurisé à l'Internet pour les personnels des établissements d'enseignement supérieur et de recherche lors de leurs déplacements en France et à l'étranger. Les utilisateurs d'un établissement membre du projet disposent ainsi d'un accès sécurisé à l'Internet dans tous les autres établissements membres et ceci en utilisant les mêmes identifiants. Le déploiement et l'utilisation de ce service est subordonné à l'acceptation de la charte eduroam et au respect des spécifications techniques. « eduroam » au Centre de Calcul et à l'IN2P3 Le CC-IN2P3 a fait le choix en 2012 de déployer un point d'accès eduroam et d'en faire un des deux moyens d'accès sans fil à Internet pour ses visiteurs. Ce choix était également fortement motivé par la volonté d'offrir une large possibilité de connexions à Internet aux personnels du CC-IN2P3 en déplacement. L'expérience fut concluante à

### ■ Préservation de données



#### PREDON : La préservation et l'exploitation des données scientifiques à long terme

Les données scientifiques, dont la croissance quantitative est illustrée en figure 1, sont produites avec des investissements humain et financier importants et sont de plus en plus complexes, souvent uniques et avec un potentiel qui dépasse le périmètre de recherche initial. Par conséquent, la préservation de ces données à long terme peut conduire à une augmentation significative du retour sur l'investissement ainsi qu'à la mise en place de nouvelles opportunités scientifiques à bas coût à travers la réutilisation de ces données. Du point de vue technologique, la conservation

### ■ Agenda

#### EGI Community Forum, 19-23 mai - Helsinki

Le prochain Community Forum d'EGI (European Grid Infrastructure) se tiendra du 19 au 23 mai (...)

[en savoir plus](#)

#### HEPIX, 19-23 mai - LAPP

The Local Organizing Committee is pleased to announce that the registration and abstract (...)

[en savoir plus](#)

#### [normanDEV] Nouveau réseau des développeurs de Normandie

C'est avec grand plaisir que nous vous annonçons la création du réseau régional normanDEV, le réseau (...)

[en savoir plus](#)

#### HEP Software Collaboration Meeting, 3-4 avril - CERN Council Chamber

La réunion de collaboration 'HEP Software' aura lieu du jeudi 3 avril 9 heures au vendredi 4 avril (...)

[en savoir plus](#)

#### JoSy "Cloud privé dans l'enseignement et la recherche", 19-20 mai - Strasbourg

Resinfo renoue avec les JoSy, ces journées consacrées à un thème intéressant les administrateurs (...)

[en savoir plus](#)

#### Ecole informatique de l'IN2P3 2014, 1-5 juillet - Lyon

L'école informatique de l'IN2P3 2014 a pour thème : "Maîtriser le cloud". Elle s'adresse aux (...)

[en savoir plus](#)

### ■ Archives

continue des media de stockage (ce qui induit des migrations matérielles) et les évolutions technologiques au niveau logiciel (ce qui implique (...))

[lire la suite](#)

faire appel à un autre module pour un fonctionnement hiérarchique. Plusieurs (...)

[lire la suite](#)

la fois localement pour les visiteurs mais aussi pour les nomades du Centre de Calcul de l'IN2P3 en (...)

[lire la suite](#)

des données scientifiques pose un problème nouveau car il n'existe pas de système standardisé capable de pérenniser un ensemble complexe de données, logiciels et métadonnées. En plus, la démarche ne se résume pas à son côté technologique : une méthodologie cohérente doit accompagner tout le processus depuis l'acquisition jusqu'au stockage à long terme, en passant par l'exploitation intensive et par la mise à disposition adéquate aux (...)

[lire la suite](#)

### **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).



© 2014 CCIN2P3



### **Equipe**

**Directeur de la publication** : Alain Fuchs.

**Responsables éditoriaux** : Giovanni Lamanna et Pierre-Etienne Macchi.

**Comité de rédaction** : Virginie Delebarre Dutruel, Dirk Hoffmann et Gaëlle Shifrin.



n°27  
Avril  
2014

# La lettre IN2P3 Informatique

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

A LA  
UNE

## Top départ pour la salle informatique « vallée » de Virtual Data !

Infrastructure



Virtual Data est le groupe de travail du labex P2IO (réseau de tous les laboratoires du Sud de l'île de France impliqués dans la physique de l'infiniment petit, de l'infiniment grand et de l'étude des conditions d'apparition de la vie) dédié à la réflexion sur l'évolution des besoins informatiques des laboratoires du labex. Après des mois de réhabilitation d'un bâtiment de la vallée de l'université Paris-Sud, une longue phase de planification et de très nombreux allers et retours entre les salles informatiques historiques des laboratoires P2IO et la nouvelle salle « vallée » de VirtualData, un grand nombre de serveurs ont rejoint leurs nouveaux emplacements, concluant ainsi brillamment le déménagement d'une grande partie de l'infrastructure informatique de nos laboratoires de la vallée ! Ces machines utilisent maintenant presque totalement 18 racks sur les 30 prévus (pour une puissance totale de 400 kW) installés dans les 100 m<sup>2</sup> de la première tranche de « Virtual Data », un projet porté par le Labex, mené à bien, dans les temps et dans le budget prévu, par un groupe de travail d'informaticiens des différents laboratoires de P2IO avec l'aide du service infrastructure du LAL.

Ce projet a pour but de regrouper les salles informatiques de huit laboratoires du campus au sein d'une plateforme commune, répartie sur une salle dans la vallée et une salle sur le plateau, pour optimiser les coûts d'infrastructure et de fonctionnement (bâtiment, système de climatisation, électricité, etc...) tout en

limitant l'impact environnemental de l'installation et en améliorant la disponibilité des services fournis. Ainsi, les racks utilisés dans la nouvelle salle « vallée » sont réfrigérés avec des portes froides, actives ou passives selon la densité de machines en service à l'intérieur. Deux groupes froids installés derrière le bâtiment alimentent deux pompes qui maintiennent à la bonne pression l'eau à 12 degrés à l'entrée des portes des racks. Ce système redondant (un seul groupe et une seule pompe suffisent pour refroidir l'ensemble de la salle) assure un fonctionnement 24h/24 et 7j/7, et est couplé à un système de « free cooling » permettant d'utiliser la température extérieure, lorsque celle-ci est suffisamment basse, pour refroidir l'eau sans utiliser les groupes. L'ensemble de la salle est monitoré (température des racks, température de la salle, pression d'eau, groupe froids, consommation électrique, fuite d'eau, ...) pour une surveillance optimale à distance. Enfin, côté plateau, l'Ecole Polytechnique est en train de rénover sa salle principale. P2IO devrait y disposer d'un espace réservé à l'automne, ce qui permettra de mettre en place entre les deux sites la redondance nécessaire aux services les plus critiques.

Et maintenant ? Ce projet ambitieux comporte trois phases. Si le financement nécessaire est débloqué, la salle vallée couvrira au final une surface de 220 m<sup>2</sup> et offrira une puissance totale de 1,5 MW. L'Université Paris-Sud ne disposant pas encore d'une telle infrastructure, plusieurs laboratoires extérieurs au Labex se sont montrés intéressés pour installer une partie de leurs ressources informatiques dans cette salle. Enfin, il convient de rappeler qu'un autre point fort de ce projet est d'avoir initié une collaboration entre des informaticiens de tous les laboratoires P2IO (CPHT, CSNSM, IAS, IMNC, IPHT, IPN, IRFU, IRSD, LAL, LLR, LPT, SERMA), ce qui permet des échanges d'expertise et de jeter les bases d'un travail en réseau. C'est maintenant à ces équipes de s'approprier cette nouvelle plateforme, de la faire vivre et de renforcer l'intérêt qu'elle suscite au sein de P2IO et sur le campus.

Valérie GIVAUDAN (LAL)



n°27  
Avril  
2014

# La lettre IN2P3 Informatique

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



## "Assurer notre excellence dans le domaine du calcul pour servir nos communautés scientifiques reste une priorité"

**Ursula Bassler, directrice adjointe scientifique de l'IN2P3 en charge de la physique des particules et du calcul**



**- Ursula, pourriez-vous vous présenter en quelques mots ?**

Comme beaucoup d'Allemandes, je suis arrivée en France comme jeune fille au-pair pour une année à l'étranger après le bac. Mais j'ai voulu rester et j'ai fait du coup mes études supérieures à l'Université Pierre et Marie Curie. Après un doctorat en physique des particules, j'ai poursuivi mes recherches au LPNHE (Laboratoire de Physique nucléaire et des Hautes Energies) à Paris sur la structure du proton auprès de HERA (Hadron Electron Ring Anlage) à DESY (Deutsches Elektronen Synchrotron), Hambourg, l'unique collisionneur électron-proton jamais construit. En 1998, j'ai fait partie d'un groupe de chercheurs de l'IN2P3 qui ont rejoint l'expérience DØ auprès du Tevatron à FNAL (Fermi National Accelerator Laboratory) près de Chicago. Ce collisionneur proton-antiproton qui produisait les collisions les plus énergétiques avant que le LHC n'en prenne la relève. A DØ, j'ai travaillé sur le calorimètre, en particulier le système de calibration en ligne et la physique du top. En 2005, j'ai été nommée directrice adjointe au LPNHE et, entre 2007 et 2013, chef du service de physique des particules à l'Irfu au CEA. Depuis janvier, j'ai donc pris mes nouvelles fonctions en tant que DAS pour la physique des particules et le calcul à l'IN2P3.

**- Comment s'organisent les responsabilités du calcul entre vous et Giovanni Lamanna ?** Giovanni fait tout le travail, et moi je dis oui ☺. Plus sérieusement, c'est ensemble avec Giovanni que nous réfléchissons à la stratégie à adopter pour le calcul à l'IN2P3, son insertion dans le contexte national et international et les différentes évolutions possibles. Pour cela, nous sommes bien sûr en contact très étroit, avant tout avec Pierre-Etienne Macchi, le directeur du Centre de Calcul, mais aussi avec Vincent Breton, le directeur de l'Institut de Grilles et du Cloud et Fairouz Malek, responsable de LCG-France. De part son expérience, Giovanni a déjà tous les liens avec les personnes sur le terrain, de mon côté j'apporte les contacts institutionnels.

**- Quelles sont les priorités en matière de calcul pour les prochaines années ?**

La première priorité est bien sûr d'assurer le calcul pour le LHC, et cela veut dire également de se préoccuper des nouvelles technologies et des possibilités du calcul parallèle dans nos domaines de recherche. Ensuite, il y a de nouveaux défis à venir avec le traitement des données de LSST, d'EUCLID et de CTA, pour lesquels les modèles de calcul restent encore à préciser. Ces expériences sont actuellement en train de définir leurs modèles de calcul et leurs besoins. Tout cela va devoir se faire dans un contexte budgétaire qui restera tendu pour les prochaines années, mais dans lequel assurer notre excellence dans le domaine du calcul pour servir nos communautés scientifiques restera une priorité à l'IN2P3. Au-delà, il est important de s'organiser au niveau européen pour pouvoir répondre efficacement aux appels H2020. L'initiative EU-T0 portée par Giovanni, qui a pour but de fédérer les centres de Tier 1 en Europe, s'inscrit dans cet effort. Dans ce sens, nous participerons également à revoir les compétences de chacun des organismes en France, afin d'avoir des structures claires, avec un minimum de déperdition. Dans ce contexte, il est important pour nous d'assurer la place du Centre de calcul comme un des piliers de la stratégie française en matière de calcul.

**PROPOS RECUEILLIS PAR GAELE SHIFRIN**





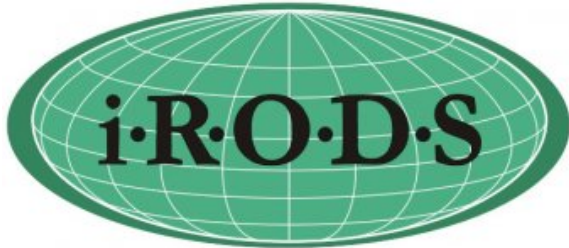
n°27  
Avril  
2014

# La lettre IN2P3 Informatique

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



## Le service iRODS à l'IN2P3



### Integrated Rule-Oriented Data System

Depuis plus d'une décennie, l'IN2P3 comme beaucoup d'autres acteurs du monde scientifique, est le témoin d'une explosion des quantités de données produites par les projets scientifiques qu'il soutient. La gestion de ces masses de données est un enjeu de première importance : il faut à la fois assurer la disponibilité et la pérennité de l'accès à celles-ci. Ces deux tâches sont rendues particulièrement délicates dans un environnement de travail qui est complexe à plus d'un titre :

- Les collaborations scientifiques sont généralement dispersées géographiquement, ce qui nécessite de pouvoir partager, diffuser et gérer de façon efficace et sécurisée, les données réelles et les simulations numériques produites par les expériences, et ce à travers des sites de traitement eux aussi distants les uns des autres.
- Les supports et systèmes de stockage employés peuvent différer grandement d'un centre de traitement à l'autre. A cette hétérogénéité, viennent s'ajouter l'évolution continue des media de stockage (ce qui induit des migrations matérielles) et les évolutions technologiques au niveau logiciel (ce qui implique éventuellement des changements au niveau du nommage des données ou des protocoles d'accès à ces dernières).

#### Qu'est-ce que iRODS ?

De tels environnements nécessitent des intergiciels permettant la gestion et la distribution des données dans un environnement hétérogène, notamment en virtualisant le stockage, c'est-à-dire en masquant la complexité et la diversité des systèmes de stockage sous-jacents tout en les fédérant.

iRODS (pour *integrated Rule based Data System*) développé principalement par le groupe DICE [1] bi-localisé à l'Université de Californie à San Diego et l'Université de Caroline du Nord à Chapel Hill, répond à ce besoin en offrant une vue logique unique des données indépendantes de l'organisation et de l'emplacement physique de ces dernières. De plus, iRODS fournit une palette riche d'aide à la gestion de données (extraction de métadonnées, intégrité des données etc...). De par sa flexibilité, il peut s'interfacer avec un nombre quasi illimité de technologies de stockage (systèmes de stockage de masse, systèmes de fichiers distribués, bases de données relationnelles, Amazon S3, Hadoop etc...).

Mais ce qui rend iRODS particulièrement attractif est son moteur de règles qui n'a pas d'équivalent parmi ses concurrents. Ce moteur de règles permet d'effectuer des tâches complexes au niveau de la gestion des données. Ces politiques de gestion se trouvent déportées du côté serveur : par exemple quand des données sont déposées dans iRODS, des tâches de fond

peuvent être déclenchées automatiquement côté serveur comme par exemple des répliquions sur plusieurs sites, des contrôles d'intégrité des données, des post-traitements sur celles-ci (extraction de métadonnées...) sans action particulière du côté client. Du coup, la politique de gestion des données est virtualisée. Cette virtualisation permet d'assurer une application stricte des règles établies par les utilisateurs, quelle que soit la localisation des données ou de l'application qui accède à iRODS.

#### Le service iRODS du CC-IN2P3

Le Centre de Calcul de l'IN2P3 offre un service iRODS à l'IN2P3 depuis 2008. Ce service est ouvert à l'ensemble des expériences qui souhaitent l'utiliser. Actuellement, 34 groupes dans les domaines de la physique des particules (BaBar, dChooz...), la physique nucléaire (Indra, Fasia...), les astroparticules et l'astrophysique (AMS, Antares, Auger, Virgo, LSST...), les Sciences Humaines et Sociales (Huma-Num) et la biologie, utilisent le service iRODS du CC-IN2P3 pour la gestion et la diffusion de leurs données. Le CC-IN2P3 fournit aussi l'hébergement du catalogue central du service iRODS nouvellement créé de France Grille, ainsi qu'un support aux administrateurs de France Grille sur l'utilisation de cette technologie.

Le service iRODS possède ses propres serveurs de disques et se trouve interfacé avec notre système de stockage de masse HPSS (stockage sur bandes magnétiques) : il gère actuellement plus de 8 Po de données, ce qui en fait le service le plus important en volumétrie recensé au niveau international.

Le service est fédéré avec d'autres services iRODS comme celui du SLAC par exemple. Dans cette perspective, il est aussi tout à fait possible de fédérer des serveurs de stockages disponibles dans les laboratoires avec le service iRODS du CC-IN2P3.

Il est appelé à se développer encore davantage de par son adoption par de nouveaux projets scientifiques.

#### L'implication du CC-IN2P3 dans iRODS

Depuis 2003 et la mise en production de **SRB** à Lyon, le CC-IN2P3 est en étroite contact avec le groupe DICE. Le CC-IN2P3 participe au projet de développement iRODS en tant que collaborateur depuis le démarrage du projet en 2006. Le CC-IN2P3 est aussi très actif au sein de la communauté des utilisateurs et administrateurs iRODS, en fournissant une aide et des conseils sur cette technologie. De par son expérience et son activité, le CC-IN2P3 est considéré comme une référence au sein de cette communauté très diversifiée de par les thématiques scientifiques et technologiques couvertes.

#### Comment utiliser le service iRODS du CC-IN2P3 ?

Pour toutes questions relatives à iRODS et à son utilisation au CC-IN2P3, vous pouvez contacter le **support utilisateur** qui répondra à vos demandes.

Jean-Yves NIEF (CC-IN2P3)

[1] Data Intensive Cyber Environments



## Pyrame, un framework de prototypage rapide pour systèmes online



# PYRAME

Pyrame est un framework léger développé au Laboratoire Leprince-Ringuet (IN2P3/École Polytechnique) qui permet de prototyper rapidement et facilement des systèmes online de complexité arbitraire.

Pyrame est basé sur une série de modules, gérant chacun un matériel spécifique ou une fonctionnalité système. Ils communiquent entre eux au moyen de connections réseau standards (TCP) et d'un protocole très simple et ouvert, basé sur XML. Cette structuration permet un déploiement très facile des modules dans un environnement réseau, même sur des sites différents.

Le module de commande est une machine virtuelle Python intégrant un serveur réseau et un parseur XML. C'est un composant passif qui exécute ses contrôles-commandes à chaque fois qu'il reçoit une requête. Toute la complexité réseau et protocolaire est masquée au développeur qui peut se concentrer sur le contrôle-commande proprement dit. Il dispose de deux primitives qui permettent de renvoyer un résultat ou de faire appel à un autre module pour un fonctionnement hiérarchique.

Plusieurs matériels, courants sur nos bancs de tests, sont déjà implémentés : alimentations basses et hautes tensions (Agilent, CAEN, Hameg...), générateur de fonctions (Agilent), table de déplacement (Newport), adaptateurs de bus (Prologix, Cometh...). Pyrame permet d'unifier l'accès à ces différents matériels de génération et de technologies différentes.

Une chaîne d'acquisition générique à haut débit est également proposée. Elle permet d'acquérir, de décoder et de dispatcher les données en temps réel. Les médias actuellement supportés sont Ethernet et UDP.

Afin d'intégrer ces contrôles-commandes dans des systèmes existants, nous fournissons des bindings (interfaces) pour les langages de programmation courants (C, C++, Python, R), et il est très facile d'en implémenter de nouveaux. Il suffit d'un support socket TCP et d'un parseur XML basique. Ces bindings permettent d'interfacer Pyrame au SCADAs courants (Tango, OPC-UA, Xdaq...) facilitant leur intégration dans des projets existants.

Pour les petits bancs, l'achat d'un ordinateur peut se révéler superflu, c'est pourquoi nous avons porté notre framework sur de petites plateformes embarquées, en particulier le Raspberry-Pi et l'Arduino. Cela permet d'avoir un contrôle-commande évolué à moindre coût.

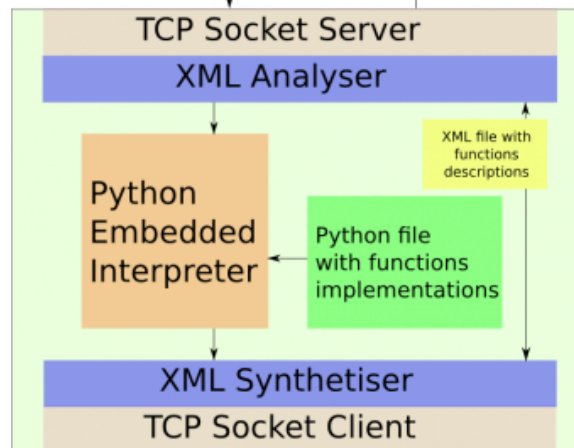
Le framework inclut également des services centralisés. Il propose un module permettant de partager des variables et leurs valeurs entre tous les modules du système. Un autre permet de collecter des statistiques parmi les différents modules et de les retransmettre (par exemple au GUIs) avec un timing adapté. Un autre module permet de repérer des signaux particuliers par leur nom afin de pouvoir activer les générateurs correspondants sans avoir à connaître leur adresse.

On peut, en interfaçant les modules entre eux, construire des systèmes complexes. Nous avons ainsi écrit le contrôle-commande du SiW-Ecal/ILD en agrégeant 25 modules entre eux. Naturellement, la configuration d'un tel système est délicate et c'est pourquoi nous proposons dans Pyrame un système de configuration globale générique au moyen de fichiers XML.

Sur les bancs de tests pour l'électronique HEP, c'est généralement Labview qui est utilisé. Cet outil présente des avantages (programmation graphique, importante bibliothèque de composants) et des inconvénients (maintenabilité, stabilité, interfaçage). Pyrame se positionne sur le même créneau en apportant des solutions à ces inconvénients : c'est un outil très stable. La programmation Python est à la fois simple et maintenable. La simplicité de son protocole garantit une facilité d'interfaçage avec tous les systèmes et tous les langages.

Pour plus d'informations, merci de [visiter cette page internet](#).

1 incoming request      4 outgoing result



2 outgoing request      3 incoming result

Frédéric MAGNIETTE (LLR – IN2P3/CNRS – École Polytechnique)



n°27  
Avril  
2014

# La lettre IN2P3 Informatique

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



## Accès nomade à Internet des personnels IN2P3 via « [eduroam.in2p3.fr](http://eduroam.in2p3.fr) »



Le **service eduroam** vise à offrir un accès sans fil sécurisé à l'Internet pour les personnels des établissements d'enseignement supérieur et de recherche lors de leurs déplacements **en France** et à **l'étranger**. Les utilisateurs d'un établissement membre du projet disposent ainsi d'un accès sécurisé à l'Internet dans tous les autres établissements membres et ceci en utilisant les mêmes identifiants.

Le déploiement et l'utilisation de ce service est subordonné à l'acceptation de **la charte eduroam** et au respect des spécifications techniques.

### « eduroam » au Centre de Calcul et à l'IN2P3

Le CC-IN2P3 a fait le choix en 2012 de déployer un point d'accès eduroam et d'en faire un des deux moyens d'accès sans fil à Internet pour ses visiteurs. Ce choix était également fortement motivé par la volonté d'offrir une large possibilité de connexions à Internet aux personnels du CC-IN2P3 en déplacement. L'expérience fut concluante à la fois localement pour les visiteurs mais aussi pour les nomades du Centre de

Calcul de l'IN2P3 en mission dans les établissements scientifiques français et étrangers.

L'accès à eduroam est conditionné par la fourniture d'un point d'accès sans-fil dans les laboratoires, mais ce déploiement de « hotspots » est une opération qui peut s'avérer lourde.

C'est pourquoi, afin d'alléger la procédure, le groupe réseau du CC-IN2P3 propose un service d'authentification global à l'Institut et un support aux laboratoires souhaitant mettre en œuvre un point d'accès eduroam via une procédure également allégée.

Concrètement pour les personnels IN2P3, tout détenteur d'un certificat délivré par l'autorité de certification CNRS, rattaché à un laboratoire de l'IN2P3, dispose dès aujourd'hui d'un accès au réseau eduroam à travers le domaine « **eduroam.in2p3.fr** », que son laboratoire propose ou non un point d'accès eduroam.

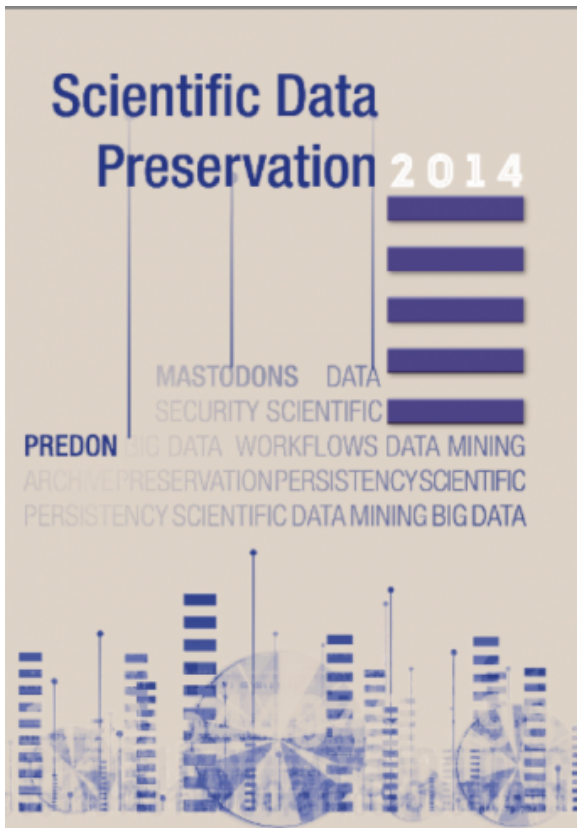
Pour la configuration de cet accès sur son ordinateur portable, chaque utilisateur devra contacter le support informatique de son laboratoire.

Pour des renseignements complémentaires sur les documentations de configuration et le déploiement de points d'accès eduroam dans les laboratoires, les services informatiques devront solliciter le groupe réseau du CC-IN2P3 via **le portail de support**.

**Benoit DELAUNAY (CC-IN2P3)**



## PREDON : La préservation et l'exploitation des données scientifiques à long terme



Le document « Scientific Data Preservation 2014 » disponible sur <http://predon.org> (version papier sur demande, dans la limite des stocks).

Les données scientifiques, dont la croissance quantitative est illustrée en figure 1, sont produites avec des investissements humains et financiers importants et sont de plus en plus complexes, souvent uniques et avec un potentiel qui dépasse le périmètre de recherche initial. Par conséquent, la préservation de ces données à long terme peut conduire à une augmentation significative du retour sur l'investissement ainsi qu'à la mise en place de nouvelles opportunités scientifiques à bas coût à travers la réutilisation de ces données.

Du point de vue technologique, la conservation des données scientifiques pose un problème nouveau car il n'existe pas de système standardisé capable de pérenniser un ensemble complexe de données, logiciels et métadonnées. En plus, la démarche ne se résume pas à son côté technologique : une méthodologie cohérente doit accompagner tout le processus depuis l'acquisition jusqu'au stockage à long terme, en passant par l'exploitation intensive et par la mise à disposition adéquate aux buts scientifiques.

Les aspects technologique et organisationnel de la préservation de données sont très similaires dans plusieurs domaines scientifiques. En effet, la croissance des données scientifiques n'est qu'un aspect de l'explosion des données digitales, un phénomène connu sous le nom de « big data ». L'augmentation de la quantité des données est assortie de la capacité croissante à effectuer des croisements, à sonder et à reclasser

les données pour extraire des connaissances nouvelles. Les « observatoires de données » conduisent à une méthode de recherche basée sur l'utilisation massive des données, appelée aussi le « quatrième paradigme scientifique » (après les méthodes empirique, théorique et computationnelle), un concept introduit par Jim Gray (Microsoft).

Afin de répondre à ce nouveau défi scientifique, la Mission Interdisciplinarité du CNRS a lancé en 2012 le programme MASTODONS, afin de stimuler la coopération des différentes disciplines autour du concept « big data ». Suivant l'initiative d'un petit groupe constitué de chercheurs de l'IN2P3, le projet PREDON a été inclus dans le plan de travail du défi MASTODONS avec la mission de fédérer les initiatives au niveau national dans le domaine de la préservation des données scientifiques. Le projet propose une approche nouvelle basée sur les capacités scientifique, technique et organisationnelle d'unités de recherche, collaborations internationales et grands centres de calcul. Le groupe PREDON est en liaison étroite avec des initiatives similaires au niveau national et international, notamment avec le panel de l'ICFA pour la préservation des données dans la physique des hautes énergies (DPHEP).

Le groupe de travail PREDON a produit récemment un document de synthèse (« Scientific Data Preservation 2014 ») qui résume les contributions des participants aux ateliers de travail. Le document est structuré en trois parties qui reflètent les aspects complémentaires de la préservation des données scientifiques : potentiel scientifique, méthodologie et technologies.

Pour en savoir plus :

- **Préservation des données scientifiques PREDON**
- **Data Preservation in High Energy Physics DPHEP**
- Défi « big data » de la mission interdisciplinarité du CNRS, MASTODONS

- **Description**
- **Colloque 2014**

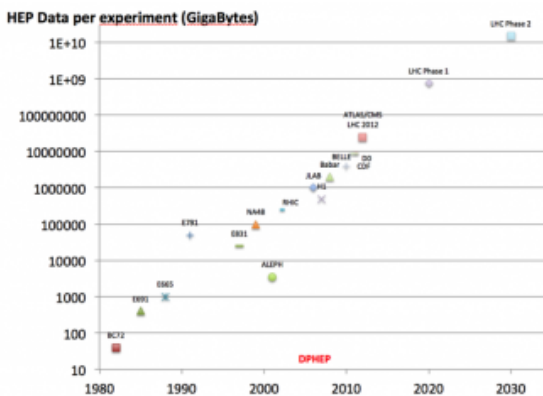


Figure 1. Illustration de l'augmentation exponentielle du volume des données correspondant aux programmes de physique des hautes énergies : un facteur 100 environ tous les 10 ans.

Cristinel DIACONU (CPPM)





n°27  
Avril  
2014

# La lettre IN2P3 Informatique

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



## EGI Community Forum, 19-23 mai - Helsinki

Le prochain Community Forum d'EGI (European Grid Infrastructure) se tiendra du 19 au 23 mai prochain à Helsinki.

Plus d'informations et inscription sur **le site de l'événement**.



## HEPIX, 19-23 mai - LAPP

The Local Organizing Committee is pleased to announce that the registration and abstract submission for the Spring 2014 HEPiX meeting are now open.

This meeting will be in Annecy-le-Vieux (France), hosted by LAPP (Laboratoire d'Annecy-le-Vieux de Physique des Particules, Université de Savoie, CNRS/IN2P3) from the 19th to the 23rd of May.

Links to abstract submission and registration are available from **the web site**.

Please register before the 15th of April to get the early-bird rate. We also strongly recommend that you reserve your hotel as early as possible.

As important, we invite you to submit abstracts to the HEPiX Program Committee. New participation has always been encouraged leading to rich and very interesting programs of HEPiX semi-annual meetings since 1991. See <http://www.hepix.org> for further details.

Looking forward to meeting you all in Annecy-le-Vieux in May !

Frédérique Chollet for the Local Organizing Committee [hepix-spring2014@lapp.in2p3.fr](mailto:hepix-spring2014@lapp.in2p3.fr)



## [normanDEV] Nouveau réseau des développeurs de Normandie

C'est avec grand plaisir que nous vous annonçons la création du réseau régional normanDEV, le réseau des développeurs de Normandie affilié au réseau national DEVLOG (Réseau du DEVeloppement LOGiciel inter-établissements CNRS, INRA, INRIA).

NormanDEV a pour objectif principal de réunir toutes les personnes ayant une activité de développement et de déploiement du logiciel au sein d'établissements d'Enseignement Supérieur et de Recherche en Normandie (ITAs, chercheurs, enseignants, doctorants).

Le but de ce réseau est d'une part d'échanger autour des techniques, des bonnes pratiques, des outils et des méthodes pour favoriser la veille technologique et la communication entre entités, et d'autre part, d'échanger et de partager nos retours d'expériences sur le développement et la gestion de projets logiciels.

Un premier événement d'une journée portant sur l'Agilité est programmée pour la semaine du 7 au 11 avril 2014. Cette journée sera découpée en deux ateliers d'une 1/2 journée chacun : Légo4SCRUM et Coding Dojo (détails à venir ultérieurement).

Vous pouvez suivre toute l'actualité du réseau sur cette page : <http://normandev.cnrs.fr>.

Si vous êtes intéressés par le développement logiciel, nous vous invitons à vous abonner dès maintenant à **la liste de diffusion de normanDEV**.

Si vous souhaitez rejoindre le Comité de Pilotage, merci d'adresser votre demande par courriel à [normandev-comite@services.cnrs.fr](mailto:normandev-comite@services.cnrs.fr).

Enfin, n'hésitez pas à transmettre et à diffuser ce message à votre entourage professionnel.

En espérant bientôt vous compter parmi les membres du réseau.

Cordialement,

Le Comité de Pilotage de normanDEV : *Thierry Chaventré, Jérôme Poincheval, David Tschumperlé, Alexis Lebon, Jean-Luc Lambert, Patrick Taillandier et Aurélie Plisson.*



## HEP Software Collaboration Meeting, 3-4 avril - CERN Council Chamber

La réunion de collaboration 'HEP Software' aura lieu du jeudi 3 avril 9 heures au vendredi 4 avril à midi, dans la Chambre du Conseil du CERN. Le colloque sera tenu en anglais.

**Plus d'informations et inscriptions.**



## **JoSy "Cloud privé dans l'enseignement et la recherche", 19-20 mai - Strasbourg**

Resinfo renoue avec les JoSy, ces journées consacrées à un thème intéressant les administrateurs système et réseau de la communauté Enseignement supérieur / Recherche.

La prochaine JoSy se tiendra à Strasbourg les 19 et 20 mai prochains.

L'IN2P3 étant particulièrement actif dans le domaine de la mise en œuvre, voire du développement de clouds, il n'est pas étonnant de retrouver des orateurs de notre institut dans le programme.

Voici l'annonce de cette journée (répartie sur deux jours...) :

La fédération de réseau Resinfo organise une journée Josy sur le thème du cloud : "Cloud privé dans l'enseignement et la recherche".

Cet événement se déroulera les 19 mai après-midi et 20 mai matin à Strasbourg, dans l'amphithéâtre du bâtiment 25 de l'IPHC.

Il sera diffusé sur le web grâce à la cellule Webcast du Centre de Calcul de l'IN2P3.

Toutes les informations concernant cet évènement sont accessibles via <http://www.resinfo.cnrs.fr/spip.php...>

L'inscription à cet évènement est gratuite mais obligatoire. La date limite des inscriptions est le 5 mai 2014.

Pour ceux qui le souhaitent, un dîner (non pris en charge par RESINFO) sera organisé le soir du 19 mai. Merci de préciser en complétant le formulaire d'inscription, si vous souhaitez participer à ce dîner.

*Le comité d'organisation : le réseau régional X/Stra ([xstra-contact@unistra.fr](mailto:xstra-contact@unistra.fr))*



## **Ecole informatique de l'IN2P3 2014, 1-5 juillet - Lyon**

L'école informatique de l'IN2P3 2014 a pour thème : "Maîtriser le cloud". Elle s'adresse aux développeurs et aux administrateurs système et réseau.

Elle se tiendra du 1er au 5 juillet 2014, à Lyon (Centre Jean Bosco et CC-IN2P3).

**Programme et plus d'informations.**

Les inscriptions se font par le canal habituel de la formation permanente de votre unité.