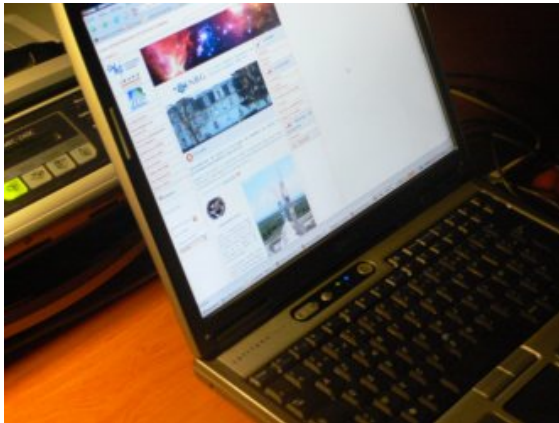




A LA
UNE

Le nomadisme : un véritable enjeu pour le CENBG



Au CENBG, le nomadisme a été considéré comme un enjeu du laboratoire dans sa Politique de Sécurité des Systèmes d'Information (PSSI). Cela signifie, que sans cette possibilité, la capacité du laboratoire à réaliser ses objectifs scientifiques peut être altérée. Il ne s'agit plus de la simple consultation des mails mais de travailler

depuis l'extérieur, comme on le fait depuis l'intérieur.

Si cette évolution est prise en compte dans une PSSI c'est parce que la mise à disposition de cette capacité se heurte évidemment à des aspects de sécurité importants. Il faut donc que le système d'information évolue, tant dans ses aspects techniques que dans sa composante humaine (sensibilisation), pour fournir un cadre de sécurité proportionné qui permettra de favoriser le développement du nomadisme.

Dispositif pour les accès distants

VPN - 1

Depuis deux ans, nous mettons en œuvre un serveur VPN basé sur OpenVPN. L'objectif était de permettre aux utilisateurs de se connecter dans leur environnement de travail habituel (c'est-à-dire dans le sous-réseau de leur groupe de travail), de disposer d'une authentification forte et de la possibilité de filtrer finement les autorisations (jusqu'à l'individu). Ces seules dispositions ne sont pas suffisantes. En effet, le développement du nomadisme (voire ...)

[lire la suite](#)

Interview Fairouz Malek : "La grille, cette utopie technologique dont le LHC avait besoin"



[lire l'interview](#)

Agenda

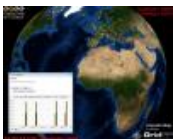
3 au 5 déc. 2008 - FUTURVIEW 2008

Il reste deux semaines avant la clôture de l'appel à contributions. Vous n'avez pas eu l'occasion (...)
[en savoir plus](#)

29 sept. au 2 oct. 2008 - 6e édition des JIs de l'IN2P3 et de l'IRFU

Les Journées Informatique de l'IN2P3 et de l'IRFU sont organisées tous les deux ans et la 6e (...)
[en savoir plus](#)

EGEE



Le Sénégal, nouvelle terre d'escalade pour EGEE

Juillet 2007 marque les débuts de l'extension de la grille EGEE à l'Afrique subsaharienne : le premier nœud EGEE a démarré et a été validé à l'université Cheick Anta Diop de Dakar (Sénégal). Cette réalisation fait suite au colloque de Montpellier en décembre 2007 organisé par la fondation « Partager le savoir » sur le développement d'Internet et des grilles de calcul en Afrique. L'Institut des grilles (IN2P3/CNRS) a été au cœur de la dynamique mise en place à cette occasion, d'une part dans le démarrage d'un partenariat avec l'Afrique du Sud pour démarrer plusieurs sites et organiser une Grid School à l'intention des futurs utilisateurs durant l'automne 2008, d'autre part pour aider le démarrage du site de l'université de Dakar. Le site de Dakar est exemplaire d'une démarche pragmatique portée par des acteurs locaux motivés. L'université de Dakar est la

Développement



TSkim : un outil pour découper les arbres ROOT

Comme de multiples expériences de physique, le Fermi Large Area Telescope (Fermi-LAT, anciennement GLAST-LAT) stocke ses données sous forme d'arbres ROOT. Une activité récurrente des physiciens consiste à établir des critères pour définir quels événements sont intéressants, et à effectuer des coupes sur les arbres ROOT afin d'extraire les données attachées à ces événements particuliers. TSKim a été développé pour les y aider. Initialement un simple couple de scripts PERL et ROOT, l'outil est à présent devenu suffisamment élaboré pour prétendre servir à d'autres expériences que FERMI. Pour un même événement physique, l'utilisateur peut souhaiter récupérer des données issues de plusieurs sources, correspondant à des représentations ou des stades différents dudit événement : simulation de la collision initiale de particules, simulation de la réponse du détecteur, vraies mesures du

Grid Access



JSAGA : pour un accès uniforme à des grilles hétérogènes

De nombreux utilisateurs ont besoin d'accéder de façon uniforme à des grilles d'envergures différentes (locale, régionale, nationale, internationale) et s'adressant à différentes communautés (académiques/industrielles, recherche/production). Ces grilles ont différents atouts et lacunes, et le choix de l'utilisateur de soumettre sur l'une ou l'autre dépend de différents critères comme les caractéristiques et la quantité de ressources matérielles requises, les contraintes de licence des logiciels utilisés, le niveau de confidentialité requis, le degré d'urgence, la charge de ces grilles, etc. A moyen terme, un accès uniforme sera possible entre certaines grilles académiques d'envergure nationale ou internationale, grâce à divers projets d'interopérabilité en cours. Cependant, les approches choisies pour ces projets (modifications côté

Laboratoire



Le CC-IN2P3 s'exporte en Corée

Cette année a été créé le Laboratoire International Associé (LIA) France - Korea Particle Physics Laboratory (FKPPL) par l'IN2P3-CNRS, l'Ecole Polytechnique, l'Université Paris-Sud et l'Université Blaise-Pascal de Clermont-Ferrand, en partenariat avec 7 établissements coréens. L'objectif de ce LIA est de renforcer les collaborations dans les domaines de la physique des hautes énergies et des grilles informatiques. Ce LIA est issu de la volonté du CNRS et de l'IN2P3 de tisser des liens plus étroits avec les établissements de recherche en Asie. Son objectif est de dynamiser les échanges scientifiques entre la France et la Corée dans le domaine de la physique des hautes énergies et des grilles informatiques en s'appuyant sur les collaborations déjà établies entre les équipes de recherche. Le premier comité

Archives

Abonnement

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

Proposer un article

Vous souhaitez proposer un article ? Envoyez un mail à LettreInformatique@in2p3.fr.



© 2008 CCIN2P3

principale université du pays avec 60 000 étudiants dans de nombreuses disciplines et ce projet de site grille connecté à EGEE a été conçu comme une ressource modeste au départ mais permettant l'accès à une ressource grille locale (...)

[lire la suite](#)

détecteur, reconstruction de base, calcul d'un jeu de grandeurs physiques caractéristiques permettant de classifier l'événement, etc. Un arbre contient les données de toute une séquence d'événements, mais pour (...)

[lire la suite](#)

serveurs) ne permettent pas d'envisager une utilisation uniforme de ces grilles et des grilles en dehors de ces projets d'interopérabilité. Dans le cadre du projet IGTM financé par l'ANR, le Centre de Calcul de (...)

[lire la suite](#)

de pilotage du LIA s'est tenu à Séoul le 21 juillet 2008. Celui-ci a passé en revue et s'est prononcé sur le financement des premiers projets scientifiques proposés dans le cadre du FKPL. Les travaux entrepris par les équipes des deux pays dans l'exploitation des grilles pour la recherche de nouveaux médicaments (...)

[lire la suite](#)

Le nomadisme : un véritable enjeu pour le CENBG



Techniquement, le serveur VPN est une machine Linux qui agit comme un garde-barrière dédié aux accès distants. Toute sa configuration est générée par le logiciel Labwall, développé localement depuis 2001 (pour notre garde-barrière/routeur), tant pour sa configuration réseau (VLANs et routage), sa configuration de filtrage (Netfilter) que pour la configuration OpenVPN (un daemon par VLAN). L'utilisateur dispose d'un client OpenVPN qui se connecte au moyen d'un HTTP-Connect sur le serveur VPN. Celui-ci authentifie le certificat de l'utilisateur et interroge le serveur Freeradius pour obtenir une adresse IP (fixe) et déterminer si l'utilisateur est bien autorisé sur le VLAN qu'il demande (figure 1). D'un point de vue fonctionnel, tout se passe comme s'il y avait, sur chaque sous-réseau, un garde-barrière virtuel derrière lequel se trouvent connectés les postes distants. Ce garde-barrière assurant le filtrage entre ces postes et les postes connectés en interne sur le VLAN et, plus généralement, avec tout le reste du réseau (figure 2).

Support à distance

L'éloignement fréquent et de longue durée des utilisateurs pose le problème de leur "solitude" face à un problème et la difficulté pour les administrateurs de les "secourir". Pour résoudre en partie ce problème, nous testons le logiciel UltraVNC. Il permet de prendre le contrôle à distance du poste de l'utilisateur suivant le protocole VNC. Cette prise de contrôle se fait à l'initiative de l'utilisateur à travers une connexion VPN. De ce fait, il n'y a pas de serveur VNC qui tourne en permanence côté utilisateur, évitant ainsi les problèmes de sécurité liés (pas de mot de passe, pas de filtrage nécessaire dans le pare-feu de l'utilisateur). De son côté, l'administrateur reçoit "l'appel" et n'a pas de problème lié à la localisation de l'adresse IP du poste.

Accueil des visiteurs

Les visiteurs ont la possibilité de se connecter sur un sous-réseau dédié. La connexion et l'authentification se font par l'intermédiaire d'un portail captif (Chillispot aujourd'hui qui deviendra prochainement Coova). Les visiteurs de longue durée (> 24 heures) sont enregistrés et ont un mot de passe individuel. Les visiteurs de passage (moins de 24 heures) peuvent se connecter moyennant la connaissance d'un mot de passe générique. En fait, il y a un auto-enregistrement dans le serveur Freeradius qui bloquera automatiquement la connexion 24 heures après ce délai (qui est paramétrable).

Contenu des postes nomades

Il est souvent difficile pour nous (ASR) de mesurer la sensibilité réelle des données présentes sur les portables. Mais, par exemple, de plus en plus d'équipes travaillent en partenariats avec l'industrie privée avec des contrats contenant des clauses de confidentialité. Il serait étonnant que des données liées ne résident pas à un moment ou l'autre sur des portables. On parle de crypter totalement les postes de travail. On peut aussi penser à crypter les clés USB. En effet, compte tenu de leur capacité de plus en plus importante, ces clés deviennent très pratiques et il est clair qu'elles peuvent contenir des données sensibles ou même personnelles. On peut même se demander si le risque ne sera pas plus là que sur le disque dur du portable. Les réponses à cette situation ne pourront qu'être progressives et proportionnées aux besoins pour éviter un rejet brutal qui serait contre-productif. Une première idée consisterait à fournir aux utilisateurs une clé USB contenant une partie chiffrée (biométriquement, par exemple), son certificat et/ou un dispositif d'authentification par mot de passe jetable. On peut même penser à équiper cette clé d'applications portables telles que Thunderbird ou Firefox, permettant ainsi leur utilisation sur n'importe quel PC, sans installation, et avec les options personnelles de l'utilisateur (compte de mail, filtres, marques pages, etc...).

Au CENBG, le nomadisme a été considéré comme un enjeu du laboratoire dans sa Politique de Sécurité des Systèmes d'Information (PSSI). Cela signifie, que sans cette possibilité, la capacité du laboratoire à réaliser ses objectifs scientifiques peut être altérée. Il ne s'agit plus de la simple consultation des mails mais de travailler depuis l'extérieur, comme on le fait depuis l'intérieur.

Si cette évolution est prise en compte dans une PSSI c'est parce que la mise à disposition de cette capacité se heurte évidemment à des aspects de sécurité importants. Il faut donc que le système d'information évolue, tant dans ses aspects techniques que dans sa composante humaine (sensibilisation), pour fournir un cadre de sécurité proportionné qui permettra de favoriser le développement du nomadisme.

Dispositif pour les accès distants

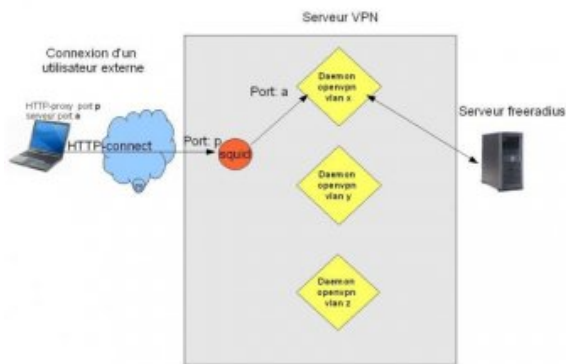


Figure 1

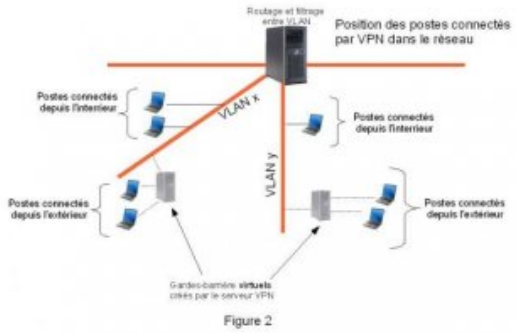
VPN - 1

Depuis deux ans, nous mettons en œuvre un serveur VPN basé sur OpenVPN. L'objectif était de permettre aux utilisateurs de se connecter dans leur environnement de travail habituel (c'est-à-dire dans le sous-réseau de leur groupe de travail), de disposer d'une authentification forte et de la possibilité de filtrer finement les autorisations (jusqu'à l'individu). Ces seules dispositions ne sont pas suffisantes. En effet, le développement du nomadisme (voire du télétravail à domicile) nous laisse entrevoir une sorte de retournement de nos réseaux où les utilisateurs seront plus à l'extérieur qu'à l'intérieur. Le risque d'intrusion est (et sera) accru. Pour cela, la mise en place de dispositifs de filtrage et d'authentification doit s'accompagner de dispositifs de détection et de limitation des intrusions. Pour cela, nous avons actuellement deux réponses. D'une part, la segmentation du

réseau à raison d'un sous-réseau par groupe ou service permet de limiter les risques de propagation d'une intrusion ; d'autre part, dans les mois prochains, une détection d'intrusion (SNORT) devrait être opérationnelle. Elle aura pour objectif de mieux connaître les flux qui circulent sur le réseau et d'identifier ceux qui sont indésirables.

Les technologies nomades évoluent très vite et elles intéressent forcément nos utilisateurs pour qui le nomadisme est intrinsèquement lié à leur métier. Tout cela nous oblige à gérer de nouveaux problèmes que l'on pourrait qualifier, pour être positif, de nouveaux défis.

Serge Borderes (CENBG)



VPN - 2



n°2
Septembre
2008

La lettre IN2P3 Informatique

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



Fairouz Malek : "La grille, cette utopie technologique dont le LHC avait besoin"



- Quel est votre rôle au sein du projet LCG-France ?

LCG-France est la partie française de la grille de calcul du Large Hadron Collider (LHC). Dans ce cadre, mon rôle principal est celui de manager. Il s'agit de coordonner les activités des sites LCG-France, d'organiser et d'animer des rencontres ou des ateliers scientifiques, de communiquer à l'intérieur de cette communauté et vers l'extérieur dans le cadre de W-LCG, de travailler sur les plans budgétaires et de recrutement de postes pour le projet, de veiller à ce que l'exécution du budget soit conforme aux prévisions et aux besoins, d'en référer auprès des agences de moyens et autorités scientifiques concernées, etc. Un travail de chef d'entreprise fort heureusement secondé par des collègues experts et compétents.

- Pourquoi avoir choisi les technologies de grille pour traiter les données du LHC ?

Le LHC s'est creusé la tête longtemps sur la manière de gérer tant de données ... centralement ? de manière distribuée ? où ? comment ? qui paie quoi ? En 1996, la communauté était convaincue de la nécessité de concevoir un modèle de calcul distribué. A cette époque, l'idée d'avoir 2 ou 3 centres régionaux hors du CERN avait déjà fait son chemin. L'idée de la grille, telle que nous la concevons aujourd'hui, s'est vite imposée dès 1998 après le succès du livre de Ian Foster. La grille de calcul qui permet de se brancher de n'importe où et de consommer les ressources là où elles se situaient sans les transporter, cette

utopie technologique, c'est bien de cela dont le calcul LHC avait besoin. Il est bien évident que nous n'en sommes pas exactement à fonctionner de cette manière, cependant la grille actuelle est un véritable défi technologique et un exploit mondial.

- A quoi va servir le LHC ?

A répondre à des questions fondamentales comme 'comment la matière acquiert sa masse ?'. Certes, ce n'est pas une question qu'on se pose tous les matins en se regardant dans le miroir, mais scientifiquement, c'est une connaissance qui contribue à comprendre la genèse de notre univers.

- Quelles ont été les principales difficultés dans le déploiement de cette grille ?

Je dirais l'augmentation rapide du matériel, quasi exponentielle, le manque de maturité des logiciels, la mise à l'échelle des logiciels et des outils. Enfin, le manque de personnel et, en particulier, l'implication faible des physiciens.

- A une semaine du démarrage du LHC, comment vous sentez-vous ?

Etrangement ... très sereine. Je considère que le calcul LHC est un sous-détecteur de chacune des quatre expériences LHC. Et comme pour chaque sous-détecteur, il va devoir être mis en route avec précaution, réglé, calibré, éprouvé encore pour quelques semaines voire quelques mois, avant de donner son maximum d'efficacité à son fonctionnement et laisser la place à la physique de se sublimer.

Propos recueillis par GAELLE SHIFRIN



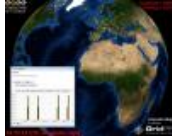
n°2
Septembre
2008

La lettre IN2P3 Informatique

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



Le Sénégal, nouvelle terre d'escale pour EGEE



Juillet 2007 marque les débuts de l'extension de la grille EGEE à l'Afrique sub-saharienne : le premier nœud EGEE a démarré et a été validé à l'université Cheick Anta Diop de Dakar (Sénégal). Cette réalisation fait suite au colloque de Montpellier en décembre 2007 organisé par la fondation « Partager le savoir » sur le développement d'Internet et des grilles de calcul en Afrique. L'Institut des grilles (IN2P3/CNRS) a été au cœur de la dynamique mise en place à cette occasion, d'une part dans le démarrage d'un partenariat avec l'Afrique du Sud pour démarrer plusieurs sites et organiser une Grid School à l'intention des futurs utilisateurs durant l'automne 2008, d'autre part pour aider le démarrage du site de l'université de Dakar.

Le site de Dakar est exemplaire d'une démarche pragmatique portée par des acteurs locaux motivés. L'université de Dakar est la principale université du pays avec 60 000 étudiants dans de nombreuses disciplines et ce projet de site grille connecté à EGEE a été conçu comme une ressource modeste au départ mais permettant l'accès à une ressource grille locale complète pour faciliter la sensibilisation et la formation à ces technologies ainsi que l'accès à la ressource grille globale pour les besoins des utilisateurs sénégalais. La ressource actuelle permet d'exécuter 10 jobs simultanés et devrait grossir d'ici la fin de l'année pour offrir une capacité de 50 jobs simultanés.

La mise en place du site s'est faite à l'occasion d'une mission d'une semaine sur place durant laquelle une première formation des collègues sénégalais a eu lieu. Le site, rattaché au ROC (Regional Operations Center) français et parrainé techniquement par GRIF, est géré par plusieurs personnes localement qui découvrent progressivement mais avec beaucoup de sérieux les spécificités de la gestion d'un nœud grille, de la communication et de la prise en charge des incidents d'exploitation...

Un des objectifs de la grille est de donner un accès « opportuniste » (sans garantie de ressources ou de priorités) à de petites communautés à faibles moyens. C'est pour cela qu'une VO (Virtual Organization) des utilisateurs de l'université Cheick Anta Diop a été créée pour faciliter l'accès des étudiants, chercheurs et enseignants à la grille. Elle a actuellement accès aux ressources de GRIF mais elle demandera probablement dans les mois qui viennent un accès à d'autres sites pour rendre un peu plus concrète l'idée de grille. Espérons que les sites français sauront faire un bon accueil à une telle demande. La grille est une technologie majeure pour permettre la réduction de la « fracture numérique » en permettant à la fois l'accès à des ressources considérables de façon transparente et la mise en œuvre de ressources locales nécessaires au développement des compétences locales.

Michel JOUVIN (LAL)



TSkim : un outil pour découper les arbres ROOT



Comme de multiples expériences de physique, le **Fermi Large Area Telescope** (Fermi-LAT, anciennement GLAST-LAT) stocke ses données sous forme d'arbres ROOT. Une activité récurrente des physiciens consiste à établir des critères pour définir quels événements sont intéressants, et à effectuer des coupes sur les arbres **ROOT** afin d'extraire les données attachées à ces événements particuliers. TSkim a été développé pour les y aider. Initialement un simple couple de scripts PERL et ROOT, l'outil est à présent devenu suffisamment élaboré pour prétendre servir à d'autres expériences que FERMI.

Pour un même événement physique, l'utilisateur peut souhaiter récupérer des données issues de plusieurs sources, correspondant à des représentations ou des stades différents dudit événement : simulation de la collision initiale de particules, simulation de la réponse du détecteur, vraies mesures du détecteur, reconstruction de base, calcul d'un jeu de grandeurs physiques caractéristiques permettant de classifier l'événement, etc. Un arbre contient les données de toute une séquence d'événements, mais pour seulement l'un des types d'information précédents. Les arbres utilisés pour un type d'information donné ont tous la même structure, mais cette structure diffère d'un type à l'autre. En effet, elles ont été définies par des équipes différentes, à des périodes différentes, avec plus ou moins de concertation, et pour répondre à des contraintes techniques différentes. Au bout du compte, d'un type à l'autre, les noms de fichier ne respectent pas les mêmes règles, les arbres n'ont pas les mêmes noms ni les mêmes branches, et les morceaux d'un même événement ne sont pas forcément placés au même rang dans les différents arbres.

Un utilisateur ordinaire de ROOT peut écrire sans trop de difficultés un script capable de traiter une séquence d'arbres d'un type donné (par exemple les mesures réelles). Par contre, la tâche devient plus ardue lorsqu'il faut le faire simultanément pour plusieurs types d'arbres (par exemple les mesures réelles et la reconstruction de base), ce qui constitue la très grande majorité des cas d'utilisation. Si on ajoute que la manipulation des arbres et des indexes de ROOT demande un certain savoir-faire, on comprendra qu'il était naturel de développer un outil pour faciliter cette activité.

Gestion de l'hétérogénéité des données

L'objectif premier de TSkim est de gérer l'hétérogénéité des données. Pour ce faire, il s'appuie sur un fichier de description des types de données, personnalisable, qui définit pour chaque type, entre autres : le nom de l'arbre à rechercher dans les fichiers, le nom de la branche qui contient les numéros d'événements, le nom de la bibliothèque de définition de données C++ qu'il faut charger en mémoire (inutile pour les arbres plats, ou "tuples", mais souvent indispensable pour les structures plus complexes). Avant de lancer une exécution de TSkim, l'utilisateur doit fournir la liste des fichiers à traiter, et des critères de coupe définis pour l'un des types de fichiers. TSkim procède alors en deux étapes. Dans un premier temps , il

enregistre dans une liste les numéros des événements intéressants, pour un type d'arbre donné et selon les critères établis par l'utilisateur. Dans un second temps, il utilise cette liste pour retrouver dans tous les fichiers les données associées à ces mêmes numéros, les lire et les recopier dans de nouveaux fichiers de sortie. Depuis peu, TSkim peut aussi, au lieu de recopier les données, écrire un fichier spécial qui se contente de lister les éléments sélectionnés (technique inspirée des KanEvents de BaBar). Ce dernier fichier pourra être réinjecté dans TSkim comme entrée pour une nouvelle découpe. On peut procéder ainsi à des coupes cumulatives, sans jamais copier la moindre donnée, ce qui peut épargner à l'utilisateur du temps et surtout de la place disque.

En ce qui concerne le savoir-faire nécessaire à la manipulation des arbres ROOT, il s'agit à la fois de profiter des optimisations fournies par ROOT lorsque le cas d'utilisation le permet, par exemple la fusion rapide de données, mais aussi de détecter et d'interdire des pratiques connues comme néfastes, par exemple l'utilisation d'indices d'événements trop grands pour être compatibles avec l'implémentation actuelle des TTreeIndex de ROOT. A propos de l'implémentation de TSkim lui-même, un grand virage a été pris lorsque nous avons décidé de transformer tous les scripts ROOT en programmes C++ compilés, suite à plusieurs déconvenues avec l'interpréteur C++ de ROOT. Cela se paie à l'installation, car il faut recompilier l'outil pour toutes les versions de ROOT qui peuvent être nécessaires aux utilisateurs, mais cela nous évite les problèmes trop fréquents liés à l'interpréteur.

N'ayant été intensivement utilisé que dans le contexte d'une seule expérience, TSkim a besoin aujourd'hui de trouver un nouveau public afin de comprendre et de corriger ce qui reste trop spécifique à la problématique du Fermi-LAT. Cependant, nous ne doutons pas du potentiel de l'outil et de l'intérêt qu'il présente dès à présent pour beaucoup d'utilisateurs de ROOT. C'est l'exemple typique d'un développement très spécifique à notre discipline, réalisé dans un coin pour une seule collaboration (et qui ne générera jamais le moindre revenu), mais qui peut assez facilement profiter à une communauté d'utilisateurs plus large, d'autant plus vite que nous réussirons à mobiliser d'autres développeurs et à en faire un produit collectif de notre institut. Il ne s'agit pas seulement de rendre notre outil plus flexible pour répondre aux besoins de telle ou telle expérience. Il y a également de nouvelles idées à explorer : la gestion automatique de l'évolution dans le temps des structures de données, la reconnaissance des événements en utilisant des "timestamps" plutôt que des numéros, la connexion avec les catalogues de données des expériences, etc.

Utilisateurs de ROOT ou développeurs, nous vous donnons rendez-vous aux Journées Informatique et sur le **nouveau site web de TSkim**. A l'heure où nous écrivons ces lignes, ce nouveau site est en construction, mais il devrait s'enrichir très rapidement dans les prochains jours.

David Chamont (LLR)



n°2
Septembre
2008

La lettre IN2P3 Informatique

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



JSAGA : pour un accès uniforme à des grilles hétérogènes



De nombreux utilisateurs ont besoin d'accéder de façon uniforme à des grilles d'envergures différentes (locale, régionale, nationale, internationale) et s'adressant à différentes communautés (académiques/industrielles, recherche/production). Ces grilles ont différents atouts et lacunes, et le choix de l'utilisateur de soumettre sur l'une ou l'autre dépend de différents critères comme les caractéristiques et la quantité de ressources matérielles requises, les contraintes de licence des logiciels utilisés, le niveau de confidentialité requis, le degré d'urgence, la charge de ces grilles, etc.

A moyen terme, un accès uniforme sera possible entre certaines grilles académiques d'envergure nationale ou internationale, grâce à divers projets d'interopérabilité en cours. Cependant, les approches choisies pour ces projets (modifications côté serveurs) ne permettent pas d'envisager une utilisation uniforme de ces grilles et des grilles en dehors de ces projets d'interopérabilité.

Dans le cadre du projet IGTM financé par l'ANR, le Centre de Calcul de l'IN2P3/CNRS (CC-IN2P3) développe JSAGA ; une API qui suit une approche complémentaire à celles de ces projets d'interopérabilité. Cette approche (encapsulation côté client) permet un accès uniforme à potentiellement n'importe quelle grille. De plus, elle permet d'avoir une solution sur le court terme puisqu'elle s'affranchit des contraintes de déploiement des infrastructures de grille, et elle bénéficie sur le long terme des avancées des autres projets afin de conserver un niveau de complexité interne raisonnable.

JSAGA est actuellement intégrée dans divers outils : le portail de grilles industrielles et académiques Elis@ (CS-SI), l'outil de soumission de collections de jobs JJS (CC-IN2P3) et l'explorateur de fichiers JUX (CC-IN2P3).

Hétérogénéité des interfaces ...

L'obstacle le plus évident à l'utilisation uniforme de plusieurs grilles est l'hétérogénéité des interfaces de leurs intergiciels. Il était donc nécessaire d'encapsuler ces interfaces dans une interface commune. Parmi les différentes alternatives possibles, la spécification SAGA (Simple API for Grid Applications) a été choisie. Cette spécification est définie par l'Open Grid Forum, structure internationale dont le rôle est d'élaborer et préconiser des standards pour les grilles informatiques. Comme son nom l'indique, SAGA est une API simple d'utilisation. Elle n'est associée à aucun des intergiciels existants et peut servir d'interface commune à ceux-ci.

JSAGA est une implémentation de la spécification SAGA en Java (d'où le "J" de JSAGA, vous l'aurez deviné...). Cette implémentation permet une utilisation uniforme des mécanismes et composants de sécurité, de gestion des données et de gestion de l'exécution, provenant de différents intergiciels de grille (Globus Toolkit, Unicore, gLite, etc.) et de technologies plus classiques (certificats X509, PKI, HTTPS, SFTP, SSH, etc.). La liste des technologies supportées par JSAGA est extensible par l'ajout de plug-ins. Les plug-ins actuellement proposés bénéficient d'une contribution significative de CS-SI / British Telecom.

Enfin, les API clientes des intergiciels ont des approches différentes pour résoudre un même problème (e.g. transfert de données, suivi de l'exécution). Ces différences conduisent souvent à une perte d'efficacité lorsque l'on contraint ces API à être utilisées au travers d'une interface commune. La conception de JSAGA et la définition des interfaces de ses plug-ins prennent en compte les différentes approches possibles afin d'utiliser efficacement les différents intergiciels, et ainsi de ne pas être sujet à cette perte d'efficacité.

... et des infrastructures

JSAGA s'affranchit de l'hétérogénéité des intergiciels grâce à son implémentation de la spécification SAGA, mais il ne se résume cependant pas à cette implémentation... Généralement sous-estimé, un autre obstacle important à l'utilisation uniforme de plusieurs grilles est l'hétérogénéité de leurs infrastructures. Les infrastructures en place se distinguent, par exemple, par les autorités de certification supportées, les politiques de filtrage réseau, les commandes déployées sur les nœuds de calcul, les variables d'environnement initialisées sur ces nœuds, etc. Ainsi, même si elles sont basées sur le même intergiciel, les infrastructures de deux grilles distinctes ne pourront généralement pas être utilisées de la même façon.

Il faut donc permettre, entre autres, une utilisation transparente de plusieurs contextes de sécurité, de plusieurs versions ou modes d'utilisation des technologies déployées, d'environnements d'exécution possédant des possibilités et contraintes très différentes. Par exemple, le chemin emprunté pour transporter jusqu'au nœud de calcul les données nécessaires à l'exécution du job peut dépendre de nombreux paramètres liés à la grille ou au site ciblés (e.g. commandes déployées sur les nœuds de calcul, règles de filtrage réseau), aux technologies employées (e.g. modes d'accès du protocole, possibilité de transfert tierce partie, de délégation des droits d'accès) et aux données elles-mêmes (e.g. taille des fichiers, partage des fichiers entre plusieurs jobs).

La particularité de JSAGA est de prendre en compte ces paramètres afin de s'affranchir de l'hétérogénéité des infrastructures. JSAGA s'appuie sur son implémentation de la spécification SAGA pour permettre la soumission efficace d'une collection de jobs sur différentes infrastructures de grilles (e.g. < a href="http://www.eu-egee.org/" target="_blank">EGEE, OSG, DEISA, NAREGI), à partir d'une description unique de cette collection. Cette description peut être écrite dans différents langages, dont le JSDL (Job Submission Description Language), standard défini par l'Open Grid Forum.

Pour en savoir plus, voir le site web de JSAGA.

Sylvain Reynaud (CC-IN2P3)



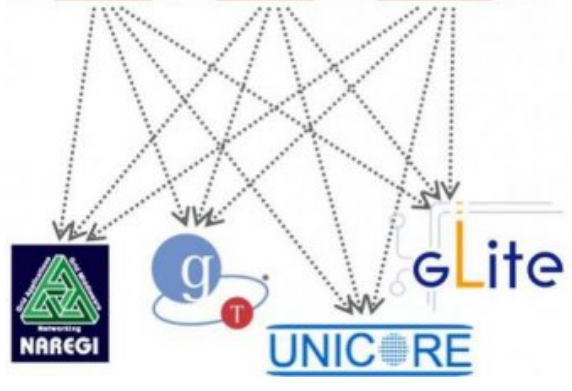
calculs

JSAGA

sécurité

données

exécution





n°2
Septembre
2008

La lettre IN2P3 Informatique

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



Le CC-IN2P3 s'exporte en Corée



ressources informatiques coréennes, pour la mise à disposition de ressources informatiques massives en Corée en vue de l'analyse des données de l'expérience ALICE sur le LHC, grâce à la technologie des grilles informatiques.

Rappelons que le LHC Computing Grid est une infrastructure internationale de grille, composée de centaines de sites hiérarchisés, qui vise à analyser et stocker les immenses quantités de données générées par le LHC. Ce puissant accélérateur, qui sera mis en production le 10 septembre prochain, devrait ainsi engendrer près de 15 millions de milliards d'octets de données par an.

Dans le cadre du LIA avec la Corée, le CC-IN2P3 collabore depuis plusieurs mois avec le KISTI afin de renforcer sa participation à la grille LCG en tant que site de niveau 2, spécifiquement dédié à l'analyse des données de l'expérience ALICE.

L'expérience ALICE a pour objectif l'étude de la matière nucléaire dans un état extrême de température et de densité, appelé le plasma de quarks et de gluons qui aurait existé quelques microsecondes après le Big-Bang. Cette expérience pourra ainsi apporter des éclairages nouveaux sur les questions fondamentales telles que l'organisation ultime de la matière soumise à l'interaction forte et l'état de la matière dans les premiers instants de l'Univers.

Dans la même optique, le CC-IN2P3 fait également bénéficier le KISTI de son expertise dans le domaine de l'opération de l'infrastructure européenne de grille EGEE, qui supporte la grille LCG.

L'une des premières actions concrètes de la collaboration entre le CC-IN2P3 et le KISTI est de mettre sur pied et de déployer une organisation virtuelle (VO) dédiée au FKPPL. Celle-ci permettra, entre autres, aux étudiants coréens qui ont suivi l'école d'été sur les grilles organisée à Séoul du 14 au 25 juillet 2008, de mettre en application les techniques qu'ils ont apprises sur une architecture internationale. Cette nouvelle VO devrait être opérationnelle dans les prochaines semaines.

Gaëlle Shifrin et Dominique Boutigny (CC-IN2P3)

Cette année a été créé le Laboratoire International Associé (LIA) France – Korea Particle Physics Laboratory (**FKPPL**) par l'IN2P3-CNRS, l'Ecole Polytechnique, l'Université Paris-Sud et l'Université Blaise-Pascal de Clermont-Ferrand, en partenariat avec 7 établissements coréens. L'objectif de ce LIA est de renforcer les collaborations dans les domaines de la physique des hautes énergies et des grilles informatiques. Ce LIA est issu de la volonté du CNRS et de l'IN2P3 de tisser des liens plus étroits avec les établissements de recherche en Asie. Son objectif est de dynamiser les échanges scientifiques entre la France et la Corée dans le domaine de la physique des hautes énergies et des grilles informatiques en s'appuyant sur les collaborations déjà établies entre les équipes de recherche.

Le premier comité de pilotage du LIA s'est tenu à Séoul le 21 juillet 2008. Celui-ci a passé en revue et s'est prononcé sur le financement des premiers projets scientifiques proposés dans le cadre du FKPPL.

Les travaux entrepris par les équipes des deux pays dans l'exploitation des grilles pour la recherche de nouveaux médicaments avaient déjà débouché en 2007 sur l'identification de molécules prometteuses pour combattre la malaria et la grippe aviaire.

Aujourd'hui, une forte collaboration s'établit entre le Centre de Calcul de l'IN2P3/CNRS et le **KISTI**, Centre national de



n°2
Septembre
2008

La lettre IN2P3 Informatique

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



3 au 5 déc. 2008 - FUTURVIEW 2008

Il reste deux semaines avant la clôture de l'appel à contributions. Vous n'avez pas eu l'occasion de présenter vos propres réalisations sous LabVIEW lors de la première édition de cet événement en 2003 ? N'hésitez pas à envoyer une proposition de communication pour la seconde édition de cette conférence utilisateurs, ouverte au plus grand nombre.

FuturVIEW 2008 sera un lieu de rencontres et d'échanges parmi la communauté sans cesse grandissante des utilisateurs de LabVIEW® qui englobe aussi bien les ingénieurs et les scientifiques, les étudiants et les enseignants mais aussi les industriels, tous désireux de développer des applications destinées à l'instrumentation et aux contrôles industriels. Lors de cette seconde édition, seront présentées des applications de mesure et d'instrumentation, mais aussi des applications émergentes de communication, tirant parti d'Internet, des réseaux, des nouvelles architectures multicoeurs... Ces journées francophones permettront non seulement aux experts de partager leur expérience et leur savoir-faire, mais elles donneront également aux utilisateurs novices l'occasion de découvrir des facettes originales de l'environnement LabVIEW® en leur proposant une demi-journée de formation, préalable à la conférence.

La soumission des communications

Les auteurs sont invités à faire parvenir leur proposition de communication sous la forme de deux pages maximum au format A4 (environ 1000 mots). Le texte définitif prêt à imprimer sera limité à 6 pages format A4 avec des consignes de présentation précisées ultérieurement. Le texte devra indiquer le titre de la communication et le nom des différents auteurs, en identifiant un auteur chargé du contact avec les organisateurs (adresse postale, numéro de téléphone, numéro de télécopie et adresse électronique). Les propositions seront examinées chacune par deux rapporteurs. Chaque communication sélectionnée sera présentée soit oralement pendant 20 minutes, soit sous la forme de poster pendant des sessions spécifiques. Dans les deux cas, le texte complet sera publié dans les Actes de la conférence FuturVIEW 2008 avec une référence ISBN.

Les textes doivent être envoyés avant le 3 septembre 2008 par courrier électronique au format Microsoft Word (.doc) ou PDF à l'adresse suivante : FV08ORG-L@lpsc.in2p3.fr.

Les consignes de présentation des articles définitifs, destinés au livre des actes de la conférence, seront publiées sur le site après sélection des propositions de communication (second semestre 2008).

Le format de présentation des posters sera publié sur ce site après sélection des propositions de communication (second semestre 2008).



29 sept. au 2 oct. 2008 - 6e édition des JIs de l'IN2P3 et de l'IRFU

Les Journées Informatique de l'IN2P3 et de l'IRFU sont organisées tous les deux ans et la 6e édition se tiendra à Obernai, près de Strasbourg, du 29 septembre au 2 octobre prochain. Elles sont un lieu de rencontre des acteurs de l'informatique de l'IN2P3 et de l'IRFU et abordent tous les aspects de la mise en œuvre et de l'utilisation de l'informatique pour la recherche en physique des particules et astroparticules et en physique nucléaire.

Première occasion de rencontre de la communauté des informaticiens IN2P3/Irfu depuis la mise en place du réseau RI3, ce sera là une occasion d'évoquer le plan d'action du tout nouveau CCRI, son comité de coordination.

Les inscriptions sont closes au 5 septembre mais les participants peuvent modifier leur inscription jusqu'au 12 septembre, date à laquelle les réservations de nuitées, de repas et de places dans les navettes seront figées.

Vous pourrez également suivre en direct les Journées Informatique de l'IN2P3 et de l'IRFU en vous connectant à l'adresse <http://webcast.in2p3.fr/JI08/>.