



A LA UNE

Le système central d'acquisition de données AUGER (CDAS)



A perte de vue, de minuscules fleurs jaunes tapissent le sol sablonneux et irrégulier de la Pampa Amarilla. Pas un arbre, une colline ou une maison alentour. Seule la Cordillère des Andes se détache au loin de l'immensité plane, écrasée par le soleil, de cette région du centre de l'Argentine. Cahotante et poussiéreuse, la route de terre qui relie Malargüe, la ville la plus proche avec ses 25 000 habitants, à cinq heures du premier aéroport, celui de Mendoza, semble interminable.

Quand soudain, à quelques mètres du chemin, apparaît une cuve cylindrique en plastique ocre pâle. A peine moins haute qu'un homme, elle est surmontée d'antennes hertzienne et GPS et d'un panneau solaire. En scrutant l'horizon, on distingue au loin d'autres cuves. Incongrue, leur présence ne

le reste pas longtemps pour qui sait que cette terre aride n'est pas seulement un lieu d'élevage extensif, mais accueille aussi le plus grand observatoire au monde de "rayons cosmiques ultra-énergétiques", l'observatoire Pierre Auger.

Un peu de physique et d'histoire...

Pierre Auger Pierre Auger Les rayons cosmiques ? Notre planète en est sans cesse bombardée. Chaque seconde, plusieurs milliers de ces invisibles rayons pénètrent chaque mètre carré de notre atmosphère. Une fois n'est pas coutume, ce sont les Anglo-Saxons qui ont proposé de nommer l'observatoire (...)

Interview "L'extension du CC-IN2P3 est lancée"

Dominique Boutigny - Directeur du CC-IN2P3

[lire l'interview](#)

Archives

Abonnement

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

Proposer un article

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

[lire la suite](#)



© 2009 CCIN2P3

■ Calcul



Le projet PetaQCD

L'étude d'une très ambitieuse machine pétaflopique consacrée à la simulation QCD sur réseau a débuté depuis quelques mois. Elle est le fruit de plusieurs années de recherches collaboratives dans ce domaine, et regroupe, au sein du projet ANR intitulé PetaQCD, un consortium pluridisciplinaire de 9 partenaires issus de l'IN2P3, du CNRS, de l'INRIA plus 2 petites entreprises privées. Ce projet devrait, sur la période 2009-2011, faire la preuve qu'une machine de ce type, capable de simuler des réseaux de l'ordre de 256x1283 à l'aide des logiciels de la famille HMC, est réalisable à un coût raisonnable avec un nombre limité de processeurs (entre 1000 et 4000). Cette preuve prendra la forme d'une maquette basée autant que faire se peut sur des éléments standards, et 2 axes particulièrement prometteurs seront explorés en priorité : l'utilisation des processeurs Cell (IBM) et celle des cartes GP-GPU (Nvidia). La machine sera donc massivement parallèle, et très probablement hétérogène, et les problèmes de communication entre processeurs, intrinsèques au problème traité plus encore qu'à l'algorithme utilisé, seront certainement l'une des clefs de l'étude. (...)

■ Méthodologie



Bonnes pratiques ASR et ITIL

Les Administrateurs des Systèmes et des Réseaux (ASR) sont responsables de la bonne marche des services informatiques. Or, que ce soit au Centre de Calcul ou dans les laboratoires, l'ensemble des activités de l'Institut est de plus en plus dépendant de la disponibilité du réseau, des données, des applications. Toute panne ou dysfonctionnement est susceptible de paralyser un établissement ou d'avoir des conséquences sérieuses. Il est donc essentiel que les ASR appliquent des méthodes de travail visant à renforcer la disponibilité et la solidité des moyens informatiques, d'en maîtriser les évolutions et d'assurer leur alignement avec les besoins des utilisateurs. La plupart des ASR en sont conscients et appliquent déjà de manière informelle des éléments d'une démarche qualité allant dans ce sens. ITIL est un ensemble de meilleures pratiques initialement conçu et développé par le gouvernement britannique à la fin des années 1980 afin d'améliorer l'efficacité des services informatiques dans le domaine public. Les Journées Informatiques d'Obernai ont permis d'associer dans une même session une présentation des (...)

■ Grille



Institut des Grilles : bilan après un an d'existence

Voilà un peu plus d'un an que l'Institut des Grilles a été créé, voici donc un premier bilan de l'année écoulée. Pour rappel, l'IdG a pour mission de consolider les infrastructures existantes dans le domaine des grilles de calcul et de permettre la meilleure synergie possible entre les différents acteurs. Le développement des grilles de production tant au niveau national qu'au niveau européen s'est poursuivi avec ardeur. Signé le 18 décembre 2008 par le Directeur Général de la Recherche de l'Innovation, le protocole d'accord, entre tous les grands organismes de recherche en France (CNRS, CEA, INRIA, INRA, CPU, RENATER et Ministère) institue en particulier le comité de pilotage national dont la mission est de conduire le développement et de pérenniser de la grille de production tant sur le plan national qu'europpéen. Les premières discussions sur le plan des embauches 2009-2010 afin de stabiliser les experts recrutés jusqu'à présent sur des contrats temporaires européens, et sur un plan de création de

■ Développement



Faut-il passer d'un système de gestion de versions centralisé à un système décentralisé

Tous ceux qui manipulent des fichiers, en souhaitant conserver l'historique de leur travail, trouveront un intérêt à s'appuyer sur un VCS. L'un de leurs principaux intérêts est de permettre le développement mutualisé, avec une gestion des conflits. Pourtant en pratique, il s'avèrera particulièrement utile, même pour un utilisateur isolé. Son choix parmi les nombreux outils disponibles (centralisé ou réparti) aura nécessairement un impact fort sur le modèle de développement suivi. Qu'est-ce qu'un 'Système de gestion de versions' ? Un VCS (Version System Control) est un outil informatique capable de gérer les évolutions et l'historique d'un ensemble de fichiers et de ressources. Stockant les informations dans un dépôt, il permet de récupérer toutes les versions intermédiaires, ainsi que les différences entre les versions. Il permet à plusieurs développeurs de travailler sur un même projet et effectue la fusion des modifications non conflictuelles en protégeant contre celles (...)

[lire la suite](#)

[lire la suite](#) nouveaux nœuds dans les
grandes villes (...)

[lire la suite](#)

[lire la suite](#)



n°4
Janvier
2009

La lettre IN2P3 Informatique

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

A LA
UNE

Le système central d'acquisition de données AUGER (CDAS)

Acquisition



A perte de vue, de minuscules fleurs jaunes tapissent le sol sablonneux et irrégulier de la Pampa Amarilla. Pas un arbre, une colline ou une maison alentour. Seule la Cordillère des Andes se détache au loin de l'immensité plane, écrasée par le soleil, de cette région du centre de l'Argentine. Cahotante et poussiéreuse, la route de terre qui relie Malargüe, la ville la plus proche avec ses 25 000 habitants, à cinq heures du premier aéroport, celui de Mendoza, semble interminable.

Quand soudain, à quelques mètres du chemin, apparaît une cuve cylindrique en plastique ocre pâle. A peine moins haute qu'un homme, elle est surmontée d'antennes hertziennes et GPS et d'un panneau solaire. En scrutant l'horizon, on distingue au loin d'autres cuves. Incongrue, leur présence ne le reste pas longtemps pour qui sait que cette terre aride n'est pas seulement un lieu d'élevage extensif, mais accueille aussi le plus grand observatoire au monde de "rayons cosmiques ultraénergétiques", l'observatoire Pierre Auger.

Un peu de physique et d'histoire...

Pierre Auger Les rayons cosmiques ? Notre planète en est sans cesse bombardée. Chaque seconde, plusieurs milliers de ces invisibles rayons pénètrent chaque mètre carré de notre atmosphère. Une fois n'est pas coutume, ce sont les Anglo-Saxons qui ont proposé de nommer l'observatoire du nom d'un Français ! Un hommage au physicien Pierre Victor Auger (1899-1993), mondialement connu dans les domaines de la physique nucléaire et des rayons cosmiques avant la Seconde Guerre mondiale. Pionnier, il a été le premier, dès 1938, à prédire l'existence de grandes "gerbes atmosphériques", des gerbes de particules que provoqueraient les rayons cosmiques de très hautes énergies frappant notre atmosphère.



Pierre Auger

Au lendemain de 1945, il joua un rôle déterminant dans le développement de grandes institutions internationales, tels le Cern (le Laboratoire européen pour la physique des particules, à Genève) ou l'Unesco. C'est donc avec cet observatoire unique au monde que l'on pourrait enfin détecter les grandes gerbes atmosphériques engendrées par ces particules d'énergies beaucoup plus élevées que celles que peuvent leur conférer les accélérateurs construits par l'homme et ça grâce à un réseau de 1600 détecteurs de particules Cherenkov à eau, espacés d'1.5 kilomètres, qui s'étendent sur une surface de 3000 km² (30 fois la superficie de Paris). En plus, 24 télescopes observent la lumière fluorescente produite par la gerbe lors de son passage dans l'atmosphère dans une nuit sans lune. Cette combinaison de détecteurs permet une étude optimale et très précise de ces rayons cosmiques aux énergies >10¹⁹ électronvolts (eV).

L'Observatoire Pierre Auger est le plus grand instrument consacré à l'étude des rayons cosmiques. Il regroupe des chercheurs de 17 pays et de plus de 70 institutions, dont une partie importante sont des chercheurs du CNRS et des universités françaises (de cinq laboratoires de l'IN2P3 et d'un laboratoire de l'Insu). L'observatoire a été construit par plus de

Chaque SU (unité de transmission/réception radio au niveau de chaque cuve) « discute » avec le BSU (unité de transmission/réception radio/E1 au niveau de chaque tour) le plus proche et peut envoyer par seconde 1 seul packet de 1200bits. Dans chaque tour, il y a entre 5 et 6 BSUs qui couvrent un réseau d'une soixante de SU et peuvent chacun envoyer un broadcast aux SU (1-2 paquet par seconde), soit environ 2-3Mbps. Chaque tour envoie les données des réseaux détecteurs (télescopes et les cuves) à la tour centrale de l'observatoire par TCP/IP.

CDAS : le chef d'orchestre

Le système central d'acquisition de données (CDAS) d'AUGER est installé à l'observatoire à Malargüe fonctionnant avec une ferme de serveurs de calcul et des baies de stockage de données. Le CDAS consiste principalement en la gestion des "triggers" et des événements, le contrôle des processus et le stockage des données des détecteurs d'AUGER, le tout fonctionnant dans une infrastructure de système d'informations utilisant des matériels commerciaux et des applications de haut-niveau développées localement. Le fait que le CDAS ne soit pas un système d'acquisition de temps-réel a permis un développement d'un système à la fois simple, rapide et robuste dans ce type d'application. Le CDAS a un rôle central celui de combiner les informations "trigger" émises par les détecteurs de surfaces (les 1600 cuves) et ceux des "yeux" de la fluorescence (les 24 télescopes) pour former un "central trigger" (T3-trigger), puis de demander aux détecteurs qui ont enregistré les données utiles (ceux qui sont dans le "run") formant ce "trigger" (T3data) pour finalement construire selon des modèles physiques de correspondance et stocker ces données avec une signature temporelle bien précise sous forme d'un événement AUGER appelé un "shower event". Parmi les processus du CDAS, on pourrait citer principalement trois processus d'entre eux qui forment le coeur même du système d'acquisition :

- le "Postmaster" (Pm), celui qui joue le rôle d'aiguilleur de paquets (entre le réseau des détecteurs et le CDAS) et de gestionnaire d'événements ;
- le "Central Trigger" (Ct), celui qui va analyser et combiner en vol les informations "trigger" des deux réseaux de détecteurs (T2) pour former le "trigger" central (T3) après application d'une des combinaisons "cross trigger" pour validation ;
- le "Event Builder" (Eb), celui qui va construire et stocker l'évènement AUGER ("shower event") final après la combinaison et le traitement des différentes informations venant de Pm, de Ct et du réseau de la fluorescence ; AUGER : Schéma de la DAQAUGER : Schéma de la DAQ

La topologie système et réseau

Le CDAS fait travailler en continue une ferme de quatre serveurs de calcul, deux baies de stockage de données de plus 6To en RAID-5 et un réseau gigabit avec des routeurs et des switches de type CISCO qui interconnecte les réseaux de détecteurs avec l'intranet de l'observatoire. Quotidiennement les données brutes locales AUGER sont répliquées automatiquement plusieurs fois par jour sur deux sites dans le monde, un au FERMILAB (USA) et l'autre au centre de calcul IN2P3 (Lyon, FRANCE), en plus d'être sauvegardées localement.

Un système complexe et autonome de gestion de surveillance et d'alertes du CDAS fonctionne en permanence et une synergie de travail des experts (locaux et extérieurs) pour la maintenance permettent de garantir une continuité des prises de données et une stabilité des activités en production.

Quelques dates clés

370 chercheurs et ingénieurs, pour un montant total d'environ 40 millions d'euros. Il est situé dans la province de Mendoza, en Argentine. Le CNRS est le principal organisme français de financement de l'observatoire. Si l'Américain Jim Cronin (prix Nobel) et l'Anglais Alan Watson sont à l'origine de l'observatoire, Murat Boratav, professeur au Laboratoire de physique nucléaire et des hautes énergies (LPNHE -CNRS/IN2P3, Universités Paris 6 et 7) en est le pionnier en France et les chercheurs français sont nombreux à participer à la collaboration Auger.

L'infrastructure du système d'informations

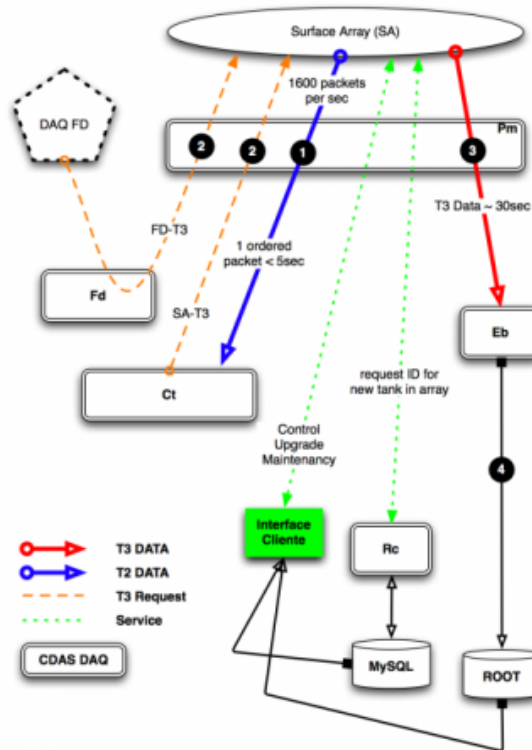
Un système de communication sur mesure

AUGER : Un des 4 tours de communication
 Le système de communication a été conçu par l'Université de Leeds (Département Electronique/Radio) et offre deux types d'interconnexions de communication entre les tours, les cuves et l'observatoire :

- Connexion WLAN (1.2Kbps – 915MHz ISM band) entre les cuves et les 4 tours du site ;
- Connexion haut-débit (16 canaux de 2.048Mbps chacun – 7GHz band) de type E1 micro-onde « point-to-point » entre les 4 tours sites et la tour centrale de l'observatoire. 4 pour les détecteurs de fluorescence (FD) et 2 pour les détecteurs de surface (SD) et le service/maintenance.



AUGER : Une des 4 tours de communication



AUGER : Schéma de la DAQ

- 1991 : le Projet Pierre Auger a vu le jour lors de International Cosmic Ray Conference à Dublin, initié par Alan Watson et Jim Cronin (Nobel Prix en 1980)
- 1996 : le site de Malargüe a été choisi pour l'hémisphère sud
- 1998 : un engineering array de 32 cuves pour la R&D pour la valider la faisabilité du projet
- Février 2001 : 1ère connexion réseau de télécommunication E1 (Cisco 2600) avec un des 4 tours du site.
- 12 Août 2001 : 1er évènement provenant des détecteurs de surface (4 cuves)
- 2002 : début de l'installation et de la production des cuves
- Janvier 2004 : 1ère prise de données, > million de rayons cosmiques mais 27 seulement sont >60EeV
- Été 2005 : 1er résultat physique Auger
- Juin 2008 : Dernier cuve installé avec l'eau et l'électronique, nous sommes à 1660 installés.

Pour en savoir plus

- www.auger.org : Official Auger Web Portal
- lpnhe-auger.in2p3.fr : LPNHE's Auger France Web Page
- www.auger.org.ar : Southern Site Web Page (Malargue, Argentina)
- www.augernorth.org : Northern Site Web Page (Colorado, USA)
- auger.cnrs.fr : CNRS's Auger France Web Page

Richard Randriatomanana (LPNHE)



n°4
Janvier
2009

La lettre IN2P3 Informatique

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



"L'extension du CC-IN2P3 est lancée"

Dominique Boutigny - Directeur du CC-IN2P3

Quels ont été les temps forts au CC-IN2P3 en 2008 ?

L'événement le plus important en 2008 pour le CC-IN2P3 a été lorsque l'Université Claude Bernard Lyon 1 a donné son accord pour attribuer au CNRS la portion de terrain nécessaire à la construction de la nouvelle salle informatique. Cette extension qui entrera en service fin 2010 est absolument indispensable pour absorber la forte croissance actuelle du centre, essentiellement liée au projet LHC. Grâce à cette attribution, l'extension du CC-IN2P3 est donc désormais lancée.

Un autre événement important est l'affectation d'un Chargé de Recherche STII au CC-IN2P3. Autour de lui va se constituer, petit à petit, une équipe de recherche en informatique essentiellement tournée vers les Grilles. Cette embauche est parfaitement cohérente avec l'une des missions de l'Institut des Grilles du CNRS qui vise à rapprocher la communauté des Grilles de production avec celle des Grille de recherche.

Bien entendu, une grande partie des activités du CC-IN2P3 a été consacrée au développement de l'infrastructure matérielle et logicielle pour le LHC en phase avec le projet LCG-France. Un investissement très important a été réalisé afin de faire évoluer l'architecture du réseau de la salle informatique du CC-IN2P3.

Comment a été gérée l'arrivée des données du LHC ?

Malheureusement, l'incident survenu sur le LHC a fait que très peu de données ont pu être acquises par les détecteurs avant la panne. Il s'agissait d'ailleurs de données de bruit de fond puisqu'aucune collision n'a eu lieu. Tout a parfaitement fonctionné mais le fait de recevoir les données au CC-IN2P3 n'était pas un challenge.

Les données d'ATLAS ont été répliquées sans incidents vers les Tiers-2 et les Tiers-3 du « nuage » français c'est-à-dire : la France, le Japon, la Chine et la Roumanie. Les données de CMS ont également été distribuées conformément au modèle de calcul de l'expérience.

En mai et juin a eu lieu un test combiné du calcul des expériences LHC (CCRC08). Il s'agissait d'une répétition générale afin de vérifier les performances du système informatique de LCG. Les résultats du CCRC08 lui-même ont été très bons ; par contre, des tests ultérieurs de retraitement des données ont mis en évidence des points nécessitant des améliorations sur lesquelles le projet LCG dans son ensemble travaille.

Excepté le LHC, quels sont les projets les plus importants du CC-IN2P3 ?

Le CC-IN2P3 supporte l'ensemble des activités menées à l'IN2P3 et à l'Irfu. La structure en projet de LCG-France permet justement de bien séparer les budgets LHC et non-LHC ce qui préserve les investissements en calcul hors LHC. Parmi les plus gros utilisateurs du CC-IN2P3 figurent les expériences BaBar et D0, ceci est très satisfaisant car cela montre que le CC-IN2P3 remplit parfaitement son rôle pour des expériences qui fournissent de très nombreux résultats scientifiques.

Le domaine des astroparticules est lui aussi très actif, on peut citer le rôle de Tier-0 du CC-IN2P3 pour AUGER et HESS ainsi que l'implication forte au niveau des bases de données d'ANTARES et d'OPERA. La mise en service très rapide du satellite GLAST / Fermi va se concrétiser très bientôt par une utilisation intensive des ressources du CC-IN2P3.

Ce ne sont là que quelques exemples pris dans la longue liste des expériences qui bénéficient du CC-IN2P3, on pourrait également citer la QCD sur réseau, NEMO, PLANCK, etc.

Au niveau de l'ouverture du CC-IN2P3, je tiens à souligner la création d'une grille régionale Rhône-Alpes en collaboration avec le projet CIRA (Calcul Intensif en Rhône Alpes) ainsi que la mise en place en collaboration avec le CINES, d'un projet de mise en ligne et d'archivage pérennes des données SHS. Ce projet s'inscrit dans le cadre du Très Grand Équipement ADONIS.

Quels services ont été développés cette année ?

Le CC-IN2P3 doit jouer un rôle également au niveau des laboratoires. La mise en place de services mutualisés pour l'IN2P3 et l'Irfu est une priorité forte. Ceci se traduit par l'achat centralisé de licences pour des logiciels coûteux déployés dans les laboratoires (GPFS, LabView, Matlab, etc.). Nous proposons aussi un système de sauvegarde centralisée qui décharge fortement les services informatiques des laboratoires. Enfin, l'APC a migré récemment tout son mail sur les serveurs du CC-IN2P3, rejoignant ainsi quelques autres unités. Cet effort va se poursuivre et se développer dans le futur.

Quelle stratégie va adopter le CC-IN2P3 pour l'année 2009 ?

Le retard lié à l'incident du LHC oblige à revoir la stratégie de déploiement des ressources. Nous allons profiter de ce malheureux délai pour renforcer tous les points faibles et augmenter les redondances et les capacités de reprises après incidents. En particulier, un investissement conséquent va être fait pour améliorer les performances et la robustesse du système de stockage AFS.

Les limites énergétiques de la salle informatique nous obligent à traquer le moindre kW de puissance gaspillé. L'arrivée fin janvier de 252 serveurs IBM i-Dataplex montés dans des racks refroidis par eau est une innovation importante pour le CC-IN2P3. Les tests de ce type de matériel vont conditionner l'évolution future du parc de machines du CC-IN2P3.

Enfin, l'ouverture vers la région et l'implantation dans le milieu universitaire lyonnais, tout en préservant notre cœur d'activité pour le support de la physique à l'IN2P3 et à l'Irfu, est un élément clé de la stratégie future du CC-IN2P3. A l'heure ou l'organisation de la recherche en France subit de profonds changements, la réussite de cette ouverture est probablement une question de survie.

Propos recueillis par Gaëlle SHIFRIN



Le projet PetaQCD



L'étude d'une très ambitieuse machine pétaflopique consacrée à la simulation QCD sur réseau a débuté depuis quelques mois. Elle est le fruit de plusieurs années de recherches collaboratives dans ce domaine, et regroupe, au sein du projet ANR intitulé PetaQCD, un consortium pluridisciplinaire de 9 partenaires issus de l'IN2P3, du CNRS, de l'INRIA plus 2 petites entreprises privées. Ce projet devrait, sur la période 2009-2011, faire la preuve qu'une machine de ce type, capable de simuler des réseaux de l'ordre de 256x1283 à l'aide des logiciels de la famille HMC, est réalisable à un coût raisonnable avec un nombre limité de processeurs (entre 1000 et 4000). Cette preuve prendra la forme d'une maquette basée autant que faire se peut sur des éléments standards, et 2 axes particulièrement prometteurs seront explorés en priorité : l'utilisation des processeurs Cell (IBM) et celle des cartes GP-GPU (Nvidia). La machine sera donc massivement parallèle, et très probablement hétérogène, et les problèmes de communication entre processeurs, intrinsèques au problème traité plus encore qu'à l'algorithme utilisé, seront certainement l'une des clefs de l'étude.

Les solutions historiques

Il est maintenant traditionnel d'étudier et d'utiliser des machines spécifiques (par exemple les machines apeNEXT ou QCDOC) pour réaliser les calculs nécessités par la simulation de QCD sur réseau (LQCD pour Lattice QCD). Ces études ont parfois permis de déboucher sur des machines d'intérêt commercial, comme le super-calculateur IBM BlueGene, rejeton de la filière QCDOC (QCD On Chip). L'explosion du HPC (High Performance Computing), tant du côté des constructeurs que des utilisateurs – qui ont découvert depuis une dizaine d'années les économies qu'ils pouvaient réaliser en basant la conception des produits complexes sur la simulation (avionique, études pharmaceutiques ...) met actuellement à notre portée plusieurs gammes de processeurs très attirants par leur capacité et leurs performances : calculs flottants en Double Précision (64 bits) rapides, multi-cœurs mieux adaptés au traitement parallèle.

Or les simulations LQCD sont de plus en plus gourmandes en temps de calcul, et la taille des réseaux à traiter ne cesse de croître. En effet, il faut que ce réseau couvre le volume géométrique dans lequel évoluent les quarks dans le noyau (c.à.d. quelques fermis) avec une granularité suffisamment fine pour ne pas trop moyenniser les effets des fluctuations qui contiennent la physique, et on pense à une maille de 0,02 à 0,05 fm. D'où des réseaux qui atteignent 229 à 230 cellules.

Des programmes de plus en plus exigeant

Cependant, les gros calculateurs existants se plient mal aux exigences de nos programmes de simulation et à leur structure. Bien qu'ils réunissent souvent quelques-unes des caractéristiques requises, les études récentes montrent que pour l'instant aucun ne répond à toutes les conditions nécessaires qui sont :

- Une puissance de traitement en calculs flottants Double Précision élevée – on cherche à atteindre le pétaflop pour éviter qu'une simulation donnée dure plus de quelques semaines, par exemple.
- Un parallélisme massif pour pouvoir décomposer le réseau à traiter en unités suffisamment petites pour pouvoir tenir sur un processeur - on cherche à limiter le nombre total de processeurs à mettre en œuvre à quelques milliers.
- Des communications très rapides entre ces processeurs, car

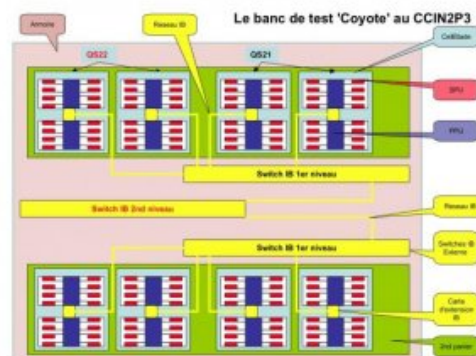
alors trop élevé. Il est rare que les 3 paramètres (puissance CPU, taille mémoire du coprocesseur, bande passante pour les accès mémoire) soient harmonieusement équilibrés, surtout quand la puissance CPU délivrée est élevée.

- Avoir un cluster qui soit le moins gourmand possible en énergie (électricité, refroidissement) et qui occupe le plus petit volume possible. Ce dernier paramètre est aussi essentiel pour la vitesse des communications par le réseau, et il est intimement lié au nombre de processeurs à mettre en œuvre ;
- On ne serait sûrement pas exhaustif si l'on mettait sous le tapis les problèmes de fiabilité, car on s'attend à ce qu'une machine (probablement hétérogène) comportant plusieurs milliers de processeurs présente au moins une panne par jour sur l'un de ses constituants. Et il est peu probable que l'un des constructeurs soit capable de nous procurer un système de checkpointing capable de faire face à ces pannes simplement et à faible coût.

Des solutions non satisfaisantes

Banc de test 'Coyote' Plusieurs types de processeurs sont actuellement candidats pour cette machine un peu spéciale, mais on ne tient pas encore le bon candidat, et il sera probablement nécessaire de revoir de fond en comble l'algorithme de simulation pour espérer obtenir les performances souhaitées. Citons pour le moment :

- Le processeur IBM Cell, qui a été initialement conçu essentiellement pour les besoins graphiques de la console de jeux Sony PS3, et qui offre une puissance CPU confortable. Cependant, son avenir n'est actuellement pas garanti (cf SC'08 à Austin-Texas). C'est un modèle avec co-processeurs embarqués (on-chip). Un banc de test, baptisé 'Coyote', financé par le CC-IN2P3, a été installé à Lyon pour y étudier en détails un petit cluster à base de Cell. Il regroupe au total 16 de ces processeurs, reliés par un réseau du type Infiniband.
- Les cartes GP-GPU du type Tesla (NVIDIA), qui sont aussi à classer dans la gamme des co-processeurs graphiques, mais cette fois off-chip, car il faut absolument les associer à un processeur maître (de type multi-cœur par exemple). Ils tiennent le haut du pavé dans le domaine du HPC ces derniers temps.
- Les processeurs multi-cœurs – on connaît déjà les quadri-cœurs qui ont tendance à devenir le standard dans nos ordinateurs de bureau un peu musclés – mais la gamme 64-cœurs de chez Intel prévue dans un ou deux ans commencerait à devenir particulièrement attirante.
- La machine IBM BlueGene (dont le modèle P à l'Ildris est aussi sous notre microscope), qui reste tout de même coûteuse, et dont certaines performances, au regard de notre application fétiche, nous laissent un peu sur notre faim.



Problématique à résoudre

l'algorithme de base requiert que les données décrivant les sites du réseau à la frontière entre chaque domaine de la décomposition soient échangées au cœur de chaque itération de la boucle de calcul, et ceci même pour le noyau de calcul qui sera le meilleur candidat pour la vectorisation de l'algorithme. Ce problème des échanges est relativement classique dans les simulations qui doivent procéder à une décomposition du domaine pour en paralléliser le traitement, comme dans les simulations météorologiques ou l'exploration sismique.

■ Une mémoire physique suffisante pour contenir les données à traiter au cœur de l'algorithme, et ceci pour les coprocesseurs vectoriels les plus aptes à nous fournir la puissance de calcul requise – si cette condition n'est pas remplie, le flux de données qui doivent être lues ou écrites pour alimenter le CPU devient

En guise de conclusion, il n'est guère possible de détailler tous les aspects délicats de chacun des systèmes mentionnés : dans la mesure où il s'agit d'un problème d'optimisation globale sur l'ensemble des paramètres (application – vitesse CPU – énergie – bande passante réseau – mémoire – prix – bande passante mémoire – fiabilité...), il nous faut concevoir des outils d'approche communs à tous ces systèmes – éventuellement en les modélisant – notamment pour l'optimisation du software (réorganisation des données et des flux de données, des boucles et des algorithmes...), pour ne pas répéter 4 fois la même étude. Et c'est bien le but du projet PetaQCD.

Gilbert Grosdidier (LAL)



n°4
Janvier
2009

La lettre IN2P3 Informatique

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



Bonnes pratiques ASR et ITIL



Les Administrateurs des Systèmes et des Réseaux (ASR) sont responsables de la bonne marche des services informatiques. Or, que ce soit au Centre de Calcul ou dans les laboratoires, l'ensemble des activités de l'Institut est de plus en plus dépendant de la disponibilité du réseau, des données, des applications. Toute panne ou dysfonctionnement est susceptible de paralyser un établissement ou d'avoir des conséquences sérieuses.

Il est donc essentiel que les ASR appliquent des méthodes de travail visant à renforcer la disponibilité et la solidité des moyens informatiques, d'en maîtriser les évolutions et d'assurer leur alignement avec les besoins des utilisateurs. La plupart des ASR en sont conscients et appliquent déjà de manière informelle des éléments d'une démarche qualité allant dans ce sens.

ITIL est un ensemble de meilleures pratiques initialement conçu et développé par le gouvernement britannique à la fin des années 1980 afin d'améliorer l'efficacité des services informatiques dans le domaine public.

Les Journées Informatiques d'Obenai ont permis d'associer dans une même session une présentation des principes d'ITIL par Geneviève Romier, une synthèse des pratiques des ASRs et des développeurs dans les laboratoires de l'IN2P3 et une table ronde débat.

Si ITIL reste une référence et un document utile pour guider une démarche qualité, il est admis que vouloir rendre un service informatique conforme à ITIL au niveau d'un laboratoire est une entreprise trop ambitieuse dans un premier temps. En revanche, adopter progressivement une culture et des bonnes pratiques proches de celles décrites par ITIL est un objectif raisonnable. Les résultats du sondage réalisé dans le cadre de l'enquête sur les bonnes pratiques ASR à l'IN2P3 montrent que certains laboratoires ont déjà mis en oeuvre des principes d'organisation et des outils. On retiendra en particulier :

- l'utilisation de systèmes de gestion des demandes des utilisateurs et les moyens de prendre connaissance des besoins des utilisateurs afin de se rapprocher de la vue centrée sur l'utilisateur préconisée par ITIL
- la maîtrise des configurations via les outils de gestion de parc et l'inventaire,
- la documentation des changements permettant à tous les intervenants de reconstituer l'historique pour un service ou une machine,
- la recherche de la pro activité par la détection de tous les indices de dysfonctionnement à travers des mécanismes de surveillance et d'alerte.

La présentation de Geneviève Romier et la synthèse de l'enquête sur les bonnes pratiques ASR sont disponibles sur le site des JI 2008 : <http://ji.in2p3.fr/JI08/Accueil.html> (cliquer sur Sixième Journées Informatiques pour accéder au programme)

Jean-Michel Barbet (SUBATECH)



n°4
Janvier
2009

La lettre IN2P3 Informatique

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



Institut des Grilles : bilan après un an d'existence



Voilà un peu plus d'un an que l'Institut des Grilles a été créé, voici donc un premier bilan de l'année écoulée. Pour rappel, l'IdG a pour mission de consolider les infrastructures existantes dans le domaine des grilles de calcul et de permettre la meilleure synergie possible entre les différents acteurs. Le développement des grilles de production tant au niveau national qu'au niveau européen s'est poursuivi avec ardeur. Signé le 18 décembre 2008 par le Directeur Général de la Recherche de l'Innovation, le protocole d'accord, entre tous les grands organismes de recherche en France (CNRS, CEA, INRIA, INRA, CPU, RENATER et Ministère) institue en particulier le comité de pilotage national dont la mission est de conduire le développement et de pérenniser de la grille de production tant sur le plan national qu'europpéen. Les premières discussions sur le plan des embauches 2009-2010 afin de stabiliser les experts recrutés jusqu'à présent sur des contrats temporaires européens, et sur un plan de création de nouveaux nœuds dans les grandes villes universitaires non encore connectées à la grille de production nationale, ont commencé fin décembre. Les différents statuts possibles de la future structure pérenne en charge de la Grille de production française ont aussi été évoqués. Sur le plan européen, le comité de pilotage a marqué son soutien fort à l'initiative EGI (European Grid Initiative) visant à pérenniser la grille de production européenne, en soutenant le dépôt d'une candidature française pour l'accueil de la structure centrale EGI.org. La candidature de Lyon a été retenue et son dépôt doit s'effectuer d'ici le 7 Janvier 2009. Le choix du site d'EGI.org sera annoncé en mars 2009. Le document décrivant en détail EGI appelé « Blueprint » est disponible sur <http://www.eu-egi.eu/blueprint.pdf>

Il devrait être approuvé le 20 janvier 2009 lors de la réunion du Policy Board, l'assemblée où tous les pays participant à EGI seront représentés. Ce calendrier serré vise à garantir une transition efficace et sans heurts entre la mise en place complète de la structure EGI et la fin du projet EGEE-III en avril 2010.

Guy Wormser, directeur de l'Institut des Grilles, et Jean-Louis Borloo, ministre de l'Ecologie, de l'Energie, du Développement durable et de l'Aménagement du territoire, à la Ville Européenne des Sciences, en novembre dernier à Paris. (© IDG/CNRS) Guy Wormser, directeur de l'Institut des Grilles, et Jean-Louis Borloo, ministre de l'Ecologie, de l'Energie, du Développement durable et de l'Aménagement du territoire, à la Ville Européenne des Sciences, en novembre dernier à Paris. (© IDG/CNRS)

Sous l'égide du Comité de pilotage national, et avec la participation de tous les organismes concernés, l'Institut des Grilles du CNRS a organisé le colloque de prospective nationale pour recenser les besoins liés aux grilles de calcul dans tous les domaines scientifiques. Il s'est tenu à Paris les 6 et 7 octobre 2008. (cf. le site du colloque). S'appuyant sur le travail de 8 groupes thématiques –chacun correspondant à une discipline scientifique- et de 6 groupes de travail transverses, ce colloque a permis de mesurer, à travers les résultats de sondage ayant touché plus de 3000 chercheurs, l'impact actuel des grilles de

calcul et l'« appétit » des scientifiques pour cet outil dans le futur. Les conclusions de ce colloque seront publiées dans un livre blanc qui sera rendu public en janvier. On peut d'ores et déjà conclure que la grille de production est devenu un outil incontournable et très apprécié dans nombre de domaines scientifiques et là où elle est encore peu utilisée et méconnue (ce qui reste encore le cas le plus fréquent d'où la nécessité d'un grand effort d'information et de formation que nous entreprendrons en 2009), elle suscite un grand intérêt et de nombreuses idées de projets ambitieux ont été mises en avant.

Les deux instances statutaires de l'Institut des Grilles, Comité de Pilotage et Conseil Scientifique, se sont respectivement réunies fin juin et début septembre 2008. Elles ont apporté un fort soutien à la politique et aux actions menées jusqu'à présent. Le Conseil Scientifique a par ailleurs recommandé d'intensifier le dialogue scientifique auprès des deux partenaires naturels de la grille de production : les supercalculateurs qui s'organisent eux-mêmes en grille (projets DEISA et PRACE) et la grille de recherche (GRID5000), en privilégiant des actions bien ciblées menées en commun. Une suggestion pour 2009 serait d'organiser une journée d'échange et de débats sur les passerelles existantes et à créer entre les grilles de production et les grilles de recherche autour de l'Ecole de Printemps de Nancy. Cécile Germain (LRI) a accepté d'explorer cette piste pour le compte de l'IdG.

Une des spécificités de l'Institut des Grilles du CNRS est de recevoir délégation du Directeur Général pour la gestion de certains programmes européens. Quatre grands programmes sont actuellement gérés par l'Institut des Grilles et sa gestionnaire très dynamique, Mélanie Pellen : EGEE-III, EGI_DS, EELA2 et EDGES. La gestion centralisée de ces programmes, grâce à l'aide de la Délégation Régionale 4 (Gif s'Yvette), permet une grande efficacité. Une nouvelle vague de projets est en cours d'approbation (EPIKH, Euroenbryome,...)

Communication et coopération sont deux axes forts de la politique de l'IDG. Sur le plan de la communication et notamment vis-à-vis du grand public, l'IDG a participé de façon très visible à l'opération « Ville Européenne des Sciences » du 14 au 16 novembre 2008 en y accueillant un nombreux public ainsi que deux ministres, Mme Valérie Pécresse et Monsieur Jean-Louis Borloo. Dans la foulée, Hélène Kérec et son équipe se sont rendus à Lyon pour participer fin novembre à la conférence majeure ICT08 organisée par la Commission Européenne pour y présenter de nombreuses démonstrations qui ont rencontré un vif succès. Côté coopération, l'IDG a décidé de se concentrer sur l'aide à l'Afrique subsaharienne avec deux portes d'entrée privilégiées : l'Afrique du Sud et le Sénégal. Une première réunion en mai à Prétoria, organisée avec l'aide de l'ambassade de France sur place et Anne Corval, représentante du CNRS pour l'Afrique, a permis de mettre en route un plan d'action comprenant de la formation d'experts africains en France (une session s'est tenue en septembre à Orsay avec la participation d'ingénieurs système d'Afrique du Sud et du Sénégal), et l'organisation avec la collaboration de nos collègues italiens de l'INFN d'une école destinée aux utilisateurs à Durban en décembre 2008. Début 2009, une importante grille sud-africaine sera mise en service et connectée à la grille européenne EGEE. En parallèle, le premier nœud de grille de l'Afrique subsaharienne a été installé à Dakar par Michel Jouvin (LAL/IDG) en juillet 2008 avec l'aide du programme HP-Unesco qui a fourni le matériel nécessaire à l'Université Cheick Anta Diop. 2009 sera donc l'année de la consolidation de cette politique avec en particulier la préparation d'un projet européen EUAfrica pour développer les applications scientifiques africaines, la participation de scientifiques africains aux grands projets internationaux, le renforcement des nœuds de grille existants et la création de nouveaux nœuds (prochain objectif : Brazzaville !).



Guy Wormser, directeur de l'Institut des Grilles, et Jean-Louis Borloo, ministre de l'Ecologie, de l'Energie, du Développement durable et de l'Aménagement du territoire, à la Ville Européenne des Sciences, en novembre dernier à Paris. (© IDG/CNRS)

L'automne 2008 aurait dû être marqué par un grand événement : le déploiement à pleine puissance de la grille mondiale W-LCG destinée à l'exploitation scientifique des données issues de l'accélérateur LHC. La panne sévère du LHC survenue le 19 septembre juste quelques jours après sa brillante mise en service ne l'a hélas pas permis. Toutefois, la grille W-LCG a prouvé lors de nombreux tests et exercices en grandeur réelle qu'elle est tout à fait prête à ce grand démarrage maintenant repoussé à l'été 2009.

Pour terminer ce panorama de nos activités 2008, je me réjouis de la grande confiance que l'AFM, IBM et le CNRS vont marquer à l'Institut des Grilles en lui « remettant les clés » de la grille Décryphon, une grille originale basée sur le logiciel DIET développé par Frédéric Desprez (LIP-ENS) et dédiée à la recherche sur les maladies génétiques.

L'année 2008 a donc été très riche en activités diverses, principalement tournées vers les grilles de production du fait de la nécessité de mettre en place tant en France qu'en Europe des structures pérennes et solides. C'est sur cette fondation forte que nous pourrons en 2009 développer nos actions de dialogue et d'animations scientifiques vers la communauté des chercheurs en informatique travaillant sur les grilles, notamment grâce aux premiers résultats de l'Observatoire des Grilles qui se met en place de façon satisfaisante.

Guy Wormser, directeur de l'Institut des Grilles



Faut-il passer d'un système de gestion de versions centralisé à un système décentralisé



Tous ceux qui manipulent des fichiers, en souhaitant conserver l'historique de leur travail, trouveront un intérêt à s'appuyer sur un VCS. L'un de leurs principaux intérêts est de permettre le développement mutualisé, avec une gestion des conflits. Pourtant en pratique, il s'avèrera particulièrement utile, même pour un utilisateur isolé.

Son choix parmi les nombreux outils disponibles (centralisé ou réparti) aura nécessairement un impact fort sur le modèle de développement suivi.

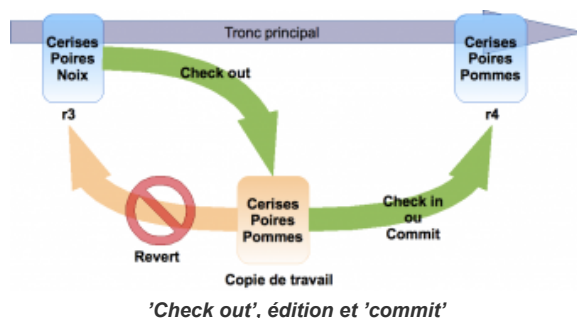
Qu'est-ce qu'un 'Système de gestion de versions' ?

Un VCS (Version System Control) est un outil informatique capable de gérer les évolutions et l'historique d'un ensemble de fichiers et de ressources. Stockant les informations dans un dépôt, il permet de récupérer toutes les versions intermédiaires, ainsi que les différences entre les versions. Il permet à plusieurs développeurs de travailler sur un même projet et effectue la fusion des modifications non conflictuelles en protégeant contre celles qui le sont.

Les mots clés de base :

- Le **dépôt** ou **référentiel** ou encore **'repository'** est l'élément de référence contenant l'ensemble des fichiers et informations associées y compris l'historique
- L'**espace de travail** ou **'working area'** ou encore **'working set'** correspond à l'espace dans lequel on développe et modifie les fichiers
- Une **révision** est l'identifiant d'un fichier versionné
- **HEAD** : la dernière révision dans le dépôt
- **'check out'** est l'opération de chargement, copie de fichier(s) depuis le dépôt dans l'espace de travail.
- **ajout** - **'add'** ajoute le(s) fichier(s) dans le mécanisme de gestion de version
- **'check in'** ou **'commit'** envoie le(s) fichier(s) - si modifié(s) - dans le dépôt ; le fichier acquiert une nouvelle révision et cette dernière version peut être rechargée par d'autres utilisateurs
- Un message de **'commit'** est un message décrivant la modification enregistrée dans le dépôt
- historique - **'log'** ou **'Changelog'** - donne à la liste des modifications
- **'update'** ou **'synch'** synchronise les fichiers locaux avec les autres développeurs ou le dépôt de référence
- **'revert'** permet de revenir sur les dernières modifications locales

Les notions de base communes à tous les VCS :



- La gestion des copies et déplacements de fichiers a été totalement revue et améliorée (déplacement direct du fichier dans l'espace de travail)
- Subversion 1.5 est beaucoup plus rapide (optimisation complète)
- Apparition du **'sparse checkout'** : on peut ne récupérer qu'une partie de projet, et non nécessairement l'intégralité
- Implémentation du **'merge tracking'** : la fusion de 2 branches devient beaucoup plus facile et ne demande plus de gestion manuelle des numéros de révision ... (voir *'svn help mergeinfo'* et *'svn help merge'*)

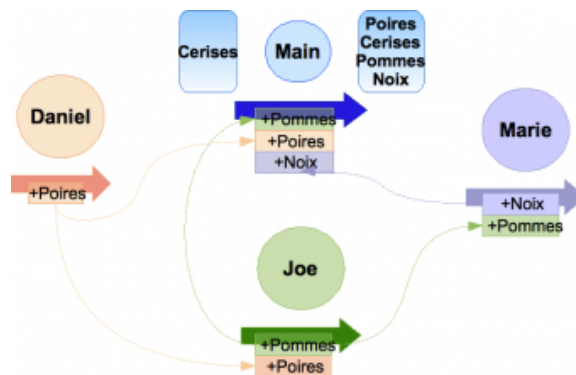
Subversion 2

Prochaine évolution majeure, elle s'inspirera d'un certain nombre de caractéristiques apportées par les DVCS :

- **'offline commits'** : permettant de gérer la granularité de ses commits dans des branches locales
- La gestion de branches locales synchronisées ultérieurement avec le dépôt central
- La centralisation de la gestion des **'metadata'** en un seul endroit de la copie locale (plus de **.svn** à chaque niveau de la hiérarchie du projet)
- Elle devrait conserver son interface très simple

B- Le modèle réparti

Contrairement au modèle centralisé, il n'impose pas techniquement de dépôt de référence ou privilégié ; chaque développeur travaille avec son dépôt qu'il synchronise avec ceux des co-développeurs et, selon le modèle de développement, avec un dépôt de référence et/ou de publication. Les principaux cités actuellement sont sans doute Git, Mercurial, Bazaar mais il en existe beaucoup d'autres ! Cette famille de VCS a été portée par les développements Open Source et correspond à l'ère du laptop dans le train ...



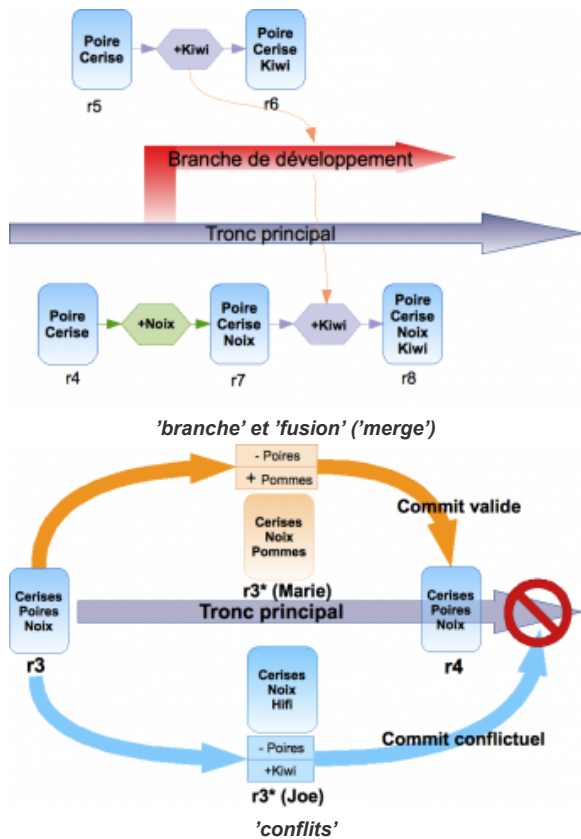
Ses points forts :

- pas de serveur à installer et à gérer
- travail déconnecté - accès rapide aux dépôts sans réseau
- modèle de développement très souple
- utilisation de branches privées très facilitée

Parmi les VCS décentralisés du moment :

Git

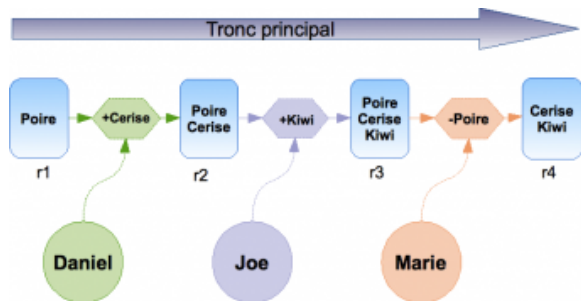
- Écrit essentiellement en C, par Linus Torvalds ...
- Très rapide, robuste et accompagné d'une documentation assez riche
- Il peut être utilisé comme un super client (bi-directionnel) de SVN - voir *svn2git*



Quel type de VCS choisir ?

A- Le modèle centralisé

Il est caractérisé par un dépôt (repository) privilégié géré par un serveur et une politique stricte d'accès ; chaque développeur travaille sur une copie et synchronise ses évolutions avec le dépôt central, le serveur détectant les conflits. CVS et son successeur Subversion (SVN) sont deux exemples d'outils très utilisés par notre communauté.



Ses points forts :

- la simplicité et la familiarité du modèle
- la maturité et la robustesse des implémentations (surtout SVN)
- le grand nombre d'outils de gestion et d'interfaces, graphiques ou non, disponibles

Subversion :

Principal compétiteur dans notre environnement, c'est le successeur, par conception même, de l'historique CVS. Il bénéficie d'une énorme base installée et d'un grand nombre d'hébergements, CERN y compris. Il est accompagné d'une excellente documentation.

Subversion 1.5

En plus de nombreux bugs corrigés, SVN 1.5 apporte un certain nombre de nouvelles fonctionnalités, dont certaines très attendues :

- Adopté par de nombreux projets (Kernel Linux, Android (Google), Ruby on Rails, VLC, ...)
- Mais complexe et présentant une interface compliquée et moins élégante que celle de SVN
- Mais pas encore de version Windows native

Mercurial

- Écrit en Python
- Interface simple et élégante
- Bonnes performances
- Bonne documentation
- Choisi par OpenSolaris, Mozilla

Autres ...

- **Bazaar** (choisi par MySQL)
- **Darcs**
- **Monotone**

Pourquoi et comment choisir ?

Le modèle centralisé est séduisant :

Le modèle est assez simple ; on le connaît, on le pratique, on le maîtrise bien Subversion est pratiquement fourni par défaut et il existe de nombreuses intégrations avec les principaux éditeurs et IDE Enfin, Subversion évolue en intégrant l'expérience acquise avec les DVCS

Le modèle réparti est également séduisant :

Il n'y a plus de serveur à installer et à maintenir et on s'affranchit largement des contraintes du réseau dans le travail quotidien Il permet de démarrer plus facilement et plus soupagement une contribution à un projet L'implémentation de mécanismes performants de fusion entre branches simplifie la mise en oeuvre de nouveaux modèles de développement ('workflow') facilitant le travail collaboratif et la remontée de correctifs ou d'améliorations Une utilisation 'dégradée' très proche de celle du modèle centralisé et avec un modèle de développement similaire est possible pour minimiser la marche à franchir ... On peut même utiliser certains VCS répartis comme de super clients sur Subversion, les dépôts répartis jouant le rôle d'espace de travail.

Les craintes face au nouveau paradigme :

Dans le modèle centralisé, il n'y a qu'un seul dépôt. Par défaut, tout le travail est public et il est publié dans le seul dépôt partagé. Il est techniquement difficile de travailler de façon privée : cela signifie gérer un gros 'patch', sans historique et créer un 'fork' est découragé. Dans le modèle réparti, chaque développeur démarre de façon tout à fait légitime un 'fork' du projet. En fait, le dépôt 'officiel' est purement conventionnel ; par défaut tout le travail du développeur est privé et caché ; la publication ne peut être que volontaire et nécessite un certain effort ! La facilité à créer des branches peut faciliter, voire encourager, le développement 'obscur', sans échanges dans l'équipe et rendant le 'code review' difficile sinon impossible. Si on considère que l'absence ou le manque de communication au sein d'un projet est un problème de cohésion et de gestion d'équipe et non pas un problème technique dû aux outils, y compris le système de gestion de versions, alors la souplesse et la richesse de fonctionnalités qu'apportent les VCS répartis peuvent valoir l'effort d'adaptation.

Conclusion Aujourd'hui, les diverses implémentations du nouveau modèle réparti semblent matures au point qu'un grand nombre de projets ('open source') ou d'organisations ont déjà effectué la migration.

Le modèle réparti de gestion de versions apporte certainement une plus grande souplesse de travail. On peut vouloir attendre Subversion 2.0 ou bien, dès maintenant, utiliser un VCS réparti comme client d'un serveur SVN.

Antoire Pérus et Laurent Garnier (LAL)